Descent on Elliptic Curves

Martin Leslie

An essay submitted in partial fulfillment of
the requirements for the degree of
B.Sc. (Honours)

Department of Mathematics
University of Queensland

July 2007

Contents

Acknowledgements

Thank you to my supervisor Dr Victor Scharaschkin, without whose knowledge and patience this thesis would never have been completed. Thank you to my family, friends and housemates who have always supported me.

Thank you to everyone in the maths honours room, especially Greg for some useful homological and topological discussions and to Greg, Leesa, Pete, Geoff and Andrew for many pleasurable hours at the Red Room.

Introduction

We consider elliptic curves as part of Diophantine Geometry, a combination of Geometry and Number Theory. In Number Theory we are interested in finding solutions to equations in rational numbers. However these equations describe geometric objects and studying them from this point of view has proven to be fruitful.

Elliptic curves are the next step up from conics, a well understood field of study (conics are defined by quadratic equations, elliptic curves are defined by cubics). The difficulties in taking this step can be seen when we consider the Hasse–Minkowski Theorem. This says that a quadratic polynomial in two variables with rational coefficients has a rational solution if and only if it has real solutions and solutions in $\mathbb{Q}_p$ for every prime p. This correspondence is sometimes called the local-global principle — to find information about an equation globally, consider it locally at each prime.

However the analogue of the Hasse–Minkowski Theorem does not hold for cubic equations. Despite this local methods are still very useful but we are now interested in cases where this kind of method fails.

An elliptic curve can be thought of as the points on a curve $E\colon y^2 = x^3 + ax + b$. A geometric group law can be defined on these points turning them into an abelian group. As we are doing number theory we are interested in studying the points with rational coordinates, denoted $E(\mathbb{Q})$. The Mordell–Weil theorem says that this group is finitely generated. By the fundamental theorem of finitely generated abelian groups this means that

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

where $E_{\text{tors}}(\mathbb{Q})$ is the set of points of finite order. We can fairly easily find $E_{\text{tors}}(\mathbb{Q})$. The real challenge is finding r, the *rank*. This process of finding r is called descent and is the main focus of this thesis.

We can show from the equation just given that finding r is equivalent to finding $E(\mathbb{Q})/mE(\mathbb{Q})$ for some positive integer m. We can find an exact sequence

$$0 \longrightarrow E(\mathbb{Q})/mE(\mathbb{Q}) \longrightarrow S^{(m)} \longrightarrow Ш[m] \longrightarrow 0.$$

where $S^{(m)}$ is computable. The thing that prevents us from calculating $E(\mathbb{Q})/mE(\mathbb{Q})$ is $Ш[m]$ which measures the failure of the local–global principle. Thus we are interested in understanding $S^{(m)}$, the *Selmer group*, and Ш, the *Tate–Shafarevich group*. Much of our work is finding useful ways to think about and work with these groups.

We assume basic knowledge of abstract algebra, especially fields and abelian groups. Exposure to Galois theory and Algebraic number theory would be useful although we provide appendices that outline the main definitions and results we need. There is a small amount of topology used although this is not necessary if the reader is willing to take some results on faith.

The first two chapters give the necessary background in Galois cohomology and algebraic curves. In the third chapter we introduce elliptic curves and study their geometry. The next two chapters give us the background we require in p–adic numbers and local methods for studying elliptic curves. Finally in chapter six we study descent, as we have discussed above. Then chapter seven completes the proof of the Mordell–Weil theorem by looking at heights, a concept of 'size' of points on an elliptic curve. In chapter eight we explore two practical methods to calculate the rank of elliptic curves in a special case.

Chapter 1

Some group cohomology

Before moving onto the more geometric aspects of our study we first need some algebra. This section will be used occasionally throughout the thesis but is used mostly in the section about twists and later on when we discuss descent.

Homological algebra began in algebraic topology but is now used in nearly all parts of algebra and number theory. The basic object is the exact sequence.

Definition 1.1. *An* exact sequence *is a set of objects (for example abelian groups) with morphisms (for us group homomorphisms) between them*

$$\cdots \longrightarrow A_{i-1} \xrightarrow{f_{j-1}} A_i \xrightarrow{f_j} A_{i+1} \longrightarrow \cdots$$

such that $\operatorname{Im} f_{j-1} = \operatorname{Ker} f_j$ *for all* j.

A short exact sequence *is an exact sequence of the form*

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0\ .$$

Note that in a short exact sequence f is injective, g is surjective and $\operatorname{Im} f = \operatorname{Ker} g$

1.1. Finite group cohomology – H^0 and H^1

We now give the basic definitions of the two cohomology groups we are interested in, following **[12]**, appendix B.

Note that since we are working with abelian groups we write our groups additively; however there will be occasions when we need to write our groups multiplicatively.

Definition 1.2. *Let* G *be a finite group that acts on an abelian group* M*. Denote the action of* $\sigma \in G$ *on* $m \in M$ *by* $m \mapsto m^\sigma$*. Then* M *is a* (right) G–module *if for all* $m, m' \in M$ *and* $\sigma, \tau \in G$,

$$m^1 = m \quad (m + m')^\sigma = m^\sigma + m'^\sigma \quad (m^\sigma)^\tau = m^{\sigma\tau}.$$

Definition 1.3. *If* M *and* N *are* G*–modules, a* G–homomorphism *is a homomorphism* $\phi\colon M \to N$ *of abelian groups commuting with the*

action of G; that is, for all $m \in M$ and $\sigma \in G$.

$$\phi(m^\sigma) = \phi(m)^\sigma.$$

Definition 1.4. *The* 0^{th}–cohomology group *of the G–module M is the group of G–invariant elements of M*

$$H^0(G, M) = M^G = \{m \in M \mid m^\sigma = m \text{ for all } \sigma \in G\}.$$

We define H^1 as the quotient of two other groups.

Definition 1.5. *Let M be a G–module. The* group of 1–cocycles from G to M, *denoted $Z^1(G, M)$, is the group of maps $\xi \colon G \to M$ under addition satisfying the cocycle condition*

$$\xi(\sigma\tau) = \xi(\sigma)^\tau + \xi(\tau).$$

The group of 1–coboundaries from G to M, *denoted $B^1(G, M)$, is the group of maps $\xi \colon G \to M$ under addition where there exists an $m \in M$ such that, for all $\sigma \in G$*

$$\xi(\sigma) = m^\sigma - m.$$

We can check that $B^1(G, M) \subseteq Z^1(G, M)$ and then we define

$$H^1(G, M) = \frac{Z^1(G, M)}{B^1(G, M)}.$$

We will use $\{x\}$ to denote the coset $x + B^1(G, M)$ throughout this thesis.

Example 1.6. If the action of G on M is trivial, that is $m^\sigma = m$ for all $\sigma \in G$, then we have $H^0(G, M) = M^G = M$. Also, cocycles are simply homomorphisms and the only coboundary is the zero map. Thus $H^1(G, M) = \mathrm{Hom}(G, M)$.

Example 1.7. Let L/K be a finite Galois extension with Galois group $G = \mathrm{Gal}(L/K)$. Then we have the *Galois action* of G on L defined by, for all $x \in L$ and $\sigma \in G$,

$$x^\sigma = \sigma^{-1}(x).$$

Then we can check that for $x, y \in L$ and $\sigma, \tau \in G$,

a) $x^1 = 1^{-1}(x) = x$.
b) $(x + y)^\sigma = \sigma^{-1}(x + y) = \sigma^{-1}(x) + \sigma^{-1}(y) = x^\sigma + y^\sigma$.
c) $x^{\sigma\tau} = (\sigma\tau)^{-1}(x) = \tau^{-1}(\sigma^{-1}(x)) = \tau^{-1}(x^\sigma) = (x^\sigma)^\tau$.

Thus L is a G–module.

Example 1.8. If M and N are G–modules then the set of G–module homomorphism from M to N is a G–module under the action $f \mapsto f^\sigma$ where f^σ is defined by

$$f^\sigma(m) = f(m^{\sigma^{-1}})^\sigma$$

for all $m \in M$. To check this we see that

a) $f^1(m) = f(m^{1^{-1}})^1 = f(m)$ for all $m \in M$ so $f^1 = f$.
b) $(f+g)^\sigma(m) = (f+g)(m^{\sigma^{-1}})^\sigma = f(m^{\sigma^{-1}})^\sigma + g(m^{\sigma^{-1}})^\sigma = f^\sigma(m) + g^\sigma(m)$ for all $m \in M$ so $(f+g)^\sigma = f^\sigma + g^\sigma$.
c) $f^{\sigma\tau}(m) = f(m^{(\sigma\tau)^{-1}})^{\sigma\tau} = (f(m^{\tau^{-1}\sigma^{-1}})^\sigma)^\tau = f^\sigma(m^{\tau^{-1}})^\tau = (f^\sigma)^\tau(m)$ for all $m \in M$ so $f^{\sigma\tau} = (f^\sigma)^\tau$.

1.2. Useful results

Group cohomology allows us to express the ideas of descent in their most efficient form. Many of the concepts in this thesis were developed before the machinery of group cohomology was invented but we feel that this is the best way to show them.

The next theorem shows the power of this approach, going from a short exact sequence to a long exact sequence with lots of new information.

Theorem 1.9. *Let*

$$0 \longrightarrow P \xrightarrow{\phi} M \xrightarrow{\psi} N \longrightarrow 0$$

be an exact sequence of G–modules. Then there is a long exact sequence

$$0 \longrightarrow H^0(G,P) \longrightarrow H^0(G,M) \longrightarrow H^0(G,N)$$

$$\xrightarrow{\delta} H^1(G,P) \longrightarrow H^1(G,M) \longrightarrow H^1(G,N)$$

where δ is the connecting homomorphism.

Proof. First note that if $p \in P^G$ then $\phi(p)^\sigma = \phi(p^\sigma) = \phi(p)$. Thus $\phi(p) \in M^G$. Similarly if $m \in M^G$ then $\psi(m) \in N^G$ so we can define our maps in the top row just to be restrictions of ϕ and ψ, denoted $\phi|_{P^G} = \phi'$ and $\psi|_{M^G} = \psi'$.

For the maps on the bottom row define

$$\phi'' \colon H^1(G,P) \to H^1(G,M) \quad \text{by} \quad \{\xi\} \mapsto \{\phi \circ \xi\}$$

and

$$\psi'' \colon H^1(G,M) \to H^1(G,N) \quad \text{by} \quad \{\xi\} \mapsto \{\psi \circ \xi\}.$$

The connecting homomorphism is defined by

$$\delta\colon n \mapsto \{\sigma \mapsto \phi^{-1}(m^\sigma - m)\}$$

where we choose $m \in M$ such that $\psi(m) = n$. Now $n \in N^G$ so $n^\sigma = n$, implying $\psi(m)^\sigma = \psi(m)$. Now ψ is a G–module homomorphism so $\psi(m^\sigma - m) = 0$ so $m^\sigma - m \in \operatorname{Ker}\psi = \operatorname{Im}\phi$ and we can apply ϕ^{-1} as in the definition. To check δ is well defined note that if there exists $m' \in M$ such that $\psi(m') = n = \psi(m)$ then $\psi(m - m') = 0$. Thus $m - m' \in \operatorname{Ker}\psi = \operatorname{Im}\phi$ and therefore there exists $p \in P$ such that $m - m' = \phi(p)$. Then

$$\begin{aligned}\phi^{-1}(m^\sigma - m) - \phi^{-1}(m'^\sigma - m') &= \phi^{-1}((m - m')^\sigma - (m - m')) \\ &= \phi^{-1}((\phi(p)^\sigma - \phi(p)) \\ &= p^\sigma - p\end{aligned}$$

meaning that

$$\{\sigma \mapsto \phi^{-1}(m^\sigma - m\} = \{\sigma \mapsto \phi^{-1}(m'^\sigma - m'\}.$$

We rewrite our diagram

$$0 \longrightarrow P^G \xrightarrow{\phi'} M^G \xrightarrow{\psi'} N^G$$

$$\xrightarrow{\delta} H^1(G, P) \xrightarrow{\phi''} H^1(G, M) \xrightarrow{\psi''} H^1(G, N)$$

and give the proof of exactness in five parts:

a) Exactness at P^G.

Injectivity of ϕ' follows from injectivity of ϕ.

b) Exactness at M^G.

For $\operatorname{Im}\phi' \subseteq \operatorname{Ker}\psi'$ note that $\psi' \circ \phi' = 0$ since $\psi \circ \phi = 0$. For the other inclusion, if $m \in \operatorname{Ker}\psi'$ then $m \in \operatorname{Ker}\psi = \operatorname{Im}\phi$ so $m = \phi(p)$ for some $p \in P$. We need to show that $p \in P^G$. But $m^\sigma = m$ so $\phi(p)^\sigma = \phi(p)$ and since ϕ is a G–module homomorphism $\phi(p^\sigma) = \phi(p)$. Then by the injectivity of ϕ, $p^\sigma = p$ and thus $p \in P^G$ so $m \in \operatorname{Ker}\psi'$.

c) Exactness at N^G.

Note that for $m \in M^G$, $(\delta \circ \psi')(m) = \{\sigma \mapsto \phi^{-1}(m^\sigma - m)\}$. Now $m^\sigma - m = 0$ since $m \in M^G$ so $(\delta \circ \psi')(m) = \{0\}$. For $\operatorname{Ker}\delta \subseteq \operatorname{Im}\psi'$ take $n \in \operatorname{Ker}\delta \subseteq N^G$ and find m such that $\psi(m) = n$. Then $\delta(n) = \{\sigma \mapsto \phi^{-1}(m^\sigma - m)\} = \{0\}$ so $m^\sigma - m = \phi(p^\sigma - p)$ for some $p \in P$. Now form $m' = m + \phi(p)$ and calculate $\psi(m') = \psi(m) + \psi \circ \phi(p) = \psi(m) = n$. From

above $(m-\phi(p))^\sigma = (m-\phi(p))$ for all $\sigma \in G$ so $m' \in M^G$ and $n = \psi(m') \in \operatorname{Im}\psi'$.

d) Exactness at $H^1(G,P)$.

For $n \in N^G$, $(\phi'' \circ \delta)(n) = \{\sigma \mapsto m^\sigma - m\} = \{0\}$. For $\operatorname{Ker}\phi'' \subseteq \operatorname{Im}\delta$, take $\{\xi\} \in \operatorname{Ker}\phi'' \subseteq H^1(G,P)$. This means that for some $m \in M$, $\phi(\xi(\sigma)) = m^\sigma - m$ for all $\sigma \in G$. So we can see that $\psi(m)^\sigma = \psi(m^\sigma) = \psi(\phi(\xi(\sigma))) + \psi(m) = \psi(m)$ giving us $\psi(m) \in N^G$. Then $\{\xi\} = \{\sigma \mapsto \phi^{-1}(m^\sigma - m)\} = \delta(\psi(m)) \in \operatorname{Im}\delta$.

e) Exactness at $H^1(G,M)$.

To see that $\operatorname{Im}\phi'' \subseteq \operatorname{Ker}\psi''$ note that $\psi \circ \phi = 0$ implies that $\{\psi \circ \phi \circ \xi\} = \{0\}$. For the other inclusion take $\{\xi\} \in \operatorname{Ker}\psi'' \subseteq H^1(G,M)$. Then there exists an $n \in N$ such that $\psi(\xi(\sigma)) = n^\sigma - n$ for all $\sigma \in G$. Now ψ is surjective so there exists $m \in M$ such that $n = \psi(m)$ and then $\psi(\xi(\sigma)) = \psi(m)^\sigma - \psi(m)$ so $\psi(\xi(\sigma) - m^\sigma + m) = 0$. Now define $\xi' \colon G \to M$ by $\xi'(\sigma) = \xi(\sigma) - m^\sigma + m$ for all $\sigma \in G$. We can see that $\{\xi'\} = \{\xi\}$ and $\psi(\xi'(\sigma)) = 0$ for all $\sigma \in G$. Thus $\xi'(\sigma) \in \operatorname{Ker}\psi = \operatorname{Im}\phi$ so we can define $\xi'' \colon G \to P$ by $\xi''(\sigma) = \phi^{-1}(\xi'(\sigma))$. Then since ϕ is injective ξ'' inherits all of ξ''s properties and $\{\xi''\} \in H^1(G,P)$. Then $\{\xi\} = \{\xi'\} = \{\phi \circ \xi''\} \in \operatorname{Im}\phi''$.

□

Our next theorem allows us to obtain a nice correspondence between related cohomology groups.

Definition 1.10. *Let H be a subgroup of G. Then a G–module M is clearly a H–module. Define a* restriction homomorphism

$$\operatorname{Res}\colon H^1(G,M) \to H^1(H,M) \quad \textit{by}$$

$$\operatorname{Res}\colon \{\xi\} \mapsto \{\xi|_H\}.$$

If $H \lhd G$ then M^H is a G/H module under the action $m^{\sigma H} = m^\sigma$ where the right hand side is the action of G on M. Using the following diagram

$$\begin{array}{ccc} G & \longrightarrow & M \\ \downarrow{\scriptstyle i} & & \uparrow{\scriptstyle j} \\ G/H & \longrightarrow & M^H \end{array}$$

where $i\colon x \mapsto xH$ and j is the obvious inclusion we can define an inflation homomorphism

$$\operatorname{Inf}\colon H^1(G/H, M^H) \to H^1(G,M) \quad \textit{by}$$

$$\operatorname{Inf}\colon \{\xi\} \mapsto \{j \circ \xi \circ i\}.$$

Theorem 1.11. *Let M be a G–module and $H \lhd G$. Then the following sequence is exact.*

$$0 \longrightarrow H^1(G/H, M^H) \xrightarrow{\text{Inf}} H^1(G, M) \xrightarrow{\text{Res}} H^1(H, M).$$

Proof. For exactness at $H^1(G/H, M^H)$ we require Inf to be injective, that is if $\text{Inf}\{\xi\} = \{0\}$ then $\{\xi\} = 0$. Say $\xi\colon G/H \to M^H$ is a cocycle such that $\text{Inf}\{\xi\} = \{0\}$. Thus there exists $m \in M$ such that

$$(j \circ \xi \circ i)(\sigma) = m^\sigma - m \quad \text{for all } \sigma \in G.$$

Since i maps $\sigma \to \sigma H$ we have for all $\tau \in H$

$$(j \circ \xi \circ i)(\tau\sigma) = m^{\tau\sigma} - m = (j \circ \xi \circ i)(\sigma) = m^\sigma - m.$$

Thus $m^{\tau\sigma} - m^\sigma = 0$ for all $\sigma \in G$ and $\tau \in H$ so $m^\tau = m$ for all $\tau \in H$. Therefore $m \in M^H$ and for all $\sigma H \in G/H$, $\xi(\sigma H) = m^\sigma - m = m^{\sigma H} - m$ so $\{\xi\} = \{0\}$.

For exactness at $H^1(G, M)$ we first show that $\text{Res} \circ \text{Inf} = 0$. Say we have a cocycle $\xi\colon G/H \to M^H$. Then for all $\tau \in H$, $i(\tau) = H$ so $\xi(i(\tau)) = \xi(H) = 0$. Thus $\text{Res} \circ \text{Inf}\{\xi\} = \{0\}$.

Next we show that $\text{Ker}(\text{Res}) \subseteq \text{Im}(\text{Inf})$. Suppose we have a cocycle $\xi\colon G \to M$ such that $\text{Res}\{\xi\} = \{0\}$. Then there exists $m \in M$ such that $\xi(\tau) = m^\tau - m$ for all $\tau \in H$. Now define a cocycle $\xi'\colon G \to M$ by $\xi'(\sigma) = \xi(\sigma) - m^\sigma + m$ for all $\sigma \in G$. We can see that $\{\xi'\} = \{\xi\}$ and that $\xi'(\tau) = 0$ for all $\tau \in H$. Then for $\sigma \in G$ and $\tau \in H$ we have

$$\xi'(\tau\sigma) = \xi'(\tau)^\sigma + \xi'(\sigma) = \xi'(\sigma).$$

So we can see that $\xi'(\sigma)$ depends only on the class of $\sigma \in G/H$. So we can define $\xi''\colon G/H \to M^H$ by $\xi''(\sigma H) = \xi'(\sigma)$. To see ξ'' maps onto M^H we need $\xi''(\sigma H)^\tau = \xi''(\sigma H)$ for all $\tau \in H$. We have

$$\begin{aligned} \xi''(\sigma H) &= \xi''(\sigma\tau H) = \xi'(\sigma\tau) \\ &= \xi'(\sigma)^\tau + \xi(\tau) = \xi'(\sigma)^\tau \\ &= \xi''(\sigma H)^\tau. \end{aligned}$$

So finally $\text{Inf}\{\xi''\} = \{\xi'\} = \{\xi\}$ so $\text{Ker}(\text{Res}) \subseteq \text{Im}(\text{Inf})$.

□

We wish to study Galois cohomology, especially the action of $G = \text{Gal}(\bar{K}/K)$. In this section we work in finite Galois extensions and later we shall pass to the limit. First we require a well known lemma that we have taken from **[9]**.

Lemma 1.12 (Dedekind's Lemma)**.** *Let G be a group and K a field. Then any finite set of elements in* $\text{Hom}(G, K^*)$ *is linearly independent*

over K. That is, if for $c_i \in K$ and $\xi_i \colon G \to K^*$ homomorphisms, $\sum c_i\xi_i = 0$ (the zero function) then $c_i = 0$ for all i.

Proof. By contradiction. Let $\{\xi_1, \ldots, \xi_n\}$ be a minimal linearly dependent set when $\xi_i \in G^*$ are given. So there exists $c_i \in K$ not all zero such that

$$\sum_i c_i\xi_i(\sigma) = 0 \text{ for all } \sigma \in G.$$

If $c_i = 0$ then we can omit ξ_i from our linearly dependent set, contradicting minimality. Thus we can assume $c_i \neq 0$ for all i. One homomorphism $G \to K^*$ cannot be dependent so we may assume that $n \geq 2$. So $\xi_1 \neq \xi_2$, that is, there exists $\tau \in G$ such that $\xi_1(\tau) \neq \xi_2(\tau)$.

Multiplying our earlier equation by $\xi_1(\tau)$ we get

$$\sum_i (c_i\xi_1(\tau))\xi_i(\sigma) = 0 \text{ for all } \sigma \in G.$$

Since ξ_i is a homomorphism this implies

$$\sum_i c_i\xi_i(\tau\sigma) = \sum_i (c_i\xi_i(\tau))\xi_i(\sigma) = 0 \text{ for all } \sigma \in G.$$

Subtracting our two sums we have

$$\sum b_i\xi_i(\sigma) = 0 \text{ for all } \sigma \in G \text{ where } b_i = c_i(\xi_1(\tau) - \xi_i(\tau)).$$

Now $b_2 \neq 0$ so this is a nontrivial linear dependence between the ξ_i involving at most $n-1$ ξ_i's (since $b_1 = 0$). This contradicts minimality. □

Theorem 1.13. *Let L/K be a finite Galois extension and let $G = \mathrm{Gal}(L/K)$. Then*

$$H^1(G, L^*) = 0.$$

Proof. Take $\xi \colon G \to L^*$ such that $\xi(\sigma\tau) = \xi(\sigma)^\tau\xi(\tau)$. We require $a \in L^*$ such that $\xi(\sigma) = a^\sigma/a$. Following **[10]** we choose a nonzero $c \in L$ and form

$$b = \sum_{\tau \in G} \xi(\tau)c^\tau$$

Now $b \neq 0$ since then Dedekind's lemma would imply that $c^\tau = 0$ for all $\tau \in G$. Then

$$b^\sigma = \sum_{\tau \in G} \xi(\tau)^\sigma c^{\tau\sigma} = \sum_{\tau \in G} \frac{\xi(\tau\sigma)}{\xi(\sigma)} c^{\tau\sigma} = \xi(\sigma)^{-1} \sum_{\tau \in G} \xi(\tau\sigma)c^{\tau\sigma}.$$

Now multiplication by σ merely permutes the elements of G so

$$\sum_{\tau \in G} \xi(\tau\sigma)c^{\tau\sigma} = \sum_{\tau \in G} \xi(\tau)c^\tau = b.$$

Thus

$$b^\sigma = \xi(\sigma)^{-1} b \text{ so } \xi(\sigma) = \frac{(b^{-1})^\sigma}{b^{-1}}.$$

So $H^1(G, L^*) = 0$. □

1.3. Galois cohomology

Let K be a perfect field, $\bar{K}$ an algebraic closure and $G = \mathrm{Gal}(\bar{K}/K)$. To do group cohomology with the infinite group G we need new definitions to take into account its topology. Recall that the open subgroups are those fixing some finite extension of K. For the material on infinite Galois theory and limits in algebra that we require see [**9**].

Definition 1.14. *A* discrete G–module *is an abelian group M on which G acts such that the action is continuous for the profinite topology on G and the discrete topology on M.*

In [**7**] Milne claims that this condition is equivalent to requiring that

$$M = \cup_H M^H \quad \text{where } H \text{ is open in } G.$$

Example 1.15. The group $\bar{K}^*$ with the Galois action of G is a discrete G–module because

$$\bar{K}^* = \cup L^*$$

where the union is over the finite extensions L of K contained in $\bar{K}$.

We can define H^0 exactly as before but the H^1 definition must be adjusted.

Definition 1.16. *Let M be a G–module. Then define*

$$H^0(G, M) = M^G = \{m \in M \mid m^\sigma = m \text{ for all } \sigma \in G\}.$$

The group of continuous 1–cocycles from G to M, *denoted $Z^1_{\mathrm{cont}}(G, M)$, is the group of continuous maps $\xi \colon G \to M$ satisfying the cocycle condition*

$$\xi(\sigma\tau) = \xi(\sigma)^\tau + \xi(\tau).$$

We can check that $B^1(G, M) \subseteq Z^1_{\mathrm{cont}}(G, M)$ and then we define

$$H^1(G, M) = \frac{Z^1_{\mathrm{cont}}(G, M)}{B^1(G, M)}.$$

We repeat Milne's assertion that

$$H^1(G, M) = \varinjlim H^1(G/H, M^H)$$

where H runs through the normal subgroups of G.

The proof of Theorems 1.9 and 1.11 are identical for G an infinite group. Now we show how to extend the other proof from last section.

Theorem 1.17. *Denote the group of m–th roots of unity by $\boldsymbol{\mu_m} = \{\zeta \in \bar{K}^* \mid \zeta^m = 1\}$. Then*

a) *(Hilbert Theorem 90)*

$$H^1(G, \bar{K}^*) = 0.$$

b) *Assume that* $\mathrm{char}(K)$ *does not divide* m *(or* $\mathrm{char}(K) = 0$*). Then*

$$H^1(G, \boldsymbol{\mu_m}) \cong \bar{K}^*/\bar{K}^{*m}.$$

Proof.

a) $H^1(G, \bar{K}^*) = \varinjlim H^1(\mathrm{Gal}(L/K), L^*) = 0.$

b) We have a sequence

$$1 \longrightarrow \boldsymbol{\mu_m} \xrightarrow{i} \bar{K}^* \xrightarrow{m} \bar{K}^* \longrightarrow 1$$

where $i\colon \zeta \mapsto \zeta$ is inclusion (the codomain is correct since can solve $\zeta^m - 1 = 0$ in $\bar{K}^*$) and $m\colon x \mapsto x^m$. Then i is obviously injective and m is surjective since working in $\bar{K}^*$. Now for $\zeta \in \boldsymbol{\mu_m}$, $\zeta^m = 1$ so $\mathrm{Im}\,(i) \subseteq \mathrm{Ker}\,(m)$. To see the other inclusion let $x \in \mathrm{Ker}\,(m)$. Then $x^m = 1$ so $x \in \boldsymbol{\mu_m} = \mathrm{Im}\,(i)$. Thus our sequence is exact. Now take Galois cohomology to find a long exact sequence

$$0 \longrightarrow H^0(G, \boldsymbol{\mu_m}) \longrightarrow H^0(G, \bar{K}^*) \longrightarrow H^0(G, \bar{K}^*)$$

$$\xrightarrow{\alpha} H^1(G, \boldsymbol{\mu_m}) \longrightarrow H^1(G, \bar{K}^*) \longrightarrow H^1(G, \bar{K}^*).$$

The connecting homomorphism is given by

$$\alpha\colon b \mapsto \{\sigma \mapsto \beta^\sigma/\beta\}$$

where we fix $\beta \in \bar{K}^*$ such that $\beta^m = b$.

Now $H^0(G, \bar{K}^*) = (\bar{K}^*)^G = K^*$ and $H^1(G, \bar{K}^*) = 0$ by Hilbert Theorem 90. So we have an exact sequence

$$K^* \xrightarrow{m} K^* \xrightarrow{\alpha} H^1(G, \boldsymbol{\mu_m}) \longrightarrow 0.$$

By the first isomorphism theorem,

$$\frac{K^*}{\mathrm{Ker}\,(\alpha)} \cong \mathrm{Im}\,(\alpha).$$

But $\mathrm{Ker}\,(\alpha) = \mathrm{Im}\,(m) = (K^*)^m$ and α is surjective. So we have

$$\frac{K^*}{(K^*)^m} \cong H^1(G, \boldsymbol{\mu_m})$$

via the bijection

$$\alpha\colon \{b\} \mapsto \{\sigma \mapsto \beta^\sigma/\beta\}.$$

□

Chapter 2

Algebraic curves

Elliptic curves, our object of study, can be analysed using many different methods. We emphasise the algebraic methods rather than complex analysis or explicit polynomial calculation. For this reason we need some algebraic geometry. To prevent half this thesis being on algebraic geometry we omit most of the proofs. We follow **[12]** in our exposition and refer the interested reader to the standard (but notoriously difficult) **[4]** for details.

2.1. Affine varieties

We begin by studying affine space and its subsets defined by zeros of polynomials. Let K be a perfect field. Let $\bar{K}$ be an algebraic closure of K and $G = \mathrm{Gal}(\bar{K}/K)$.

Definition 2.1. Affine n–space (over K) *is the set of n–tuples*

$$\mathbb{A}^n = \mathbb{A}^n(\bar{K}) = \{\, P = (x_1, \ldots, x_n) \mid x_i \in \bar{K} \,\}.$$

Similarly, the set of K–rational points *in* $\mathbb{A}^n$ *is the set*

$$\mathbb{A}^n(K) = \{\, P = (x_1, \ldots, x_n) \mid x_i \in K \,\}.$$

Recall that the Galois group, G, acts on $\bar{K}$ and that we can characterise K by $K = \{\, x \in \bar{K} \mid x^\sigma = x \text{ for all } G \,\}$. We can extend this action to $\mathbb{A}^n$; for $\sigma \in G$ and $P \in \mathbb{A}^n$ define

$$P^\sigma = (x_1^\sigma, \ldots, x_n^\sigma).$$

Then $\mathbb{A}^n(K)$ may be characterised by

$$\mathbb{A}^n(K) = \{\, P \in \mathbb{A}^n \mid P^\sigma = P \text{ for all } \sigma \in G \,\}.$$

Let $\bar{K}[X] = \bar{K}[X_1, \ldots, X_n]$ be a polynomial ring in n variables and $I \subseteq \bar{K}[X]$ an ideal. For each I we have a subset of $\mathbb{A}^n$

$$V_I = \{\, P \in \mathbb{A}^n \mid f(P) = 0 \text{ for all } f \in I \,\}$$

This is just formalising the idea of graphing solutions to the polynomial equations. We will eventually be studying cases where we have only one curve in I – an elliptic curve.

Definition 2.2. *An* (affine) algebraic set *is any set of the form V_I. If V is an algebraic set, the* ideal of V *is given by*

$$I(V) = \{\, f \in \bar{K}[X] \colon f(P) = 0 \text{ for all } P \in V \,\}.$$

An algebraic set V is defined over K *(denoted V/K) if its ideal $I(V)$ can be generated by polynomials in $K[X]$. If V is defined over K the set of* K–rational points *is*

$$V(K) = V \cap \mathbb{A}^n(K).$$

For $f \in \bar{K}[X]$ and $P = (x_1, \ldots, x_n) \in \mathbb{A}^n$ we have the action of $\sigma \in G$ on $f(P)$ by

$$\begin{aligned} f(P)^\sigma &= (\sum c_i x_1^{r_{i,1}} \cdots x_n^{r_{i,n}})^\sigma \\ &= \sum c_i^\sigma (x_1^\sigma)^{r_{i,1}} \cdots (x_n^\sigma)^{r_{i,n}} \\ &= f^\sigma(P^\sigma). \end{aligned}$$

This concurs with our earlier discussion – $\mathbb{A}^n$ is a G–module and we have an action on the G–module homomorphisms by $f^\sigma(P) = f(P^{\sigma^{-1}})^\sigma$. Note that if $f \in K[X]$ then for $\sigma \in G$ and $P \in \mathbb{A}^n$, since the coefficients stay fixed under the action of G,

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

If V is defined over K this means that $P \in V$ implies $P^\sigma \in V$ so the action of G on $\mathbb{A}^n$ induces an action of G on V. Once again we see

$$V(K) = \{\, P \in V \colon P^\sigma = P \text{ for all } \sigma \in G \,\}.$$

Definition 2.3. *An affine algebraic set V is called an* (affine) variety *if $I(V)$ is a prime ideal in $\bar{K}[X]$. Let V/K be a variety. Then the* affine coordinate ring of V/K *is defined by*

$$K[V] = \frac{K[X]}{I(V/K)}.$$

The affine coordinate ring is an integral domain and its quotient field, denoted $K(V)$, is called the function field of V/K. *Similarly we define*

$$\bar{K}[V] = \frac{\bar{K}[X]}{I(V)}$$

and $\bar{K}(V)$ as its quotient field.

To see how G acts on $\bar{K}[V]$, note that $f \in \bar{K}[V]$ is defined up to addition of polynomials vanishing on V so is well defined as a map $f \colon V \to \bar{K}$. Now G acts on f by acting on its coefficients, so since V is defined over K, G maps $I(V)$ to itself. This gives us the action of G

on $\bar{K}[V]$ which we can extend to an action on $\bar{K}(V)$. We denote the action of σ on f by f^σ and note that

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

From the above definitions we can already see the interplay between the algebraic and geometric ideas. The variety is a geometric idea and to it we associate the algebraic structure of the affine coordinate ring. We will be using algebraic geometry to answer questions about elliptic curves over the rationals – questions of number theory. However algebraic geometry allows us to use geometric ideas such as the concepts of dimension and smoothness, which we now introduce.

Definition 2.4. *Let V be a variety. The* dimension of V*, denoted* $\dim(V)$*, is the transcendence degree of $\bar{K}(V)$ over $\bar{K}$.*

We can see that the dimension of $\mathbb{A}^n$ is n since it has $X_1, \ldots, X_n$ as algebraically independent transcendental elements. Also if $V \subseteq A^n$ is given by a single non-constant polynomial equation

$$f(X_1, \ldots, X_n) = 0,$$

then $\dim(V) = n-1$. This is because factoring out by $f(X)$ is equivalent to expressing one of the variables X_i in terms of the others so there are only $n-1$ algebraically independent transcendental elements.

Definition 2.5. *Let V be a variety, $P \in V$ and $f_1, \ldots, f_m \in \bar{K}[X]$ a set of generators for $I(V)$. Then V is* non–singular *(or* smooth*) at P if the $m \times n$ matrix*

$$\left(\frac{\partial f_i}{\partial X_j}(P)\right)_{1 \le i \le m, 1 \le j \le n}$$

has rank $n - \dim(V)$. If V is non-singular at every point, then we say that V is non–singular*.*

Let V be given by a single non-constant polynomial equation

$$f(X_1, \ldots, X_n) = 0.$$

Then, as we have seen $\dim(V) = n-1$ so $P \in V$ is a singular point if and only if

$$\frac{\partial f}{\partial X_1}(P) = \cdots = \frac{\partial f}{\partial X_n}(P) = 0.$$

2.2. Projective varieties

Projective space is used to make our theory more consistent – for example in projective space two lines always meet, parallel lines meet 'at infinity'. To define this abstractly we consider projective space as lines in affine space of one higher dimension.

Definition 2.6. Projective n–space (over K), *denoted* $\mathbb{P}^n$ *or* $\mathbb{P}^n(K)$, *is*

$$\mathbb{P}^n = \{(x_0, \ldots, x_n) \in \mathbb{A}^{n+1}\}/\sim$$

where $\sim$ *is an equivalence relation defined by*

$$(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$$

if there exists $\lambda \in \bar{K}^*$ *such that* $x_i = \lambda y_i$ *for all* i. *An equivalence class is denoted* $(x_0 \colon \cdots \colon x_n)$ *and the* x_i *are called* homogenous coordinates *for the point in* $\mathbb{P}^n$. *The* set of K–rational points in $\mathbb{P}^n$ *is the set*

$$\mathbb{P}^n(K) = \{(x_0 \colon \cdots \colon x_n) \mid x_i \in K\}.$$

Definition 2.7. *Let* $P = (x_0 \colon \cdots \colon x_n) \in \mathbb{P}^N(\bar{K})$. *The* minimal field of definition for P (over K) *is the field*

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

The Galois group acts on $\mathbb{P}^n$ by acting on homogenous coordinates. We can check that this respects the equivalence relation $\sim$ and that

$$\mathbb{P}^n(K) = \{P \in \mathbb{P}^n \mid P^\sigma = P \text{ for all } \sigma \in G\},$$

and

$$K(P) = \text{ fixed field of } \{\sigma \in G \mid P^\sigma = P\}.$$

Definition 2.8. *A polynomial* $f \in \bar{K}[X] = \bar{K}[X_0, \ldots, X_n]$ *is* homogenous of degree d *if*

$$f(\lambda X_0, \ldots, \lambda X_n) = \lambda^d f(X_0, \ldots, X_n)$$

for all $\lambda \in \bar{K}$. *An ideal* $I \subseteq \bar{K}[X]$ *is* homogenous *if it is generated by homogenous polynomials.*

Now we notice that for a homogenous polynomial f, if $(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$ then $f(x_0, \ldots, x_n) = 0$ if and only if $f(y_0, \ldots, y_n) = 0$. Thus it makes sense to ask whether $f(P) = 0$ for some $P \in \mathbb{P}^n$ and we can associate to each homogenous ideal, I, a subset of $\mathbb{P}^n$

$$V_I = \{P \in \mathbb{P}^n \mid f(P) = 0 \text{ for all } f \in I\}.$$

Definition 2.9. *A* (projective) algebraic set *is any set of the form V_I. If V is a projective algebraic set, the* (homogenous) ideal of V, *denoted $I(V)$, is the ideal in $\bar{K}[X]$ generated by*

$$\{f \in \bar{K}[X] \mid f \text{ is homogenous and } f(P) = 0 \text{ for all } P \in V\}.$$

Such a V is defined over K, *denoted V/K, if its ideal $I(V)$ can be generated by homogenous polynomials in $K[X]$. If V is defined over K the* set of K–rational points of V *is the set*

$$V(K) = V \cap \mathbb{P}^n(K).$$

Also

$$V(K) = \{P \in V \mid P^\sigma = P \text{ for all } \sigma \in G\}.$$

Definition 2.10. *A projective algebraic set is called a* (projective) variety *if its homogenous ideal $I(V)$ is a prime ideal in $\bar{K}[X]$.*

To see the connections between affine and projective space consider the inclusion

$$\phi_i \colon \mathbb{A}^n \to \mathbb{P}^n$$

$$(y_1, \ldots, y_n) \mapsto (y_1 \colon \cdots \colon y_{i-1} \colon 1 \colon y_{i+1} \colon \cdots y_n).$$

Let $U_i = \{P = (x_0 \colon \cdots \colon x_n) \in \mathbb{P}^n \mid x_i \neq 0\}$ then we have a bijection

$$\theta^{-1} \colon U_i \to \mathbb{A}^n$$

$$(x_0 \colon \cdots \colon x_n) \mapsto \left(\frac{x_0}{x_i}, \cdots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \cdots, \frac{x_n}{x_i}\right)$$

which identifies $\mathbb{A}^n$ with the set U_i in $\mathbb{P}^n$.

Now let V be a projective algebraic set with homogenous ideal $I(V) \subseteq \bar{K}[X]$. Then $V \cap \mathbb{A}^n$ is an affine algebraic set with ideal $I(V \cap \mathbb{A}^n) \subseteq \bar{K}[Y]$ given by

$$I(V \cap \mathbb{A}^n) = \{f(Y_1, \ldots, Y_{i-1}, 1, Y_i, \ldots, Y_n) \mid f(X_0, \ldots, X_n) \in I(V)\}.$$

Any projective variety V is covered by subsets $V \cap U_0, \ldots, V \cap U_n$, each of which is an affine variety. The process of replacing $f(X_0, \ldots, X_n)$ by $f(Y_1, \ldots, Y_{i-1}, 1, Y_i, \ldots, Y_n)$ is called *dehomogenisation with respect to X_i*.

For any $f(Y) \in \bar{K}[Y]$, let

$$f^*(X_0, \ldots, X_n) = X_i^d f\left(\frac{X_0}{X_i}, \cdots, \frac{X_{i-1}}{X_i}, \frac{X_{i+1}}{X_i}, \cdots, \frac{X_n}{X_i}\right),$$

where d is the degree of f. We say that f^* is the *homogenisation of f with respect to X_i*.

Definition 2.11. *Let V be an affine algebraic set with ideal $I(V)$ and consider V as a subset of $\mathbb{P}^n$ via the map*

$$V \subseteq \mathbb{A}^n \xrightarrow{\phi_i} \mathbb{P}^n.$$

The projective closure of V*, denoted $\bar{V}$, is the projective algebraic set whose homogenous ideal $I(\bar{V})$ is generated by $\{f^*(X) \mid f \in I(V)\}$.*

Theorem 2.12.

a) *Let V be an affine variety. Then $\bar{V}$ is a projective variety and $V = \bar{V} \cap \mathbb{A}^n$.*
b) *Let V be a projective variety. Then $V \cap \mathbb{A}^n$ is an affine variety and either*
$$V \cap \mathbb{A}^n = \emptyset \text{ or } V = \overline{V \cap \mathbb{A}^n}.$$
c) *If an affine (respectively projective) variety V is defined over K then V (respectively $V \cap \mathbb{A}^n$) is also defined over K.*

Proof. Omitted. See [**4**]. □

In view of these results we can now identify an affine variety V with a unique projective variety. The points of $\bar{V} - V$ are called the *points at infinity* on $\bar{V}$. We will abuse our notation as in the following example.

Example 2.13. Let V be the projective variety given by

$$V\colon Y^2 = X^3 + 17.$$

This is really the variety in $\mathbb{P}^2$ given by the homogenous equation

$$\bar{Y}^2\bar{Z} = \bar{X}^3 + 17\bar{Z}^3,$$

the identification being $X = \bar{X}/\bar{Z}$ and $Y = \bar{Y}/\bar{Z}$. This variety has one point at infinity, $(0\colon 1\colon 0)$, obtained by setting $\bar{Z} = 0$. Thus

$$V(\mathbb{Q}) = \{(x, y) \in \mathbb{A}^2(\mathbb{Q}) \mid y^2 = x^3 + 17\} \cup \{(0\colon 1\colon 0)\}.$$

In fact V is an elliptic curve and we will be taking this point of view most of the time – considering it as an affine curve with one point at infinity that we treat as a special case.

Now we define the properties of a projective variety V in terms of the affine subvariety $V \cap \mathbb{A}^n$.

Definition 2.14. *Let V/K be a projective variety and choose $\mathbb{A}^n \subseteq \mathbb{P}^n$ so that $V \cap \mathbb{A}^n \neq \emptyset$. The* dimension of V *is the dimension of $V \cap \mathbb{A}^n$. The* function field of V*, denoted $K(V)$, is the function field of $V \cap \mathbb{A}^n$ and similarly for $\bar{K}(V)$.*

The variety V is non–singular (or smooth) *at P if $V \cap \mathbb{A}^n$ is non–singular at P.*

2.3. Maps between varieties

Now we consider algebraic maps between projective varieties.

Definition 2.15. *Let V_1 and $V_2 \subseteq \mathbb{P}^n$ be projective varieties. A* rational map from V_1 to V_2 *is a map of the form*

$$\phi\colon V_1 \to V_2$$

$$\phi = (f_0, \ldots, f_n),$$

where $f_i \in \bar{K}(V_1)$ are such that for every point $P \in V_1$ at which they are all defined

$$\phi(P) = (f_0(P), \ldots, f_n(P)) \in V_2.$$

If V_1 and V_2 are defined over K then G acts on ϕ by

$$\phi^\sigma(P) = (f_0^\sigma(P), \ldots, f_n^\sigma(P)).$$

Notice that for all $\sigma \in G$ and $P \in V_1$

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

If there is some $\lambda \in \bar{K}^*$ such that $\lambda f_i \in K(V_1)$ for all i then ϕ is said to be defined over K*. It can be proven (using some group cohomology) that ϕ is defined over K if and only if $\phi = \phi^\sigma$ for all $\sigma \in G$.*

Definition 2.16. *A rational map*

$$\phi = (f_0, \ldots, f_n)\colon V_1 \to V_2$$

is regular *at $P \in V_1$ if there is a function $g \in \bar{K}(V_1)$ such that*

a) *each gf_i is regular at P; and*
b) *for some i, $(gf_i)(P) \neq 0$.*

If such a g exists, set

$$\phi(P) = ((gf_0((P), \ldots, (gf_n)(P)).$$

A rational map which is regular at every point is called a morphism*.*

Definition 2.17. *Let V_1 and V_2 be varieties. We say that V_1 and V_2 are* isomorphic *and write $V_1 \cong V_2$, if there are morphisms $\phi\colon V_1 \to V_2$ and $\psi\colon V_2 \to V_1$ such that $\psi \circ \phi$ and $\phi \circ \psi$ are the identity maps on V_1 and V_2 respectively. We say V_1/K and V_2/K are* isomorphic over K *if such ϕ and ψ can be defined over K.*

2.4. Curves

A *curve* is a projective variety of dimension one. Our first result is that for smooth curves, a rational map is always defined at every point.

Theorem 2.18. *Let C be a curve, $V \subseteq \mathbb{P}^N$ a variety, $P \in C$ a smooth point, and $\phi\colon C \to V$ a rational map. Then ϕ is regular at P. In particular, if C is smooth then ϕ is a morphism.*

Proof. Omitted. See [**12**] II.2.1. □

Theorem 2.19. *Let $\phi\colon C_1 \to C_2$ be a morphism of curves. Then ϕ is either constant or surjective.*

Proof. Omitted. See [**4**] II.6.8. □

Let C_1/K and C_2/K be curves and $\phi\colon C_1 \to C_2$ a non-constant rational map defined over K. Then composition with ϕ induces an injection of function fields fixing K,

$$\phi^*\colon K(C_2) \to K(C_1)$$

$$\phi^* f = f \circ \phi.$$

Definition 2.20. *Let $\phi\colon C_1 \to C_2$ be a map of curves defined over K. If ϕ is constant define the* degree *of ϕ to be 0; otherwise define*

$$\deg(\phi) = [K(C_1)\colon \phi^* K(C_2)].$$

We will require the following two results in a later chapter.

Corollary 2.21. *Let C_1 and C_2 be smooth curves and let $\phi\colon C_1 \to C_2$ be a map of degree 1. Then ϕ is an isomorphism.*

Proof. Omitted. See [**12**] 2.4.1. □

Theorem 2.22. *Let $\phi\colon C_1 \to C_2$ be a non-constant map of smooth curves defined over a perfect field K. Then for all but finitely many $Q \in C_2$,*

$$\left|\phi^{-1}(Q)\right| = \deg(\phi).$$

Proof. Omitted. See [**4**], II.6.8. □

2.5. Twists of a curve

This section is necessary for when we discuss principal homogenous spaces, which give a geometric interpretation of a particular cohomology group. We discuss it here as we are working for a general field K, whereas later on we specialise.

Definition 2.23. *Let C/K be a smooth curve. The* isomorphism group of C, *denoted* $\mathrm{Isom}(C)$, *is the group of isomorphisms from C to itself (defined over $\bar{K}$).*

Definition 2.24. *A* twist of C/K *is a smooth curve C'/K which is isomorphic to C over $\bar{K}$. We identify two twists if they are isomorphic over K. The set of twists of C/K, modulo K–isomorphism, is denoted* $\mathrm{Twist}(C/K)$.

In the next theorem we discuss $H^1(G, \mathrm{Isom}(C))$. For an elliptic curve C, the group $\mathrm{Isom}(C)$ is non-abelian so we require a non-abelian group cohomology theory. We have not provided one, but read **[12]** B.3 for the details.

Theorem 2.25. *Let C/K be a smooth curve. For each twist C'/K of C/K, choose an isomorphism $\phi\colon C' \to C$ and define a map $\xi\colon G \to \mathrm{Isom}(C)$ by $\xi(\sigma) = \phi^\sigma \phi^{-1}$. Then*

a) *The map ξ is a 1–cocycle (that is, $\xi(\sigma\tau) = \xi(\sigma)^\tau \xi(\tau)$).*
b) *The cohomology class $\{\xi\}$ is determined by the K–isomorphism class of C', independent of the choice of ϕ. We thus obtain an injective map*
$$\mathrm{Twist}(C/K) \to H^1(G, \mathrm{Isom}(C)).$$
c) *This map is a bijection.*

Proof.

a) Calculate $\xi(\sigma\tau) = \phi^{\sigma\tau}\phi^{-1} = (\phi^\sigma\phi^{-1})^\tau(\phi^\tau\phi^{-1}) = \xi(\sigma)^\tau\xi(\tau)$.
b) Let C''/K be a twist of C/K which is K–isomorphic to C'. Choose a $\bar{K}$–isomorphism $\psi\colon C'' \to C$ and K–isomorphism $\theta\colon C'' \to C'$. We must show that the cocycles $\phi^\sigma\phi^{-1}$ and $\psi^\sigma\psi^{-1}$ are cohomologous.

   Use the following diagram to define $\alpha = \phi\theta\sigma^{-1}$

$$\begin{array}{ccc} C'' & \xrightarrow{\theta} & C' \\ & \searrow^{\psi} & \downarrow{\scriptstyle \phi} \\ & & C \end{array}$$

   then calculate $\alpha^\sigma(\psi^\sigma\psi^{-1}) = \phi^\sigma\theta^\sigma\psi^{-\sigma}\psi^\sigma\psi^{-1} = \phi^\sigma\theta^\sigma\psi^{-1}$. Now θ is defined over K so $\alpha^\sigma(\psi^\sigma\psi^{-1}) = \phi^\sigma\theta\psi^{-1} = (\phi^\sigma\phi^{-1})(\phi\theta\psi^{-1}) = \phi^\sigma\phi^{-1}\alpha$. Thus $\phi^\sigma\phi^{-1} = \alpha^\sigma(\psi^\sigma\psi^{-1})\alpha^{-1}$ and our cocycles are homologous.

   To see injectivity, suppose C'/K and C''/K give the same cohomology class in $H^1(G, \mathrm{Isom}(C))$. Since these are twists of C we have $\bar{K}$–isomorphisms $\phi\colon C' \to C$ and $\psi\colon C'' \to C$. The

twists C' and C'' being cohomologous implies that there exists an $\alpha \in \mathrm{Isom}(C)$ such that

$$\alpha^\sigma(\phi^{-1}\psi^{-1}) = (\phi^\sigma\phi^{-1})\alpha$$

for all $\sigma \in G$. Now define a map $\theta\colon C'' \to C$ by $\theta = \phi^{-1}\alpha\psi$. By construction this map is a $\bar{K}$–isomorphism but we want to show that it is actually defined over K. To see this calculate

$$\theta^\sigma = (\phi^\sigma)^{-1}(\alpha^\sigma\psi^\sigma) = (\phi^\sigma)^{-1}(\phi^\sigma\phi^{-1}\alpha\psi) = \phi^{-1}\alpha\psi = \theta.$$

Therefore C'' and C' are K–isomorphic so are the same element of $\mathrm{Twist}(C/K)$.

c) Omitted. See **[12]** to see how, given $\xi\colon G \to \mathrm{Isom}(C)$, to construct the function field $\bar{K}(C)_\xi$ and from this the curve in $\mathrm{Twist}(C/K)$.

□

Chapter 3

Elliptic curves

3.1. Elliptic curves as algebraic curves

The *genus* is a quantity that is useful for the classification of algebraic curves. Genus zero curves are those which are equivalent to the projective line, for example conics. These have either no rational points or infinitely many and are well understood. Curves with genus greater than one have only a finite number of rational points by a deep result conjectured by Mordell and proved by Falting. For more discussion see [**12**].

We study curves of genus one with a distinguished point. Although we haven't actually defined genus we only need the results from the following theorem.

Definition 3.1. *An* elliptic curve *E is a curve of genus 1 with a distinguished point, $\mathcal{O} \in E$. The elliptic curve is* defined over *K, written E/K, if E is defined over K as a curve and $\mathcal{O} \in E(K)$.*

Theorem 3.2. *Let E be an elliptic curve defined over K.*

a) *There exist functions $x, y \in K(E)$ such that the map*

$$\phi\colon E \to \mathbb{P}^2 \qquad \phi = (x, y, 1)$$

gives an isomorphism of E/K onto a curve given by a Weierstrass equation

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

with $a_i \in K$, such that $\phi(\mathcal{O}) = (0\colon 1\colon 0)$.

b) *Any two Weierstrass equations for E are related by a linear change of variable of the form*

$$x = u^2x' + r$$

$$y = u^3y' + su^2x' + t$$

with $u, r, s, t \in K, u \neq 0$.

c) *Conversely, every smooth cubic curve C given by a Weierstrass equation is an elliptic curve defined over K with origin $\mathcal{O} = (0\colon 1\colon 0)$.*

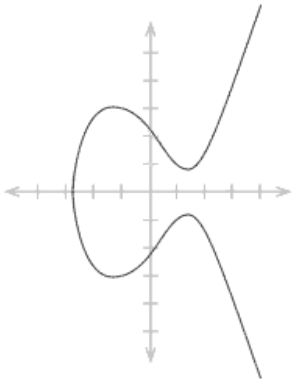

FIGURE 1. An elliptic curve, drawn over the reals.

Proof. Omitted. See [**12**]. □

3.2. Weierstrass equations

We have asserted that any elliptic curve can be given by a *Weierstrass equation*

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

where $a_i \in \bar{K}$, together with $\mathcal{O}$, the "point at infinity". We also require that the curve is non-singular, so has no cusps or self intersections. An example is shown in Figure 1, where we consider the point at infinity to be the point infinitely far to the top and bottom of the graph.

Note that if the field we are working over has characteristic not equal to two then we can complete the square, replacing y by $\frac{1}{2}(y - a_1x - a_3)$. We are following [**12**] and will do all algebra by computer in appendix A. This gives us

$$E\colon y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6,$$

where

$$b_2 = a_1^2 + 4a_2, \quad b_4 = 2a_4 + a_1a_3, \quad b_6 = a_3^2 + 4a_6.$$

We also define

$$b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2,$$

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6.$$

If the characteristic of our field is not two or three we can make a further change of variables,

$$(x, y) \mapsto \left(\frac{x - 3b_2}{36}, \frac{y}{108}\right)$$

giving the equation

$$E\colon y^2 = x^3 - 27c_4x - 54c_6,$$

where

$$c_4 = b_2^2 - 24b_4 \quad \text{and} \quad c_6 = b_2^3 + 36b_2b_4 - 216b_6.$$

If we have $\mathrm{char}(K) \neq 2, 3$ then we may assume that an elliptic curve has Weierstrass equation

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

For this curve we can calculate that

$$\Delta = -16(4A^3 + 27B^2).$$

Now we show the significance of the discriminant.

Theorem 3.3. *A curve given by a Weierstrass equation is nonsingular if and only if $\Delta \neq 0$.*

Proof. Let E be given by

$$E\colon f(x, y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 = 0.$$

To see that the point at infinity is never singular consider the projective curve

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

at the point $\mathcal{O} = (0\colon 1\colon 0)$. Since

$$\frac{\partial F}{\partial Z}(\mathcal{O}) = 1 \neq 0,$$

we see that $\mathcal{O}$ is a non-singular point on E.

Now suppose that E is singular, say at $P_0 = (x_0, y_0)$. We can translate this point to the origin by a change of coordinates

$$x = x' + x_0 \quad y = y' + y_0$$

that leaves Δ unchanged. Now note that $f(0, 0) = -a_6$, $\frac{\partial f}{\partial x}(0, 0) = -a_4$ and $\frac{\partial f}{\partial y}(0, 0) = a_3$. So if $(0, 0) \in E$ is a singular point then E must have a Weierstrass equation of the form

$$f(x, y) = y^2 + a_1xy - a_2x^2 - x^3 = 0.$$

Direct calculation shows that $\Delta = 0$.

Now suppose that E is non-singular. We aim to show $\Delta \neq 0$. To simplify calculations assume $\mathrm{char}(K) \neq 2$ (a proof in this case can be found in appendix A of **[12]**). So we can assume that E is given by

$$f(x, y) = y^2 - 4x^3 - b_2x^2 - 2b_4x - b_6.$$

Now if this equation were to be singular, say at $P_0 = (x_0, y_0)$ then:

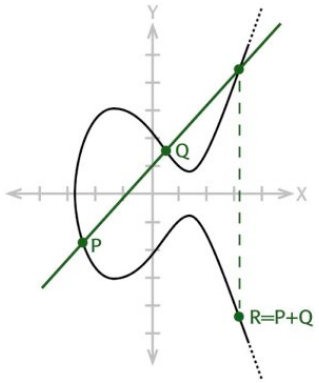

FIGURE 2. The group law on an elliptic curve.

a) $\frac{\partial f}{\partial y}(x_0, y_0) = 2y_0 = 0$. Thus $y_0 = 0$ and then x_0 must be a root of $4x^3 + b_2x^2 + 2b_4x + b_6$.
b) $\frac{\partial f}{\partial x}(x_0, y_0) = 12x_0^2 + 2b_2x_0 + 2b_4 = 0$.

So we require x_0 to be a double root of $g(x) = 4x^3 + b_2x^2 + 2b_4x + b_6$. But a polynomial has a multiple root if and only its discriminant is zero. We can calculate that $\text{disc}(g) = 16\Delta$ so E is non-singular implies that Δ is non-zero. □

3.3. The group law

Definition 3.4. *Let E be an elliptic curve given by a Weierstrass equation. Given $P, Q \in E$, draw the line from P to Q until you hit the curve again. This forms another point on the curve. Now draw a line from the point at infinity, $\mathcal{O}$, through this new point. The point where this line intersects the elliptic curve again is $P + Q$.*

This definition relies on a line in $\mathbb{P}^2$ intersecting an elliptic curve E at exactly three points (counting multiplicity). This follows from Bezout's Theorem (see [**4**]) however as we provide explicit formulae below we can simply take this as our definition. If $P = Q$ or one of P and Q equals $\mathcal{O}$ then our definition takes some interpreting. Because we only have one point at infinity we can usually get away with considering $\mathcal{O}$ as a special case.

If we are trying to find $2P$, then P and Q are the same point. In this case we take the line between P and Q to be the tangent line at P and proceed in the same manner as above. If the line from P to Q doesn't intersect the curve anywhere on the finite plane (in our figures this means the line is vertical) then we say that it intersects the elliptic curve at $\mathcal{O}$ – this is why we needed to include this point on our curve.

Then the line from $\mathcal{O}$ to $\mathcal{O}$, the "line at infinity", intersects the curve at $\mathcal{O}$. Thus $P + Q = \mathcal{O}$.

Note that this definition is symmetrical, the line between P and Q is the same as the line between Q and P so $P + Q = Q + P$. Also note that if we want to calculate $P + \mathcal{O}$ then the line between these two points is the same as the line between $\mathcal{O}$ and the point where the first line intersects the curve. This means $P + \mathcal{O} = P$ and $\mathcal{O}$ is the identity of our group.

From our geometric group law we can provide an explicit algorithm to add points on an elliptic curve. These formulae come from [**12**] where they are developed from the above ideas using conceptually simple but somewhat tedious coordinate geometry.

Algorithm 3.5. Let E be an elliptic curve given by

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

(a) If $P_0 = \mathcal{O}$ then $-\mathcal{O} = \mathcal{O}$. Otherwise let $P_0 = (x_0, y_0) \in E$. Then

$$-P_0 = (x_0, -y_0 - a_1x_0 - a_3).$$

(b) If one of P_1 or P_2 equals $\mathcal{O}$ then use $P + \mathcal{O} = \mathcal{O} + P = P$, if $x_1 = x_2$ and $y_1 + y_2 + a_1x_2 + a_3 = 0$ then $P_1 + P_2 = \mathcal{O}$. Otherwise let

$$P_1 + P_2 = P_3 \text{ with } P_i = (x_i, y_i) \in E.$$

then calculate

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1},\ \nu = \frac{y_1x_2 - y_2x_1}{x_2 - x_1} \text{ if } x_1 \neq x_2;$$

$$\lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} \text{ and}$$

$$\nu = \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3} \text{ if } x_1 = x_2.$$

(So $y = \lambda x + \nu$ is the line through P_1 and P_2, or tangent to E if P_1=P_2.)

Then $P_3 = (x_3, y_3)$ where

$$x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2$$

$$y_3 = -(\lambda + a_1)x_3 - \nu - a_3.$$

Theorem 3.6. *The addition law on an elliptic curve, E, as given in Algorithm 3.5, has the following properties:*

a) $P + \mathcal{O} = P$ *for all* $P \in E$.
b) $P + Q = Q + P$ *for all* $P, Q \in E$.

c) *Let $P \in E$. There is a point of E, denoted $-P$, so that*

$$P + (-P) = \mathcal{O}.$$

d) *Let $P, Q, R \in E$. Then*

$$(P + Q) + R = P + (Q + R).$$

In other words, the addition law makes E into an abelian group with identity $\mathcal{O}$. Furthermore:

e) *Suppose E is defined over K, that is, given by*

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \text{ with } a_i \in K.$$

Then the points on the curve with coordinates in K

$$E(K) = \{(x, y) \in K^2 \colon y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6\} \cup \{\mathcal{O}\}$$

form a subgroup of E.

Proof.

a) By definition.
b) The addition formulae given above are symmetrical. If you swap P_1 and P_2 then λ and ν remain the same.
c) $-P$ is defined above and has this property.
d) Omitted. Can draw some pictures or calculate some sums to convince yourself. The most illuminating proof is by the Riemann–Roch Theorem as in **[12]**.
e) If $P_1, P_2 \in E(K)$ then we can see that $P_1 + P_2 = (x_3, y_3)$ where the coordinates are given by rational functions (quotients of polynomials) in the variables x_i, y_i, a_i, all of which are in the field K. This means that $x_3, y_3 \in K$, that is $P_1 + P_2 \in E(K)$. Therefore $E(K)$ is a subgroup of E.

□

Theorem 3.7. *Let E/K be an elliptic curve. Then the equations giving the group law on E define morphisms*

$$+ \colon E \times E \to E \qquad \text{and} \qquad - \colon E \to E$$
$$(P_1, P_2) \mapsto P_1 + P_2 \qquad\qquad P \mapsto -P.$$

Proof. First, the subtraction map

$$(x, y) \mapsto (x, -y - a_1x - a_3)$$

is a rational map of a smooth curve E so is a morphism. Next choose $Q \neq \mathcal{O}$ and define a translation map $\tau_Q \colon E \to E$ by

$$\tau_Q \colon P \mapsto P + Q.$$

By the addition formula given earlier this is a rational map so is a morphism. Furthermore it has an inverse, τ_{-Q}, so is an isomorphism. Consider the general addition map $+\colon E \times E \to E$. From our addition formula it is a rational map except possibly at points of the form $(P,P), (P,-P), (P,\mathcal{O})$ and $(\mathcal{O},P)$.

To deal with these cases we use translation maps. Write τ_i for τ_{Q_i} and consider the composition

$$\phi\colon E \times E \xrightarrow{\tau_1\times\tau_2} E \times E \xrightarrow{+} E \xrightarrow{\tau_1^{-1}} E \xrightarrow{\tau_2^{-1}} E.$$

Since the group law on E is commutative and associative these maps act by

$$(P_1,P_2) \mapsto (P_1{+}Q_1, P_2{+}Q_2) \mapsto P_1{+}Q_1{+}P_2{+}Q_2 \mapsto P_1{+}P_2{+}Q_2 \mapsto P_1{+}P_2.$$

Thus the rational map ϕ agrees with the addition map whenever they are both defined.

Since the τ_i's are isomorphism, ϕ is an isomorphism except possibly at points of the form

$$(P{-}Q_1, P{-}Q_2) \quad (P{-}Q_1, -P{-}Q_2) \quad (P{-}Q_1, -Q_2) \quad (-Q_1, P{-}Q_2).$$

But we can choose any Q_1, Q_2 that we please. Thus we can find a finite set of rational maps

$$\phi_1, \phi_2, \ldots, \phi_n\colon E \times E \to E$$

such that

a) ϕ_1 is the addition map.
b) For each $(P_1,P_2) \in E \times E$, some ϕ_i is defined at (P_1,P_2).
c) If ϕ_i and ϕ_j are both defined at (P_1,P_2), then $\phi_i(P_1,P_2) = \phi_j(P_1,P_2)$.

Then addition is defined on all of $E \times E$ so is a morphism. □

Although we are primarily interested in non–singular curves when we study these curves we consider them 'modulo p' and sometimes have to consider curves that possibly have singular points.

Definition 3.8. *Let E be a curve given by a Weierstrass equation. Then the* non–singular part of the curve*, denoted E_{ns}, is the set of non–singular points of E.*

Theorem 3.9. *Let E be a curve given by a Weierstrass equation with $\Delta = 0$. Then the group law makes E_{ns} into an abelian group*

Proof. Omitted. See [**12**] for the two cases:

a) $c_4 \neq 0$ implies E has a node and $E_{ns} \cong \bar{K}^*$.
b) $c_4 = 0$ implies E has a cusp and $E_{ns} \cong \bar{K}^+$.

□

There is an effective method for calculating the torsion subgroup of an elliptic curve defined over $\mathbb{Q}$. We present the following theorem which, although not the method used in practice, is good enough for our purposes.

Theorem 3.10 (Lutz–Nagell Theorem)**.** *Let $E/\mathbb{Q}$ be an elliptic curve with discriminant Δ. Suppose $P \in E(\mathbb{Q})$ is a non-zero torsion point. Then $x(P), y(P) \in \mathbb{Z}$ and either $2P = \mathcal{O}$ or $y(P)^2$ divides Δ.*

Proof. Omitted. See **[12]** for a fancy proof using formal groups or **[13]** for an elementary proof using coordinate geometry. □

We give a sketch of the calculation for an elliptic curve that we shall study later.

Example 3.11. Let $E/\mathbb{Q}$ be given by

$$y^2 = x^3 + 2x^2 - 3x = x(x-1)(x+3).$$

We calculate $\Delta = 2^8 3^2 = 2304$. We can easily see that $E(\mathbb{Q})[2] = \{\mathcal{O}, (0,0), (-3,0), (1,0)\}$ since if $2P = \mathcal{O}$ then $P = -P$ so we must have $y = 0$. To find other points we know that $y^2 \mid \Delta$ and y is an integer so we can conclude that the only possibilities for y^2 are $1, 2^2, 2^4, 2^6, 2^8, 3^2, 2^2 3^2, 2^6 3^2, 2^8 3^2$. For each of these we have the equation

$$x^3 + 2x^2 - 3x - y^2 = 0$$

for which we only have to check a finite number of possible integer solutions (those dividing y^2). For example $y^2 = 2^2 3^2 = 36$ gives $x = 3$ so $(x, y) = (3, \pm 6)$ gives us two more points on $E(\mathbb{Q})$. Following through this example gives us

$$E_{\text{tors}}(\mathbb{Q}) = \{\mathcal{O}, (0,0), (-3,0), (1,0), (3, \pm 6), (-1, \pm 2)\}.$$

We can calculate that $2E_{\text{tors}}(\mathbb{Q}) = \{\mathcal{O}, (1,0)\}$ so that $|E_{\text{tors}}(\mathbb{Q})/2E_{\text{tors}}(\mathbb{Q})| = 4$ (this fact will be important later).

Chapter 4

p–adic numbers and Hensel's lemma

A useful method of discovering whether an equation has solutions in integers is consider the equation modulo some prime. The numbers we are about to define generalise this. For example saying that an equation is solvable 3–adically means that it has a solution modulo 3^n for all n.

4.1. p–adic numbers

We define the p–adic integers as an inverse limit and then construct the p–adic numbers from them, following [**11**]. For each $n \geq 1$ form $A_n = \mathbb{Z}/p^n\mathbb{Z}$. Then we have a map $\phi_n \colon A_n \to A_{n-1}$ defined by $\phi_n \colon a \mapsto a \pmod{p^{n-1}}$. Note that (A_n, ϕ_n) forms an inverse system indexed by the integers.

Definition 4.1. *The p–adic integers are the ring*

$$\mathbb{Z}_p = \varprojlim A_n.$$

Recall that this means that $x \in \mathbb{Z}_p$ is $x = (\ldots, x_n, \ldots, x_1)$ where, for all $n \geq 1$, $x_n \in A_n$ and $\phi_{n+1}(x_{n+1}) = x_n$. We consider $\mathbb{Z}_p \subseteq \prod_{n\geq 1} A_n$ so addition and multiplication are defined coordinate wise. Notice that we have an inclusion $i \colon \mathbb{Z} \to \mathbb{Z}_p$ by $i \colon a \mapsto (\ldots, a \pmod{p^n}, \ldots, a \pmod{p})$.

Theorem 4.2. *Let $[p^n]$ denote the map 'multiplication by p^n' and ϵ_n map x to its nth component. Then the following sequence is exact.*

$$0 \longrightarrow \mathbb{Z}_p \xrightarrow{[p^n]} \mathbb{Z}_p \xrightarrow{\epsilon_n} A_n \longrightarrow 0$$

Proof. To see that $\operatorname{Ker}[p^n] = 0$, if $px = 0$ then $(\ldots, p, \ldots, p, 0) \times (\ldots, x_n, \ldots, x_2, x_1) = 0$, that is $px_{n+1} = 0$ for all $n \geq 1$. This implies that $x_{n+1} = p^n y_{n+1}$ for some $y_{n+1} \in A_{n+1}$. Then $x_n = \phi_{n+1}(x_{n+1}) = 0$ for all $n \geq 1$ implying $x = 0$.

Note that $\epsilon(p^n x) = p^n x_n \equiv 0 \pmod{p^n}$ so $\operatorname{Im}[p^n] \subseteq \operatorname{Ker}\epsilon_n$. For the other inclusion, if $\epsilon_n(x) = 0$ then $x_n \equiv 0 \pmod{p^n}$. Thus $x_{n+m} \equiv 0 \pmod{p^n}$ for all $m \geq 0$ and therefore $x_{n+m} = p^n y_{n+m}$ for some

$y_{n+m} \in A_{n+m}$. Then $x = (\ldots, x_{n+1}, x_n, 0, \ldots, 0) = p^n(\ldots, y_{n+1}, y_n, y_n \pmod{p^{n-1}}, \ldots, y_n \pmod{p}) \in \operatorname{Im}[p^n]$.

For surjectivity, for all $a \in A_n$, take $x = (\ldots, a, \ldots, a, a \pmod{p^{n-1}}, \ldots, a \pmod{p}) \in \mathbb{Z}_p$. Then $\epsilon_n(x) = a$. □

From this short exact sequence we can conclude that $\mathbb{Z}_p/p^n\mathbb{Z}_p \cong A_n = \mathbb{Z}/p^n\mathbb{Z}$.

Theorem 4.3.

a) *An element $x \in \mathbb{Z}_p$ is invertible if and only if p does not divide x.*

b) *Every nonzero element of $\mathbb{Z}_p$ can be written uniquely in the form $p^n u$ with $u \in \mathbb{Z}_p^*$ and $n \geq 0$.*

Proof.

a) We first prove the statement for A_n. If $x \in A_n$ is not divisible by p then $x \notin pA_n$ so its image in A_1 is not zero. Since $A_1 = \mathbb{Z}/p\mathbb{Z}$ is a field then this is invertible. Therefore there exists $y \in A_1$ such that $xy = 1 \pmod{p}$, that is $xy = 1 - pz$ for some $z \in \mathbb{Z}$. Then noting that

$$\begin{aligned}\frac{1}{1-pz} &= 1 + pz + p^2z^2 + \ldots \\ &\equiv 1 + pz + p^2z^2 + \ldots + p^{n-1}z^{n-1} \pmod{p^n},\end{aligned}$$

we see that $xy(1 + pz + \ldots + p^{n-1}z^{n-1}) \equiv 1 \pmod{p^n}$ and thus x is invertible.

To prove our statement for $\mathbb{Z}_p$, note that if $p \mid x_i$ for any $i \in \mathbb{Z}$ then since $x_n \equiv x_i \pmod{p}$, $p \mid x_n$ for all $n \in \mathbb{Z}$. Thus $p \nmid x$ implies $p \nmid x_n$, meaning each x_n is invertible and we can form $x^{-1} = (\ldots, x_n^{-1}, \ldots, x_1^{-1})$.

b) If $x \in \mathbb{Z}_p$ is nonzero then there exists a largest n such that $x_n = 0$ (otherwise all coordinates of x are zero). Then $p^n \mid x_{n+1}$ since $\phi_{n+1}(x_{n+1}) = 0$. Thus $x_{n+1} = p^n u_{n+1}$ for some $u_{n+1} \in A_{n+1}$ such that $p \nmid u_{n+1}$ (otherwise $x_{n+1} = 0$). Now for all $m \geq n+1$, $x_m \equiv x_{n+1} \pmod{p^{n+1}}$ so then $p_n \mid x_m$ and $x_m = p^n u_m$ for some $u_m \in A_m$. Thus

$$x = (\ldots, x_{n+1}, 0, \ldots, 0) = p^n(\ldots, u_{n+1}, u_n, \ldots, u_1)$$

where $u_i = u_{n+1} \pmod{p^i}$ for all $i \leq n$.

We have $u_i \equiv u_{n+1} \pmod{p}$ for all i so $p \nmid u_{n+1}$ implies that $p \nmid u_i$ for all i. That is $p \nmid u$ and u is a unit.

To see that $x = p^n u$ is a unique expression of this form, note that n is uniquely determined by the above process and that if

$x = p^n u = p^n v$ for $u, v \in \mathbb{Z}_p^*$ we can simply cancel p^n to see that $u = v$.

□

Definition 4.4. *For a nonzero $x \in \mathbb{Z}_p$ write $x = p^n u$. The* p–adic valuation of x, *denoted $v_p(x)$, is n. Set $v_p(0) = \infty$.*

With this valuation we can show that $\mathbb{Z}_p$ is an integral domain.

Theorem 4.5. *For $x, y \in \mathbb{Z}_p$,*

a) $v_p(xy) = v_p(x) + v_p(y)$.
b) $v_p(x, y) \geq \min(v_p(x), v_p(y))$.
c) $\mathbb{Z}_p$ *is an integral domain.*

Proof.

a) If $x = 0$ or $y = 0$ then both $v_p(xy)$ and $v_p(x) + v_p(y)$ are infinite so equality holds. Otherwise say $x = p^n u$ and $y = p^m v$ with $n, m \geq 0$ and $u, v \in \mathbb{Z}_p^*$. Then

$$v_p(xy) = v_p(x^{n+m} uv) = n + m = v_p(x) + v_p(y)$$

since uv is a unit.

b) If $x = 0$ or $y = 0$ then $v_p(x + y) = \min(v_p(x), v_p(y))$. Otherwise again say $x = p^n u$ and $y = p^m v$ with $n, m \geq 0$ and $u, v \in \mathbb{Z}_p^*$. We have three cases:

(i) If $n < m$ then

$$\begin{aligned} v_p(x + y) &= v_p(p^n u + p^m v) = v_p(p^n(u + p^{m-n} v)) \\ &= n = \min(v_p(x), v_p(y)). \end{aligned}$$

This follows since $p \nmid u$ implies that $p \nmid (u + p^{m-n} v)$.

(ii) Similarly, if $m < n$ then

$$v_p(x + y) = m = \min(v_p(x), v_p(y)).$$

(iii) If $m = n$ then $x + y = p^n(u + v)$. Now $u + v \in \mathbb{Z}_p$ so $u + v = p^l w$ for some $l \geq 0$ and $w \in \mathbb{Z}_p^*$. Then

$$v_p(x + y) = v_p(p^{n+l} w) = n + l \geq n = \min(v_p(x), v_p(y)).$$

c) Say $a, b \in \mathbb{Z}_p$ such that $ab = 0$. Take the p–adic valuation of both sides and we see that

$$v_p(a) + v_p(b) = \infty.$$

Thus either $a = 0$ or $b = 0$ and $\mathbb{Z}_p$ is an integral domain.

□

We have the following topological facts about $\mathbb{Z}_p$.

Theorem 4.6.

a) $\mathbb{Z}_p$ *is compact.*
b) *The topology on* $\mathbb{Z}_p$ *is the same as the topology generated by* $d(x,y) = |x-y|$ *where* $|x| = p^{-v_p(x)}$ *for all* $x \in \mathbb{Z}_p$.
c) $\mathbb{Z}$ *is dense inside* $\mathbb{Z}_p$.

Proof. We give A_n the discrete topology and $\prod_{n=1}^{\infty} A_n$ the product topology (the coarsest topology for which the projection maps $p_i \colon \prod A_n \to A_i$ are continuous).

a) Note that $\mathbb{Z}_p = \bigcap_{i=1}^{\infty} B_i$ where

$$B_i = \{(x_j) \in \prod A_n \mid \phi_{i+1}(x_{i+1}) = x_i\}.$$

We will show that the B_i are closed. Then noting that the A_n are compact (every cover is already finite because A_n is finite) and recalling that the product of compact sets is compact (Tychonoff's Theorem) and that a closed subset of a compact set is compact we can conclude that $\mathbb{Z}_p$ is compact.

To see B_i is closed we take $(y_j) \notin B_i$ and construct an open set around (y_j) that does not intersect B_i. We have $\phi_{i+1}(y_{i+1}) = z \neq y_i$. Then we can form the sets $p_i^{-1}(z)$ and $p_{i+1}^{-1}(\phi_{i+1}^{-1}(y_i))$ both of which are open. Thus the intersection of these two sets in $\prod A_n$ is open. We can see that (y_j) is in this intersection and that it is disjoint with B_i.

b) First we show that the collection $\{p^n\mathbb{Z}_p\}$ forms a basis of neighbourhoods of zero. If U is an open set containing zero then it is a union of sets of the form

$$\prod_{i=1}^{\infty} Y_i$$

where $Y_i \subseteq A_i$ with equality for all but finitely many i. Let n be the largest integer such that $Y_n \neq A_n$. Then if $x \in p^n\mathbb{Z}_p$, $x = (\ldots, a, 0, \ldots, 0) \in U$ where the last n digits are zero. Thus $p^n\mathbb{Z}_p \subseteq \prod Y_i \subseteq U$ and $\{p^n\mathbb{Z}_p \mid n \in \mathbb{N}\}$ forms a basis for neighbourhoods of zero. Thus if U is an open set around zero then

$$U = \bigcup_{n \in I} p^n\mathbb{Z}_p$$

for some $I \subset \mathbb{N}$. Now if $i \geq j$ then $p^i\mathbb{Z}_p \subset p^j\mathbb{Z}_p$ so

$$U = \bigcup_{n \in I} p^n\mathbb{Z}_p = p^i\mathbb{Z}_p$$

where i is the minimum of the n's. Now if U is an open set around a point x then $U - x$ is an open set around zero so $U = p^i\mathbb{Z}_p + x$. Now a general open set is a union of open sets around points so the collection $\{x+p^n\mathbb{Z}_p \mid x \in \mathbb{Z}_p, n \in \mathbb{N}\cup\{0\}\}$ is a basis for the induced topology.

Now note that $x \in p^n\mathbb{Z}_p$ if and only if $|x| \leq p^{-n}$. The metric topology has basis given by open balls $B(x,r) = \{y \in \mathbb{Z}_p \mid |x-y| < r\}$. We show that the two bases are equal thus showing all open sets are equal and the topologies are equal. First note that for an open ball $B(x,r)$, the fact that $|x|$ only takes on values of p^{-n} for integer n means that $B(x,r) = B(x,p^{-n})$ where $p^{-n-1} < r \leq p^{-n}$. Then we can see that $x + p^n\mathbb{Z}_p = B(x,p^{-n})$ and that $B(x,r) = B(x,p^{-n}) = x + p^n\mathbb{Z}_p$.

c) For $\mathbb{Z}$ to be dense in $\mathbb{Z}_p$ we require for each $x \in \mathbb{Z}_p$ a sequence of integers that converges to x. If $x = (\ldots, x_n, \ldots, x_1)$ then the sequence of integers (x_n) satisfies this as can be seen by

$$|x_n - x| = |(\ldots, x_n - x_{n+1}, 0, \ldots, 0)| \leq p^{-n}.$$

□

Definition 4.7. *The field of p–adic numbers, denoted $\mathbb{Q}_p$, is defined to be the field of fractions of the ring $\mathbb{Z}_p$.*

We have the following immediate facts.

Theorem 4.8.

a) $\mathbb{Q}_p = \mathbb{Z}_p[p^{-1}]$.
b) *Every $x \in \mathbb{Q}_p^*$ can be written uniquely as $x = p^n u$ where $n \in \mathbb{Z}$ and $u \in \mathbb{Z}_p^*$. Use this to define the* p–adic valuation on $\mathbb{Q}_p$ *by $v_p(p^n u) = n$ and $v_p(0) = \infty$.*
c) *For $x \in \mathbb{Q}_p$, $v_p(x) \geq 0$ if and only if $x \in \mathbb{Z}_p$.*

Proof.

a) First note that zero is in both rings. If $x \in \mathbb{Q}_p$ is not zero then $x = \frac{p^n u}{p^m v}$ where $n, m \geq 0$ and $u, v \in \mathbb{Z}_p^*$. Thus $x = p^n(uv^{-1})(p^{-1})^m \in \mathbb{Z}_p[p^{-1}]$. Conversely, if $x \in \mathbb{Z}_p[p^{-1}]$ is not zero then for some $n, m \geq 0$ and $u \in \mathbb{Z}_p^*$, $x = p^n u(p^{-1})^m = \frac{p^n u}{p^m 1} \in \mathbb{Q}_p$.
b) If $x \in \mathbb{Q}_p^*$ then there exists $n, m \geq 0$ and $u, v \in \mathbb{Z}_p^*$ such that $x = \frac{p^n u}{p^n v} = p^{n-m}(uv^{-1})$. Now say that also for $s, t \geq 0$ and $w, z \in \mathbb{Z}_p^*$, $x = \frac{p^s w}{p^t z} = p^{s-t}(wz^{-1})$. However since $\frac{p^n u}{p^n v} = \frac{p^s w}{p^t z}$ we can conclude that $p^{m+s}wv = p^{n+t}uz$ and thus that $p^{n-m}(uv^{-1}) = p^{s-t}(wz^{-1})$. Thus the representation of x is unique.

c) For $x \in \mathbb{Q}_p$, if $v_p(x) \geq 0$ then for $n \geq 0$ and $u \in \mathbb{Z}_p^*$, $x = \frac{p^n u}{1} \in \mathbb{Z}_p$. Conversely if $x \in \mathbb{Z}_p$, for some $n \geq 0$ and $u \in \mathbb{Z}_p^*$, $x = \frac{p^n u}{1}$ so $v_p(x) \geq 0$.

□

Definition 4.9. *Define the* p–adic absolute value $|\cdot| : \mathbb{Q}_p \to \mathbb{R}$ *by*

$$|x| = p^{-v_p(x)}.$$

Theorem 4.10. *The p–adic absolute value satisfies the following properties for all $x, y \in \mathbb{Q}_p$*

a) $|x| = 0$ *if and only if* $x = 0$.
b) $|xy| = |x| \cdot |y|$.
c) $|x + y| \leq \max\{|x|, |y|\}$ *with equality if* $|x| \neq |y|$.

Proof.

a) If $|x| = 0$ then $v_p(x) = \infty$ so $x = 0$. Conversely $|0| = p^{-\infty} = 0$.
b) If x or y are zero then trivially true. Otherwise let $x = p^n u$ and $y = p^m v$ for $n, m \in \mathbb{Z}$ and $u, v \in \mathbb{Z}_p^*$. Then
$$|xy| = |p^{n+m} uv| = p^{-(n+m)} = p^{-n} p^{-m} = |x||y|.$$
c) If x or y are zero then trivially true. Otherwise let $x = p^n u$ and $y = p^m v$ for $n, m \in \mathbb{Z}$ and $u, v \in \mathbb{Z}_p^*$. Then if $n \neq m$, without loss of generality take $m < n$ and we have
$$|x + y| = |p^m (u + p^{n-m})| = p^{-m} = \max\{|x|, |y|\}$$
since $p \nmid u$ implies $p \nmid (u + p^{n-m} v)$. If $m = n$ then
$$|x + y| = |p^n (u + v)| \leq p^{-n} = \max\{|x|, |y|\}.$$

□

This theorem shows that $|\cdot|$ is a norm on $\mathbb{Q}_p$. The third condition is the *ultrametric inequality* which is stronger than the usual triangle inequality. It can be shown that $\mathbb{Q}_p$ is the completion of $\mathbb{Q}$ with respect to the p–adic absolute value. In fact the only completions of $\mathbb{Q}$ are $\mathbb{Q}_p$, for each prime p, and $\mathbb{R}$. For this reason we use the notation $\mathbb{Q}_\infty = \mathbb{R}$, considering the real numbers as the 'infinite–adic numbers'.

Define the topology on $\mathbb{Q}_p$ to be induced by the metric $d(x, y) = |x - y|$ and then we have the following.

Theorem 4.11.

a) $\mathbb{Z}_p$ *is an open subring of* $\mathbb{Q}_p$.
b) $\mathbb{Q}$ *is dense in* $\mathbb{Q}_p$.
c) *Let* $E/\mathbb{Q}_p$ *be an elliptic curve. Then* $E(\mathbb{Q}_p)$ *is compact.*

Proof.

a) First note that $x \in \mathbb{Z}_p$ if and only if $v_p(x) \geq 0$, that is if and only if $|x| \leq 1$. For $\mathbb{Z}_p$ open in $\mathbb{Q}_p$ we need for all $x \in \mathbb{Z}_p$ an $\epsilon > 0$ such that for all $y \in \mathbb{Q}_p$, $|x - y| < \epsilon$ implies $y \in \mathbb{Z}_p$.

For $x \in \mathbb{Z}_p$ take $\epsilon = 1$. If $|x| = |y|$ then $y \in \mathbb{Z}_p$ so assume $|x| \neq |y|$. Then $|x - y| = \max\{|x|, |y|\} < 1$ means that $|y| \leq 1$ so $y \in \mathbb{Z}_p$.

b) If $x \in \mathbb{Q}_p$ we need a sequence (x_n) with $x_n \in \mathbb{Q}$ such that $x_n \to x$. Now $x = p^m u$ for some $m \in \mathbb{N}$ and $u \in \mathbb{Z}_p^*$. Now $u \in \mathbb{Z}_p$ and $\mathbb{Z}$ is dense in $\mathbb{Z}_p$ so there exists a sequence (u_n) with $u_n \in \mathbb{Z}$ such that $u_n \to u$. Let $x_n = p^m u_n$ and then

$$|x - x_n| = |p^m(u_n - u)| = p^{-m}|u_n - u| \to 0 \text{ as } n \to \infty.$$

c) We give $\mathbb{P}^2(\mathbb{Q}_p)$ the quotient topology defined by the map $q\colon \mathbb{Q}_p^3 \setminus \{(0,0,0)\} \to \mathbb{P}^2(\mathbb{Q}_p)$ (the finest topology for which q is continuous). By multiplying by a suitable power of p we can choose a representative of a point in $\mathbb{P}^2(\mathbb{Q}_p)$ to be $(x_1 : x_2 : x_3)$ with $x_i \in \mathbb{Z}_p$ and at least one x_i not divisible by p (so $x_i \in \mathbb{Z}_p^*$). Note that $\mathbb{Z}_p^*$ is a closed subset of a compact set and is therefore compact. Thus

$$\mathbb{P}^2(\mathbb{Q}_p) = q(\mathbb{Z}_p^* \times \mathbb{Z}_p \times \mathbb{Z}_p) \cup q(\mathbb{Z}_p \times \mathbb{Z}_p^* \times \mathbb{Z}_p) \cup q(\mathbb{Z}_p \times \mathbb{Z}_p \times \mathbb{Z}_p^*)$$

is a union of continuous images of compact sets and is therefore compact. If we let $f(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$ then $E(\mathbb{Q}_p) = f^{-1}(\{\mathcal{O}\})$ is a preimage of a closed set under a continuous map and so is closed and therefore compact.

□

4.2. Hensel's lemma

We prove Hensel's lemma, following [**7**].

Lemma 4.12. *Let $f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$, and let $a \in \mathbb{Z}^n$ such that for some $m \geq 0$ and $r \geq 1$,*

$$f(a) \equiv 0 \pmod{p^{2m+r}}$$

but for some i,

$$\left(\frac{\partial f}{\partial X_i}\right)(a) \not\equiv 0 \pmod{p^{m+r}}.$$

Then there exists $b \in \mathbb{Z}^n$ such that

$$b \equiv a \pmod{p^{m+r}}$$

and

$$f(b) \equiv 0 \pmod{p^{2m+r+1}}.$$

Proof. We use the Taylor expansion

$$f(X) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial X_i}(a)(X_i - a_i) + \text{terms of higher degree in } (X_i - a_i).$$

Set $b_i = a_i + h_i p^{m+r}$ for some $h_i \in \mathbb{Z}$ to be determined. Then

$$f(b) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial X_i}(a) h_i p^{m+r} + \text{terms divisible by } p^{2m+2r}.$$

Now by assumption there exists $k \leq m + r - 1$ such that p^k divides $\frac{\partial f}{\partial X_i}(a)$ for all i but there exists some j such that p^{k+1} does not divide $\frac{\partial f}{\partial X_j}(a)$. Then p^{k+m+1} divides both $f(a)$ and $\frac{\partial f}{\partial X_i}(a)p^{m+r}$. For all $i \neq j$ set h_i to zero and using the fact that $p \nmid \frac{\partial f}{\partial X_j}(a)/p^k$ we can set $h_j = -\frac{f(a)}{p^{k+m+1}} / \frac{\frac{\partial f}{\partial X_j}(a)}{p^k} \pmod{p^{m-k+r}}$. Then

$$\frac{f(a)}{p^{k+m+1}} + \sum \frac{\frac{\partial f}{\partial X_i}(a)}{p^k} h_i \equiv 0 \pmod{p^{m-k+r}}$$

and

$$f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial X_i}(a) h_i p^{m+1} \equiv 0 \pmod{p^{2m+r+1}}.$$

Thus $f(b) \equiv 0 \pmod{p^{2m+r+1}}$. □

Theorem 4.13 (Hensel's Lemma)**.** *Let $f(X_1, \ldots, X_n) \in \mathbb{Z}[X_1, \ldots, X_n]$, and let $a \in \mathbb{Z}^n$ such that for some $m \geq 0$,*

$$f(a) \equiv 0 \pmod{p^{2m+1}}$$

but for some i,

$$\left(\frac{\partial f}{\partial X_i}\right)(a) \not\equiv 0 \pmod{p^{m+1}}.$$

Then there exists $b \in \mathbb{Z}_p^n$ such that $f(b) = 0$ and $b \equiv a \pmod{p^{m+1}}$.

Proof. Applying the above lemma with $r = 1$, we find a_{2m+2} such that $a_{2m+2} \equiv a \pmod{p^{m+1}}$ and $f(a_{2m+2}) \equiv 0 \pmod{p^{2m+2}}$. Applying the lemma again with $r = 2$ we find a_{2m+3} such that $a_{2m+3} \equiv a_{2m+2} \pmod{p^{m+2}}$ and $f(a_{2m+3}) \equiv 0 \pmod{p^{2m+3}}$. Continuing in this way we find a sequence $a, a_{2m+2}, a_{2m+3}, \ldots$. Write

$$a = a_{2m+1} = (a_{2m+1,1}, a_{2m+1,2}, \ldots, a_{2m+1,n})$$

and similarly for a_{2m+2} and so on. Then for $i = 1, 2, \ldots, n$ define

$$a_{(i)} = (\ldots, a_{2m+2,i}, a_{2m+1,i}) \in \mathbb{Z}_p.$$

Then define $b = (a_{(1)}, \ldots, a_{(n)})$ and we can see that $b \equiv a \pmod{p^{m+1}}$ and

$$f(b) = f\left(\lim_{r\to\infty} a_{2m+r}\right) = \lim_{r\to\infty} f(a_{2m+r}) = 0.$$

□

Example 4.14. Let $f(X) \in \mathbb{Z}[X]$ and let $\bar{f}(X) \in \mathbb{F}_p(X)$ be the polynomial formed by reducing it coefficients modulo p. Let $a \in \mathbb{Z}$ be such that $\bar{a} \in \mathbb{F}_p$ is a simple root of $\bar{f}(X)$. Then $\frac{d\bar{f}}{dX}(\bar{a}) \neq 0$ so $\frac{df}{dX}(a) \neq 0$ and by Hensel's Lemma there exists $b \in \mathbb{Z}_p$ such that $f(b) = 0$.

4.3. Krasner's lemma and $\mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$

We can show that $\overline{\mathbb{Q}}_p/\mathbb{Q}_p$ is a Galois extension and thus form $G_p = \mathrm{Gal}(\overline{\mathbb{Q}}_p/\mathbb{Q}_p)$. Later on we shall require an injective map from G_p into $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. To prove this map exists we require Krasner's Lemma. For this we require some development of absolute values on extensions of $\mathbb{Q}_p$.

Definition 4.15. *For a field K, a map $|\cdot|: K \to \mathbb{R}$ is called an* absolute value on K *if for all $x, y \in K$*

a) $|x| \geq 0$ *with equality if and only if $x = 0$,*
b) $|xy| = |x| \cdot |y|$ *and*
c) $|x + y| \leq |x| + |y|$.

Theorem 4.16. *Let K be complete with respect to an absolute value $|\cdot|_K$ and let L be a (possibly infinite) separable algebraic extension of K. Then $|\cdot|_K$ extends uniquely to an absolute value $|\cdot|_L$ on L. If L/K is a finite extension then L is complete with respect to the extended absolute value.*

Proof. See [**8**]. □

Lemma 4.17 (Krasner's Lemma)**.** *Let $\alpha, \beta \in \overline{\mathbb{Q}}_p$. If*

$$|\beta - \alpha| < |\alpha' - \alpha|$$

for all conjugates $\alpha' \neq \alpha$ of α then $\alpha \in \mathbb{Q}_p(\beta)$.

Proof. Let $K/\mathbb{Q}_p(\beta)$ be the normal closure of the field extension $\mathbb{Q}_p(\alpha, \beta)/\mathbb{Q}_p(\beta)$. Let $\tau \in \mathrm{Gal}(K/\mathbb{Q}_p(\beta))$. Now $\mathbb{Q}_p(\beta)$ is a finite extension of $\mathbb{Q}_p$ therefore has a complete extended valuation. Now the above theorem gives us an absolute value $|\cdot|$ on K and we can define another absolute value

on K by $|a|' = |\tau(a)|$. Now $|\cdot|' = |\cdot|$ on $\mathbb{Q}_p(\beta)$ so the above theorem tells us that $|\cdot| = |\cdot|'$ on K so $|\tau(a)| = |a|$ for all $a \in K$. Then

$$|\beta - \tau(\alpha)| = |\tau(\beta - \alpha)| = |\beta - \alpha| < |\alpha' - \alpha|.$$

Thus

$$\begin{aligned} |\alpha - \tau(\alpha)| &= |\alpha - \beta + \beta - \tau(\alpha)| \\ &\leq \max\{|\alpha - \beta|, |\beta - \tau(\alpha)|\} \\ &< |\alpha - \alpha'|. \end{aligned}$$

So we can see that since $\tau(\alpha)$ is a conjugate of α we must have $\tau(\alpha) = \alpha$ and thus $\alpha \in \mathbb{Q}_p(\beta)$. □

For polynomials $f = \sum a_i x^i \in K[x]$ define $||f|| = \max\{|a_i|\}$. This allows us to have a concept of closeness for polynomials. We prove the triangle inequality: if $f = \sum a_i x^i$ and $g = \sum b_i x^i$ then

$$\begin{aligned} ||f + g|| &= \max\{|a_i + b_i|\} \\ &\leq \max\{|a_i| + |b_i|\} \\ &\leq \max\{|a_i|\} + \max\{|b_i|\} \\ &= ||f|| + ||g|| \end{aligned}$$

Corollary 4.18. *Let $f(x)$ be a degree n monic irreducible polynomial in $\mathbb{Q}_p[x]$ and let α be a root of f. Then any monic degree n polynomial g sufficiently close to f is also irreducible and has a root β such that $\alpha \in \mathbb{Q}_p(\beta)$.*

Proof. Firstly note that if

$$g(x) = x^n + \sum_{i=1}^{n-1} b_i x^i$$

and β is a root of g then

$$\beta^n = -\sum_{i=1}^{n-1} b_i \beta^i$$

so $|\beta|^n \leq \max\{|b_j \beta^j|\}$. Thus $|\beta|^{n-j} \leq |c_j|$ for $j < n$ so $|\beta| \leq |b_j|^{1/(n-j)} \leq ||g||$.

Now

$$f(x) = \prod_{i=1}^{n} (x - \alpha_i).$$

Say $||f - g|| < \epsilon$ then

$$||g|| = ||f + g - f|| \leq ||f|| + ||g - f|| < ||f|| + \epsilon.$$

If β' is a root of g then writing $f = \sum a_i x^i$ we have

$$\begin{aligned} |f(\beta')| &= |(f-g)(\beta')| \\ &\leq \sum |a_i - b_i||\beta'|^i \\ &\leq \epsilon \max\{1, |\beta'|^n\} \\ &\leq \epsilon \max\{1, (||f|| + \epsilon)^n\}. \end{aligned}$$

We also have $|f(\beta')| = \prod |\beta' - \alpha_i|$ so we know that for at least one i,

$$|\beta' - \alpha_i| < \epsilon^{1/n} \max\{1, (||f|| + \epsilon)\}$$

(otherwise our bound is violated). So if we take ϵ sufficiently small we have $|\beta' - \alpha_i| < |\alpha_i - \alpha_j|$ for all $j \neq i$. Now let $K/\mathbb{Q}_p$ be the normal closure of $\mathbb{Q}_p(\alpha_i, \beta')/\mathbb{Q}_p$ there exists $\sigma \in \mathrm{Gal}(K/\mathbb{Q}_p)$ such that $\sigma(\alpha_i) = \alpha$. Then $\sigma(\beta') = \beta$ where β is some conjugate of β'. Then

$$\begin{aligned} |\beta - \alpha| &= |\sigma(\beta' - \alpha_i)| \\ &= |\beta' - \alpha_i| \\ &< |\alpha_i - \alpha_j| \quad \text{for all } j \\ &= |\sigma(\alpha_i - \alpha_j)| \\ &= |\alpha - \alpha_k| \quad \text{for all } k. \end{aligned}$$

Then by Krasner's lemma $\alpha \in \mathbb{Q}_p(\beta)$. □

Corollary 4.19. *If $\sigma \in G_p$ fixes $\mathbb{Q}_p$ and $\overline{\mathbb{Q}}$ then it fixes $\overline{\mathbb{Q}}_p$.*

Proof. Take $\alpha \in \overline{\mathbb{Q}}_p$. Then α has minimum polynomial

$$f(x) = \prod (x - \alpha_i) \in \mathbb{Q}_p[x].$$

Since $\mathbb{Q}$ is dense in $\mathbb{Q}_p$ we can choose a monic polynomial $g \in \mathbb{Q}[x]$ with rational coefficients arbitrarily close to f and a root $\beta \in \overline{\mathbb{Q}}$ such that $\alpha \in \mathbb{Q}_p(\beta)$. Then the condition that σ fixes $\mathbb{Q}_p$ and $\overline{\mathbb{Q}}$ implies that σ fixes α. □

Theorem 4.20. *The map from $G_p \to G$ defined by $\sigma \mapsto \sigma|_{\overline{\mathbb{Q}}}$ is well defined and injective.*

Proof. Every element of $\overline{\mathbb{Q}}$ has a minimum polynomial in $\mathbb{Q}[X]$ which by the definition of $\overline{\mathbb{Q}}$ splits in $\overline{\mathbb{Q}}$. Since Galois maps take elements to other roots of their minimum polynomial this means that

$$\sigma|_{\overline{\mathbb{Q}}} \colon \overline{\mathbb{Q}} \to \overline{\mathbb{Q}}$$

and thus our map from G_p to G is well defined.

For injectivity consider $\sigma, \tau \in \overline{\mathbb{Q}}_p$. Then both these maps fix $\mathbb{Q}_p$ so $\sigma\tau^{-1}$ fixes $\mathbb{Q}_p$. Now if $\sigma|_{\overline{\mathbb{Q}}} = \tau|_{\overline{\mathbb{Q}}}$ then $\sigma\tau^{-1}$ fixes $\overline{\mathbb{Q}}$. The above

corollary tells us that $\sigma\tau^{-1}$ fixes all of $\overline{\mathbb{Q}}_p$ so $\sigma = \tau$ and our map is injective. $\square$

Chapter 5

Elliptic curves over $\mathbb{Q}_p$

5.1. Reduction modulo p

Let $E/\mathbb{Q}_p$ be an elliptic curve with Weierstrass equation

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

We can clear out the denominators by making a change of variables

$$(x, y) \mapsto \left(\frac{x}{u^2}, \frac{y}{u^3}\right)$$

and thus find a Weierstrass equation with $a_i \in \mathbb{Z}_p$. Then $\Delta \in \mathbb{Z}_p$ since it is a combination of the a_i's. Then noticing that v_p maps onto $\mathbb{Z}$ we can find the equation with $v_p(\Delta)$ as small as possible.

Definition 5.1. *Let $E/\mathbb{Q}_p$ be an elliptic curve. A Weierstrass equation is called a* minimal equation for E at p *if $v_p(\Delta)$ is minimised subject to the condition that all $a_i \in \mathbb{Z}_p$.*

We denote the operation 'reduction modulo p' from $\mathbb{Z}_p \to \mathbb{Z}_p/p\mathbb{Z}_p \cong \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ by $a \mapsto \bar{a}$. Having chosen a minimal Weierstrass equation for $E/\mathbb{Q}_p$ we can reduce its coefficients modulo p to obtain a curve over $\mathbb{F}_p$:

$$\bar{E}\colon y^2 + \bar{a}_1xy + \bar{a}_3y = x^3 + \bar{a}_2x^2 + \bar{a}_4x + \bar{a}_6.$$

The curve $\bar{E}/\mathbb{Q}_p$ is called the *reduction of E modulo p.*

We can write a point $P \in E(\mathbb{Q}_p)$ as $P = (x_0\colon y_0\colon z_0)$ where $x_0, y_0, z_0 \in \mathbb{Z}_p$ and $\gcd(x_0, y_0, z_0) = 1$. Then we have a map 'reduction modulo p' from $E(\mathbb{Q}_p) \to \bar{E}(\mathbb{F}_p)$ defined by $P \mapsto \bar{P} = (\bar{x}_0\colon \bar{y}_0\colon \bar{z}_0)$.

Definition 5.2. *Let $E/\mathbb{Q}_p$ be an elliptic curve with reduction $\bar{E}/\mathbb{Q}_p$. We say E has* good reduction over $\mathbb{Q}_p$ *if $\bar{E}$ is non–singular.*

Theorem 5.3. *Let $E/\mathbb{Q}_p$ be an elliptic curve with discriminant Δ. If $p \nmid \Delta$ then E has good reduction over $\mathbb{Q}_p$.*

Proof. The discriminant is defined as a polynomial in the a_i's. Since reduction modulo p is a homorphism we can calculate the discriminant of $\bar{E}$ to be $\Delta \pmod p$. Thus $\bar{E}$ is nonsingular if $p \nmid \Delta$. □

We require one example of a curve that does not have good reduction.

Example 5.4. The plane projective curve

$$E_0 : Y^2Z = X^3$$

has a cusp at $S = (0 : 0 : 1)$. There exists an isomorphism $E_0(\mathbb{F}_p) \setminus \{S\} \to \mathbb{F}_p$.

Proof. See [**7**]. □

5.2. p–adic filtration

Let E have good reduction over $\mathbb{Q}_p$. We define groups

$$E(\mathbb{Q}_p) \subset E^1(\mathbb{Q}_p) \subset E^2(\mathbb{Q}_p) \subset \cdots$$

and consider their quotients.

Define

$$E^1(\mathbb{Q}_p) = \{P \in E(\mathbb{Q}_p) \mid \bar{P} = \bar{\mathcal{O}}\}.$$

We can see that $E^1(\mathbb{Q}_p)$ consists of $(x\colon y\colon z)$ such that p divides x and z but does not divide y.

Then go on to define

$$E^n(\mathbb{Q}_p) = \left\{P \in E^1(\mathbb{Q}_p) \mid \frac{x(P)}{y(P)} \in p^n\mathbb{Z}_p\right\}$$

for $n > 1$.

Theorem 5.5. *The filtration*

$$E(\mathbb{Q}_p) \subset E^1(\mathbb{Q}_p) \subset \cdots$$

has the following properties:

a) *The map $P \mapsto \bar{P}$ defines an isomorphism $E(\mathbb{Q}_p)/E^1(\mathbb{Q}_p) \to \bar{E}(\mathbb{F}_p)$.*
b) *For $n \geq 1$, $E^n(\mathbb{Q}_p)$ is a subgroup of $E(\mathbb{Q}_p)$ and the map $P \mapsto p^{-n}\frac{x(P)}{y(P)} \pmod p$ is an isomorphism $E^n(\mathbb{Q}_p)/E^{n+1}(\mathbb{Q}_p) \to \mathbb{F}_p$.*
c) *The filtration is exhaustive, that is $\cap_n E^n(\mathbb{Q}_p) = \{\mathcal{O}\}$.*

Proof.

a) We show that the sequence

$$0 \longrightarrow E^1(\mathbb{Q}_p) \longrightarrow E(\mathbb{Q}_p) \longrightarrow \bar{E}(\mathbb{F}_p) \longrightarrow 0$$

is exact. The reduction map $\mathbb{P}^2(\mathbb{Q}_p) \to \mathbb{P}^2(\mathbb{F}_p)$ maps lines to lines so the map $E(\mathbb{Q}_p) \to \bar{E}(\mathbb{F}_p)$ is a homomorphism.

We have inclusion $E^1(\mathbb{Q}_p) \to E(\mathbb{Q}_p)$ so the first map is injective. Exactness at $E(\mathbb{Q}_p)$ follows from the definition of $E^1(\mathbb{Q}_p)$ since $P \in E(\mathbb{Q}_p)$ maps to $\bar{\mathcal{O}}$ if and only if $P \in E^1(\mathbb{Q}_p)$.

To see surjectivity of the reduction map we use Hensel's Lemma. Let

$$f(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6 = 0$$

be a minimal Weierstrass equation, $\bar{f}(x,y)$ the corresponding polynomial reduced modulo p and $\bar{P} = (\alpha, \beta) \in \bar{E}_{ns}(\mathbb{F}_p)$. Since $\bar{P}$ is non–singular, we know either

$$\frac{\partial \bar{f}}{\partial x}(\bar{P}) \neq 0 \quad \text{or} \quad \frac{\partial \bar{f}}{\partial y}(\bar{P}) \neq 0.$$

Without loss of generality say the former is true. Choose $y_0 \in \mathbb{Z}_p$ with $\bar{y_0} = \beta$ and consider

$$f(x, y_0) = 0.$$

Reduced modulo p, this equation has α as a simple root since

$$\frac{\partial \bar{f}}{\partial x}(\alpha, \bar{y_0}) \neq 0.$$

Then by Hensel's Lemma we can lift α to an $x_0 \in \mathbb{Z}_p$ such that $\bar{x_0} = \alpha$ and $f(x_0, y_0) = 0$. Then $P = (x_0, y_0) \in E^0(\mathbb{Q}_p)$ reduces to $\bar{P}$.

b) Since the characteristic of $\mathbb{Q}_p$ is zero we can take our Weierstrass equation to be $y^2 = x^3 + ax + b$. We proceed by induction. We know that $E^1(\mathbb{Q}_p)$ is a subgroup of $E(\mathbb{Q}_p)$ so assume that $E^n(\mathbb{Q}_p)$ is a subgroup. Say $P = (x : y : z) = (x/z : y/z : 1) \in E^1(\mathbb{Q}_p)$. Say $x/z = p^{-m}x_0$ and $y/z = p^{-m'}y_0$ where $m, m' \in \mathbb{Z}$ and $x_0, y_0 \in \mathbb{Z}$. Then since P is on E we have

$$p^{-2m'}y_0^2 = p^{-3m}x_0^3 + ap^{-m}x_0 + b.$$

Taking the p–adic valuation we see that $-2m' = -3m$. Since m, m' are integers there must exist $k \in \mathbb{Z}$ such that $m = 2k$ and $m' = 3k$. In fact $v_p(x/y) = v_p(x/z) - v_p(y/z) = -2k - (-3k) = k$.

Now if $P = (x : y : z) \in E^n(\mathbb{Q}_p) \setminus E^{n+1}(\mathbb{Q}_p)$ then $k = v_p(x/y) = n$ so we can write

$$P = (p^{-2n}x_0 : p^{-3n}y_0 : 1) = (p^n x_0 : y_0 : p^{3n}).$$

In fact for any $P \in E^n(\mathbb{Q}_p)$ either the above is true or $P \in E^{n+1}(\mathbb{Q}_p)$ so

$$P = (p^{n+1}x_0' : y_0 : p^{3n+3}) = (p^n x_0 : y_0 : p^{3n}z_0).$$

Thus for all $P \in E^n(\mathbb{Q}_p)$ we can write $P = (p^n x_0 : y_0 : p^{3n} z_0)$ where $v_p(y_0) = 0$ and $x_0, z_0 \in \mathbb{Z}_p$.

Then P being on E implies that

$$p^{3n} y_0^2 z_0 = p^{3n} x_0^3 + a p^{7n} x_0 z_0^2 + b p^{9n} z_0^3.$$

Dividing by p^{3n} and defining $P_0 = (\bar{x_0} : \bar{y_0} : \bar{z_0})$ we see that P_0 lies on the curve $E_0/\mathbb{F}_p$ defined by

$$E_0 : Y^2 Z = X^3.$$

We have a map from $E^n(\mathbb{Q}_p) \to E_0(\mathbb{F}_p)$ given by $P \mapsto P_0$ that can be seen to be a homomorphism by the geometric group law (here we use our inductive hypothesis). We can see that $\mathrm{Ker}(P \mapsto P_0) = E^{n+1}(\mathbb{Q}_p)$ since $P = (p^n x_0 : y_0 : p^{3n} z_0) \in E^n(\mathbb{Q}_p)$ maps to $(0 : 1 : 0)$ if and only if $p \mid x_0$ and $p \mid z_0$ but $p \nmid y_0$. From this we see that $E^{n+1}(\mathbb{Q}_p)$ is a subgroup of $E(\mathbb{Q}_p)$.

To see that $\mathrm{Im}(P \mapsto P_0) = E_0(\mathbb{F}_p) \setminus \{S\}$ take a nonsingular point $(\bar{x_0} : \bar{y_0} : \bar{z_0})$ on $E_0(\mathbb{F}_p)$. By Hensel's lemma it lifts to a point $(x : y : z)$ on

$$p^{3n} y_0^2 z_0 = p^{3n} x_0^3 + a p^{7n} x_0 z_0^2 + b p^{9n} z_0^3$$

which is the image of a point $(p^n x_0 : y_0 : p^{3n}) \in E^n(\mathbb{Q}_p)$.

Then the composition

$$P \mapsto P_0 \mapsto \frac{x(P_0)}{y(P_0)}$$

given by

$$P \mapsto \frac{p^{-n} x(P)}{y(P)}$$

is the required isomorphism.

c) If $P = (x : y : z) \in \cap_{n=1}^{\infty} E^n(\mathbb{Q}_p)$ then $x/y \in p^n \mathbb{Z}_p$ for all $n \geq 1$ which implies that $x = 0$ and $y \neq 0$. This tells us that either $x = 0$ or $y^2 = bz^2$. However $P \in E^1(\mathbb{Q}_p)$ implies that $p \mid z$ and $p \nmid x$ but if $y^2 = bz^2$ then $p \mid z$ implies $p \mid x$. Thus $x = 0$ and $P = (x : y : z) = (0 : y : 0) = \mathcal{O}$.

□

Chapter 6

The theory of descent

A natural number–theoretic question is to discover the structure of an elliptic curve over the rationals, and more generally over a number field, K. We will eventually prove a special case of the following theorem, proved for $K = \mathbb{Q}$ by Mordell in 1922 and for general K (and also for abelian varieties, not just elliptic curves) by Weil in 1928.

Theorem 6.1 (Mordell-Weil Theorem)**.** *Let K be a number field and E/K an elliptic curve. Then the group $E(K)$ is finitely generated.*

We will see in Chapter 8 that to prove this we need the following result.

Theorem 6.2 (Weak Mordell-Weil Theorem)**.** *Let K be a number field, E/K an elliptic curve, and $m \geq 2$ an integer. Then $E(K)/mE(K)$ is a finite group.*

It is also true that if, for one m, we can find the generators for $E(K)/mE(K)$ then we have an effective procedure to find the generators for $E(K)$. Unfortunately an effective procedure to find generators for $E(K)/mE(K)$ is not known, although we will present a method that often works.

We begin by making the simplifying assumptions that $K = \mathbb{Q}$ and $m = 2$ so we are studying 2–descent for elliptic curves defined over the rational numbers.

6.1. The Kummer sequence

Firstly we set some notation. If A is an abelian group, let $A[2] = \{\, a \in A \mid 2a = 0 \,\}$ be the 2–torsion subgroup of A. Let $G = \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ and let $E/\mathbb{Q}$ be an elliptic curve. We use $\{x\}$ to denote the coset of x for elements of our cohomology groups.

To find 2–torsion points we have to solve $2P = \mathcal{O}$ which amounts to solving polynomials with rational coefficients. So we have the 2–torsion points $E(\overline{\mathbb{Q}})[2]$ and the rational 2–torsion points $E(\mathbb{Q})[2]$. Define the following sequence

$$0 \longrightarrow E(\overline{\mathbb{Q}})[2] \xrightarrow{f} E(\overline{\mathbb{Q}}) \xrightarrow{[2]} E(\overline{\mathbb{Q}}) \longrightarrow 0$$

where $f\colon P \mapsto P$ and $[2]\colon P \mapsto 2P$. We can easily see that f is injective so $\operatorname{Ker}(f) = 0$. Also for all $P \in E(\overline{\mathbb{Q}})$ we can find $Q \in E(\overline{\mathbb{Q}})$ such that $2Q = P$ by solving equations with rational coefficients. This means that $[2]$ is surjective so $\operatorname{Im}([2]) = E(\overline{\mathbb{Q}})$. Finally, if $P \in E(\overline{\mathbb{Q}})[2]$ then $[2]P = \mathcal{O}$ so $\operatorname{Im}(f) \subseteq \operatorname{Ker}([2])$. Also if $P \in \operatorname{Ker}([2])$ so $2P = \mathcal{O}$ then $P \in \operatorname{Im}(f)$. Therefore $\operatorname{Im}(f) = \operatorname{Ker}([2])$ and our sequence is exact.

Now we take Galois cohomology to find this long exact sequence

$$0 \longrightarrow H^0(G, E(\overline{\mathbb{Q}})[2]) \longrightarrow H^0(G, E(\overline{\mathbb{Q}})) \longrightarrow H^0(G, E(\overline{\mathbb{Q}}))$$

$$\xrightarrow{\delta} H^1(G, E(\overline{\mathbb{Q}})[2]) \longrightarrow H^1(G, E(\overline{\mathbb{Q}})) \longrightarrow H^1(G, E(\overline{\mathbb{Q}})).$$

Now drop the first two terms and note that $H^0(G, E(\overline{\mathbb{Q}})) = E(\overline{\mathbb{Q}})^G = E(\mathbb{Q})$. This gives us a long exact sequence

$$E(\mathbb{Q}) \xrightarrow{[2]} E(\mathbb{Q}) \xrightarrow{\delta} H^1(G, E(\overline{\mathbb{Q}})[2]) \xrightarrow{f'} H^1(G, E(\overline{\mathbb{Q}})) \xrightarrow{2'} H^1(G, E(\overline{\mathbb{Q}}))$$

where

$$[2]\colon P \mapsto 2P,$$
$$\delta\colon P \mapsto \{\sigma \mapsto N^\sigma - N\}$$

where we fix $N \in E(\overline{\mathbb{Q}})$ such that $2N = P$. Also

$$f'\colon \{\xi\} \mapsto \{\xi'\}$$

where ξ' is just an extension of the codomain of ξ (i.e. given $\xi\colon G \to E(\overline{\mathbb{Q}})[2]$ we define $\xi'\colon G \to E(\overline{\mathbb{Q}})$ by $\xi'(\sigma) = \xi(\sigma)$ for all $\sigma \in G$). Finally

$$2'\colon \{\xi\} \mapsto \{2\xi\}$$

where 2ξ is defined by $(2\xi)(\sigma) = 2(\xi(\sigma))$ for all $\sigma \in G$.

Now we can use the first isomorphism theorem to shorten this sequence. We know that $E(\mathbb{Q})/\operatorname{Ker}(\delta) \cong \operatorname{Im}(\delta)$ and since $\operatorname{Ker}(\delta) = \operatorname{Im}([2]) = 2E(\mathbb{Q})$ we can define a map $\mu\colon E(\mathbb{Q})/2E(\mathbb{Q}) \to H^1(G, E(\overline{\mathbb{Q}})[2])$ by

$$\mu\colon \{P\} \mapsto \delta(P).$$

On the other side we replace $H^1(G, E(\overline{\mathbb{Q}}))$ by $\operatorname{Ker}(2') = \operatorname{Im}(f') = H^1(G, E(\overline{\mathbb{Q}}))[2]$. Both these operations preserve the exactness and give us the following theorem.

Theorem 6.3. *Let $E/\mathbb{Q}$ be an elliptic curve. Then the following sequence, the* Kummer Sequence*, is exact.*

$$0 \longrightarrow E(\mathbb{Q})/2E(\mathbb{Q}) \xrightarrow{\mu} H^1(G, E(\overline{\mathbb{Q}})[2]) \xrightarrow{f'} H^1(G, E(\overline{\mathbb{Q}}))[2] \longrightarrow 0.$$

6.2. Principal homogenous spaces

We now give a geometric interpretation of the group $H^1(G, E(\overline{\mathbb{Q}}))$ by associating to each element a twist of E called a *homogenous space*. First we define a homogenous space and we will come back to showing the connection to cohomology. This section is heavily based on the exposition from [**12**]. Let K be a perfect field and $G = \mathrm{Gal}(\bar{K}/K)$.

Definition 6.4. *Let E/K be an elliptic curve. A* (principal) homogenous space for E/K *is a smooth curve C/K together with a morphism $\mu\colon C \times E \to C$ defined over K such that*

a) $\mu(p, \mathcal{O}) = p$ *for all* $p \in C$.
b) $\mu(\mu(p, P), Q) = \mu(p, P + Q)$ *for all* $p \in C$ *and* $P, Q \in E$.
c) *For all* $p, q \in C$ *there is a unique* $P \in E$ *satisfying* $\mu(p, P) = q$.

We denote $\mu(p, P)$ by $p \oplus P$. This means that b) is 'just' the associative law,

$$(p \oplus P) \oplus Q = p \oplus (P + Q),$$

but also distinguishes between addition on the elliptic curve, $+$, and the action of C on E, $\oplus$.

Because in c) we have a unique element P such that $p \oplus P = q$ we can define a *subtraction map* on C by

$$\nu\colon C \times C \to E$$

$$\nu(q, p) \text{ is the unique } P \in E \text{ such that } \mu(p, P) = q.$$

We will denote $\nu(q, p)$ by $q \ominus p$.

We can now prove that the $\oplus$ and $\ominus$ notations provide us with the correct intuition.

Lemma 6.5. *Let C/K be a homogenous space for E/K. Then for all $p, q \in C$ and $P, Q \in E$:*

a) $\mu(p, \mathcal{O}) = p$ *and* $\nu(p, p) = \mathcal{O}$.
b) $\mu(p, \nu(q, p)) = q$ *and* $\nu(\mu(p, P), p) = P$.
c) $\nu(\mu(q, Q), \mu(p, P)) = \nu(q, p) + (Q - P)$.

In other words:

a) $p \oplus \mathcal{O} = p$ *and* $p \ominus p = \mathcal{O}$.
b) $p \oplus (q \ominus p) = q$ *and* $(p \oplus P) \ominus p = P$.
c) $(q \oplus Q) \ominus (p \oplus P) = (q \ominus p) + (Q - P)$.

Proof.

a) Firstly, $\mu(p, \mathcal{O}) = p$ is part of the definition of homogenous space. Next note that $\nu(p, p)$ is the unique element that gives $\mu(p, \nu(p, p)) = p$ but we have $\mu(p, \mathcal{O}) = p$. So by uniqueness

$\nu(p,p) = \mathcal{O}$.

b) By the definition of the subtraction map, $\mu(p, \nu(q,p)) = q$. Replacing the first argument of ν with $\mu(p,P)$ we have

$$\mu(p, \nu(\mu(p,P),p)) = \mu(p,P).$$

Again by the uniqueness of ν we have $\nu(\mu(p,P),p) = P$.

c) By the definition of ν we have $q = \mu(p, \nu(q,p))$. Now adding Q to both sides we have

$$\mu(q,Q) = \mu(\mu(p,\nu(q,p),Q).$$

By b) in the definition of homogenous space we have

$$\mu(\mu(p,\nu(q,p),Q) = \mu(p,\nu(q,p)+Q).$$

Now add and subtract P and use the definition b) again so

$$\begin{aligned}\mu(q,Q) &= \mu(p, P+\nu(q,p)+Q-P) \\ &= \mu(\mu(p,P),\nu(q,p)+Q-P).\end{aligned}$$

Now by the uniqueness of ν we have

$$\nu(\mu(q,Q),\mu(p,P)) = \nu(q,p)+Q-P.$$

□

Next we can show that a homogenous space for E is always a twist of E (recall that means is isomorphic over $\bar{K}$).

Theorem 6.6. *Let E/K be an elliptic curve and C/K be a homogenous space for E/K. Fix a point $p_0 \in C$ and define a map*

$$\theta\colon E \to C \qquad \theta(P) = p_0 \oplus P.$$

Then:

a) *The map θ is an isomorphism defined over $K(p_0)$. In particular C/K is a twist of E/K.*
b) *For all $p \in C$ and $P \in E$*

$$p \oplus P = \theta(\theta^{-1}(p) + P).$$

c) *For all $p, q \in C$*

$$q \ominus p = \theta^{-1}(q) - \theta^{-1}(p).$$

d) *The subtraction map ν is a morphism defined over K.*

Proof.

a) Note that the action of E on C is, by definition, defined over K. This means that $(p \oplus P)^\sigma = p^\sigma \oplus P^\sigma$. To see that θ is defined over $K(p_0)$ note that $\mathrm{Gal}(\overline{K(p_0)}/K(p_0)) = \{\, \sigma \in G \mid p_0^\sigma = p_0 \,\}$. So for $\sigma \in \mathrm{Gal}(\overline{K(p_0)}/K(p_0))$, we have

$$\theta(P)^\sigma = (p_0 \oplus P)^\sigma = p_0^\sigma \oplus P^\sigma = p_0 \oplus P^\sigma = \theta(P^\sigma).$$

Since $\theta(P)^\sigma = \theta^\sigma(P^\sigma)$ this implies that $\theta^\sigma = \theta$ and θ is defined over $K(p_0)$. We can find $\theta^{-1}(q) = q \ominus p_0$. Then for all $q \in C, |\theta^{-1}(q)| = 1$ so θ has degree 1 and therefore is an isomorphism (using theorems 2.6c and 2.4.1 from **[12]**).

b) Since θ is an isomorphism it has an inverse. Then $\theta^{-1}(p)$ is the unique point of E that when added to p_0 gives p. Thus,

$$\theta(\theta^{-1}(p) + P) = (p_0 \oplus \theta^{-1}(p)) \oplus P = p \oplus P.$$

c) We add and subtract p_0 (recall this is actually using the earlier Lemma) to see,

$$\theta^{-1}(q) - \theta^{-1}(p) = (p_0 \oplus \theta^{-1}(q)) \ominus (p_0 \oplus \theta^{-1}(p)) = q \ominus p.$$

d) The subtraction map is a morphism since by (c), $q \ominus p = \theta^{-1}(q) - \theta^{-1}(p)$. Thus we can see that it is the composition of the morphisms θ^{-1}, addition on an elliptic curve and taking the negative of a point on an elliptic curve. To check that ν is defined over K, for any $\sigma \in G$,

$$(q \ominus p)^\sigma = (\theta^{-1}(q) - \theta^{-1}(p))^\sigma = \theta^{-1}(q)^\sigma - \theta^{-1}(p)^\sigma,$$

since subtraction on E is defined over K. Then, since μ is defined over K we can add and subtract p_0^σ and rearrange to

$$(q \ominus p)^\sigma = [p_0 \oplus \theta^{-1}(q)]^\sigma \ominus [p_0 \oplus \theta^{-1}(p)]^\sigma = q^\sigma \ominus p^\sigma$$

where the last equality is by definition of θ^{-1}. From this we can see that $\nu^\sigma = \nu$ for all $\sigma \in G$ and thus ν is defined over K.

□

Definition 6.7. *Two homogenous spaces C/K and C'/K for E/K are* equivalent *if there exists an isomorphism $\theta\colon C \to C'$ defined over K such that for all $p \in C$ and $P \in E$*

$$\theta(p \oplus P) = \theta(p) \oplus P.$$

We can easily check that this is an equivalence relation on the set of homogenous spaces.

Definition 6.8. *The collection of homogenous spaces of E, modulo equivalence, is called the* Weil–Chatalet group for E/K *and is denoted* $WC(E/K)$. *The equivalence class containing E is called the* trivial class.

Theorem 6.9. *Let C/K be a homogenous space for E/K. Then C/K is in the trivial class if and only if $C(K)$ is not empty.*

Proof. If C/K is in the trivial class then there exists an equivalence (K–isomorphism) $\theta\colon E \to C$. Now we claim that $\theta(\mathcal{O})$ is a K–rational point on C. For this we require $\theta(\mathcal{O})^\sigma = \theta(\mathcal{O})$ for all $\sigma \in G$. But

$$\theta(\mathcal{O})^\sigma = \theta^\sigma(\mathcal{O}^\sigma) = \theta(\mathcal{O})$$

since θ is defined over K and $\mathcal{O} \in E(K)$. Thus $\theta(\mathcal{O}) \in C(K)$ so $C(K)$ is nonempty.
Conversely, suppose $p_0 \in C(K)$ then

$$\theta\colon E \to C \quad \text{defined by} \quad \theta(P) = p_0 \oplus P$$

is an isomorphism over $K(p_0) = K$. For θ to be an equivalence we require for all $P, Q \in E$,

$$\theta(P+Q) = \theta(P) \oplus Q$$

$$\text{i.e. } p_0 \oplus (P+Q) = (p_0 \oplus P) \oplus Q.$$

But this is part of the definition of homogenous space so is satisfied. □

Our main aim in this section is to show a bijection between the Weil–Chatalet group and a cohomology group. The result we have just proved shows that this will give us a tool to help us study the diophantine problem of checking whether a curve has any rational points.

Lemma 6.10. *Let $\theta\colon C/K \to C'/K$ be an equivalence. Then for all $p, q \in C$*

$$\theta(q) \ominus \theta(p) = q \ominus p.$$

Proof. We add and subtract $p \ominus q$,

$$\theta(q) \ominus \theta(p) = ([\theta(q) \oplus (p \ominus q)] \ominus \theta(p)) + (q \ominus p).$$

Then by the compatibility of θ, $\theta(q) \oplus (p \ominus q) = \theta(q \oplus (p \ominus q)) = \theta(p)$. Thus

$$\theta(q) \ominus \theta(p) = (\theta(p) \ominus \theta(p)) + (q \ominus p) = \mathcal{O} + q \ominus p = q \ominus p.$$

□

Theorem 6.11. *Let E/K be an elliptic curve. Then there is a bijection*

$$WC(E/K) \to H^1(G, E(\bar{K}))$$

defined by choosing a point $p_0 \in C$ and then sending

$$\{C/K\} \mapsto \{\sigma \mapsto p_0^\sigma \ominus p_0\}.$$

Proof. Firstly note that $\{\sigma \mapsto p_0^\sigma \ominus p_0\}$ is a cocycle since

$$\begin{aligned} p_0^{\sigma\tau} \ominus p_0 &= (p_0^{\sigma\tau} \ominus p_0^\tau) + (p_0^\tau \ominus p_0) \\ &= (p_0^\sigma \ominus p_0)^\tau + (p_0^\tau \ominus p_0). \end{aligned}$$

To see that our map is well defined suppose C'/K is equivalent to C/K by a K–isomorphism $\theta\colon C \to C'$. Let $p_0' \in C'$. Then

$$\begin{aligned} p_0^\sigma \ominus p_0 &= \theta(p_0^\sigma) \ominus \theta(p_0) \\ &= (p_0'^\sigma \ominus p_0') + [(\theta(p_0) \ominus p_0')^\sigma - (\theta(p_0) \ominus p_0')]. \end{aligned}$$

Thus $p_0^\sigma \ominus p_0$ and $p_0'^\sigma \ominus p_0'$ are cohomologous as we required.

To see that our map is injective suppose $p_0^\sigma \ominus p_0$ and $p_0'^\sigma \ominus p_0'$ are cohomologous so there exists $P_0 \in E$ such that

$$p_0^\sigma \ominus p_0 = (p_0'^\sigma \ominus p_0') + (P_0^\sigma - P_0).$$

Define a map $\theta\colon C \to C'$ by

$$\theta(p) = p_0' \oplus (p \ominus p_0) + P_0.$$

Clearly θ is an isomorphism and $\theta(p \oplus P) = \theta(p) \oplus P$ for all $p \in C$ and $P \in E$. To see that θ is defined over K check

$$\begin{aligned} \theta(p)^\sigma &= p_0'^\sigma \oplus (p^\sigma \ominus p_0^\sigma) + P_0^\sigma \\ &= p_0' \oplus (p^\sigma \ominus p_0) + P_0 + [(p_0'^\sigma \ominus p_0') \oplus P_0^\sigma - P_0 - (p_0^\sigma \ominus p_0)] \\ &= \theta(p^\sigma) \end{aligned}$$

where the term in brackets is zero because $p_0^\sigma \ominus p_0$ and $p_0'^\sigma \ominus p_0'$ are cohomologous.

To see surjectivity let $\{\xi\} \in H^1(G, E(\bar{K})$. Now embed $E(\bar{K})$ in $\mathrm{Isom}(E)$ by sending $P \mapsto \tau_P$ where $\tau_P\colon Q \mapsto Q + P$. Consider $\{\xi\}$ as being in $H^1(G, \mathrm{Isom}(E))$ and then by Theorem 2.25 there exists a curve C/K and a $\bar{K}$–isomorphism $\phi\colon C \to E$ such that for all $\sigma \in G$

$$\phi^\sigma \circ \phi^{-1} = \tau_{-\xi_\sigma}.$$

Define $\mu\colon C \times E \to C$ by

$$\mu(p, P) = \phi^{-1}(\phi(p) + P).$$

We aim to show that C/K is a principal homogenous space over E/K with cohomology class $\{\xi\}$. Firstly we show that for $p, q \in C$ there is a unique $P \in E$ such that $\mu(p, P) = q$. If $\mu(p, P) = q$ then $\phi^{-1}(\phi(p) +$

$P) = q$ so $P = \phi(q) - \phi(p)$ is uniquely determined. To see that μ is defined over K, we let $\sigma \in G$ and compute

$$\begin{aligned}\mu(p,P)^\sigma &= (\phi^{-1})^\sigma(\phi^\sigma(p^\sigma) + P^\sigma) \\ &= \phi^{-1}([\phi(p^\sigma) - \xi(\sigma) + P^\sigma] + \xi(\sigma)) \\ &= \mu(p^\sigma, P^\sigma)\end{aligned}$$

Thirdly to find the cohomology class associated to C/K choose $p_0 = \phi^{-1}(\mathcal{O})$ (recall that we can choose *any* $p_0 \in C$). Then

$$\begin{aligned}p_0^\sigma - p_0 &= (\phi^\sigma)^{-1}(\mathcal{O}) - \phi^{-1}(\mathcal{O}) \\ &= \phi^{-1}(\mathcal{O} + \xi(\sigma)) - \phi^{-1}(\mathcal{O}) \\ &= \xi(\sigma)\end{aligned}$$

□

6.3. Selmer and Tate–Shafarevich group

We can carry out a similar process as in our earlier 2–descent theory but with the p–adic numbers, $\mathbb{Q}_p$, instead of $\mathbb{Q}$. Let $G_p = \mathrm{Gal}(\overline{\mathbb{Q}}_p, \mathbb{Q}_p)$ and for $p = 2, 3, 5, \ldots, \infty$ we have a short exact sequence

$$0 \longrightarrow E(\mathbb{Q}_p)/2E(\mathbb{Q}_p) \xrightarrow{\mu_p} H^1(G_p, E(\overline{\mathbb{Q}}_p)[2]) \xrightarrow{f'_p} H^1(G_p, E(\overline{\mathbb{Q}}_p))[2] \longrightarrow 0.$$

We have the inclusion $i\colon \mathbb{Q} \to \mathbb{Q}_p$ and by Krasner's Lemma we have an injective map $j\colon G_p \to G$. For a map $\xi\colon G \to E(\overline{\mathbb{Q}})$ we can use the following commutative diagram to define $\xi'\colon G_p \to E(\overline{\mathbb{Q}}_p)$.

$$\begin{array}{ccc} G & \xrightarrow{\xi} & E(\overline{\mathbb{Q}}) \\ {\scriptstyle j}\uparrow & & \downarrow{\scriptstyle i} \\ G_p & \xrightarrow{\xi'} & E(\overline{\mathbb{Q}}_p) \end{array}$$

Putting together our exact sequence for the rationals and the product of the exact sequence for each p we get the following commutative diagram with exact rows (all products are over $p = 2, 3, 5, \ldots, \infty$).

$$\begin{array}{ccccccccc} 0 & \longrightarrow & E(\mathbb{Q})/2E(\mathbb{Q}) & \xrightarrow{\mu} & H^1(G, E(\overline{\mathbb{Q}})[2]) & \xrightarrow{f'} & H^1(G, E(\overline{\mathbb{Q}}))[2] & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \theta_1} & & \downarrow{\scriptstyle \theta_2} & & \downarrow{\scriptstyle \theta_3} & & \\ 0 & \longrightarrow & \prod E(\mathbb{Q}_p)/2E(\mathbb{Q}_p) & \xrightarrow{\prod \mu_p} & \prod H^1(G_p, E(\overline{\mathbb{Q}}_p)[2]) & \xrightarrow{\prod f'_p} & \prod H^1(G_p, E(\overline{\mathbb{Q}}_p))[2] & \longrightarrow & 0 \end{array}$$

The new maps are defined by

$$(\theta_1)_p\colon \{P\} \mapsto \{P\},$$

$$(\theta_2)_p \colon \{\xi\} \mapsto \{j \circ \xi \circ i\}$$

and

$$(\theta_3)_p \colon \{\xi\} \mapsto \{j \circ \xi \circ i\}.$$

Now consider the map $H^1(G, E(\overline{\mathbb{Q}})) \to \prod H^1(G_p, E(\overline{\mathbb{Q}}_p))$ defined by $\{\xi\} \mapsto (\{j \circ \xi \circ i\})_{p=2}^{\infty}$. We call the kernel of this map the *Tate–Shafarevich group*, Ш. We can easily see that $Ш[2] = \mathrm{Ker}\,(\theta_3)$.

A useful way to think about the Tate–Shafarevich group is to recall that $H^1(G, E(\overline{\mathbb{Q}})) \cong WC(E/\mathbb{Q})$. Thus we can think of the elements of Ш as homogenous spaces that have a $\mathbb{Q}_p$–rational point for every prime p (that is are locally trivial). A nontrivial element of Ш is a homogenous space that is locally trivial but has no rational point.

Define the *2–Selmer group* by $S^{(2)} = \mathrm{Ker}\,(\theta_3 \circ f')$ and we can see that we have a short exact sequence (both these kernels are in the top row so just need to check that the maps fit together nicely). To summarise:

Theorem 6.12. *Let $E/\mathbb{Q}$ be an elliptic curve. Then the following sequence (the* descent sequence*) is exact.*

$$0 \longrightarrow E(\mathbb{Q})/2E(\mathbb{Q}) \longrightarrow S^{(2)} \longrightarrow Ш[2] \longrightarrow 0.$$

In the next section we will prove that $S^{(2)}$ is finite. This implies that $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite, the Weak Mordell–Weil Theorem. It also implies that $Ш[2]$ is finite. Although we can show that $Ш[m]$ is finite for all m it has not been proven that Ш is finite.

We can also see from this exact sequence that the Selmer group contains information on both of the groups we are interested in – there is an isomorphic copy of $E(\mathbb{Q})/2E(\mathbb{Q})$ inside $S^{(2)}$ and any point that is not in this copy maps to a nonzero point in $Ш[2]$. The hard part is finding out where a point in the Selmer group comes from.

6.4. Finiteness of the Selmer group

In this section we shall prove the finiteness of the Selmer group, $S^{(2)} \subseteq H^1(G, E(\overline{\mathbb{Q}})[2])$, following **[7]**. This proof is quicker than the standard proof in **[12]**.

Firstly we work up to a lemma proved in **[7]**, using material from reduction of elliptic curves over $\mathbb{Q}_p$ and a theorem from algebraic number theory.

Lemma 6.13. *Let $E/\mathbb{Q}_p$ be an elliptic curve with good reduction and let n be an integer such that p does not divide n. Then for $P \in E(\mathbb{Q}_p)$ there exists $Q \in E(\mathbb{Q}_p)$ such that $P = nQ$ if and only if $\bar{P} = n\bar{Q}$ for some $\bar{Q} \in E(\mathbb{F}_p)$.*

Proof. If $P = nQ$ then $\bar{P} = n\bar{Q}$.

Conversely, recall the filtration

$$E(\mathbb{Q}_p) \subset E^1(\mathbb{Q}_p) \subset E^2(\mathbb{Q}_p) \subset \cdots$$

where

$$\frac{E(\mathbb{Q}_p)}{E^1(\mathbb{Q}_p)} \cong \bar{E}(\mathbb{F}_p)$$

and, for all $n \geq 1$

$$\frac{E^n(\mathbb{Q}_p)}{E^{n+1}(\mathbb{Q}_p)} \cong \mathbb{F}_p.$$

So by assumption there exists a $Q_0 \in E(\mathbb{Q}_p)$ such that

$$nQ_0 \equiv P \pmod{E^1(\mathbb{Q}_p)}.$$

This means that $P - nQ_0 \in E^1(\mathbb{Q}_p)$. Then since

$$\frac{E^1(\mathbb{Q}_p)}{E^2(\mathbb{Q}_p)} \cong \mathbb{F}_p$$

and p does not divide n we can see that multiplication by n is an isomorphism, therefore there exists $Q_1 \in E^1(\mathbb{Q}_p)$ such that

$$p - nQ_0 \equiv nQ_1 \pmod{E^2(\mathbb{Q}_p)}.$$

Continuing on in this way we can find $Q_i \in E^i(\mathbb{Q}_p)$ such that

$$P - n\sum_{i=0}^{m} Q_i \in E^{m+1}(\mathbb{Q}_p).$$

Then by the compactness of $E(\mathbb{Q}_p)$,

$$\sum_{i=0}^{\infty} Q_i = Q \in E(\mathbb{Q}_p) \quad \text{and} \quad nQ = P.$$

□

Lemma 6.14. *For any finite extension k of $\mathbb{F}_p$, there exists an extension K of $\mathbb{Q}_p$ with the following properties:*

a) $[K : \mathbb{Q}_p] = [k : \mathbb{F}_p]$.
b) *The integral closure R of $\mathbb{Z}_p$ in K is a principal ideal domain with p as its only prime element (up to associates) and $R/pR = \mathbb{F}_p$.*

Lemma 6.13 remains true with $\mathbb{Q}_p$ replaced by such a K.

Proof. See **[5]** Theorem II.3.9 and Propositions II.2.4 and II.2.7. □

Lemma 6.15. *Let $E/\mathbb{Q}$ be an elliptic curve with discriminant Δ. For any $\gamma \in S^{(2)}$ and prime p_0 not dividing 2Δ there exists a finite unramified extension, K, of $\mathbb{Q}_{p_0}$ such that γ maps to zero in $H^1(G_{\bar{K}/K}, E(\bar{K})[2])$.*

Proof. For $\gamma \in S^{(2)} \subset H^1(G, E(\overline{\mathbb{Q}})[2])$, we can see by the diagram used to define $S^{(2)}$ that $\theta_2(\gamma)$ maps to zero in $H^1(G_p, E(\overline{\mathbb{Q}}_p))$. Then by exactness of the bottom row there exists a $\{P\} \in E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)$ such that $\{P\} \mapsto \theta_2(\gamma)$. Now E has good reduction at p so there exists a field K (by Lemma 6.14) such that $P \in 2E(K)$ (by Lemma 6.13). Then by the following commutative diagram

$$\begin{array}{ccc} E(\mathbb{Q})/2E(\mathbb{Q}) & \longrightarrow & H^1(G, E(\overline{\mathbb{Q}})[2]) \\ \downarrow & & \downarrow \\ E(\mathbb{Q}_p)/2E(\mathbb{Q}_p) & \longrightarrow & H^1(G_p, E(\overline{\mathbb{Q}}_p)[2]) \\ \downarrow & & \downarrow \\ E(K)/2E(K) & \longrightarrow & H^1(G_{\bar{K}/K}, E(\overline{K})[2]) \end{array}$$

we can see that $\gamma \in S^{(2)}$ maps to zero since $\{P\}$ is in the kernel of the vertical arrow on the left hand side. □

Now we can prove the finiteness of the Selmer group in a special case.

Theorem 6.16. *Let $E/\mathbb{Q}$ be an elliptic curve with $E(\overline{\mathbb{Q}})[2] \subseteq E(\mathbb{Q})$. Then $S^{(2)}$ is finite.*

Proof. We know that since $E(\overline{\mathbb{Q}})[2]$ is a group of order four with no element of order four that

$$E(\overline{\mathbb{Q}})[2] \cong \boldsymbol{\mu_2}^2.$$

as groups. Since all points of order two are rational points then G acts trivially on both sides so this actually an isomorphism of G–modules.

So,

$$H^1(G, E(\overline{\mathbb{Q}})[2]) \cong H^1(G, \boldsymbol{\mu_2} \times \boldsymbol{\mu_2}) \cong H^1(G, \boldsymbol{\mu_2}) \times H^1(G, \boldsymbol{\mu_2}).$$

Now by Theorem 1.17.b

$$H^1(G, \boldsymbol{\mu_2}) \cong \mathbb{Q}^*/\mathbb{Q}^{*2}.$$

So we can see that

$$H^1(G, E(\overline{\mathbb{Q}})[2]) \cong (\mathbb{Q}^*/\mathbb{Q}^{*2})^2.$$

Now take $\gamma \in S^{(2)}$. By multiplying by the square of a suitable rational number we can consider γ as mapping to

$$\left((-1)^{\epsilon(\infty)}\prod p^{\epsilon(p)}, (-1)^{\epsilon'(\infty)}\prod p^{\epsilon'(p)}\right)$$

where $\epsilon(p), \epsilon'(p) = 0$ or 1.

Now, by the above lemma, for any prime p_0 not dividing 2Δ there exists a finite unramified extension, K, of $\mathbb{Q}_{p_0}$ such that γ is in the kernel of the vertical arrows.

$$\begin{array}{ccc} H^1(G, E(\overline{\mathbb{Q}})[2]) & \xrightarrow{\cong} & (\mathbb{Q}^*/\mathbb{Q}^{*2})^2 \\ \downarrow & & \vdots \\ H^1(G_{\bar{K}/K}, E(\bar{K})[2]) & \xrightarrow{\cong} & (K^*/K^{*2})^2 \end{array}$$

We know that the element corresponding to γ is in the kernel of the map from $(\mathbb{Q}^*/\mathbb{Q}^{*2})^2 \to (K^*/K^{*2})^2$, so $(-1)^{\epsilon(\infty)}\prod p^{\epsilon(p)}$ is a square in K. This implies that $\epsilon(p_0) = 0$ since we know that p_0 is not a square in K. But we can see that $\operatorname{ord}_{p_0}((-1)^{\epsilon(\infty)}\prod p^{\epsilon(p)}) = \epsilon(p_0)$.

Therefore the order of any p_0 not dividing 2Δ is zero and thus $\gamma = \pm p_1^{\epsilon_1} p_2^{\epsilon_2} \dots p_n^{\epsilon_n}$ where $\epsilon_i = 0$ or 1 and the p_i are the primes that divide 2Δ. Thus there is only a finite number of choices for γ and $S^{(2)}$ must be finite. □

This approach can be extended to the case where the points of order two are not necessarily rational. We don't quite have the background to do this rigorously but if we take the extension of some preceding theorems on faith we can sketch the approach. Find a finite extension L of $\mathbb{Q}$ such that $E(\overline{\mathbb{Q}})[2] \subset E(L)$. The next lemma shows that to prove $S^{(2)}(E/\mathbb{Q})$ finite it suffices to show that $S^{(2)}(E/L)$ is finite ($S^{(2)}(E/L)$ is defined analogously to $S^{(2)}(E/\mathbb{Q})$).

Lemma 6.17. *For a finite Galois extension L of $\mathbb{Q}$ the kernel of*

$$S^{(2)}(E/\mathbb{Q}) \to S^{(2)}(E/L)$$

is finite.

Proof. It suffices to prove that the kernel of the map

$$\text{Res}\colon H^1(G_{\overline{\mathbb{Q}}/\mathbb{Q}}, E(\overline{\mathbb{Q}})[2]) \to H^1(G_{\bar{\mathbb{Q}}/L}, E(\bar{\mathbb{Q}})[2])$$

is finite.

First define a homomorphism

$$i\colon G_{\overline{\mathbb{Q}}/\mathbb{Q}} \to G_{L/\mathbb{Q}} \quad \text{by}$$
$$i\colon \sigma \mapsto \sigma|_L.$$

Now $\sigma|_L = 1$ if and only if $\sigma \in G_{\overline{\mathbb{Q}}/L}$. Thus $\mathrm{Ker}\,(i) = G_{\overline{\mathbb{Q}}/L} \lhd G_{\overline{\mathbb{Q}}/\mathbb{Q}}$. Also i is surjective since for all $\sigma \in G_{L/\mathbb{Q}}$ we can define a $\sigma' \in G_{\overline{\mathbb{Q}}/\mathbb{Q}}$ by

$$\sigma'(x) = \begin{cases} \sigma(x) & \text{if } x \in L \\ x & \text{if } x \notin L \end{cases}$$

such that $i(\sigma') = \sigma$. Then by the first isomorphism theorem

$$\frac{G_{\overline{\mathbb{Q}}/\mathbb{Q}}}{G_{\overline{\mathbb{Q}}/L}} \cong G_{L/\mathbb{Q}}.$$

So then the inflation–restriction exact sequence (Theorem 1.11) with $G = G_{\overline{\mathbb{Q}}/\mathbb{Q}}$, $H = G_{\overline{\mathbb{Q}}/L}$ and $M = E(\overline{\mathbb{Q}})[2]$ is

$$0 \longrightarrow H^1(G_{L/\mathbb{Q}}, E(L)[2]) \xrightarrow{\mathrm{Inf}} H^1(G_{\overline{\mathbb{Q}}/\mathbb{Q}}, E(\overline{\mathbb{Q}})[2]) \xrightarrow{\mathrm{Res}} H^1(G_{\overline{\mathbb{Q}}/L}, E(\overline{\mathbb{Q}})[2]).$$

Thus we see that $\mathrm{Ker}\,(\mathrm{Res}) = \mathrm{Inf}\,(H^1(G_{L/\mathbb{Q}}, E(L)[2]))$ and thus the finiteness of $G_{L/\mathbb{Q}}$ and $E(L)[2]$ implies that the kernel of Res is finite. □

The other thing that we need to emulate our proof in the special case is the following lemma.

Lemma 6.18. *Let S be a finite set of primes. Let N be the kernel of the map*

$$a \mapsto (ord_{\mathfrak{p}}(a) \pmod 2)\colon L^*/L^{*2} \to \oplus_{\mathfrak{p} \notin S}\mathbb{Z}/2\mathbb{Z}.$$

Then there is an exact sequence

$$0 \longrightarrow U_S/U_S^2 \longrightarrow N \longrightarrow C_S[2].$$

Proof. For $a \in \mathbb{O}_L$ define $v_{\mathfrak{p}}(a)$ for prime ideals $\mathfrak{p}$ of $\mathbb{O}_L$ by

$$\langle a \rangle = \prod_{\mathfrak{p}} \mathfrak{p}^{v_{\mathfrak{p}}(a)}.$$

Then extend this to $a/b \in L$ by $v_{\mathfrak{p}}(a/b) = v_{\mathfrak{p}}(a) - v_{\mathfrak{p}}(b)$. For a finite set S of prime ideals of L we have the following exact sequence (see appendix)

$$0 \longrightarrow U_S \longrightarrow L^* \xrightarrow{a \mapsto (v_{\mathfrak{p}}(a))} \oplus_{\mathfrak{p} \notin S}\mathbb{Z} \longrightarrow C_S \longrightarrow 0.$$

This gives us inclusion $U_S \to L^*$ which induces inclusion $U_S/U_S^2 \to L^*/L^{*2}$. In fact $U_S/U_S^2 \to N$ since $x \in U_S$ maps to zero in $\oplus_{\mathfrak{p} \notin S}\mathbb{Z}$. For $\alpha \in N$ we know that $2 \mid v_{\mathfrak{p}}(\alpha)$ so we can define the map to $C_S[2]$ by $\alpha \mapsto c = (v_{\mathfrak{p}}(\alpha))$. Then $2c = (v_{\mathfrak{p}}(\alpha)) = 0$ in C_S by the exact sequence.

For exactness we require $\mathrm{Ker}(N \to C_S[2]) = U_S/U_S^2$. If $\alpha \in \mathrm{Ker}(N \to C_S[2])$ then $\alpha \mapsto c = (v_{\mathfrak{p}}(\alpha)/2) = 0$. Thus there exists

$\beta \in L^*$ such that $v_{\mathfrak{p}}(\beta) = v_{\mathfrak{p}}(\alpha)/2$. Then $v_{\mathfrak{p}}(\alpha/\beta^2) = v_{\mathfrak{p}}(\alpha) - 2v_{\mathfrak{p}}(\beta) = 0$ so $\alpha/\beta^2 \in U_S$ and $\{\alpha\} \in U_S/U_S^2$ maps to c. □

From this we can see that

$$\frac{N}{U_S/U_S^2} \subseteq C_S[2].$$

We know (see appendix C) that U_S is finitely generated (so U_S/U_S^2 is finite) and C_S is finite. Thus a quotient of N by a finite group is contained in a finite group so N is finite.

Let S be the set of primes dividing Δ. Now since $G_{\overline{\mathbb{Q}}/L}$ acts trivially on $E(\overline{\mathbb{Q}})[2] \subset E(L)$,we have

$$H^1(G_{\overline{\mathbb{Q}}/L}, E(\overline{\mathbb{Q}})[2]) \cong H^1(G_{\overline{\mathbb{Q}}/L}, \boldsymbol{\mu_2})^2 \cong L^*/L^{*2}.$$

Then $\gamma \in S^{(2)}(E/L)$ corresponds to an element of N so $S^{(2)}(E/L)$ is finite.

Thus we have proven

Theorem 6.19 (Weak Mordell–Weil Theorem for $m = 2$ and $K = \mathbb{Q}$)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Then $E(\mathbb{Q})/2E(\mathbb{Q})$ is a finite group.*

Proof. By the descent sequence, there is an injective map from $E(\mathbb{Q})/2E(\mathbb{Q})$ to $S^{(2)}$. But we have just proved that $S^{(2)}$ is finite. Therefore $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite. □

Chapter 7

The Mordell–Weil theorem

7.1. The descent procedure

So far we know that $E(\mathbb{Q})/mE(\mathbb{Q})$ is finite. This is not enough to show that $E(\mathbb{Q})$ is finite. We need a concept of 'size' of points on an elliptic curve so that we can show that there are only a finite number of points of small size. In this section we give the conditions we would like this 'size' function (called a *height function*) to have and show that the existence of such a function implies the finiteness of $E(\mathbb{Q})$.

Theorem 7.1 (Descent Theorem)**.** *Let A be an abelian group. Suppose we have a function*

$$h\colon A \to \mathbb{R}$$

satisfying:

a) *Let $Q \in A$. There is a constant C_1, depending on A and Q, so that for all $P \in A$,*

$$h(P+Q) \leq 2h(P) + C_1.$$

b) *There is an integer $m \geq 2$ and a constant C_2, depending on A, such that for all $P \in A$,*

$$h(mP) \geq m^2 h(P) - C_2.$$

c) *For every constant C_3,*

$$\{P \in A \mid h(P) \leq C_3\}$$

is a finite set.

Then we call h a height function for A. *Further, suppose for the integer m in (ii), the quotient group A/mA is finite. Then A is finitely generated.*

Proof. Choose points $Q_1, \ldots, Q_r \in A$ to represent the finite number of cosets in A/mA. Let $P \in A$ and write

$$P = mP_1 + Q_{i_1} \quad \text{for some} 1 \leq i_1 \leq r.$$

Continue in this way writing

$$P_{n-1} = mP_n + Q_{i_n} \text{ for all } 1 \leq n \leq r.$$

Now for any j,

$$\begin{aligned} h(P_j) &\le \frac{1}{m^2}(h(mP_j) + C_2) \text{ from (b)} \\ &= \frac{1}{m^2}(h(P_{j-1} - Q_{i_j}) + C_2) \\ &\le \frac{1}{m^2}(2h(P_{j-1}) + C_1' + C_2) \text{ from (a)}, \end{aligned}$$

where we take $C_1' =$ maximum of the constants from (a) for $Q = -Q_i$ for $1 \le i \le r$. Now

$$\begin{aligned} h(P_n) &\le \frac{1}{m^2}\left(2h(P_{n-1}) + C_1' + C_2\right) \\ &\le \frac{2}{m^2}(\frac{1}{m^2}\left(2h(P_{n-2}) + C_1' + C_2\right) + \frac{1}{m^2}\left(C_1' + C_2\right) \\ &= \frac{4}{m^4}h(P_{n-2}) + \left[\frac{1}{m^2} + \frac{2}{m^4}\right](C_1' + C_2) \\ &\le \dots \\ &\le \left(\frac{2}{m^2}\right)^n h(P) + \left[\frac{1}{m^2} + \frac{2}{m^4} + \dots + \frac{2^{n-1}}{m^{2n}}\right](C_1' + C_2) \\ &\le \left(\frac{2}{m^2}\right)^n h(P) + \frac{C_1' + C_2}{m^2 - 2} \quad (\bigstar) \\ &\le 2^{-n}h(P) + (C_1' + C_2)/2 \text{ since } m \ge 2. \end{aligned}$$

The inequality ($\bigstar$) follows since

$$\begin{aligned} \frac{1}{m^2} + \frac{2}{m^4} + \dots + \frac{2^{n-1}}{m^{2n}} &\le \sum_{j=1}^{\infty} \frac{2^{j-1}}{m^{2j}} \\ &= \frac{1}{m^2 - 2}. \end{aligned}$$

Now taking n sufficiently large we have

$$h(P_n) \le 1 + (C_1' + C_2)/2.$$

From our definition of the P_i we can write

$$P = m^n P_n + \sum_{j=1}^{n} m^{j-1} Q_{i_j},$$

then every $P \in A$ is a linear combination of points in the set

$$\{Q_1, \dots, Q_r\} \cup \{Q \in A \mid h(Q) \le 1 + (C_1' + C_2)/2\}$$

which is finite by (c). □

7.2. Heights on $E(\mathbb{Q})$

Take an elliptic curve $E/\mathbb{Q}$. We have seen that we can fix a Weierstrass equation of the form

$$E\colon y^2 = x^3 + Ax + B$$

with $A, B \in \mathbb{Z}$. Now we define a height function on $E(\mathbb{Q})$.

Definition 7.2. *For $t \in Q$, write $t = p/q$ as a fraction in lowest terms. The* height *of t, denoted $H(t)$, is defined by*

$$H(t) = \max\{|p|, |q|\}.$$

The height on $E(\mathbb{Q})$ *is the function*

$$h_x\colon E(\mathbb{Q}) \to \mathbb{R}$$

$$h_x(P) = \begin{cases} \log H(x(P)) & \text{if } P \neq \mathcal{O} \\ 0 & \text{if } P = \mathcal{O} \end{cases}.$$

Now we show that this is a height function as we discussed in the last section.

Theorem 7.3.

a) *Let $P_0 \in E(\mathbb{Q})$. There is a constant C_1, depending on P_0, A, B, such that for all $P \in E(\mathbb{Q})$,*

$$h_x(P + P_0) \leq 2h_x(P) + C_1.$$

b) *There is a constant C_2, depending on A, B, such that for all $P \in E(\mathbb{Q})$,*

$$h_x(2P) \geq 4h_x(P) - C_2.$$

c) *For every constant C_3, the set*

$$\{P \in E(\mathbb{Q}) \mid h_x(P) \leq C_3\}$$

finite.

Proof.

a) We may assume $P_0 \neq \mathcal{O}$ and $P \neq \mathcal{O}, \pm P_0$ by taking $C_1 > \max\{h_x(P_0), h_x(2P_0)\}$. Then writing

$$P = (x, y) = \left(\frac{a}{d^2}, \frac{b}{d^3}\right) \quad P_0 = (x_0, y_0) = \left(\frac{a_0}{d_0^2}, \frac{b_0}{d_0^3}\right)$$

in lowest terms (we showed that you could do this in the proof of Theorem 5.5.b) the addition formula gives

$$\begin{aligned} x(P+P_0) &= \left(\frac{y-y_0}{x-x_0}\right)^2 - x - x_0 \\ &= \frac{(xx_0+A)(x+x_0)+2B-2yy_0}{(x-x_0)^2} \\ &= \frac{(aa_0+Ad^2d_0^2)(ad_0^2+a_0d^2)+2Bd^4d_0^4-2bdb_0d_0}{(ad_0^2-a_0d^2)^2}. \end{aligned}$$

Now cancellation only decreases the height and the denominator is an integer so

$$|x(P+P_0)| \le |(aa_0+Ad^2d_0^2)(ad_0^2+a_0d^2)+2Bd^4d_0^4-2bdb_0d_0|$$

and using the triangle inequality we can say that

$$H(x(P+P_0)) \le C_1' \max\{|a|^2, |d|^4, |bd|\},$$

where C_1' only depends on A, B, a_0, b_0, d_0. Since P is on the curve $y^2 = x^3 + Ax + B$ we see that

$$b^2 = a^3 + Aad^4 + Bd^6,$$

so

$$|b| \le C_1'' \max\{|a|^{3/2}, |d|^3\}.$$

Thus

$$H(x(P+P_0)) \le C_1''' \max\{|a|^2, |d|^4\} = C_1''' H(x(P))^2.$$

Taking logarithms gives

$$h_x(P+P_0) \le 2h_x(P) + C_1$$

where $C_1 = \log(C_1''')$ depends only on A, B, a_0, b_0, d_0.

b) We may assume that $2P \neq \mathcal{O}$ by choosing $C_2 \ge 4h_x(T)$ for each $T \in E(\mathbb{Q})[2]$. Then writing $P = (x, y)$ the duplication formula gives

$$x(2P) = \frac{x^4 - 2Ax^2 - 8Bx + A^2}{4x^3 + 4Ax + 4B}.$$

We define

$$\begin{aligned} F(X,Z) &= X^4 - 2AX^2Z^2 - 8BXZ^3 + A^2Z^4 \\ G(X,Z) &= 4X^3Z + 4AXZ^3 + 4BZ^4. \end{aligned}$$

Then if $x = a/b$ in lowest terms

$$x(2P) = \frac{F(a,b)}{G(a,b)}.$$

Now using Maple (see appendix) we can see that $\gcd(F, G) = 1$ and thus we can find $f_1, g_1, f_2, g_2 \in \mathbb{Q}[X, Z]$ such that

$$f_1(X,Z)F(X,Z) + g_1(X,Z)G(X,Z) = 4\Delta Z^7$$

$$f_2(X,Z)F(X,Z) + g_2(X,Z)G(X,Z) = 4\Delta X^7.$$

Using Maple we can find

$$\begin{aligned}
f_1(X,Z) &= 12X^2Z + 16AZ^3 \\
g_1(X,Z) &= -3X^3 + 5AXZ^2 + 27BZ^3 \\
f_2(X,Z) &= 4(4A^3 + 27B^2)X^3 - 4A^2BX^2Z \\
&\quad + 4A(A^3 + 22B^2)XZ^2 + 12B(A^3 + 8B^2)Z^3 \\
g_2(X,Z) &= A^2BX^3 + A(5A^3 + 32B^2)X^2Z \\
&\quad + 2B(13A^3 + 96B^2)XZ^2 - 3A^2(A^3 + 8B^2)Z^3.
\end{aligned}$$

Let $\delta = \gcd(F(a,b), G(a,b))$ then from the above equations we can see that δ divides $4\Delta b^7$ and $4\Delta a^7$ so since $\gcd(a,b) = 1$ we see that $\delta \mid 4\Delta$. Therefore $|\delta| \leq |4\Delta|$, and so

$$H(x(2P)) \geq \frac{\max\{F(a,b), G(a,b)\}}{|4\Delta|}.$$

The same identities give

$$|4\Delta b^7| \leq 2\max\{f_1(a,b), g_1(a,b)\}\max\{F(a,b), G(a,b)\},$$

$$|4\Delta a^7| \leq 2\max\{f_2(a,b), g_2(a,b)\}\max\{F(a,b), G(a,b)\}.$$

Also, looking at the definition of f_1, f_2, g_1 and g_2 we see that

$$\max\{f_1(a,b), g_1(a,b), f_2(a,b), g_2(a,b)\} \leq C\max\{|a|^3, |b|^3\},$$

where C is a constant depending on A and B. Combining our inequalities we have

$$\begin{aligned}
&\max\{|4\Delta a^7|, |4\Delta b^7|\} \\
&\leq 2C\max\{|a|^3, |b|^3\}\max\{F(a,b), G(a,b)\}
\end{aligned}$$

Now cancelling $\max\{|a|^3, |b|^3|\}$ and noting $\max\{|a|^4, |b|^4\} \geq \max\{|a|, |b|\}$ we see that

$$\frac{\max\{F(a,b), G(a,b)\}}{|4\Delta|} \geq (2C)^{-1}\max\{|a|, |b|\}.$$

Since $H(x(P)) = \max\{|a|, |b|\}$ we have

$$H(x(2P)) \geq (2C)^{-1}H(x(P)).$$

c) For a constant C the set

$$\{t \in \mathbb{Q} \mid H(t) \leq C\}$$

is finite. Given any x there are at most two values of y for which (x, y) is a point of E. Therefore

$$\{P \in E(\mathbb{Q}) \mid h_x(P) \leq C_3\}$$

is also finite.

□

From this we can now prove:

Theorem 7.4 (Mordell–Weil Theorem over $\mathbb{Q}$)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. Then the group $E(\mathbb{Q})$ is finitely generated.*

Proof. Firstly $E(\mathbb{Q})$ is an abelian group. By Theorem 6.19 $E(\mathbb{Q})/2(E\mathbb{Q})$ is finite and we have just seen that there is a height function h_x on $E(\mathbb{Q})$. So by Theorem 7.1, $E(\mathbb{Q})$ is finitely generated. □

7.3. Rank of an elliptic curve

From the Mordell–Weil theorem and the fundamental theorem of finitely generated abelian groups we can conclude that

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

for a nonnegative integer r, called the *rank* of E.

This implies that

$$\frac{E(\mathbb{Q})}{2E(\mathbb{Q})} \cong \frac{E_{\text{tors}}(\mathbb{Q})}{2E_{\text{tors}}(\mathbb{Q})} \oplus \frac{\mathbb{Z}^r}{2\mathbb{Z}^r}$$

and since $\mathbb{Z}^r/2\mathbb{Z}^r \cong (\mathbb{Z}/2\mathbb{Z})^r$ we can then calculate

$$2^r = \frac{|E(\mathbb{Q})/2E(\mathbb{Q})|}{|E_{\text{tors}}(\mathbb{Q})/2E_{\text{tors}}(\mathbb{Q})|}.$$

Then the descent sequence gives us

$$2^r = \frac{\left|S^{(2)}\right|}{|E_{\text{tors}}(\mathbb{Q})/2E_{\text{tors}}(\mathbb{Q})| \times |Ш[2]|} \leq \frac{\left|S^{(2)}\right|}{|E_{\text{tors}}(\mathbb{Q})/2E_{\text{tors}}(\mathbb{Q})|}.$$

Thus we can use the Selmer group to bound the rank, and if we can find the Tate–Shafarevich group, calculate it exactly.

It is conjectured that there are elliptic curves of arbitrarily large rank. As of 2006 the highest rank curve known is at least rank 28 (see

[**2**]), with Weierstrass equation

$$y^2 + xy + y = x^3 - x^2 - 2006776241557552658503$$
$$32082093385427509302303121789565 02x +$$
$$3448161179503055646703298569039072037485594435931918 0$$
$$3612660082962919394487322434 29.$$

Chapter 8

Complete 2–descent in practice

The process of 2–descent is calculating generators for $E(\mathbb{Q})/2E(\mathbb{Q})$ when $Ш[2] = 0$ (if $Ш[2] \neq 0$ then our methods only give us a bound for the rank although it may be possible to carry out a higher descent). There are three situations with different methods for each.

a) Complete 2-descent. Used when $E(\overline{\mathbb{Q}})[2] \subseteq E(\mathbb{Q})$.
b) Descent via two-isogeny. Used when $E(\mathbb{Q})[2] \neq 0$.
c) General 2–descent. Used for any $E(\overline{\mathbb{Q}})[2]$.

We only study the first method, giving the standard approach based on homogenous spaces and an alternative due to Flynn.

8.1. Complete 2–descent

Firstly say that $E(\mathbb{Q})[2] = \{\mathcal{O}, (e_1, 0), (e_2, 0), (e_3, 0)\}$ so $y^2 = (x - e_1)(x - e_2)(x - e_3)$. Define $S = \{p \mid p \text{ divides } 2\Delta\}$ and $\mathbb{Q}(S, 2) = \{b \in \mathbb{Q}^*/\mathbb{Q}^{*2} \mid v_p(b) = 0 \mod 2 \text{ for all } p \notin S\}$. Thus $\mathbb{Q}(S, 2)$ is the elements that can be written as $\pm p_1^{\epsilon_1} p_2^{\epsilon_2} \dots p_n^{\epsilon_n}$ where $\epsilon_i = 0$ or 1 and the p_i are the primes that divide 2Δ. Note that we have the following commutative diagram where f and α are isomorphisms and δ is an injective homomorphism:

$$\begin{array}{ccc} E(\mathbb{Q})/2E(\mathbb{Q}) & \xrightarrow{\phi} & (\mathbb{Q}^*/\mathbb{Q}^{*2})^2 \\ \downarrow \delta & & \downarrow \alpha \\ H^1(G, E(\overline{\mathbb{Q}})[2]) & \xrightarrow[f]{} & H^1(G, \boldsymbol{\mu_2})^2 \end{array}$$

with the maps defined as follows. Define

$$\delta \colon \{P\} \mapsto \{\sigma \mapsto M^\sigma - M\}$$

where we fix $M \in E(\overline{\mathbb{Q}})$ such that $2M = P$.

Next

$$f \colon \{\xi\} \mapsto (\{\xi_1\}, \{\xi_2\})$$

where

$$\xi_1(\sigma) = \begin{cases} 1 & \text{if } \xi(\sigma) = \mathcal{O} \text{ or } (e_1, 0) \\ -1 & \text{if } \xi(\sigma) = (e_2, 0) \text{ or } (e_3, 0) \end{cases}$$

and

$$\xi_2(\sigma) = \begin{cases} 1 & \text{if } \xi(\sigma) = \mathcal{O} \text{ or } (e_2, 0) \\ -1 & \text{if } \xi(\sigma) = (e_1, 0) \text{ or } (e_3, 0) \end{cases}.$$

Finally

$$\alpha \colon \{b\} \mapsto \left\{ \sigma \mapsto \frac{\sqrt{b}^{\sigma}}{\sqrt{b}} \right\}.$$

Now looking at the above proof we see that $\phi = \alpha^{-1} \circ f \circ \delta$ is an injective homomorphism into $\mathbb{Q}(S,2)^2$. If we could describe ϕ explicitly then we could give an algorithm to find $E(\mathbb{Q})/2E(\mathbb{Q})$. However finding α^{-1} explicitly seems hard. Instead, from **[12]** we can take the following explicit description which can be proved to be an injective homomorphism by direct computation.

$$\phi(x,y) = \begin{cases} (x - e_1, x - e_2) & \text{if } x \neq e_1, e_2 \\ (\frac{e_1 - e_3}{e_1 - e_2}, e_1 - e_2) & \text{if } x = e_1 \\ (e_2 - e_1, \frac{e_2 - e_3}{e_2 - e_1}) & \text{if } x = e_2 \\ (1,1) & \text{if } (x,y) = \mathcal{O} \end{cases}.$$

From this we can find the image of the points in $E(\mathbb{Q})[2]$. For $(b_1, b_2) \in \mathbb{Q}(S,2)^2$ that are not the image of one of these points, to check if they are the image of any point in $E(\mathbb{Q})/2E(\mathbb{Q})$ we need to find $(x,y) \in E(\mathbb{Q})$ and $z_1, z_2 \in \mathbb{Q}^*$ such that $\phi(x,y) = (b_1 z_1^2, b_2 z_2^2)$. This gives us the equations

$$b_1 z_1^2 = x - e_1, \quad b_2 z_2^2 = x - e_2 \text{ and}$$
$$y^2 = (x - e_1)(x - e_2)(x - e_3).$$

Now let $z_3 = y/(b_1 b_2 z_1 z_2)$ so

$$z_3^2 = \frac{y^2}{b_1^2 b_2^2 z_1^2 z_2^2} = \frac{b_1 z_1^2 b_2 z_2^2 (x - e_3)}{b_1^2 b_2^2 z_1^2 z_2^2} = \frac{x - e_3}{b_1 b_2}.$$

Thus we now have another equation $b_1 b_2 z_3^2 = x - e_3$. This can be combined with the two similar equations to give us

$$b_1 z_1^2 - b_2 z_2^2 = e_2 - e_1 \text{ and}$$
$$b_1 z_1^2 - b_1 b_2 z_3^2 = e_3 - e_1.$$

We can see that (b_1, b_2) is the image of a point in $E(\mathbb{Q})/2E(\mathbb{Q})$ if and only if we can find solutions $(z_1, z_2, z_3) \in \mathbb{Q}^* \times \mathbb{Q}^* \times \mathbb{Q}$. If we can find a solution then working backwards we have a point $(x,y) = (b_1 z_1^2 + e_1, b_1 b_2 z_1 z_2 z_3)$ on our curve E.

The equations $b_1 z_1^2 - b_2 z_2^2 = e_2 - e_1$ and $b_1 z_1^2 - b_1 b_2 z_3^2 = e_3 - e_1$ describe a curve in $\mathbb{P}^3$ which is actually the homogenous space corresponding to the pair (b_1, b_2).

So we have the following process for carrying out complete 2–descent. It is not a true algorithm as sometimes we can't be sure whether some of these equations have solutions.

Algorithm 8.1 (Complete 2–descent)**.** Let $E/\mathbb{Q}$ be an elliptic curve given by

$$y^2 = (x-e_1)(x-e_2)(x-e_3) \quad \text{with } e_1, e_2, e_3 \in \mathbb{Q}.$$

a) Calculate the discriminant, Δ, and find $\mathbb{Q}(S,2) = \{\pm p_1^{\epsilon_1} p_2^{\epsilon_2} \dots p_n^{\epsilon_n} \mid \epsilon_i = 0 \text{ or } 1 \text{ and } p_i \text{ divides } 2\Delta\}$.
b) Find the images of $\mathcal{O}, (e_1,0), (e_2,0), (e_3,0)$.
c) For all $(b_1,b_2) \in \mathbb{Q}(S,2)^2$ that are not one of these images check if it is the image of a point in $E(\mathbb{Q})/2E(\mathbb{Q})$, that is find solutions of the above equations. We have various tools for this:
   (i) Check if it has a solution over $\mathbb{R}$.
   (ii) Check if it has a solution over $\mathbb{Q}_p$.
   (iii) Use the homomorphism property of the map. Adding a point on the curve corresponds to multiplying by its (b_1,b_2).

8.2. Examples

Firstly we study an elliptic curve that we can perform descent upon successfully. This example is exercise 10.13 in [**12**].

Example 8.2. Let $E/\mathbb{Q}$ be given by $y^2 = x(x-1)(x+3) = x^3+2x^2-3x$.

We have seen that $\Delta = 2304 = 2^8 \times 3^2$ and that

$$E_{\text{tors}}(\mathbb{Q}) = \{\mathcal{O}, (-3,0), (0,0), (1,0), (3,6), (3,-6), (-1,2), (-1,-2)\} \cong \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

Following our algorithm we set $S = \{2,3\}$ and $\mathbb{Q}(S,2) = \{\pm 1, \pm 2, \pm 3, \pm 6\}$. Next we check where the points of order two map to. We have

$$\mathcal{O} \mapsto (1,1) \quad (0,0) \mapsto (-3,-1)$$
$$(1,0) \mapsto (1,4) = (1,1) \quad (-3,0) \mapsto (-3,-4) = (-3,-1).$$

Note that some of our points map to the same (b_1,b_2). This doesn't violate the injectivity of ϕ because they differ by an element of $2E(\mathbb{Q})$. For example $(-3,0) = (0,0) + (1,0)$ and $(1,0) = 2(3,6)$.

Now we list how we found or ruled out solutions, for each $(b_1,b_2) \in \mathbb{Q}(S,2)^2$, to the equations

$$b_1 z_1^2 - b_2 z_2^2 = 1 \tag{8.3}$$
$$b_1 z_1^2 - b_1 b_2 z_3^2 = -3 \tag{8.4}$$

a) Have $\mathcal{O} \mapsto (1,1)$ and $(0,0) \mapsto (-3,-1)$.
b) If $b_1 < 0$ and $b_2 > 0$ then (8.3) has no solution in $\mathbb{R}$.

c) If $b_1 > 0$ and $b_2 < 0$ then (8.4) has no solution in $\mathbb{R}$.
d) If b_1 is even then (8.4) gives us $0 \equiv 1 \pmod 2$ so no solution in $\mathbb{Q}_2$.
e) If $b_1 \equiv 0 \pmod 3$ and $b_2 \equiv 0 \pmod 3$ then (8.3) gives us $0 \equiv 1 \pmod 3$ so no solution in $\mathbb{Q}_3$.
f) If $b_1 = 1$ and $b_2 = 2$ then we have equations $z_1^2 - 2z_2^2 = 1$ and $z_1^2 - 2z_3^2 = -3$. Subtracting we have $z_3^2 - z_2^2 = 2$. By Gauss' lemma any solution will be in integers and we know that there are no integers whose squares differ by 2. Therefore there are no solutions.
g) If $b_1 = 1$ and $b_2 \equiv 0 \pmod 3$ then (8.3) gives us $z_1^2 \equiv 1 \pmod 3$ and (8.4) gives us $z_1^2 \equiv 0 \pmod 3$, thus there is no solution in $\mathbb{Q}_3$.
h) If $b_1 = 3$ and $b_2 = 1$ then we can see that (8.3) gives us $z_2^2 \equiv 2 \pmod 3$ which has no solution, thus there is no solution in $\mathbb{Q}_3$.
i) If $b_1 = 3$ and $b_2 = 2$ then we have the equations $3z_1^2 - 2z_2^2 = 1$ and $3z_1^2 - 6z_3^2 = -3$ which has the solution $z_1 = 1, z_2 = 1, z_3 = 1$. This gives us the point on the elliptic curve $(b_1 z_1^2 + e_1, b_1 b_2 z_1 z_2 z_3) = (3, 6)$.
j) If $b_1 = -1$ and $b_2 = -1$ we have $-z_1^2 + z_2^2 = 1$ and $-z_1^2 - z_3^2 = -3$. By Gauss' lemma any solution will be in integers and we know that the only integers whose squares differ by 1 are 1 and 0. However we require $z_1, z_2 \in \mathbb{Q}^*$ so there are no solutions.
k) If $b_1 = -1$ and $b_2 = -2$ then we have the equations $-z_1^2 + 2z_2^2 = 1$ and $-z_1^2 - 3z_3^2 = -3$ which has the solution $z_1 = 1, z_2 = 1, z_3 = 1$. This gives us the point on the elliptic curve $(-1, -2)$.
l) If $b_1 = -1$ and $b_2 \equiv 0 \pmod 3$ then we can see that (8.3) gives us $z_1^2 \equiv 2 \pmod 3$ which has no solution, thus there is no solution in $\mathbb{Q}_3$.
m) If $b_1 = -3$ and $b_2 = -2$ then we can see that (8.3) gives us $2z_2^2 \equiv 1 \pmod 3$ which has no solution, thus there is no solution in $\mathbb{Q}_3$.

Now this tells us that, choosing representatives, $E(\mathbb{Q})/2E(\mathbb{Q}) = \{\mathcal{O}, (0,0), (3,6), (-1,2)\}$. Then

$$2^r = \frac{|E(\mathbb{Q})/2E(\mathbb{Q})|}{E_{\text{tors}}(\mathbb{Q})/2E_{\text{tors}}(\mathbb{Q})|}\frac{4}{4} = 1.$$

Thus $r = 0$ and $E(\mathbb{Q}) = E_{\text{tors}}(\mathbb{Q})$. We can also see that all our locally trivial homogenous spaces had rational points so $Ш[2] = 0$.

Our next example is the first homogenous space that we see that is locally trivial but does not have a rational point, that is, a nontrivial element of the Tate–Shafarevich group.

Example 8.5. The elliptic curve $X^3+Y^3+60Z^3 = 0$ has a homogenous space $3X^3 + 4Y^3 + 5Z^3 = 0$. This curve has a point in $\mathbb{Q}_p$ for every p but has no rational point.

Proof. To see that $X^3 + Y^3 + 60Z^3 = 0$ is an elliptic curve note that $(1, -1, 0)$ is a rational point. For the proof that it has homogenous space $3X^3 + 4Y^3 + 5Z^3 = 0$ see **[6]**. That it has a point for every $\mathbb{Q}_p$ is exercise 10.12 in **[12]**. The hard part of the proof, proving there is no rational point, requires arithmetic in $\mathbb{Q}(\sqrt[3]{60})$ and can be found in **[1]**. □

8.3. Complete 2–descent without homogenous spaces

In this section we discuss an approach due to Flynn of performing complete 2–descent without homogenous spaces. We need the following fact proven using formal groups in chapter 7 of **[3]**.

Theorem 8.6. *Let $E/\mathbb{Q}_p$ be an elliptic curve. Then*

$$\left|\frac{E(\mathbb{Q}_p)}{2E(\mathbb{Q}_p)}\right| = \frac{|E(\mathbb{Q}_p)[2]|}{|2|_p}$$

where $|\cdot|_p$ is the p–adic valuation.

Since we are interested in complete 2–descent we know that $E(\overline{\mathbb{Q}})[2] \subseteq E(\mathbb{Q})$ so $|E(\mathbb{Q}_p)[2]| = 4$ and we have

$$\left|\frac{E(\mathbb{Q}_p)}{2E(\mathbb{Q}_p)}\right| = \begin{cases} 2 & \text{if } p = \infty \\ 4 & \text{if } p \neq 2, \infty \\ 8 & \text{if } p = 2 \end{cases}$$

Now consider the diagram

$$\begin{array}{ccc} E(\mathbb{Q})/2E(\mathbb{Q}) & \xrightarrow{\phi} & \mathbb{Q}(S,2)^2 \\ \downarrow{\scriptstyle i_p} & & \downarrow{\scriptstyle j_p} \\ E(\mathbb{Q}_p)/2E(\mathbb{Q}_p) & \xrightarrow{\phi_p} & (\mathbb{Q}_p^*/\mathbb{Q}_p^{*2})^2. \end{array}$$

where

$$\phi(x,y) = \begin{cases} (x - e_1, x - e_2) & \text{if } x \neq e_1, e_2 \\ (\frac{e_1-e_3}{e_1-e_2}, e_1 - e_2) & \text{if } x = e_1 \\ (e_2 - e_1, \frac{e_2-e_3}{e_2-e_1}) & \text{if } x = e_2 \\ (1,1) & \text{if } (x,y) = \mathcal{O} \end{cases}$$

and i_p and j_p are induced by $\mathbb{Q} \to \mathbb{Q}_p$ (they are not necessarily injective).

To understand the groups $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$ we need a little more work.

Theorem 8.7. *A p–adic number $x \in \mathbb{Q}_p^*$ is a square if and only if*

a) $p = \infty$ *case:* $x > 0$.

b) $3 \leq p < \infty$ *case:* $x = p^n u$ *where n is even and* $\left(\frac{u_0}{p}\right) = 1$.

c) $p = 2$ *case:* $x = 2^n u$ *where n is even and* $u \equiv 1 \pmod 8$.

Proof.

a) $\mathbb{Q}_\infty = \mathbb{R}$ and the real non-zero squares are exactly the positive reals.

b) If x is a square then $x = p^n u = (p^m v)^2 = p^{2m} v^2$. By the uniqueness of this representation $n = 2m$ is even and u is a square so $\left(\frac{u_0}{p}\right) = 1$.

Conversely if $x = p^n u$ where n is even and $\left(\frac{u_0}{p}\right) = 1$ let $f(x) = x^2 - u$. Then $\left(\frac{u_0}{p}\right) = 1$ implies there exists $v_0^2 \equiv u_0 \not\equiv 0 \pmod p$ and then $f'(v_0) = 2v_0 \not\equiv 0 \pmod p$ since p is odd. Then by Hensel's lemma v_0 lifts to a root $v \in \mathbb{Z}_p$. Then $x = p^n u = p^{2m} v^2 = (p^m v)^2$.

c) See [**11**].

□

Theorem 8.8. *Let p be a prime number. Then if $v \in \mathbb{Z}_p^*$ such that* $\left(\frac{v_0}{p}\right) = -1$

$$\mathbb{Q}_p^*/\mathbb{Q}_p^{*2} = \begin{cases} \{\pm 1\} & \text{if } p = \infty \\ \{1, p, v, pv\} & \text{if } 3 \leq p < \infty \\ \{\pm 1, \pm 2, \pm 3, \pm 6\} & \text{if } p = 2 \end{cases}$$

Proof. If $x \in \mathbb{R}$ then modulo $|x| = \sqrt{|x|}^2$, $x \equiv \pm 1 \pmod{\mathbb{R}^{*2}}$.

If $x = p^n u \in \mathbb{Q}_p^*$ for $3 \leq p < \infty$ then either u is a square or $\left(\frac{u_0}{p}\right) = -1$. In the latter case

$$\left(\frac{u_0 v_0}{p}\right) = \left(\frac{u_0}{p}\right)\left(\frac{u_0}{p}\right) = -1 \times -1 = 1.$$

Thus uv is a square so $u^{-1}v = (u^{-1})^2(uv)$ is a square. Then $x = p^n u \equiv (1 \text{ or } p) \times (1 \text{ or } v) \pmod{\mathbb{Q}_p^{*2}}$ so $x \in \{1, p, v, pv\}$.

For $u \in \mathbb{Z}_2^*$ we know that u is a square if and only if $u \equiv 1 \pmod 8$. Now if $u \equiv 3 \pmod 8$ then since $3 \in \mathbb{Q}_2^*$ we have $3^{-1} \in \mathbb{Q}_2^*$ and then

$3^{-1}u \equiv 1 \pmod 8$ and $3^{-1}u = v^2$ for some $v \in \mathbb{Q}_2^*$. Then $u = 3v^2 \equiv 3 \pmod{\mathbb{Q}_2^{*2}}$. Similarly if $u \equiv 5 \pmod 8$ then $u \equiv -3 \pmod{\mathbb{Q}_2^{*2}}$ and if $u \equiv 7 \pmod 8$ then $u \equiv -1 \pmod{\mathbb{Q}_2^{*2}}$. Thus, since we have $2^2 \in \mathbb{Q}_2^{*2}$, $x = 2^n u \equiv (1 \text{ or } 2) \times (\text{ one of } \pm 1, \pm 3) \pmod{\mathbb{Q}_p^{*2}}$. So $x \in \{\pm 1, \pm 2, \pm 3, \pm 6\}$. □

Consider the inclusion map $j_p \colon \mathbb{Q}^*/\mathbb{Q}^{*2} \to \mathbb{Q}_p^{*2}/\mathbb{Q}_p^{*2}$. See that kernel of j_p is exactly elements that are not squares in $\mathbb{Q}^*$ but are squares in $\mathbb{Q}_p^*$. Thus $\operatorname{Ker}(j_p) = \mathbb{Q} \cap (\mathbb{Q}_p^{*2}/\mathbb{Q}^{*2})$. So for the restricted map $j_p \colon \mathbb{Q}(s,2)^2 \to (\mathbb{Q}_p^{*2}/\mathbb{Q}_p^{*2})^2$ we have

$$\operatorname{Ker}(j_p) = \mathbb{Q}(S,2)^2 \cap (\mathbb{Q}_p^{*2}/\mathbb{Q}^{*2})^2 \tag{8.9}$$

In Flynn's approach we have members of $E(\mathbb{Q})$ that we suspect generate $E(\mathbb{Q})/2E(\mathbb{Q})$. We let H be the image of these points under ϕ and then for each $p \in S$ calculate $j_p^{-1}(\phi_p(E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)))$. To find $E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)$ we consider our candidate points and find any other points by searching – we know when we've found them all by Theorem 8.6. Hopefully knowing that $\operatorname{Im}\phi \le j_p^{-1}(\phi_p(E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)))$ for each p will give us enough information to conclude that $\operatorname{Im}\phi = H$.

Algorithm 8.10 (Complete 2–descent without homogenous spaces)**.** Let $E/\mathbb{Q}$ be an elliptic curve given by

$$y^2 = (x-e_1)(x-e_2)(x-e_3) \quad \text{with } e_1, e_2, e_3 \in \mathbb{Q}.$$

Say we have a set $A \subset E(\mathbb{Q})$ that we think generates $E(\mathbb{Q})/2E(\mathbb{Q})$.

a) Calculate the discriminant, Δ, and let $S = \{\text{primes } p \mid p \text{ divides } \Delta\} \cup \{2, \infty\}$ and $\mathbb{Q}(S,2) = \{\pm p_1^{\epsilon_1} p_2^{\epsilon_2} \dots p_n^{\epsilon_n} \mid \epsilon_i = 0 \text{ or } 1 \text{ and } p_i \text{ divides } 2\Delta\}$.
b) Calculate $H = \phi(A)$.
c) For each p in S (or until we can conclude that $H = \operatorname{Im}\phi$).
   (i) Find $\mathbb{Q}_p^*/\mathbb{Q}_p^{*2}$.
   (ii) Find $\operatorname{Ker} j_p = \mathbb{Q}(S,2)^2 \cap (\mathbb{Q}_p^{*2}/\mathbb{Q}^{*2})^2$ (just need to consider whether each element of $\mathbb{Q}(S,2)^2$ is a square in $\mathbb{Q}_p^*$).
   (iii) Test whether A generates $E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)$ using
   $$|E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)| = 4/|2|_p.$$
   If A doesn't suffice, find points on $E(\mathbb{Q}_p)$ that will generate $E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)$.
   (iv) Calculate $j_p^{-1}(\phi_p(E(\mathbb{Q}_p)/2E(\mathbb{Q}_p))) = \langle \operatorname{Ker} j_p, H, \text{any other points found} \rangle$ and use the fact that $\operatorname{Im}\phi \le j_p^{-1}(\phi_p(E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)))$ to refine our knowledge of what $\operatorname{Im}\phi$ could be.

We compute $E(\mathbb{Q})2E(\mathbb{Q})$ for the same example as we did earlier.

Example 8.11. Let $E/\mathbb{Q}$ be given by $y^2 = x(x-1)(x+3) = x^3 + 2x^2 - 3x$.

We can calculate that $\Delta = 2304 = 2^8 \times 3^2$. Also we can use the Nagell–Lutz theorem to find

$$E_{\text{tors}}(\mathbb{Q}) = \{\mathcal{O}, (-3,0), (0,0), (1,0), (3,6), (3,-6), (-1,2), (-1,-2)\} = \langle (0,0), (3,6) \rangle.$$

Following our algorithm we set $S = \{2, 3, \infty\}$ and $\mathbb{Q}(S,2) = \{\pm 1, \pm 2, \pm 3, \pm 6\}$ so $\mathbb{Q}(S,2)^2 = \langle (-1,1), (2,1), (3,1), (1,-1), (1,2), (1,3) \rangle$. We aim to show that $A = \{(0,0), (3,6)\}$ generates $E(\mathbb{Q})/2E(\mathbb{Q})$.

Note that $\phi\colon (0,0) \mapsto (-3,-1)$ and $\phi\colon (3,6) \mapsto (3,2)$ so $H = \phi(A) = \langle (-3,-1), (3,2) \rangle$. We now consider each prime in S.

a) For $p = \infty$ we calculate

$$\mathbb{Q}_p^*/\mathbb{Q}_p^{*2} = \mathbb{R}^*/\mathbb{R}^{*2} = \{\pm 1\} \quad \text{and}$$

$$\operatorname{Ker} j_\infty = \mathbb{Q}(S,2)^2 \cap (\mathbb{R}^{*2}/\mathbb{Q}^{*2})^2 = \langle (2,1), (3,1), (1,2), (1,3) \rangle.$$

We know $|E(\mathbb{R})/2E(\mathbb{R})| = 2$ so have $E(\mathbb{R})/2E(\mathbb{R}) = \langle (0,0) \rangle$ since $(-3,-1) \neq (1,1)$. Therefore we have

$$\begin{aligned} \operatorname{Im} \phi &\leq j_\infty^{-1}(\phi_\infty(E(\mathbb{R})/2E(\mathbb{R}))) \\ &= \langle \operatorname{Ker} j_\infty, H \rangle \\ &= \langle (2,1), (3,1), (1,2), (1,3), (-3,-1), (3,2) \rangle \end{aligned}$$

b) For $p = 3$ we see that -1 is not a square modulo 3 so

$$\mathbb{Q}_3^*/\mathbb{Q}_3^{*2} = \{\pm 1, \pm 3\}^2 = \langle (1,-1), (1,3), (-1,1), (3,1) \rangle \quad \text{and}$$

$$\operatorname{Ker} j_3 = \mathbb{Q}(S,2)^2 \cap (\mathbb{Q}_3^{*2}/\mathbb{Q}^{*2})^2 = \langle (-2,1), (1,-2) \rangle.$$

We know $|E(\mathbb{Q}_3)/2E(\mathbb{Q}_3)| = 4$ so have $E(\mathbb{Q}_3)/2E(\mathbb{Q}_3) = \langle (0,0), (3,6) \rangle$ since $(-3,-1) \neq (3,2) \neq (1,1)$. Therefore we have

$$\begin{aligned} \operatorname{Im} \phi &\leq j_3^{-1}(\phi_3(E(\mathbb{Q}_3)/2E(\mathbb{Q}_3))) \\ &= \langle \operatorname{Ker} j_3, H \rangle \\ &= \langle (-2,-1), (1,-2), (-3,-1), (3,2) \rangle \end{aligned}$$

c) For $p = 2$ we have

$$\mathbb{Q}_2^*/\mathbb{Q}_2^{*2} = \{\pm 1, \pm 2, \pm 3, \pm 6\}^2$$

so j_2 is injective and $\operatorname{Ker} j_2 = \langle (1,1) \rangle$. We know $|E(\mathbb{Q}_2)/2E(\mathbb{Q}_2)| = 8$ and we have two independent generators (since $(3,2) \neq (-1,-2) \neq (1,1)$) so we need to find another generator. We choose small integers x and check whether $y^2 = x(x+3)(x-1)$ gives us a square in $\mathbb{Q}_2^*$ that is not a rational square. Setting $x = 7$ gives

us $y^2 = 7 \times 10 \times 6 = 2^2(105)$. Now $105 \equiv 1 \pmod 8$ so there exists $\gamma \in \mathbb{Q}_2^*$ such that $(7, \gamma) \in E(\mathbb{Q}_2)$. Then

$$\phi_2 \colon (7, \gamma) \mapsto (7, 6) = (-1, 6)$$

since $-7 \equiv 1 \pmod 8$ so $-7 \in \mathbb{Q}_2^{*2}$. Therefore

$$\begin{aligned} \operatorname{Im} \phi &\leq j_2^{-1}(\phi_2(E(\mathbb{Q}_2)/2E(\mathbb{Q}_2))) \\ &= \langle \operatorname{Ker} j_2, H, (-1, 6) \rangle \\ &= \langle (-3, -1), (3, 2), (-1, 6) \rangle \end{aligned}$$

From these three cases (actually only need the third and one of the first two) we can conclude that $\operatorname{Im} \phi = H$ and thus that $E(\mathbb{Q})/2E(\mathbb{Q}) = \langle A \rangle = \{\mathcal{O}, (0, 0), (3, 6), (-1, 2)\}$. Then as before we conclude that $r = 0$.

This example shows us (compare length of Example 8.2 with that of Example 8.11) that when we suspect that we know generators for $E(\mathbb{Q})/2E(\mathbb{Q})$ it can be more efficient to carry out descent without homogenous spaces. However homogenous spaces are still useful for finding points.

Appendix A

Maple calculations

First we show that the change of variables

$$y \mapsto \bar{y} = \frac{1}{2}(y - a_1x - a_3)$$

converts

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

to

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$$

where $b_2 = a_1^2 + 4a_2$, $b_4 = 2a_4 + a_1a_3$ and $b_6 = a_3^2 + 4a_6$.

```
> ybar:=1/2*(y-a1*x-a3):
> latex(collect(expand(-4*(ybar^2+a1*x*ybar+a3*ybar-x^3-a2*x^2-a4*x-a6)),x));
```

$$4\,x^3 + \left(a1^{\,2} + 4\,a2\right)x^2 + \left(2\,a1\;a3 + 4\,a4\right)x - y^2 + a3^{\,2} + 4\,a6$$

Next, if we send

$$(x, y) \mapsto (\bar{x}, \bar{y}) = \left(\frac{x - 3b_2}{36}, \frac{y}{108}\right)$$

we convert

$$y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$$

to

$$y^2 = x^3 - 27c_4x - 54c_6$$

where $c_4 = b_2^2 - 24b_4$ and $c_6 = b_3^2 + 36b_2b_4 - 216b_6$.

```
> xbar:=(x-3*b2)/36:
> ybar:=y/108:
> latex(collect(expand(-11664*(ybar^2-4*xbar^3-b2*xbar^2-2*b4*xbar-b_6)),x));
```

$$x^3 + \left(648\,b4 - 27\,b2^2\right)x - y^2 - 1944\,b4\;b2 + 11664\,b\_6 + 54\,b2^3$$

This next calculation shows that a curve with Weierstrass equation $y^2 = x^3 + Ax + B$ has $\Delta = -16(4A^3 + 27B^2)$.

```
> a1:=0:
> a2:=0:
> a3:=0:
> a4:=A:
> a6:=B:
> b2:=a1^2+4*a2:
```

```
> b4:=2*a4+a1*a3:
> b6:=a3^2+4*a6:
> b8:=a1^2*a6+4*a2*a6-a1*a3*a4+a2*a3^2-a4^2:
> Delta:=-b2^2*b8-8*b4^3-27*b6^2+9*b2*b4*b6:
> latex(Delta);
```

$$4\,x^3+\left(a1^2+4\,a2\right)x^2+\left(2\,a1\;a3+4\,a4\right)x-y^2+a3^2+4\,a6$$

Now for transformation $(x,y)\mapsto(\bar{x},\bar{y})=(x+r,y+sx+t)$ we see that Δ is invariant

```
> b2:=a1^2+4*a2:
> b4:=2*a4+a1*a3:
> b6:=a3^2+4*a6:
> b8:=a1^2*a6+4*a2*a6-a1*a3*a4+a2*a3^2-a4^2:
> Delta:=-b2^2*b8-8*b4^3-27*b6^2+9*b2*b4*b6:
> xbar:=x+r:
> ybar:=y+s*x+t:
> latex(collect(expand((ybar^2+a1*xbar*ybar+a3*ybar-xbar^3-a2*xbar^2-a4*xbar-a6
```

$$-x^3+\left(-a2-3\,r+a1\,s+s^2\right)x^2+\left(\left(2\,s+a1\right)y+2\,st-a4+a1\,rs+a1\,t-2\,a2\,r-3\,r^2+a3\,s\right.$$

```
> m1:=2*s+a1:
> m3:=a3+a1*r+2*t:
> m2:=a2+3*r-s^2-a1*s:
> m4:=-a3*s+a4-2*s*t+2*a2*r-a1*t+3*r^2-a1*r*s:
> m6:=a6-a3*t+r^3-a1*r*t+a2*r^2+a4*r-t^2:
> n2:=m1^2+4*m2:
> n4:=2*m4+m1*m3:
> n6:=m3^2+4*m6:
> n8:=m1^2*m6+4*m2*m6-m1*m3*m4+m2*m3^2-m4^2:
> Delta2:=-n2^2*n8-8*n4^3-27*n6^2+9*n2*n4*n6:
> simplify(Delta2-Delta);
```

$$0$$

For a Weierstrass equation of the form $y^2+a_1xy=x^3+a_2x^2$ we show that $\Delta=0$.

```
> a3:=0:
> a4:=0:
> a6:=0:
> b2:=a1^2+4*a2:
> b4:=2*a4+a1*a3:
> b6:=a3^2+4*a6:
> b8:=a1^2*a6+4*a2*a6-a1*a3*a4+a2*a3^2-a4^2:
> Delta:=-b2^2*b8-8*b4^3-27*b6^2+9*b2*b4*b6;
                              Delta := 0
```

Our next calculation shows that for $g(x) = 4x^3 + b_2x^2 + 2b_4x + b_6$, $\mathrm{disc}(g) = 16\Delta$.

```
> b2:=a1^2+4*a2:
> b4:=2*a4+a1*a3:
> b6:=a3^2+4*a6:
> b8:=a1^2*a6+4*a2*a6-a1*a3*a4+a2*a3^2-a4^2:
> Delta:=-b2^2*b8-8*b4^3-27*b6^2+9*b2*b4*b6:
> g:=4*x^3+b2*x^2+2*b4*x+b6:
> simplify(discrim(g,x)-16*Delta);
                                    0
```

Finally, for two polynomials F and G we find f_1, g_1, f_2, g_2 such that $f_1F + g_1G = 4\Delta Z^7$ and $f_2F + g_2G = 4\Delta X^7$.

```
> F:=x^4-2*A*x^2*z^2-8*B*x*z^3+A^2*z^4:
> G:=4*x^3*z+4*A*x*z^3+4*B*z^4:
> gcd(F,G);
                                    1
> Delta:=4*A^3+27*B^2:
> gcdex(F,G,4*Delta*z^7,x,'f1','g1'):
> latex(f1);
```

$$12\,zx^2 + 16\,Az^3$$

```
> latex(g1);
```

$$-3\,x^3 + 5\,Axz^2 + 27\,Bz^3$$

```
> gcdex(F,G,4*Delta*x^7,x,'f2','g2');
> latex(f2);
```

$$4(4A^3 + 27B^2)x^3 - 4A^2Bx^2z + 4A(A^3 + 22B^2)xz^2 + 12B(A^3 + 8B^2)z^3$$

```
> latex(g2);
```

$$A^2Bx^3 + A(5A^3 + 32B^2)x^2z + 2B(13A^3 + 96B^2)xz^2 - 3A^2(A^3 + 8B^2)z^3$$

Appendix B

Some Galois theory

We have included this chapter as a quick introduction to infinite Galois theory. The details can be found in [**9**]. First we define a Galois group.

Definition B.1. *If L/K is a field extension, a* K–automorphism of L *is an isomorphism $\sigma\colon L \to L$ such that $\sigma(a) = a$ for all $a \in K$. The collection of all K–automorphisms of L, with group operation composition, is called the* Galois group of L/K *and denoted* $\mathrm{Gal}(L/K)$.

Example B.2. It can be shown that $\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q}) = \{1, \sigma\} \cong \mathbb{Z}/2\mathbb{Z}$ where $(a + b\sqrt{2})^1 = a + b\sqrt{2}$ and $(a + b\sqrt{2})^\sigma = a - b\sqrt{2}$.

This example gives the right intuition. If $L = K(\alpha_1, \ldots, \alpha_n)$ is a finite algebraic extension of K then any $\sigma \in \mathrm{Gal}(L/K)$ is determined once we know all the $\sigma(\alpha_i)$. Furthermore if f_i is the minimum polynomial of α_i then $\sigma(\alpha_i)$ must be one of the roots of f_i in L.

Definition B.3. *Let L/K be a (finite or infinite) extension of fields with Galois group G. We say that L/K is a* Galois extension *if any of the following equivalent conditions are satisfied.*

a) *L/K is algebraic and $L^G = K$.*
b) *L/K is separable and L is a splitting field over K of a set of polynomials.*
c) *L/K is a splitting field over K of a set of separable polynomials.*

Let $\mathcal{M} = \{M \mid K \subseteq M \subseteq L, [M\colon K] < \infty \text{ and } M/K \text{ Galois}\}$. Then it can be shown that

$$G \cong \varprojlim \mathrm{Gal}(M_i/K)$$

where the M_i range over $\mathcal{M}$. Thus the Galois group of an infinite extension L/K is a projective limit of finite Galois groups and we can understand it by considering those groups.

Example B.4. Let $\overline{\mathbb{Q}}$ be an algebraic closure of $\mathbb{Q}$. Then $\overline{\mathbb{Q}}/\mathbb{Q}$ is an infinite extension since for every N there exists a polynomial over $\mathbb{Q}$ of degree N. The roots of such a polynomial are in $\overline{\mathbb{Q}}$ so $[\overline{\mathbb{Q}}\colon \mathbb{Q}] > N$ for

every N. Also, $\overline{\mathbb{Q}}/\mathbb{Q}$ is a Galois extension since it is the splitting field of all the polynomials in $\mathbb{Q}[x]$, which are all separable.

We can place a topology on $G = \mathrm{Gal}(L/K)$, called the *Krull topology*, and then the fundamental theorem of Galois theory gives a correspondence between the intermediate fields of L/K and closed subgroups of G.

Appendix C

Some algebraic number theory

The following gives the main definitions and results that we need from algebraic number theory. All this (and more) can be found in [**8**].

Definition C.1. *A* number field *K is a finite extension of $\mathbb{Q}$.*

Definition C.2. *The* ring of integers *of a number field K is defined by*

$\mathbb{O}_K = \{\alpha \in K \mid \alpha$ *satisfies a monic polynomial with integer coefficients*$\}$.

Definition C.3. *For ideals $\mathfrak{a}$ and $\mathfrak{b}$ define their product by*

$$\mathfrak{ab} = \left\{ \sum a_i b_i \mid a_i \in \mathfrak{a}, b_i \in \mathfrak{b} \right\}.$$

Theorem C.4 (Unique factorisation of ideals)**.** *Every ideal of $\mathbb{O}_K$ can be written uniquely as a product of prime ideals.*

Definition C.5. *A* fractional ideal *of $\mathbb{O}_K$ is an K–submodule $\mathfrak{a}$ for which there exists a $d \in \mathbb{O}_K \backslash \{0\}$ such that $d\mathfrak{a} \subseteq A$. We say a fractional ideal is* principal *if it can be generated as an $\mathbb{O}_K$–module by a single element.*

Definition C.6. *Denote the group of fractional ideals of $\mathbb{O}_K$ by $\mathcal{F}$ and the subgroup of principal fractional ideals by $\mathcal{P}$. The class group of $\mathbb{O}_K$ is the quotient group $Cl(\mathbb{O}_K) = \mathcal{F}/\mathcal{P}$.*

Theorem C.7 (Finiteness of the class number)**.** *The class group is finite.*

Theorem C.8 (Dirichlet Unit Theorem)**.** *The group of units of $\mathbb{O}_K$ is finitely generated.*

A more efficient way of expressing these facts is the following.

Theorem C.9. *Let K be a number field. Then the following sequence is exact where the direct sum is over all nonzero prime ideals of $\mathbb{O}_K$*

$$0 \longrightarrow \mathbb{O}_K^* \longrightarrow K^* \longrightarrow \oplus_{\mathfrak{p}} \mathbb{Z} \longrightarrow Cl(\mathbb{O}_K) \longrightarrow 0$$

with $\mathbb{O}_K^*$ finitely generated and $Cl(\mathbb{O}_K)$ finite.

To generalise this exact sequence slightly we need a way to define the class group directly by the exact sequence.

Theorem C.10. *If A and B are abelian groups and $f\colon A \to B$ is a group homomorphism then define the* cokernel of f *by*

$$\operatorname{Coker}(f) = B/\operatorname{Im}(f).$$

Then there exists an exact sequence

$$0 \longrightarrow \operatorname{Ker}(f) \longrightarrow A \longrightarrow B \longrightarrow \operatorname{Coker}(f) \longrightarrow 0$$

Proof. Follows immediately from the definitions. □

Using this idea we can generalise our above exact sequence.

Theorem C.11. *Let K be a number field and T be a finite set of prime ideals of K. Let $U_T = \operatorname{Ker}(K^* \to \oplus_{\mathfrak{p}\notin T}\mathbb{Z})$ and $C_T = \operatorname{Coker}(K^* \to \oplus_{\mathfrak{p}\notin T}\mathbb{Z})$. Then we have an exact sequence*

$$0 \longrightarrow U_T \longrightarrow K^* \longrightarrow \oplus_{\mathfrak{p}\notin T}\mathbb{Z} \longrightarrow C_T \longrightarrow 0$$

where U_T is finitely generated and C_T is finite.

Proof. The Dirichlet-Chevalley-Hasse unit theorem (see **[5]**, Theorem 8.2) tells us that U_T is finitely generated. The map $i\colon \oplus_{\mathfrak{p}\notin T}\mathbb{Z} \to \oplus_{\mathfrak{p}}\mathbb{Z}$ induces an injection from C_T to the full class group $C = \operatorname{Coker}(K^* \to \oplus_{\mathfrak{p}}\mathbb{Z})$. Thus the finiteness of C_T follows from the finiteness of C. □

References

[1] J.W.S. Cassels, *Local Fields*, London Mathematical Society Student Texts, 1986.

[2] Andrej Dujella, *History of elliptic curves rank records*, `http://web.math.hr/~duje/tors/rankhist.html`

[3] J.W.S. Cassels and E.V. Flynn, *Prolegomena to a middlebrow arithmetic of curves of genus 2*, London Mathematical Society Lecture Note Series No. 230, 1996.

[4] R. Hartshorne, *Algebraic Geometry*, Springer–Verlag Graduate texts in mathematics, 1977.

[5] Gerald J. Janusz, *Algebraic Number Fields*, Graduate Studies in Mathematics, American Mathematical Society, 1996.

[6] Anthony W. Knapp, *Elliptic Curves*, Princeton University Press, 1992.

[7] J.S. Milne, *Elliptic Curves* – online course notes. Available at `http://www.jmilne.org/math/CourseNotes/math679.html`.

[8] J.S. Milne, *Algebraic Number Theory* – online course notes. Available at `http://www.jmilne.org/math/CourseNotes/math676.html`.

[9] Victor Scharaschkin, *Galois Theory Notes*, University of Queensland, 2005.

[10] Jean-Pierre Serre, *Local Fields*, Springer–Verlag Graduate texts in mathematics, 1995.

[11] Jean-Pierre Serre, *A Course In Arithmetic*, Springer–Verlag Graduate texts in mathematics, 1973.

[12] Joseph H. Silverman, *The Arithmetic of Elliptic Curves*, Springer–Verlag Graduate texts in mathematics Vol. 106, 1986.

[13] Joseph H. Silverman and John Tate, *Rational Points on Elliptic Curves*, Undergraduate Texts in Mathematics, 1992.