Math 225AB: Elliptic Curves

Simon Rubinstein-Salzedo

Winter and Spring 2007

0.1 Introduction

These notes are based on a graduate course on elliptic curves I took from Professor Adebisi Agboola in the Winter and Spring of 2007. The textbooks were *The Arithmetic of Elliptic Curves* and *Advanced Topics in the Arithmetic of Elliptic Curves*, both by Joseph Silverman. Other recommended books were *Rational Points on Elliptic Curves* by Joseph Silverman and John Tate, *Elliptic Curves* by Anthony Knapp, *Elliptic Functions* by Serge Lang, *Introduction to Arithmetic Theory of Automorphic Functions* by Goro Shimura, *Elliptic Curves* by James Milne (available at `http://www.jmilne.org/math/CourseNotes/math679.pdf`), and *Rational Points on Modular Elliptic Curves* by Henri Darmon (available at `http://www.math.mcgill.ca/darmon/pub/Articles/Research/36.NSF-CBMS/chapter.ps`).

Chapter 1

A Crash Course on Varieties

K is a perfect field, and $\bar{K}$ an algebraic closure of K.

Definition 1.1 (Affine n-space) $\mathbb{A}^n = \mathbb{A}^n(\bar{K}) = \{(x_1, \ldots, x_n) : x_i \in \bar{K}, 1 \leq i \leq n\}$. $\mathbb{A}^n(K) = \{(x_1, \ldots, x_n) : x_i \in K\}$.

Write $\bar{K}[X] = \bar{K}[X_1, \ldots, X_n]$, and suppose that I is an ideal in $\bar{K}[X]$.

Hilbert Basis Theorem. I is finitely generated.

Definition 1.2 An **affine algebraic set** is any set of the form $V_i = \{P \in \mathbb{A}^n : f(P) = 0 \text{ for all } f \in I\}$. If V is any algebraic set, then we define $I(V) := \{f \in \bar{K}[X] : f(P) = 0 \text{ for all } P \in V\}$ — the **ideal** of V. $V(K) := V \cap \mathbb{A}^n(K)$ — the set of K-rational points of V.

We say that V is defined over K if $I(V)$ is generated by polynomials in $K[X]$. So we see that if V is defined over K with $f_1, \ldots, f_m \in K[X]$ generators of $I(V)$, then $V(K) = \{x = (x_1, \ldots, x_n) \in \mathbb{A}^n(K) : f_1(x) = \cdots = f_m(x) = 0\}$.

Examples.

(a) $V : X^n + Y^n = 1$ $(n > 2)$. Wiles showed that $V(\mathbb{Q})$ is finite.

(b) $V : Y^2 = X^3 - 2$. $V(\mathbb{Q})$ is infinite. [Fermat showed that $V(\mathbb{Z}) = \{(3, \pm 5)\}$.]

Definition 1.3 Say that an affine algebraic set is an **affine algebraic variety** if $I(V)$ is a prime ideal in $\bar{K}[X]$.

Definition 1.4 Suppose that V is an affine algebraic variety defined over K. Set $I(V/K) := I(V) \cap K[X]$. $K[V] := \frac{K[X]}{I(V/K)}$ is the affine coordinate ring of V/K — this is an integral domain. $K(V)$, the quotient field of $K[V]$, is the **function field** of V/K. Define $\bar{K}[V]$ and $\bar{K}(V)$ similarly. Each element $f \in \bar{K}[V]$ induces a function $f : V \to \bar{K}$.

Definition 1.5 The **dimension** of a variety V is $\dim(V) := \text{tr deg}(\bar{K}(V)/\bar{K})$.

Example. $\bar{K}(\mathbb{A}^n) = \bar{K}(X_1, \dots, X_n)$, so $\dim(\mathbb{A}^n) = n$.

1.1 Smoothness

Definition 1.6 Suppose that $V \subseteq \mathbb{A}^n$ is a variety, and $P \in V$. Let $f_1, \dots, f_m \in \bar{K}[X]$ be a set of generators of $I(V)$. Say that V is **smooth** at P (or **nonsingular** at P) if the matrix

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

has rank $n - \dim(V)$.

Example. If V is given by a single nonconstant polynomial equation $f(X_1, \dots, X_n) = 0$, then $\dim(V) = n - 1$. So $P \in V$ is singular iff

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

Alternatively, set $M_P := \{f \in \bar{K}[V] : f(P) = 0\}$. Then M_P is a maximal ideal of $\bar{K}[V]$, for there is an isomorphism

$$\frac{\bar{K}[V]}{M_P} \xrightarrow{\sim} \bar{K}$$

given by $f \mapsto f(P)$. Then P is nonsingular iff $\dim_{\bar{K}}(M_P/M_P^2) = \dim(V)$.

Example. Let $V_1 : Y^5 = X^4 - X$ and $V_2 : Y^4 = X^3 + X^2$ and $P = (0, 0)$. Then M_P is generated by X and Y; M_P^2 is generated by X^2, Y^2, and XY. For V_1, we have $X = Y^5 - X^5 \equiv 0 \pmod{M_P^2}$, so $\dim_K(M_P/M_P^2) = 1$, and V_1 is smooth at P. For V_2, there are no nontrivial relations between X and Y modulo M_P^2, so $\dim_{\bar{K}} M_P/M_P^2 = 2$, so V_2 is singular at P.

Definition 1.7 The local ring $\bar{K}[V]_P$ of V at P is the localization of $\bar{K}[V]$ at M_P; i.e. $\bar{K}[V]_P = \{F \in \bar{K}(V) : F = f/g \text{ with } f, g \in \bar{K}[V] \text{ and } g(P) \neq 0\}$.

1.2 Projective Varieties

Definition 1.8 (Projective n-space) $\mathbb{P}^n$ or $\mathbb{P}^n(\bar{K})$ is the set of all $(n+1)$-tuples $(x_0, \ldots, x_n) \in \mathbb{A}^{n+1}$ such that at least one $x_i \neq 0$, modulo the equivalence relation $(x_0, \ldots, x_n) \sim (\lambda x_0, \ldots, \lambda x_n)$ for all $\lambda \in \bar{K}^\times$. Write $[x_0, \ldots, x_n]$ for the equivalence class of $(x_0, \ldots, x_n)$. We call these **homogeneous coordinates** of the corresponding point in $\mathbb{P}^n$. $\mathbb{P}^n(K) := \{[x_0, \ldots, x_n] \in \mathbb{P}^n : x_i \in K, 0 \leq i \leq n\}$, the set of K-rational points of $\mathbb{P}^n$.

Definition 1.9 Say that a polynomial $f \in \bar{K}[X] = \bar{K}[X_0, \ldots, X_n]$ is homogeneous of degree d if $f(\lambda X_0, \ldots, \lambda X_n) = \lambda^d f(X_0, \ldots, X_n)$ for all $\lambda \in \bar{K}$. Say that an ideal I of $\bar{K}[X]$ is homogeneous if it is generated by homogeneous polynomials.

Definition 1.10 A projective algebraic set is any set of the form $V_I := \{P \in \mathbb{P}^n : f(P) = 0 \text{ for all homogeneous } f \in I\}$, for a homogeneous ideal I in $\bar{K}[X]$. If V is a projective algebraic set, we define $I(V) := \{f \in \bar{K}[X] : f \text{ is homogeneous and } f(P) = 0 \text{ for all } P \in V\}$, the **homogeneous ideal** of V. Say that V is a projective algebraic variety if $I(V)$ is a prime ideal of $\bar{K}[X]$.

Consider the maps $\phi_i : \mathbb{A}^n \to \mathbb{P}^n$ given by $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_i, 1, x_{i+1}, \ldots, x_n)$. If V is a projective variety, then $V \cap \phi_i(\mathbb{A}^n)$ is an affine variety.

Example. $V : X^3 + Y^2 = 1$. This gets sent to $X^3 + Y^2 Z = Z^3$.

Example. $X^2 Z + Z^3 + Y^2 = 0$. Dividing by Z^3 gives

$$\left(\frac{X}{Z}\right)^2 + 1 + \left(\frac{Y}{Z}\right)^2 = 0.$$

Definition 1.11 Suppose $V_1, V_2 \subseteq \mathbb{P}^n$ are projective varieties. A **rational map** from V_1 to V_2 is a map of the form $\phi : V_1 \to V_2$ given by $P \mapsto [f_0(P), \ldots, f_n(P)]$, where $f_0, \ldots, f_n \in \bar{K}(V_1)$, at every point $P \in V_1$ at which $f_0, \ldots, f_n$ are all defined. Say that ϕ is **regular** (or **defined**) at $P \in V_1$ if there exists $g \in \bar{K}(V_1)$ such that gf_i is regular at P *and* $gf_i(P) \neq 0$ for some i. A **morphism** is a rational map which is regular at every point. We say that $V_1 \simeq V_2$ if there are morphisms $\varphi : V_1 \to V_2$ and $\psi : V_2 \to V_1$ such that $\psi \circ \varphi = id_{V_1}$ and $\varphi \circ \psi = id_{V_2}$.

Chapter 2

A Crash Course on Algebraic Curves

Definition 2.1 For us a curve is an irreducible projective variety of dimension 1, defined over K. $K(C)$ is the function field of C. (We have a map $C \to \mathbb{P}^1$, and $K(C)/K$ has transcendence degree 1.) If $P \in C(K)$, set $M_P = \{g \in K(C) \mid g(P) = 0\}$. (Note that M_P is the maximal ideal of $K[C]_P$, the local ring of C at P.)

Given $f \in K(C)^\times$, we say that $\operatorname{ord}_P(f) = i$ if $f \in M_P^i$ and $f \notin M_P^{i+1}$. $\operatorname{Div}(C)$ is the free abelian group generated by $C(\bar{K})$, or

$$\left\{ \sum n_i P_i \,\middle|\, n_i \in \mathbb{Z},\ P_i \in C(\bar{K}) \right\}.$$

So we have a divisor $(f) = \operatorname{div}(f) = \sum_P \operatorname{ord}_P(f) \cdot P$. This gives us a map $K(C)^\times \to \operatorname{Div}(C)$.

Theorem 2.2 Let $(f) = \sum_P \operatorname{ord}_P(f) \cdot P$. Then

(1) $\deg(f) := \sum_P \operatorname{ord}_P(f) = 0$.

(2) $(f) = 0$ iff $f \in K^\times$.

Reasons.

(1) A nonconstant $f \in \bar{K}(C)$ gives a map $f : C \to \mathbb{P}^1$ given by

$$P \mapsto \begin{cases} [f(P), 1] & \text{if } f \text{ is regular at } P, \\ [1, 0] & \text{otherwise.} \end{cases}$$

Then $(f) = f^*\{\{0\} - \{\infty\}\}$, and this last divisor has degree zero.

(2) If $(f) = 0$, then f has no poles. So the map $f : C \to \mathbb{P}^1$ is not surjective, and therefore is constant (see below).

$\mathrm{Div}^0(C)$ is the set of divisors of degree zero. Set

$$\mathrm{Pic}^0_{\bar{K}}(C) = \frac{\mathrm{Div}^0(C)}{\{(f) \mid f \in \bar{K}(C)^\times\}}$$

— this carries the structure of an **abelian variety**.

More Facts. Suppose $\varphi : C_1 \to C_2$ is a rational map.

(a) If $P \in C_1$ is a smooth point, then φ is regular at P.

(b) So, if C_1 is smooth, then φ is a morphism.

Proof. Let $\varphi = [f_0, \ldots, f_n]$, $f_i \in \bar{K}(C)$. Choose a uniformizer $t \in \bar{K}(C_1)$ at P (i.e. a generator of M_P). (We can do this, as P is a smooth point, by hypothesis.) If $\alpha := \min_{0 \le i \le n}\{\mathrm{ord}_P(f_i)\}$, then

- $\mathrm{ord}_P(t^{-\alpha} f_i) \ge 0$ for all i, and
- $\mathrm{ord}_P(t^{-\alpha} f_j) = 0$ for some j.

Hence each $t^{-\alpha} f_i$ is regular at P, and $t^{-\alpha} f_j(P) \neq 0$. Therefore φ is regular at P.

(c) If φ is a morphism, then φ is either constant or surjective (see Hartshorne, Chapter II, Proposition 6.8).

From a morphism $\varphi : C_1 \to C_2$, we obtain a corresponding morphism of function fields $\varphi^* : K(C_2) \to K(C_1)$ given by $f \mapsto f \circ \varphi$. In fact, there is a 1–1 correspondence (actually an equivalence of categories)

$$\left\{\begin{matrix}\text{nonconstant morphisms} \\ \varphi : C_1 \to C_2\end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix}\text{injections } \varphi^* : K(C_2) \to K(C_1) \\ \text{fixing } K\end{matrix}\right\}.$$

Definition 2.3 The **degree** of φ is defined by

$$\deg(\varphi) = \begin{cases} [K(C_1) : \varphi^* K(C_2)] & \text{if } \varphi \text{ is nonconstant,} \\ 0 & \text{if } \varphi \text{ is constant.} \end{cases}$$

(Define the separable and inseparable degrees $\deg_s(\varphi)$ and $\deg_i(\varphi)$ similarly.) The map $\varphi_* : K(C_1) \to K(C_2)$ is defined by $\varphi_* = (\varphi^*)^{-1} \circ N_{K(C_1)/\varphi^* K(C_2)}$.

Fact. If $\varphi : C_1 \to C_2$ is a map of degree one between two smooth curves, then φ is an isomorphism.

2.1 Local Behavior

Let $\varphi : C_1 \to C_2$ be a nonconstant morphism. Let $P \in C_1$, and let $t_{\varphi(P)}$ be a local uniformizer at $\varphi(P) \in C_2$. The **ramification index** $e_\varphi(P)$ of φ at P is defined by

$$e_\varphi(P) := \operatorname{ord}_P(\varphi^* t_{\varphi(P)}).$$

(So $e_\varphi(P) \geq 1$.) Say that φ is **unramified** at P if $e_\varphi(P) = 1$. Say that φ is **unramified** if it is unramified at every point of C_1.

Theorem 2.4

(1) For all but finitely many points Q of C_2, we have $\#\varphi^{-1}(Q) = \deg_s(\varphi)$. (Here we are counting the number of points over $\bar{K}$.) (cf: only finitely many primes ramify in a finite extension L/K of number fields.)

(2) $\sum_{P \in \varphi^{-1}(Q)} e_\varphi(P) = \deg(\varphi)$. (cf: $\sum e_i f_i = [L : K]$ for number fields.)

(3) If $\psi : C_2 \to C_3$ is another nonconstant map, and $P \in C_1$, then $e_{\psi\circ\varphi}(P) = e_\varphi(P) e_\psi(\varphi(P))$. (cf: multiplicativity of ramification in towers of number fields.)

2.2 The Frobenius Morphism

Suppose now that K is perfect with $\mathrm{char}(K) = p > 0$, and set $q = p^r$. For any polynomial f, we may form the polynomial $f^{(q)}$ by raising each coefficient of f to the q^{th} power. So, given a curve C/K, we obtain the curve $C^{(q)}/K$. There is a natural map $\phi : C \to C^{(q)}$ given by $[x_0, \ldots, x_n] \mapsto [x_0^q, \ldots, x_n^q]$. ϕ is called the q^{th} power **Frobenius morphism**.

Theorem 2.5 Notation as above.

(1) $\phi^*(K(C^{(q)})) = K(C)^q = \{f^q : f \in K(C)\}$.

(2) ϕ is purely inseparable.

(3) $\deg(\phi) = q$.

(4) Suppose that $\psi : C_1 \to C_2$ is a map of smooth curves. Then ψ factors as

$$C_1 \xrightarrow{\phi} C_1^q \xrightarrow{\lambda} C_2,$$

where $q = \deg_i(\psi)$, ϕ is the q^{th} power Frobenius map, and λ is separable.

(See Silverman II, §2.)

2.3 Divisors

$\mathrm{Div}(C) = \left\{D = \sum_{P \in C} n_P(P) \,\middle|\, n_P \in \mathbb{Z} \text{ and } n_P = 0 \text{ for almost all } P\right\}$, i.e. $\mathrm{Div}(C)$ is the free abelian group generated by the points on C. The **degree** of D is $\deg(D) := \sum_{P \in C} n_P$. $\mathrm{Div}^0(C) = \{D \in \mathrm{Div}(C) \mid \deg(D) = 0\}$. $D \in \mathrm{Div}(C)$ is **principal** if $D = (f) = \mathrm{div}(f)$ for some $f \in \bar{K}(C)^\times$. Say that D_1 and D_2 are **linearly equivalent**, and write $D_1 \sim D_2$, if $D_1 - D_2$ is principal. $\mathrm{Pic}(C) := \mathrm{Div}(C)/\{\text{principal divisors}\}$ is the **Picard group** of C.

Example. Every divisor of degree zero on $\mathbb{P}^1$ is principal. For suppose $D = \sum n_P(P)$, $\deg(D) = 0$, with $P = [x_P, y_P] \in \mathbb{P}^1$. Then D is the divisor of the function

$$\prod_{P \in \mathbb{P}^1} (y_P x - x_P y)^{n_P}.$$

We have an exact sequence

$$1 \to \bar{K}^\times \to \bar{K}(C)^\times \overset{\mathrm{div}}{\to} \mathrm{Div}^0(C) \to \mathrm{Pic}^0(C) \to 0.$$

c.f.: if L is a number field, we have an exact sequence

$$1 \to \mathfrak{o}_L^\times \to L^\times \to I_L \to \mathrm{Cl}(\mathfrak{o}_L) \to 0.$$

Definition 2.6 Suppose $\varphi : C_1 \to C_2$ is a nonconstant map of smooth curves. Define the **pullback** $\varphi^* : \mathrm{Div}(C_2) \to \mathrm{Div}(C_1)$ by

$$(Q) \mapsto \sum_{P \in \varphi^{-1}(Q)} e_\varphi(P)(P)$$

and the **pushforward** $\varphi_* : \mathrm{Div}(C_1) \to \mathrm{Div}(C_2)$ by $(P) \mapsto (\varphi P)$. Extend to arbitrary divisors by $\mathbb{Z}$-linearity.

So for example if C is smooth and $f \in \bar{K}(C)$ is nonconstant, then we have $f : C \to \mathbb{P}^1$ given by

$$P \mapsto \begin{cases} [f(P), 1] & \text{if } f \text{ is regular at } P, \\ [1, 0] & \text{otherwise.} \end{cases}$$

Then $\mathrm{div}(f) = f^*((0) - (\infty))$.

Properties.

(a) $\deg(\phi^* D) = (\deg \varphi)(\deg D)$ for all $D \in \mathrm{Div}(C_2)$.

(b) $\phi^* \,\mathrm{div}(f) = \mathrm{div}(\phi^* f)$ for all $f \in \bar{K}(C_2)^\times$.

(c) $\deg(\phi_* D) = \deg(D)$ for all $D \in \mathrm{Div}(C_1)$.

(d) $\phi_* \operatorname{div}(f) = \operatorname{div}(\phi_*(f))$ for all $f \in \bar{K}(C_1)^\times$.

(e) $\phi_* \circ \phi^*$ is multiplication by $\deg(\phi)$ on $\mathrm{Div}(C_2)$.

(f) If $\psi : C_2 \to C_3$ is another map between smooth curves, then $(\psi \circ \phi)^* = \phi^* \circ \psi^*$ and $(\psi \circ \phi)_* = \psi_* \circ \phi_*$.

2.4 Differentials

Definition 2.7 Let C/K be a curve. The space of differential forms Ω_C on C is the $\bar{K}(C)$-vector space generated by the symbols $\{dx \mid x \in \bar{K}(C)\}$ subject to the relations

(a) $d(x+y) = dx + dy$ for all $x, y \in \bar{K}(C)$.

(b) $d(xy) = x\ dy + y\ dx$.

(c) $da = 0$ for all $a \in \bar{K}$.

If $\varphi : C_1 \to C_2$ is a nonconstant morphism of curves, then there is a natural map $\varphi^* : \Omega_{C_2} \to \Omega_{C_1}$ given by

$$\sum f_i\ dx_i \mapsto \sum (\varphi^* f_i)\ d(\varphi^* x_i).$$

Theorem 2.8

(1) Ω_C is a 1-dimensional $\bar{K}(C)$-vector space.

(2) $\varphi : C_1 \to C_2$ is separable iff $\varphi^* : \Omega_{C_2} \to \Omega_{C_1}$ is nonzero (and so is injective).

Suppose that $P \in C$, and let $t \in \bar{K}(C)$ be a local uniformizer at P.

(3) Suppose $\omega \in \Omega_C$. Then there exists a unique function $g \in \bar{K}(C)$ (depending on ω and t) such that $\omega = g\ dt$. Set $g := \frac{\omega}{dt}$.

(4) If $f \in \bar{K}(C)$ is regular at P, then $\frac{df}{dt}$ is regular at P also.

(5) $\mathrm{ord}_P(\omega) := \mathrm{ord}_P\left(\frac{\omega}{dt}\right)$ depends only upon ω and P and not upon t.

(6) Suppose that $x \in \bar{K}(C)$ with $\bar{K}(C)/\bar{K}(x)$ separable, with $x(P) = 0$. Then $\mathrm{ord}_P(f\ dx) = \mathrm{ord}_P(f) + \mathrm{ord}_P(x) - 1$ for all $f \in \bar{K}(C)$.

(7) $\mathrm{ord}_P(\omega) = 0$ for all but finitely many points $P \in C$.

We may attach a divisor to $\omega \in \Omega_C$ as follows:

Definition 2.9 Suppose that $\omega \in \Omega_C$. Then

$$\mathrm{div}(\omega) := \sum_{P \in C} \mathrm{ord}_P(\omega)(P).$$

ω is **regular** or **holomorphic** if $\mathrm{ord}_P(\omega) \geq 0$ for all $P \in C$. ω is **nonvanishing** if $\mathrm{ord}_P(\omega) \leq 0$ for all $P \in C$.

Definition 2.10 Suppose $\omega \in \Omega_C$, $\omega \neq 0$. The image of ω in $\mathrm{Pic}(C)$ is called the **canonical divisor class** on C. (Note that this definition makes sense because Ω_C is a 1-dimensional $\bar{K}(C)$-vector space.) Any divisor in this class is called a canonical divisor.

Example. Let $C = \mathbb{P}^1$. Suppose that t is a coordinate function on $\mathbb{P}^1$. What is $\mathrm{div}(dt)$? If $\alpha \in \bar{K}$, then $t - \alpha$ is a uniformizer at α. Then $dt = 1 \cdot d(t - \alpha)$, so $\mathrm{ord}_\alpha(dt) = 0$. At $\infty \in \mathbb{P}^1$, $1/t$ is a uniformizer. Then $dt = -t^2\ d\left(\frac{1}{t}\right)$, so $\mathrm{ord}_\infty(dt) = \mathrm{ord}_\infty\left(-t^2\ d\left(\frac{1}{t}\right)\right) = -2$. Thus $\mathrm{div}(dt) = -2(\infty)$. So if $\omega \in \Omega_P$, $\omega \neq 0$, then $\deg(\mathrm{div}(\omega)) = \deg(\mathrm{div}(dt)) = -2$. So ω is nonholomorphic.

We say that a divisor $D = \sum_P n_P(P) \in \mathrm{Div}(C)$ is **effective** or **positive**, and we write $D \geq 0$, if $n_P \geq 0$ for all $P \in C$. If $D_1, D_2 \in \mathrm{Div}(C)$, then $D_1 \geq D_2$ iff $D_1 - D_2 \geq 0$.

Example. $\mathrm{div}(f) \geq -n(P)$ means that f has a single pole of order at most n at P.

Definition 2.11 Suppose that $D \in \operatorname{Div}(C)$. Define

$$\mathscr{L}(D) := \{f \in \bar{K}(C)^\times : \operatorname{div}(f) \geq -D\} \cup \{0\}.$$

Then $\mathscr{L}(D)$ is a finite-dimensional $\bar{K}$-vector space (exercise). Set $\ell(D) := \dim_{\bar{K}} \mathscr{L}(D)$.

Proposition 2.12 Let $D \in \operatorname{Div}(C)$.

(a) If $\deg(D) < 0$, then $\mathscr{L}(D) = \{0\}$, and $\ell(D) = 0$.

(b) $\mathscr{L}(D)$ is a finite-dimensional $\bar{K}$-vector space.

(c) If $D' \sim D$, then $\mathscr{L}(D) \simeq \mathscr{L}(D')$, and $\ell(D) = \ell(D')$.

Example. Suppose that $K_C \in \operatorname{Div}(C)$ is a canonical divisor on C, with $K_C = \operatorname{div}(\omega)$, say. Then $f \in \mathscr{L}(K_C)$ iff $\operatorname{div}(f) \geq -\operatorname{div}(\omega)$ iff $\operatorname{div}(f\omega) \geq 0$ iff $f\omega$ is holomorphic. But every differential on C is of the form $f\omega$, so we have

$$\mathscr{L}(K_c) \simeq \{\omega \in \Omega_C : \omega \text{ is holomorphic}\}.$$

Theorem 2.13 (Riemann-Roch) Let C be a smooth curve and K_C a canonical divisor on C. There is an integer $g \geq 0$ (the **genus** of C) such that for every divisor $D \in \operatorname{Div}(C)$, we have $\ell(D) - \ell(K_C - D) = \deg(D) - g + 1$.

Corollary 2.14

(a) $\ell(K_C) = g$.

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

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

Proof.

(a) Take $D = 0$ in Riemann-Roch.

(b) Take $D = K_C$ in Riemann-Roch, and apply (a).

(c) Observe that from (b), we have $\deg(D) > 2g - 2$, so $\deg(K_C - D) < 0$, and so $\ell(K_C - D) = 0$. Now apply Riemann-Roch.

Example. Let $C = \mathbb{P}^1$. There are no holomorphic differentials on $\mathbb{P}^1$, so $\ell(\mathbb{P}^1) = 0$. Thus the genus of $\mathbb{P}^1$ is 0. Applying Riemann-Roch gives $\ell(D) - \ell(-2(\infty) - D) = \deg(D) + 1$. If $\deg(D) \geq -1$, then $\ell(-2(\infty) - D) = 0$, and we have $\ell(D) = \deg(D) + 1$.

Example. Suppose that $\mathrm{char}(K) \neq 2$ and that $e_1, e_2, e_3 \in \bar{K}$ are distinct. Let

$$C: \qquad y^2 = (x - e_1)(x - e_2)(x - e_3). \tag{$\dagger$}$$

Exercise: Show that C is smooth and has a single point $P_\infty = [0, 1, 0]$ at ∞. Set $P_i = (e_i, 0) \in C$ for $1 \leq i \leq 3$.

(a) For example,

$$x - e_1 = \frac{y^2}{(x - e_2)(x - e_3)},$$

and $\mathrm{div}(x - e_1) = 2(P_1) - 2(P_\infty)$. Now $\mathrm{div}(x - e_i) = 2(P_i) - 2(P_\infty)$ $(1 \leq i \leq 3)$ and ($\dagger$) give

$$\mathrm{div}(y) = (P_1) + (P_2) + (P_3) - 3(P_\infty).$$

(b) Let's compute $\mathrm{div}(dx)$. [Recall: if $\beta \in \bar{K}(C)$ with $\bar{K}(C)/\bar{K}(\beta)$ separable and $\beta(P) = 0$, then

$$\mathrm{ord}_P(\alpha\, d\beta) = \mathrm{ord}_P(\alpha) + \mathrm{ord}_P(\beta) - 1$$

for all $\alpha \in \bar{K}(C)$ (Theorem 2.7(c)).] Now we have $(1 \leq i \leq 3)$

$$dx = d(x - e_i) = -x^2\, d\left(\frac{1}{x}\right).$$

Thus $\mathrm{ord}_{P_i}(dx) = \mathrm{ord}_{P_i}(d(x - e_i)) = 1$, and

$$\begin{aligned}
\mathrm{ord}_{P_\infty}(dx) &= \mathrm{ord}_{P_\infty}\left(-x^2\, d\left(\frac{1}{x}\right)\right) \\
&= \mathrm{ord}_{P_\infty}(-x^2) + \mathrm{ord}_{P_\infty} d\left(\frac{1}{x}\right) \\
&= \mathrm{ord}_{P_\infty}(-x^2) + \mathrm{ord}_{P_\infty}\left(\frac{1}{x}\right) - 1 \\
&= -4 + 2 - 1 \\
&= -3.
\end{aligned}$$

At other points $Q \in C$, the map $x : C \to \mathbb{P}^1$ given by

$$P \mapsto \begin{cases} [x(P), 1] & \text{if } x \text{ is regular at } P, \\ [1, 0] & \text{otherwise,} \end{cases}$$

is unramified, and so $x - x(Q)$ is a uniformizer at Q. So $\operatorname{ord}_Q(dx) = \operatorname{ord}_Q(d(x - x(Q))) = 0$. Hence

$$\operatorname{div}(dx) = (P_1) + (P_2) + (P_3) - 3(P_\infty) = \operatorname{div}(y).$$

Therefore $\operatorname{div}\left(\frac{dx}{y}\right) = 0$, and so $\frac{dx}{y}$ is a nonvanishing holomorphic differential on C.

(c) $\operatorname{div}\left(\frac{dx}{y}\right) = 0$, so $K_C = 0$. Thus g, the genus of C, is equal to $\ell(K_C) = \ell(0) = 1$. Riemann-Roch tells us that $\ell(D) = \deg(D)$ if $\deg(D) \geq 1$.

Some special cases.

(i) Let $P \in C$. Then $\ell((P)) = 1$, so $\mathscr{L}((P)) = \bar{K}$ (since certainly $\bar{K} \subseteq \mathscr{L}((P))$!). So there are no functions on C that have a single simple pole.

(ii) $\ell(2(P_\infty)) = 2$. A basis for $\mathscr{L}(2(P_\infty))$ is $\{1, x\}$.

(iii) A basis for $\mathscr{L}(3(P_\infty))$ is $\{1, x, y\}$. A basis for $\mathscr{L}(4(P_\infty))$ is $\{1, x, y, x^2\}$.

(iv) Observe that $\{1, x, y, x^2, xy, y^2, x^3\} \subseteq \mathscr{L}(6(P_\infty))$. But $\ell(6(P_\infty)) = 6$, so these functions are R-linearly dependent.

Theorem 2.15 (Hurwitz Genus Theorem). Let $\varphi : C_1 \to C_2$ be a nonconstant separable map of smooth curves with g_i the genus of C_i. Then

$$2g_1 - 2 \geq \deg(\varphi)(2g_2 - 2) \geq \deg(\varphi)(2g_2 - 2) + \sum_{P \in C_1} (e_\varphi(P) - 1),$$

with equality iff either

(i) $\operatorname{char}(K) = 0$, or

(ii) $\operatorname{char}(K) = p > 0$ and $p \nmid e_\varphi(P)$ for all $P \in C_1$.

Proof. Let $\varphi : C_1 \to C_2$ be given by $P \mapsto Q := \varphi(P)$, and let $\omega \in \Omega_{C_2}$, $\omega \neq 0$. If φ is separable, then $\varphi^*\omega \neq 0$. The strategy is to compare $\operatorname{ord}_P(\varphi^*\omega)$ with $\operatorname{ord}_Q(\omega)$ and use $\deg(\operatorname{div}(\varphi^*\omega)) = 2g_1 - 2$.

Set $\omega = f\ dt$, with $t \in \bar{K}(C_2)$ a uniformizer at Q. Then $\varphi^* t = us^e$, $e := e_\varphi(P)$, s a uniformizer at P, and $U(P) \neq 0$. Then

$$\varphi^*\omega = (\varphi^* f)\ d(\varphi^* t) = (\varphi^* f)\ d(us^e) = (\varphi^* f) \cdot \left(eus^{e-1} + \frac{du}{ds} s^e \right)\ ds.$$

Now if u is regular at P, then $\frac{du}{ds}$ is regular at P, i.e. $\operatorname{ord}_P\left(\frac{du}{ds}\right) \geq 0$, so $\operatorname{ord}_P(\varphi^*\omega) \geq \operatorname{ord}_P(\varphi^* f) + e - 1$, with equality iff $e \neq 0$ in K. Also $\operatorname{ord}_P(\varphi^* f) = e_\varphi(P) \operatorname{ord}_Q(f) = e_\varphi(P) \operatorname{ord}_Q(\omega)$. Hence

$$\begin{aligned}
\deg(\operatorname{div}(\varphi^*\omega)) &\geq \sum_{P \in C_1} [e_\varphi(P) \operatorname{ord}_{\varphi(P)}(\omega) + e_\varphi(P) - 1] \\
&= \sum_{Q \in C_2} \sum_{P \in \varphi^{-1}(Q)} e_\varphi(P) \operatorname{ord}_Q(\omega) + \sum_{P \in C_1} (e_\varphi(P) - 1) \\
&= (\deg \varphi^*)(\deg(\operatorname{div}(\omega))) + \sum_{P \in C_1} (e_\varphi(P) - 1).
\end{aligned}$$

Hence

$$2g_1 - 2 \geq (\deg \varphi)(2g_2 - 2) + \sum_{P \in C_1} (e_\varphi(P) - 1).$$

Chapter 3

The Geometry of Elliptic Curves

Definition 3.1 An elliptic curve E/K is a smooth curve over K, of genus 1, with a specified point $O \in E(K)$.

Example. (Weierstraß Curves). Assume $\text{char}(K) \neq 2$ or 3. In $\mathbb{P}^2$, take the curve C (which we suppose to be smooth)

$$y^2z + a_1xyz + a_3yz^2 - x^3 - a_2x^2z - a_4xz^2 - a_6z^3 = 0,$$

$a_i \in K$ for all i. The affine equation is

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4 + a_6.$$

Set $O := [0, 1, 0]$ and $f(x, y) := y^2 + a_1xy + a_3y - (x^2 + a_2x^2 + a_4x + a_6)$. Define a differential

$$\omega = \frac{dx}{2y + a_1x + a_3} = \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y}.$$

[Note: We have equality above because the left side is $\frac{dx}{f_y(x,y)}$ and the right side is $\frac{-dy}{f_x(x,y)}$, and equality results from $f_x(x, y)\ dx + f_y(x, y)\ dy = 0$.]

We claim that ω is holomorphic and nonvanishing. If $P = (x_0, y_0)$ were a pole of ω, then we would have $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$, which is a contradiction since C is smooth. Consider the map $C \to \mathbb{P}^1$ given by $[x, y, 1] \mapsto [x, 1]$ — this map is of degree 2. Thus $\text{ord}_P(x - x_0) \leq 2$. $\text{ord}_P(x - x_0) = 2$ iff $f(x_0, y)$ has a double root iff $f_y(x_0, y_0) = 0$. Now $\omega = \frac{dx}{f_y(x,y)} = \frac{d(x-x_0)}{f_y(x,y)}$, so

$$\text{ord}_P(\omega) = \text{ord}_P(x - x_0) - \text{ord}_P(f_y) - 1 = 0.$$

We now check $P = \underline{0}$. $\text{ord}_O(x) = -2$ and $\text{ord}_O(y) = -3$. So if t is a uniformizer at O, then $x = t^{-2}F$ and $y = t^{-3}G$, where $F(O) \neq 0$ or ∞, and $G(O) \neq 0$ or ∞. So

$$\omega = \frac{dx}{f_y(x,y)} = \left(\frac{-2t^{-3}F + t^{-2}F'}{2t^{-3}G + a_1t^{-2}F + a_3} \right) dt$$

(where $F' = \frac{dF}{dt}$). Since F is regular at O, $\frac{dF}{dt}$ is also regular at O (Theorem 2.8(4)). Thus $\frac{-2F+tF'}{2G+a_1tF+a_3t^3}$ is regular and nonvanishing at O (char$(K) \neq 2$!). Thus $\text{ord}_O(\omega) = 0$ as desired. [If char$(K) = 2$, then in fact the same assertion holds, as may be seen by calculating with $\omega = \frac{-dy}{f_x(x,y)}$ instead.] Hence ω is holomorphic and nonvanishing (i.e. $(\omega) = 0$). Now apply Riemann-Roch: $\deg(\omega) = 2g - 2$, so $g = 1$.

What happens if a Weierstraß curve is singular?

Lemma 3.2 If C is defined by a Weierstraß equation and is *not* smooth, then there is a map $C \to \mathbb{P}^1$ of degree 1.

Proof. Suppose $(0,0)$ is the singular point. Then $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$. The Weierstraß equation for C is of the form $y^2 + a_1xy + x^3 + a_2x^2$. Consider the map $C \to \mathbb{P}^1$ given by $(x,y) \mapsto \frac{y}{x}$. We have

$$\left(\frac{y}{x}\right)^2 + a_1\left(\frac{y}{x}\right) = x + a_2,$$

and so there is exactly one inverse image of each $\frac{y}{x}$, as required.

Theorem 3.3 If E/K is an elliptic curve, then there exist $a_1, a_2, a_3, a_4, a_6 \in K$ such that E is isomorphic to the Weierstraß elliptic curve $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$.

Proof. Recall that $\mathscr{L}(d(O)) = \{$all functions on E with no poles except possibly a pole of order at most d at $O\}$. Riemann-Roch tells that $\ell(d(O)) = d-g+1$ if $d > 2g-2$, which is d if $d \geq 1$ (since here $g = 1$). Applying this tells us that $\mathscr{L}((O)) = \bar{K}$.

$\mathscr{L}(2(O)) = \langle 1, x\rangle$, where $x \in \mathscr{L}(2(O)) - \mathscr{L}((O))$, and so $\mathrm{ord}_O(x) = -2$, and x has no other poles. $\mathscr{L}(3(O)) = \langle 1, x, y\rangle$, where $y \in \mathscr{L}(3(O)) - \mathscr{L}(2(O))$, and so $\mathrm{ord}_O(y) = -3$, and y has no other poles. $\mathscr{L}(4(O)) = \langle 1, x, y, x^2\rangle$. $\mathscr{L}(5(O)) = \langle 1, x, y, x^2, xy\rangle$. $\mathscr{L}(6(O)) = \langle 1, x, y, x^2, xy, x^3\rangle$ or $\langle 1, x, y, x^2, xy, y^2\rangle$, and we know that x^3 and y^2 are not independent, since $\ell(6(O)) = 6$. So $\{1, x, y, x^2, xy, x^3, y^2\}$ are linearly dependent. Hence we can write $A_0y^2 + A_1xy + A_3y = A_0'x^3 + A_2x^2 + A_4x + A_6$ with $A_0A_0' \neq 0$. Without loss of generality, $A_0 = 1$. Perform the transformation $x \mapsto A_0'x$, $y \mapsto A_0'y$; then without loss of generality $A_0 = A_0' = 1$. So the equation becomes

$$C : y^2 + A_1xy + A_3y = x^2 + A_2x^2 + A_4x + A_6.$$

Define a map $\varphi : E \to C$ via $\varphi^* : x \mapsto x$, $y \mapsto y$. To show that φ is an isomorphism, it suffices to show that φ is of degree 1, and C is smooth. We have $x : E \to \mathbb{P}^1$ with $\deg(x) = 2$. So $[K(E) : K(x)] = 2$. Similarly, since $\deg(y) = 3$, $[K(E) : K(y)] = 3$. Hence $[K(E) : K(x, y)]$ divides both 2 and 3, and so $K(E) = K(x, y)$. Thus

$$\deg(\varphi) = [K(E) : \varphi^*K(C)] = [K(E) : K(x, y)] = 1.$$

Suppose now that C is not smooth. Then there exists a map $\psi : C \to \mathbb{P}^1$ of degree 1 (Lemma 3.2), and so $\psi \circ \varphi$ is an isomorphism, since both E and $\mathbb{P}^1$ are smooth. This is impossible, since $\mathbb{P}^1$ does not have genus 1. Thus C is smooth, and so φ is an isomorphism.

Corollary 3.4 The Weierstraß coordinates x and y on E are unique up to $x \mapsto u^2x' + r$ and $y \mapsto u^3y' + su^2x' + t$, with $u, r, s, t \in K$, $u \neq 0$.

Proof. Suppose $\{x, y\}$ and $\{x', y'\}$ are two sets of Weierstraß coordinates on E. Then $\mathrm{ord}_O(x) = \mathrm{ord}_O(x') = -2$, and $\mathrm{ord}_O(y) = \mathrm{ord}_O(y') = -3$, so $\{1, x\}$ and $\{1, x'\}$ are bases of $\mathscr{L}(2(O))$, and $\{1, x, y\}$ and $\{1, x', y'\}$ are bases of $\mathscr{L}(3(O))$. Thus there exist $u_1, u_2, r, s_2, t \in K$ with $u_1u_2 \neq 0$ such that $x = u_1x' + r$ and $y = u_2y' + s_2x' + t$. (x, y) and (x', y') both satisfy Weierstraß equations with coefficients of Y^2 and X^3 equal to 1, so $u_1^3 = u_2^3$. Now set $u = u_2/u_1$ and $s = s_2/u^2$ to obtain the result.

3.1 The Addition Law on E

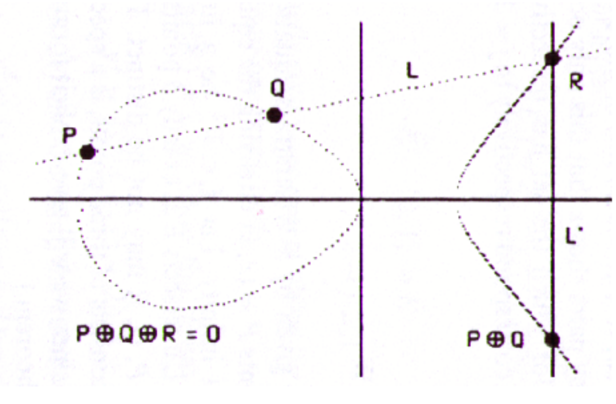
Proposition 3.5 If $D \in \mathrm{Div}_K^0(E)$, then there is a unique point $P \in E(K)$ such that $D \sim (P) - (O)$. (Recall that $A \sim B$ iff there exists $f \in K(E)$ such that $(f) = A - B$.)

Proof. It follows from the Riemann-Roch Theorem that $\ell(D + (O)) = 1$, since $\deg(D) = 0$. Thus there exists $f \in \mathscr{L}(D + (O))$ with $(f) \geq -D - (O)$. Since $\deg(f) = 0$, there exists a point P such that $(f) = -D - (O) + (P)$. Thus $D \sim (P) - (O)$, and this demonstrates the existence of P. Next, observe that if $D \sim (P') - (O)$, then there exists $g \in K(E)$ such that $(g) = -D - (O) + (P')$, so $g \in \mathscr{L}(D + (O))$, so $g = cf$ for some $c \in K^\times$, since $\ell(D + (O)) = 1$. Thus $(g) = (f)$, and so $(P) = (P')$.

Thus we have a map $\sigma : \mathrm{Pic}_K^0(E) \to E(K)$ given by $[D] \mapsto P$, where $D \sim (P) - (O)$. σ is plainly surjective. It is injective because if $\sigma(D) = O$, then $D \sim (O)$. The inverse of σ is the map $\kappa : E(K) \to \mathrm{Pic}_K^0(E)$ given by $P \mapsto [(P) - (O)]$.

3.2 Another Description of the Addition Law

We define a composition law $\oplus$ on E as follows: Let $P, Q \in E$, let L be the line connecting P and Q, and let R be the third point of intersection of L with E (Bézout's Theorem). Let L' be the line connecting R and O. $P \oplus Q :=$ the point on E such that L' intersects E at R, O, and $P \oplus Q$.



To show that this law of composition is the same as the one defined above, it suffices to show that $\kappa(P \oplus Q) = \kappa(P) + \kappa(Q)$. (Here "+" means addition of divisor classes in $\mathrm{Pic}^0_K(E)$.) Let $f = \alpha X + \beta Y + \gamma Z = 0$ be the equation of the line L' connecting R and O. Then $(f) = (P) + (Q) + (R) - 3(O)$, $(f') = (R) + (P \oplus Q) - 2(O)$ (since f and f' have no poles in the affine plane). Thus

$$\left(\frac{f'}{f}\right) = (P \oplus Q) - (P) - (Q) + (O) \sim (O),$$

and this implies that $\kappa(P \oplus Q) - \kappa(P) - \kappa(Q) = 0$.

The addition law on E is a *morphism*. We have that (see Silverman III, §2.3)

$$(x_1, y_1) + (x_2, y_2) = \left(\frac{*}{(x_2 - x_1)^3}, \frac{*}{(x_2 - x_1)^3}\right)$$

if $x_1 \neq x_2$. So the addition map is regular except possibly at (P, P), $(P, -P)$, (P, O), and (O, P). To take care of these points: For $Q \in E(\bar{K})$, consider the map $\tau_Q : E \to E$ given by $P \mapsto P + Q$; this is a morphism (even an isomorphism!). Now look at

$$E \times E \overset{(\tau_{Q_1}, \tau_{Q_2})}{\longrightarrow} E \times E \overset{+}{\longrightarrow} E \overset{\tau_{-Q_1-Q_2}}{\longrightarrow} E$$

given by

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

Choose Q_1 and Q_2 to avoid the "bad set."

3.3 Isogenies

Definition 3.6 An **isogeny** from E_1 to E_2 (elliptic curves) is a morphism $\varphi : E_1 \to E_2$ with $\varphi(O) = O$. (In particular, according to this definition, $E_1 \to O$ *is* an isogeny.) Say that E_1 and E_2 are **isogenous** if there is a *nonconstant* isogeny $\varphi : E_1 \to E_2$.

(If $\psi : E_1 \to E_2$ is a morphism, then $\tau_{-\psi(O)}$ is an isogeny.)

Theorem 3.7 If $\varphi : E_1 \to E_2$ is an isogeny, then φ is a group homomorphism, i.e. $\varphi(P+Q) = \varphi(P) + \varphi(Q)$.

Proof. This follows from the fact that the following diagram commutes:

$$\begin{tikzcd}
E_1(\bar{K}) \arrow[r, "\sim"] \arrow[d, "\varphi"'] & \mathrm{Pic}^0_{\bar{K}}(E_1) \arrow[d, "\varphi_*"] \\
E_2(\bar{K}) \arrow[r, "\sim"] & \mathrm{Pic}^0_{\bar{K}}(E_2)
\end{tikzcd}$$

and three of the arrows (i.e. all except possibly φ) are group homomorphisms:

$$\begin{tikzcd}
P \arrow[r, mapsto] \arrow[d, mapsto] & (P) - (O) \arrow[d, mapsto] & \\
\varphi(P) \arrow[r, mapsto] & (\varphi(P)) - (\varphi(O)) \arrow[r, equal] & (\varphi(P)) - (O)
\end{tikzcd}$$

Notation. Set $\mathrm{Hom}(E_1, E_2) = \{\text{isogenies } E_1 \to E_2\}$. This is a group under addition on E_2. $\mathrm{End}(E) = \mathrm{Hom}(E, E)$ is a ring under addition and composition.

Examples.

1. Let $n \in \mathbb{Z}$. $[n]$ is multiplication by n, $[n] \in \mathrm{End}(E)$.
2. If $\mathrm{char}(K) = p > 0$, then the Frobenius map $\varphi : x \mapsto x^r$ on K induces a map $\varphi : E \to E^{(p)}$ via $(x, y) \mapsto (x^p, y^p)$. If $K = \mathbb{F}_q$ (q is a power of p), then $\varphi^{(q)} : E \to E$, and $\varphi^{(q)}$ is an endomorphism of degree q.
3. Consider the curve $y^2 = x^3 - x$, and suppose $\sqrt{-1} \in K$. Then we may define a map $\varphi : E \to E$ by $(x, y) \mapsto (-x, iy)$, where $i = \sqrt{-1}$. Note that $\varphi \neq [n]$ for any n, since $\varphi^2 = [-1]$.

Theorem 3.8 $[n]$ is nonzero for all $n \neq 0$. $\mathrm{Hom}(E_1, E_2)$ is a torsionfree $\mathbb{Z}$-module. $\mathrm{End}(E)$ is an integral domain of characteristic 0. Define $\deg([0]) = 0$. Then $\deg(\psi \circ \varphi) = \deg(\psi)\deg(\varphi)$.

Proof. We make the following claims:

1. There exists $P \in E(\bar{K})$ with $2P \neq O$.
2. There exists $Q \in E(\bar{K})$, $Q \neq O$, with $2Q = O$.

Note that (1)implies that $[2] \neq 0$. If n is odd, and Q is as is (2), then $[n]Q = Q \neq O$. So $[n] \neq 0$. (This implies that $[n] \neq 0$ for all $n \neq 0$ — any map between two smooth curves is either constant or surjective.) To see that $\mathrm{End}(E)$ is an integral domain, suppose that $\varphi \circ \psi = 0$. Then $\deg(\varphi)\deg(\psi) = 0$, so $\deg \varphi = 0$ or $\deg \psi = 0$, and so $\varphi = 0$ or $\psi = 0$. Hence $\mathrm{End}(E)$ is an integral domain. A similar arguments shows that $\mathrm{End}(E)$ is of characteristic 0, and that $\mathrm{Hom}(E_1, E_2)$ is $\mathbb{Z}$-torsionfree.

We now prove the claim.

$$x(2P) = \frac{x^4 - b_4x^2 - 2b_6x - b_8}{4x^3 + b_2x^2 + 2b_4x + b_6}$$

(see III, §2.3 in Silverman). There exists an $x \in \bar{K}$ such that this function has no pole at x. Choose the corresponding $y \in \bar{K}$. Then $2(x, y) \neq O$. This proves (1). For (2), we want $x \in \bar{K}$ which *is* a pole of $x(2P)$. Check that the polynomial $4x^3 + b_2x^2 + 2b_4x + b_6$ does not divide $x^4 - b_4x^2 - 2b_6x - b_8$. Choose such an x and the corresponding y. Then $(x, y) \neq O$, but $2(x, y) = O$.

Proposition 3.9 Let $\varphi : E_1 \to E_2$ be a nonconstant isogeny. Then

1. $\#\varphi^{-1}(Q) = \deg_s \varphi$. $e_\varphi(P) = \deg_i \varphi$. ($P \in \varphi^{-1}(Q)$, say.)
2. The map $\ker(\varphi) \to \mathrm{Aut}(K'(E_1)/\varphi^* K'(E_2))$ given by $R \mapsto \tau_R^*$ is an isomorphism. (Here K' is any field big enough to contain the coordinates of all $R \in \ker(\varphi)$.)
3. If φ is separable, then φ is unramified, $\# \ker \varphi = \deg \varphi$, and $K'(E_1)/\varphi^* K'(E_2)$ is Galois.

Proof.

(1) Plainly $\#\varphi^{-1}(Q) = \#\ker\varphi$ for any Q since φ is a group homomorphism. But $\#\varphi^{-1}(Q) = \deg_s \varphi$ for almost all Q (Theorem 2.4(1)), and so it is equal to $\deg_s \varphi$ always. Next, we claim that $e_\varphi(P)$ is independent of the choice of $P \in \varphi^{-1}(Q)$. For if $R \in \ker\varphi$, then

$$e_{\varphi\circ\tau_R}(P) = e_{\tau_R}(P)e_\varphi(\tau_R(P)) = e_{\tau_R}(P)e_\varphi(P+R),$$

and $e_{\tau_R}(P) = 1$, since τ_R is an isomorphism. Now observe that $e_{\varphi\circ\tau_R}(P) = e_\varphi(P)$ since $\tau_R = \varphi$ (remember $R \in \ker\varphi$!). We have

$$\sum_{P\in\varphi^{-1}(Q)} e_\varphi(P) = \deg\varphi = \deg_i\varphi\deg_s\varphi,$$

i.e. $e_\varphi(P)\deg_s\varphi = \deg_i\varphi\deg_s\varphi$, so $e_\varphi(P) = \deg_i\varphi$.

(2) Since $\varphi \circ \tau_R = \varphi$, we have $\tau_R^*\varphi^* = \varphi^*$ (induced maps of function fields). So τ_R^* acts as the identity on $\varphi^*K'(E_2)$. Hence we have a map $\ker\varphi \to \mathrm{Aut}(K'(E_1)/\varphi^*K'(E_2))$; the left side has order $\deg_s\varphi$, while the right side has order at most $\deg_s\varphi$. So it suffices to show that the map is injective.

Suppose that $\tau_R^* = id$ on $K'(E_1)$. This implies that for all $f \in K'(E_1)$, we have $f(P+R) = f(P)$ for all $P \in E_1(\bar{K})$. In particular, $x(P+R) = x(P)$, $y(P+R) = y(P)$, so $P+R = P$, so $R = O$. Hence the map is both injective and surjective, and so is an isomorphism.

Corollary 3.10 Suppose that $\varphi : E_1 \to E_2$ and $\psi : E_1 \to E_3$ are isogenies, with φ separable and $\ker\varphi \subseteq \ker\psi$. Then there exists an isogeny $\lambda : E_2 \to E_3$ with $\lambda\circ\varphi = \psi$.

$$\begin{tikzcd}
E_1 \arrow[r, "\varphi"] \arrow[d, "\psi"'] & E_2 \arrow[dl, dashed, "\lambda"] \\
E_3 &
\end{tikzcd}$$

Proof. Let K' be a field of rationality for $\ker\psi$. Theorem 3.9 implies that

$$\varphi^*K'(E_2) = K'(E_1)^{\{\tau_R^* : R\in\ker\varphi\}}.$$

$\psi^*K'(E_3)$ is fixed by all τ_R^* with $R \in \ker\psi$. Thus we have $K'(E_1) \supseteq \varphi^*K'(E_2) \supseteq \psi^*K'(E_3)$. This implies that there exists $\lambda : E_2 \to E_3$ such that $\lambda\circ\varphi = \psi$, with

$\lambda(O) = \lambda(\varphi(O)) = \psi(O) = O$.

Theorem 3.11 Given a finite subgroup Φ of $E(\bar{K})$, there is an elliptic curve E' and a separable isogeny $\varphi : E \to E'$ with $\ker \varphi = \Phi$. Furthermore, (φ, E') is unique up to isomorphism. We write $E' = E/\Phi$.

Proof. Set $G = \{\tau_R^* : R \in \Phi\}$. Then G acts as a group of automorphisms of $\bar{K}(E)$, and, via Galois theory, we have that $[\bar{K}(E) : \bar{K}(E)^G] = \#\Phi$. Thus there exists a non-singular curve $C/\bar{K}$ and a finite morphism $\varphi : E \to C$ such that $\varphi^* \bar{K}(C) = \bar{K}(E)^G$.

We claim that $\bar{K}(E)/\bar{K}(C)$ is unramified. To see that this claim is true, suppose that $Q \in C(\bar{K})$. Then if $\varphi(P) = Q$, then we have that $\varphi(P + R) = Q$ for all $R \in \Phi$. $\#\varphi^{-1}(Q) \geq \#\Phi$, $\#\varphi^{-1}(Q) = \deg_s \varphi$. Since

$$\sum_{P \in \varphi^{-1}(Q)} e_\varphi(P) = \deg(\varphi),$$

this implies that $e_\varphi(P) = 1$ for all $P \in \varphi^{-1}(Q)$, and that φ is separable.

Now apply the Hurwitz genus formula (Theorem 2.15): $2g_E - 2 = \deg \varphi (2g_C - 2)$, so $g_C = 1$. Now define $O_C = \varphi(O_E)$. Then C is an elliptic curve, and φ is an isogeny with $\ker \varphi = \Phi$. Uniqueness follows from the fact that if $\ker \varphi \subseteq \ker \psi$, with φ separable, then we have

$$\begin{tikzcd} E \arrow[r, "\varphi"] \arrow[dr, "\psi"'] & E_1 \arrow[d, dashed, leftrightarrow, "{\lambda,\lambda^{-1}}"] \\ & E_2 \end{tikzcd}$$

(cf Corollary 3.10).

3.4 Invariant Differentials

Now let E/K be an elliptic curve with Weierstraß equation

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

Set

$$\omega = \frac{dx}{2y + a_1x + a_3} = \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y}.$$

We showed earlier that $(\omega) = 0$.

Theorem 3.12 If $Q \in E(\bar{K})$, then $\tau_Q^*\omega = \omega$.

Proof. Since Ω_E is 1-dimensional, we have that $\tau_Q^*\omega = f_Q\omega$, $f_Q \in \bar{K}(E)^\times$. Since τ_Q^* is an isomorphism, it follows that $(\tau_Q^*\omega) = 0$, whence $(f_Q) + (\omega) = 0$, so $(f_Q) = 0$, so $f_Q \in \bar{K}$. We note that $Q \mapsto f_Q \in \bar{K}^\times$ is a rational map $E \to \mathbb{P}^1$ (this is clear because we could do everything explicitly and express f_Q as a rational function of $x(Q)$ and $y(Q)$). This map is *not* surjective, since it misses 0 and ∞. This implies that $Q \mapsto f_Q$ is a constant map, and so $f_Q = f_O = 1$.

Theorem 3.13 If $\varphi, \psi : E \to E'$ are isogenies, then $\varphi^*\omega + \psi^*\omega' = (\varphi + \psi)^*\omega'$.

Proof. If $f_1, f_2 \in \bar{K}(E)$ satisfy the Weierstraß equation of E, define

$$\omega(f_1, f_2) = \frac{df_1}{2f_2 + a_1f_1 + a_3} = \frac{df_2}{3f_1^2 + 2a_2f_1 + a_4 - a_1f_2}$$

(so e.g. $\omega(x, y) = \omega$, using our earlier notation). We wish to prove that

$$\omega(\varphi^*(x', y')) + \omega(\psi^*(x', y')) = \omega((\varphi + \psi)^*(x', y')).$$

We claim that $\omega(f_1, f_2) + \omega(g_1, g_2) = \omega((f_1, f_2), (g_1, g_2))$. Now

$$\omega((f_1, f_2) + (g_1, g_2)) = F(f_1, f_2, g_1, g_2)\omega(f_1, f_2) + G(f_1, f_2, g_1, g_2)\omega(g_1, g_2), \qquad (\dagger)$$

and to establish the claim, we have to show that F and G are identically 1. Take $(g_1, g_2) = Q \in E(\bar{K})$ and $(f_1, f_2) = (x, y)$. Then

$$\omega((f_1, f_2) + (g_1, g_2)) = \tau_Q^*\omega = \omega.$$

The right side of $(\dagger)$ is $F(x, y, Q)\omega = 1 \cdot \omega$, and this holds for all $Q \in E(\bar{K})$. This implies that $F(f_1, f_2, g_1, g_2)$ is identically 1. Now choose $(f_1, f_2) = Q$, $(g_1, g_2) = (x, y)$ and use a similar argument to deduce that $G(f_1, f_2, g_1, g_2)$ is identically 1.

Theorem 3.14 $[m]^*\omega = m\omega$.

Proof. The result is true for 0 and 1, and so, by induction, it's true for $n+1$, since we have $[n+1]^*\omega = [n]^*\omega + [1]^*\omega$.

Application. $[m]$ is separable iff $(m, \operatorname{char}(K)) = 1$, or $\operatorname{char}(K) = 0$ and $m \neq 0$. [If C_1 and C_2 are curves, and $\varphi : C_1 \to C_2$ is a morphism, then φ is separable iff $\varphi^* : \Omega_{C_2} \to \Omega_{C_1}$ is nonzero (or, equivalently, injective). See Silverman II, 4.2(c).]

Consider the Frobenius isogeny $\varphi_q : E \to E^{(q)}$ given by $x \mapsto x^q$. Then $\varphi^* dx = d(x^q) = 0$, and so φ is not separable, since $\varphi^*\omega = 0$. Now suppose that $E/\mathbb{F}_q$; then $\varphi_q : E \to E$, and $1 - \varphi_q$ is separable, since $(1-\varphi_q)^*\omega = \omega$. This is useful: Observe that

$$E(\mathbb{F}_q) = E(\bar{\mathbb{F}}_q)^{\varphi_q} = \ker(1 - \varphi_q).$$

Thus $\#E(\mathbb{F}_q) = \deg(1 - \varphi_q)$.

3.5 Dual Isogenies

Suppose that we have an isogeny $E \xrightarrow{\varphi} E'$. This induces $\operatorname{Pic}^0(E) \xleftarrow{\varphi^*} \operatorname{Pic}^0(E')$. Since we may identify $\operatorname{Pic}^0(E)$ with E, we want to think of φ^* as being a map $E' \xrightarrow{\varphi^*} E$.

Theorem 3.15 Suppose that $\varphi : E \to E'$ is an isogeny of degree m. Then there exists a unique $\hat{\varphi} : E' \to E$ such that $\hat{\varphi} \circ \varphi = [m]$. Furthermore, $\hat{\varphi}$ is given by

$$E' \xrightarrow{\sim} \operatorname{Pic}^0(E') \xrightarrow{\varphi^*} \operatorname{Pic}^0(E) \xrightarrow{\sim} E.$$

Proof. We first show uniqueness. Suppose that $\hat{\varphi} \circ = \varphi = \hat{\varphi}' \circ \varphi = [m]$. Then $(\hat{\varphi} - \hat{\varphi}') \circ \varphi = 0$. Since φ is nonconstant, it follows that $\hat{\varphi} - \hat{\varphi}'$ is constant, whence

$\hat{\varphi} = \hat{\varphi}'$.

We now show existence. Suppose that φ is separable. Then $\# \ker \varphi = m = \deg \varphi$, and so it follows that $\ker \varphi \subseteq \ker[m]$. Via Corollary 3.10, we see that there is an isogeny $\hat{\varphi} : E' \to E$ such that the following diagram commutes:

$$\begin{tikzcd}
E \arrow[r, "\varphi"] \arrow[dr, "{[m]}"'] & E' \arrow[d, dashed, "\hat{\varphi}"] \\
& E
\end{tikzcd}$$

Suppose now that $\mathrm{char}(K) = p > 0$. Then if ω is an invariant differential on E, we have $[p]^*\omega = p\omega = 0$ (Theorem 3.14), and so $[p]$ is *not* separable. Hence $[p] = \lambda \circ F^e$, where λ is separable, F is the Frobenius map $x \mapsto x^p$, and $e \geq 1$ (Theorem 2.5(4)). So, we define $\hat{F} = \lambda \circ F^{e-1}$. Now observe that for *any* isogeny φ, we have $\varphi = \mu \circ F^r$, where F is Frobenius and μ is separable. Define $\hat{\varphi} = \hat{F}^r \circ \hat{\mu}$. Then

$$\hat{\varphi} \circ \varphi = (\hat{F}^r \circ \hat{\mu}) \circ (\mu \circ F^r) = \deg \mu \cdot p^r = \deg \varphi.$$

Suppose that $Q \in E'(K)$. What is $\hat{\varphi}(Q)$? First notice that $Q = \varphi(P)$ for some $P \in E(\bar{K})$, and so $\hat{\varphi}(Q) = \hat{\varphi}(\varphi(P)) = mP$. We have (under the composition described in the statement of the theorem):

$$\begin{aligned}
Q &\mapsto Q - O \\
&\mapsto \sum_{S \in \varphi^{-1}(Q)} e_\varphi(S)S - \sum_{R \in \varphi^{-1}(O)} e_\varphi(R)R \\
&= \deg_i \varphi \sum_{R \in \ker \varphi} (P + R - R) \\
&= (\deg_i \varphi)(\deg_s \varphi)P \\
&= (\deg \varphi)P.
\end{aligned}$$

Theorem 3.16 Suppose that $\varphi : E_1 \to E_2$ is an isogeny. Then

(1) $\hat{\varphi} \circ \varphi = \varphi \circ \hat{\varphi} = [\deg \varphi]$.

(2) If $\lambda : E_0 \to E_1$, then $\widehat{\varphi \circ \lambda} = \hat{\lambda} \circ \hat{\varphi}$.

(3) If $\lambda : E_1 \to E_2$, then $\widehat{\varphi + \lambda} = \hat{\varphi} + \hat{\lambda}$.

(4) $\widehat{[m]} = [m]$.

(5) $\deg \varphi = \deg \hat{\varphi}$.

(6) $\hat{\hat{\varphi}} = \varphi$.

Proof.

(1) By definition, we have $\hat{\varphi} \circ \varphi = [\deg \varphi]$. So

$$\varphi \circ \hat{\varphi} \circ \varphi = \varphi \circ [\deg \varphi] = [\deg \varphi] \circ \varphi,$$

and thus $\varphi \circ \hat{\varphi} = [\deg \varphi]$.

(2) Observe that we have

$$(\hat{\lambda} \circ \hat{\varphi}) \circ (\varphi \circ \lambda) = \hat{\lambda} \circ [\deg \varphi] \circ \lambda = [\deg \varphi][\deg \lambda] = [\deg(\varphi \circ \lambda)],$$

and now the result follows via the uniqueness of the dual isogeny.

(3) $\hat{\varphi}(Q) = \varphi^*(Q - O)$. So we need to show that $\varphi^* + \lambda^* = (\varphi + \lambda)^*$ on $\mathrm{Pic}(E_2)$. (See Silverman III, §6.2.)

(4) This is true for $m = 0$ and $m = 1$. Now observe that, using induction,

$$\widehat{[m \pm 1]} = \widehat{[m]} \pm \widehat{[1]} = [m] \pm [1] = [m \pm 1].$$

(5) First note that

$$[\deg[m]] = \widehat{[m]} \circ [m] = [m] \circ [m] = [m^2].$$

So $\deg[m] = m^2$. Now suppose $\deg \varphi = m$. Then we have

$$\begin{aligned} [m^2] &= [\deg[m]] \\ &= [\deg(\varphi \circ \hat{\varphi})] \\ &= [(\deg \varphi)(\deg \hat{\varphi})] \\ &= [m \circ \deg \hat{\varphi}]. \end{aligned}$$

Hence $\deg \hat{\varphi} = m$.

(6) Suppose that $m = \deg \varphi$. Then

$$\hat{\varphi} \circ \varphi = [m] = \widehat{[m]} = \widehat{\hat{\varphi} \circ \varphi} = \hat{\varphi} \circ \hat{\hat{\varphi}}.$$

Hence $\varphi = \hat{\hat{\varphi}}$.

So now we can describe $E_m = E[m]$, the kernel of $[m] : E(\bar{K}) \to E(\bar{K})$.

Case I. $\mathrm{char}(K) = 0$ or $\mathrm{char}(K) \nmid m$. Then

$$\#E_m = \#\ker([m]) = \deg_s[m] = \deg[m] = m^2,$$

so $E_m \simeq \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. (Look at the number of possible cyclic factors in the prime power case.)

Case II. Consider E_{p^e}, $p = \mathrm{char}(K)$. Then $\#E_{p^e} = \deg_s[p^e] = \deg_s(\hat{\varphi} \circ \varphi)^e$, where φ is the Frobenius map $x \mapsto x^p$. Then

$$\hat{\varphi} \circ \varphi = [p] = (\deg_s(\hat{\varphi} \circ \varphi))^e = \deg_s(\hat{\varphi})^e.$$

Then

$$\deg_s(\hat{\varphi}) = \begin{cases} 1 & \text{if } \hat{\varphi} \text{ is inseparable,} \\ p & \text{if } \hat{\varphi} \text{ is separable.} \end{cases}$$

So if $\hat{\varphi}$ is inseparable, then $\#E_{p^e} = 1$. If $\hat{\varphi}$ is separable, then $\#E_{p^e} = p^e$ for all e, so $E_{p^e} \simeq \mathbb{Z}/p^e\mathbb{Z}$.

We can now describe all of the possibilities for the automorphism algebra of an elliptic curve.

Properties of $\mathrm{End}(E)$**.**

- Ring with identity.
- No zero divisors.

- End(E) has an involution $\varphi \mapsto \hat{\varphi}$ which is additive and antimultiplicative, $-\varphi \circ \hat{\varphi} \in \mathbb{Z}$, and $\varphi \circ \hat{\varphi} \geq 0$, with equality iff $\varphi = 0$.

Theorem 3.17 (Hurwitz) Any ring R with the above properties is one of the following:

1. $\mathbb{Z}$.
2. An order in an imaginary quadratic field with ˆ being complex conjugation.
3. An order in a definite quaternion algebra over $\mathbb{Q}$ with ˆ being the canonical involution. [A definite quaternion algebra over $\mathbb{Q}$ is an algebra $\mathbb{Q}+\mathbb{Q}\alpha+\mathbb{Q}\beta+\mathbb{Q}\alpha\beta$, where $\alpha^2, \beta^2 \in \mathbb{Q}$, $\alpha^2, \beta^2 < 0$, and $\alpha\beta = -\beta\alpha$.]

Proof Sketch. We have $\mathbb{Z} \subseteq R$. If $\mathbb{Z} \subsetneq R$, choose $\alpha \in R$ such that $\alpha^2 \in \mathbb{Z}$, $\alpha^2 < 0$ (use reduced norms and traces to do this; $N(\alpha) = \alpha\hat{\alpha}$, $\mathrm{Tr}(\alpha) = \alpha + \hat{\alpha}$). Then $\mathbb{Z}[\alpha] \subseteq R$; if $\mathbb{Z}[\alpha]$ is of finite index in R, then we are done. If not, then find $\beta \in R$ with $\beta^2 \in \mathbb{Z}$, $\beta^2 < 0$, and $\alpha\beta = -\beta\alpha$. Then $\mathbb{Z}[\alpha, \beta, \alpha\beta] \subseteq R$. If $\operatorname{rank} R > 4$, then there exists a Cayley algebra contained in R, which is a contradiction since Cayley algebras are nonassociative [cf. J. Baez, "The Octonions," AMS Bulletin **39** (2002), 145–205].

If $\mathrm{char}(K) = 0$, then we have (1) or (2). If $\mathrm{char}(K) = p > 0$, then we have (2) or (3).

Proposition 3.18 Let E be an elliptic curve, and suppose that $D = \sum n_P(P) \in \mathrm{Div}(E)$. Then D is principal iff $\sum n_P = 0$ and $\sum [n_P]P = O$.

Proof. Recall (Proposition 3.5) that we have a map $\sigma : \mathrm{Pic}^0_K(E) \xrightarrow{\sim} E(K)$, $[D] \sim \mathrm{Div}(P) - (O) \mapsto P$. Every principal divisor has degree zero. Suppose that $D \in \mathrm{Div}^0(E)$. $D \sim 0$ iff $\sigma(D) = O$ iff $\sum [n_P]((P) - (O)) =)$ iff $\sum [n_P]P = O$.

3.6 The Weil Pairing

This is a pairing $[\cdot,\cdot]_m : E_m \times E_m \to \mu_m$, $\mathrm{char}(K) \nmid m$. It is bilinear, alternating, nondegenerate, and Galois equivariant.

Construction. Suppose that $S, T \in E_m$. Observe that the divisor $m(T) - m(O)$ is principal. Suppose $m(T) - m(O) = (f)$, say. Suppose that T' is such that $mT' = T$. Then

$$[m]^*(T) - m^*(O) = \sum_{R \in E_m} (T' + R) - (R),$$

and this is again a principal divisor equal to (g), say. Observe that

$$(f \circ [m]) = [m]^*(m(T) - m(O)) = m(g) = (g^m).$$

Therefore $f \circ [m]$ and g^m are the same up to a constant. Choose the constant implicit in the definition of f to ensure that $f \circ [m] = g^m$. Then

$$g(X+S)^m = f \circ [m](X+S) = f(mX + mS) = f(mX) = g(X)^m.$$

[So $m(g \circ \tau_S) = (f \circ \tau_{mS} \circ [m]) = (f \circ [m]) = m[g]$.] Hence we have that $\frac{g(X+S)}{g(X)} \in \mu_m \subseteq \bar{K}$, and we define

$$[S,T]_m = \frac{g(X+S)}{g(X)}.$$

This is the Weil pairing.

Bilinear in S.

$$[S_1 + S_2, T] = \frac{g(X + S_1 + S_2)}{g(X)} = \frac{g(X + S_1 + S_2)}{g(X + S_1)} \cdot \frac{g(X + S_1)}{g(X)} = [S_2, T][S_1, T].$$

Bilinear in T. Choose functions f_i and g_i with

$$\begin{aligned}
(f_1) &= m(T_1) - m(O), & (g_1) &= [m]^*(T_1) - [m]^*(O),\\
(f_2) &= m(T_2) - m(O), & (g_2) &= [m]^*(T_2) - [m]^*(O),\\
(f_3) &= m(T_1 + T_2) - m(O), & (g_3) &= [m]^*(T_1 + T_2) - [m]^*(O).
\end{aligned}$$

There exists a function h such that $(h) = (T_1 + T_2) - (T_1) - (T_2) + (O)$. We have

$$[S, T_1 + T_2] = \frac{g_3(X + S)}{g_3(X)}.$$

From the construction of h, we have that

$$\left(\frac{g_3}{g_1 g_2}\right) = [m]^*(h),$$

and so we have $\frac{g_3}{g_1 g_2} = c(h \circ [m])$; we may assume that $c = 1$.

$$\frac{g_3(X + S)}{g_3(X)} = \frac{g_1(X + S)}{g_1(X)} \cdot \frac{g_2(X + S)}{g_2(X)} \cdot \frac{h(m(X + S))}{h(mX)},$$

i.e. $[S, T_1 + T_2] = [S, T_1] \cdot [S,T_2]$.

Alternating. It suffices to show that $[T, T] = 1$. Now

$$\left(\prod_{i=0}^{m-1} f \circ \tau_{iT}\right) = \sum_{i=0}^{m-1} m(((i + 1)T) - (iT)) = 0,$$

and so the function $\prod_{i=0}^{m-1} f \circ \tau_{iT}$ is constant. Also, if $mT' = T$, then $\prod_{i=0}^{m-1} g \circ \tau_{iT'}$ is also constant, since

$$\left(\prod_{i=0}^{m-1} g \circ \tau_{iT'}\right)^m = \prod_{i=0}^{m-1} g^m \circ \tau_{iT'} = \prod_{i=0}^{m-1} f \circ [m] \circ \tau_{iT'} = \left(\prod_{i=0}^{m-1} f \circ \tau_{iT}\right) \circ [m],$$

which is constant. Hence we have

$$\left(\prod_{i=0}^{m-1} g \circ \tau_{iT'}\right)(X) = \left(\prod_{i=0}^{m-1} g \circ \tau_{iT'}\right)(X + T'),$$

so $g(X) = g(X + T)$, so $[T, T] = 1$.

Nondegeneracy. Suppose that $[S, T] = 1$ for all $S \in E_m$. Then $g(X) = g(X + S)$ for all $S \in E_m$. Recall (see Proposition 3.9(2)) that there is an isomorphism $E_m \xrightarrow{\sim} \mathrm{Aut}(\bar{K}(E)/[m]^*\bar{K}(E))$, $S \mapsto \tau_S^*$. It follows that we have $g \in [m]^*\bar{K}(E)$, i.e. $g = h \circ [m]$ for some $h \in \bar{K}(E)$. Then

$$h^m \circ [m] = (h \circ [m])^m = g^m = f \circ [m],$$

so $f = h^m$. Thus $m(h) = (f) = m(T) - m(O)$. Thus $(h) = (T) - (O)$, so $T = O$.

Galois equivariance. Suppose $\sigma \in \mathrm{Gal}(\bar{K}/K)$. If f and g are the functions corresponding to T, then f^σ and g^σ are the functions corresponding to T^σ. So

$$[S^\sigma, T^\sigma] = \frac{g^\sigma(X^\sigma + S^\sigma)}{g^\sigma(X^\sigma)} = \left(\frac{g(X+S)}{g(X)}\right)^\sigma = [S,T]^\sigma.$$

Compatibility. IF $S \in E_{mm'}$ and $T \in E_m$, then $[S,T]_{mm'} = [m'S,T]_m$. For we have $(f^{m'}) = mm'(T) - mm'(O)$. So

$$(g \circ m')^{mm'} = (f \circ [mm'])^{m'}.$$

Thus

$$[S,T]_{mm'} = \frac{g \circ [m'](X+S)}{g \circ [m'](X)} = \frac{g(m'X + m'S)}{g(m'X)} = [m'S,T]_m.$$

Proposition 3.19 There exist $S,T \in E_m$ such that $[S,T]_m$ is a primitive m^{th} root of unity. Hence if $E_m \subseteq E(K)$, then $\mu_m \subseteq K^\times$.

Proof. The set $\{[S,T]_m \mid S,T \in E_m\}$ is a subgroup μ_d of μ_m. So for all $S,T \in E_m$, we have $[S,T]_m^d = 1$, so $[dS,T]_m = 1$, so $ds = O$ (since $[\cdot,\cdot]_m$ is nondegenerate), so $d = m$ (since S is arbitrary). The final assertion follows from the Galois equivariance of the Weil pairing.

Proposition 3.20 Suppose that $\phi : E_1 \to E_2$ is an isogeny and that $S \in E_1[m]$ and $T \in E_2[m]$. Then $[S, \hat{\phi}(T)]_m = [\phi(S), T]_m$.

Proof. Choose $f, g \in \bar{K}(E_2)$ such that $(f) = m(T) - m(O)$ and $f \circ [m] = g^m$ (as described in the construction of the Weil pairing). Then

$$[\phi(S), T]_m = \frac{g(X + \phi(S))}{g(X)}.$$

Now observe that we may choose $h \in \bar{K}(E_1)$ such that

$$\phi^*((T)) - \phi^*((O)) = (\hat{\phi}(T)) - (O) + (h).$$

($\hat{\phi}(T)$ is the sum of the points of the divisor on the left side — see Theorem 3.16.) Then we have

$$\operatorname{div}\left(\frac{f \circ \phi}{h^m}\right) = \phi^*(f) - m(h) = m(\hat{\phi}(T)) - m(O)$$

and

$$\left(\frac{g \circ \phi}{h \circ [m]}\right)^m = \frac{f \circ [m] \circ \phi}{(h \circ [m])^m} = \left(\frac{f \circ \phi}{h^m}\right) = [m].$$

So

$$\begin{aligned}
[S, \hat{\phi}(T)]_m &= \frac{\left(\frac{g \circ \phi}{h \circ [m]}\right)(X + S)}{\left(\frac{g \circ \phi}{h \circ [m]}\right)(X)} \\
&= \frac{g(\phi(X) + \phi(S))}{g(\phi(X))} \cdot \frac{h([m]X)}{h([m]X + [m]S)} \\
&= [\phi(S), T]_m.
\end{aligned}$$

Consequence. Fix a prime $\ell \neq \operatorname{char}(K)$. Then the following diagram commutes:

$$\begin{tikzcd}
E_{\ell^{n+1}} \times E_{\ell^{n+1}} \arrow[r, "{[\cdot,\cdot]_{\ell^{n+1}}}"] \arrow[d, "{[\ell]\times[\ell]}"'] & \mu_{\ell^{n+1}} \arrow[d, "x \mapsto x^\ell"] \\
E_{\ell^n} \times E_{\ell^n} \arrow[r, "{[\cdot,\cdot]_{\ell^n}}"] & \mu_{\ell^n}
\end{tikzcd}$$

Via compatibilities, we obtain a pairing

$$\begin{tikzcd}
\varprojlim E_{\ell^n} \times \varprojlim E_{\ell^n} \arrow[r] \arrow[d, equal] & \varprojlim \mu_{\ell^n} \arrow[d, equal] \\
T_\ell(E) \times T_\ell(E) \arrow[r] & \mathbb{Z}_\ell(1)
\end{tikzcd}$$

3.7 The Tate Module

Let E/K be an elliptic curve and $m \geq 2$, with $(m, \mathrm{char}(K)) = 1$. Recall that $E_m \xrightarrow{\sim} \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. Since $\mathrm{Gal}(\bar{K}/K)$ acts on E_m, we obtain a representation

$$\mathrm{Gal}(\bar{K}/K) \to \mathrm{Aut}(E_m) \simeq \mathrm{GL}_2(\mathbb{Z}/m\mathbb{Z}).$$

In order to study these representations, it is (extremely!) helpful to introduce the following definition:

Definition 3.21 The ℓ-adic **Tate module** of E is $T_\ell(E) : \varprojlim E_{\ell^n}$, where the inverse limit is with respect to the maps $[\ell] : E_{\ell^{n+1}} \to E_{\ell^n}$. Then $T_\ell(E)$ is a $\mathbb{Z}_\ell$-module, and we have

$$T_\ell(E) \simeq \begin{cases} \mathbb{Z}_\ell \times \mathbb{Z}_\ell & \text{if } \ell \neq \mathrm{char}(K), \\ 0 \text{ or } \mathbb{Z}_\ell & \text{if } \ell = \mathrm{char}(K). \end{cases}$$

$T_\ell(E)$ carries a natural $\mathrm{Gal}(\bar{K}/K)$ action.

Definition 3.22 The ℓ-adic representation of $\mathrm{Gal}(\bar{K}/K)$ associated to E is the natural map

$$\rho_\ell : \mathrm{Gal}(E/K) \to \mathrm{Aut}(T_\ell(E)) \simeq \mathrm{GL}_2(\mathbb{Z}_\ell).$$

Exercise. Define $\mathbb{Z}_\ell(1) := \varprojlim \mu_{\ell^n}$. Then we have a representation

$$\chi_\ell : \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to \mathrm{Aut}(\mathbb{Z}_\ell(1)) \simeq \mathbb{Z}_\ell^\times = \mathrm{GL}_1(\mathbb{Z}_\ell).$$

Show that χ_ℓ is surjective.

Theorem 3.23 (Serre)

(a) $\mathrm{Im}(\rho_\ell)$ is of finite index in $\mathrm{GL}_2(\mathbb{Z}_\ell)$ for all ℓ.

(b) $\mathrm{Im}(\rho_\ell) = \mathrm{GL}_2(\mathbb{Z}_\ell)$ for almost all ℓ.

(See e.g. Serre's *Abelian ℓ-adic Representations and Elliptic Curves.*)

3.8 Isogenies

Suppose $\phi : E_1 \to E_2$ is an isogeny. Then ϕ induces homomorphisms $\phi : E_1[\ell^n] \to E_2[\ell^n]$ for all $n \geq 1$, which in turn induce $\phi_\ell : T_\ell(E_1) \to T_\ell(E_2)$. So we obtain a homomorphism

$$\mathrm{Hom}(E_1, E_2) \to \mathrm{Hom}(T_\ell(E_1), T_\ell(E_2))$$

given by $\phi \mapsto \phi_\ell$.

Theorem 3.24 Notation as above. The natural map

$$\mathrm{Hom}(E_1, E_2) \otimes \mathbb{Z}_\ell \to \mathrm{Hom}(T_\ell(E_1), T_\ell(E_2))$$

given by $\phi \mapsto \phi_\ell$ is injective.

Definition 3.25 Suppose that M is any abelian group. A function $d : M \to \mathbb{R}$ is a **quadratic form** if

(a) $d(m) = d(-m)$ for all $m \in M$.

(b) The pairing $M \times M \to \mathbb{R}$ given by $(m_1, m_2) \mapsto d(m_1 + m_2) - d(m_1) - d(m_2)$ is bilinear.

We say that a quadratic form is **positive definite** if

(c) $d(m) \geq 0$ for all $m \in M$, with equality iff $m = 0$.

Lemma 3.26 Suppose that E_1 and E_2 are elliptic curves. Then the degree map $\deg : \mathrm{Hom}(E_1, E_2) \to \mathbb{Z}$ is a positive definite quadratic form.

Proof. The only nontrivial point is to show that the pairing $\langle \phi, \psi \rangle = \deg(\phi + \psi) - \deg(\phi) - \deg(\psi)$ is bilinear. Now

$$\begin{aligned} [\langle \phi, \psi \rangle] &= [\deg(\phi + \psi)] - [\deg(\phi)] - [\deg(\psi)] \\ &= (\widehat{\phi + \psi}) \circ (\phi + \psi) - \hat{\phi} \circ \phi - \hat{\psi} \circ \psi \\ &= \hat{\phi} \circ \psi + \hat{\psi} \circ \phi, \end{aligned}$$

and this last expression is linear in ϕ and ψ.

Lemma 3.27 Let $M \subseteq \mathrm{Hom}(E_1, E_2)$ be any finitely generated subgroup. Define

$$M_{\mathrm{sat}} := \{\phi \in \mathrm{Hom}(E_1, E_2) \mid [m] \circ \phi \in M \text{ for some integer } m \geq 1\}.$$

Then M_{sat} is also finitely generated.

Proof. Extend the degree mapping $\deg : M \to \mathbb{Z}$ to

$$\deg : M \otimes \mathbb{R} \to \mathbb{R}, \tag{$*$}$$

where we view $M \otimes \mathbb{R}$ as a finite dimensional real vector space equipped with the topology inherited from $\mathbb{R}$. Then $(*)$ is continuous, and so $U := \{\phi \in M \otimes \mathbb{R} \mid \deg \phi < 1\}$ is an open neighborhood of the origin. Recall that $\mathrm{Hom}(E_1, E_2)$ is a torsionfree $\mathbb{Z}$-module (Theorem 3.8), and so there is a natural inclusion $M_{\mathrm{sat}} \hookrightarrow M \otimes \mathbb{R}$. Plainly $M_{\mathrm{sat}} \cap U = 0$ (since every nonzero isogeny has degree at least 1). So M_{sat} is a discrete subgroup of the finite dimensional vector space $M \otimes \mathbb{R}$, and so M_{sat} is finitely generated.

Proof of Theorem 3.24 Suppose $\phi \in \mathrm{Hom}(E_1, E_2) \otimes \mathbb{Z}_\ell$ with $\phi_\ell = 0$. Let $M \subseteq \mathrm{Hom}(E_1, E_2)$ be any finitely generated subgroup such that $\phi \in M \otimes \mathbb{Z}_\ell$. Then M_{sat} is finitely generated and torsionfree (Lemma 3.27 and Theorem 3.8), and so is free. Choose a basis $\phi_1, \ldots, \phi_t \in \mathrm{Hom}(E_1, E_2)$ of M_{sat}, and suppose that $\phi = \alpha_1\phi_1 + \cdots + \alpha_t\phi_t$, with $\alpha_i \in \mathbb{Z}_\ell$. For each $1 \leq i \leq t$, choose $a_i \in \mathbb{Z}$ such that $a_i \equiv \alpha_i \pmod{\ell^n}$, and consider the isogeny

$$\psi := [a_1] \circ \phi_1 + \cdots + [a_t] \circ \phi_t \in \mathrm{Hom}(E_1, E_2).$$

Then $\phi_\ell = 0$ implies that ψ kills $E_1[\ell^n]$, so ψ factors through $[\ell^n]$ (Corollary 3.10), i.e. there exists $\lambda \in \mathrm{Hom}(E_1, E_2)$ such that $\psi = [\ell^n] \circ \lambda$. Now $\lambda \in M_{\mathrm{sat}}$, and so there exists $b_i \in \mathbb{Z}$ such that

$$\lambda = [b_1] \circ \phi_1 + \cdots + [b_t] \circ \phi_t.$$

Since the ϕ_i's are a $\mathbb{Z}$-basis of M_{sat}, we have $a_i = \ell^n b_i$ for $1 \leq i \leq t$, so $\alpha_i \equiv 0 \pmod{\ell^n}$ for $1 \leq i \leq t$. Since n was arbitrary, it follows that $\alpha_i = 0$ for $1 \leq i \leq t$, and so $\phi = 0$.

Theorem 3.28 (Tate, Faltings). The natural map

$$\mathrm{Hom}_K(E_1, E_2) \otimes \mathbb{Z}_\ell \to \mathrm{Hom}_K(T_\ell(E_1), T_\ell(E_2))$$

is an isomorphism if K is a finite field (Tate), or if K is a number field (Faltings).

3.9 The j-invariant

Suppose that $\mathrm{char}(K) \neq 2$ or 3, and let E/K be an elliptic curve. Then the Weierstraß model of E can be put in the form $E : y^2 = x^3 + ax + b$ (see Silverman III, §1). Then

$$j(E) := \frac{4a^3}{4a^3 + 27b^2} \in K.$$

Theorem 3.29 If $E_1 \simeq E_2$, then $j(E_1) = j(E_2)$. If $j(E_1) = j(E_2)$, then $E_1 \simeq_{\bar{K}} E_2$.

Proof. Suppose that $E_1 \simeq E_2$. Then $x_2 = u^2 x$, $y_2 = u^3 x$, $u \in K$ or $\bar{K}$ (cf Corollary 3.4). Then $a_2 = u^{-4} a_1$ and $b_2 = u^{-6} b_1$, so $j(E_1) = j(E_2)$. Suppose that $j(E_1) = j(E_2)$. Then we have

$$(4a_1)^3(4a_2^3 + 27b_2^2) = 4a_2^3(4a_1^3 + 27b_1^2),$$

so $a_1^3 b_2^2 = a_2^3 b_1^2$. If a_1, b_1, a_2, and b_2 are all nonzero, then

$$\left(\frac{a_1}{a_2}\right)^3 = \left(\frac{b_1}{b_2}\right)^2 = u^{12},$$

say, i.e. $\frac{a_1}{a_2} = u^4$ and $\frac{b_1}{b_2} = u^6$, and so construct an isomorphism using this u.

Exercise. Do the other cases.

Chapter 4

Elliptic Curves over Finite Fields

Let $K = \mathbb{F}_q$, and let E/K be an elliptic curve.

Problem. Estimate the number of points in $E(K)$, i.e. estimate the number of solutions to the equation

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

with $(x, y) \in K^2$.

Lemma 4.1 Let M be an abelian group, and let $d : M \to \mathbb{Z}$ be a positive definite quadratic form. Then for all $\phi, \psi \in M$, we have

$$|d(\psi - \phi) - d(\phi) - d(\psi)| \leq 2\sqrt{d(\phi)d(\psi)}.$$

Proof. Set $L(\psi, \phi) := d(\psi - \phi) - d(\phi) - d(\psi)$. Then L is bilinear (since d is a quadratic form). As d is positive definite, we have, for all $m, n \in \mathbb{Z}$,

$$0 \leq d(m\psi - n\phi) = m^2d(\psi) + mnL(\psi, \phi) + n^2d(\phi).$$

Take $m = -L(\psi, \phi)$ and $n = 2d(\psi)$; then

$$0 \leq d(\psi)[4d(\psi)d(\phi) - L(\psi, \phi)^2],$$

and this is enough.

Theorem 4.2 (Hasse) Suppose that $K = \mathbb{F}_q$ and E/K is an elliptic curve. Then

$$|\#E(K) - q - 1| \leq 2\sqrt{q}.$$

Proof. Choose a Weierstraß equation for E/K. Let $\phi : E \to E$, $(x, y) \mapsto (x^q, y^q)$ be the q^{th} power Frobenius morphism. Now $\mathrm{Gal}(\bar{K}/K)$ is topologically generated by the q^{th} power map on $\bar{K}$. Hence if $P \in E(\bar{K})$, then $P \in E(K)$ iff $\phi(P) = P$, so $E(K) = \ker(1 - \phi)$. Since $1 - \phi$ is separable, we have

$$\#E(K) = \#\ker(1 - \phi) = \deg(1 - \phi).$$

Thus Lemma 4.2 yields

$$|\deg(1 - \phi) - \deg(\phi) - \deg(1)| \leq 2\sqrt{\deg(\phi)\deg(1)},$$

so

$$|\#E(K) - q - 1| \leq 2\sqrt{q}.$$

Example. (Estimating character sums). Suppose that $K = \mathbb{F}_q$, with q odd. Let $f(x) = ax^3 + bx^2 + cx + d \in K[x]$ be a cubic polynomial with distinct roots in $\bar{K}$. Let $\chi : K^\times \to \{\pm 1\}$ be the unique nontrivial character of order 2 (so $\chi(t) = 1$ iff t is a square in $K^\times$). Set $\chi(0) = 0$; then χ is defined on K. Use χ to count the number of K-rational points on the elliptic curve $E : y^2 = f(x)$. Each $x \in K$ gives 0, respectively 1, respectively 2 points $(x, y) \in E(K)$ if $f(x)$ is a nonsquare, respectively zero, respectively a square in K. So

$$\#E(K) = 1 + \sum_{x \in K} (\chi(f(x)) + 1) = 1 + q + \sum_{x \in K} \chi(f(x)).$$

Hence we have

$$\left| \sum_{x \in K} \chi(f(x)) \right| \leq 2\sqrt{q}.$$

This is the tip of a vast iceberg, cf for example “Sommes exponentielles,” Astérisque **79** by N. Katz.

Let $K = \mathbb{F}_q$, and set K_n to be the unique extension of K of degree n. (So $\#K_n = q^n$.) Let V/K be a projective variety. $V(K_n) :=$ the set of points of V with coordinates in K_n.

Definition 4.3 The **zeta function** of V/K is the power series

$$Z(V/K;T) = \exp\left(\sum_{n=1}^{\infty}(\#V(K_n))\frac{T^n}{n}\right).$$

(Here $\exp(F(T)) := \sum_{i=0}^{\infty} \frac{F(T)^i}{i!}$ for $F(T) \in \mathbb{Q}[[T]]$ with no constant term.)

We have

$$\#V(K_n) = \frac{1}{(n-1)!}\,\frac{d^n}{dT^n}\log(Z(V/K;T))\bigg|_{T=0}.$$

Example. Take $V = \mathbb{P}^n$. Then each point in $V(K_n)$ is given by homogeneous coordinates $[x_0 : \ldots : x_N]$ with $x_i \in K_n$, not all zero. Two sets of coordinates give the same point only if they differ by multiplication by an element of $K_n^\times$. So we have

$$\#V(K_n) = \frac{q^{n(N+1)} - 1}{q^n - 1} = \sum_{i=0}^{N} q^{ni}.$$

Hence

$$\log Z(V/K;T) = \sum_{n=1}^{\infty}\left(\sum_{i=0}^{N} q^{ni}\right)\frac{T^n}{n} = \sum_{i=0}^{N} -\log(1 - q^i T).$$

So

$$Z(V/K;T) = \frac{1}{(1-T)(1-qT)\cdots(1-q^N T)} \in \mathbb{Q}(T).$$

Remark. A similar argument shows that in general, if there are $\alpha_1, \ldots, \alpha_r \in \mathbb{C}$ such that $\#V(K_n) = \pm\alpha_1^n \pm \cdots \pm \alpha_r^n$ for all $n \in \mathbb{N}$, then $Z(V/K;T)$ will be a rational function.

Theorem 4.4 (The Weil Conjectures) Let $K = \mathbb{F}_q$, and suppose that V/K is a smooth projective variety of dimension n.

(a) Rationality: $Z(V/K; T) \in \mathbb{Q}(T)$.

(b) Functional equation: There is an integer ε such that

$$Z\left(V/K; \frac{1}{q^n T}\right) = \pm q^{n\varepsilon/2} T^{\varepsilon} Z(V/K; T).$$

(c) Riemann Hypothesis: There is a factorization

$$Z(V/K; T) = \frac{P_1(T) \cdots P_{2n-1}(T)}{P_0(T) P_2(T) \cdots P_{2n}(T)},$$

with each $P_i(T) \in \mathbb{Z}[T]$. Also $P_0(T) = 1 - T$, $P_{2n}(T) = 1 - q^n T$, and for each $1 \leq i \leq 2n - 1$, we have $P_i(T) = \prod_j (1 - \alpha_{ij} T)$, $\alpha_{ij} \in \mathbb{C}$ with $|\alpha_{ij}| = q^{i/2}$.

4.1 Proof of the Weil Conjectures for Elliptic Curves E/K

Recall that we have a map $\mathrm{End}(E) \to \mathrm{End}(T_\ell(E))$ given by $\psi \mapsto \psi_\ell$. ψ_ℓ may be written as a 2×2 matrix over $\mathbb{Z}_\ell$, so we may compute $\det(\psi_\ell), \mathrm{Tr}(\psi_\ell) \in \mathbb{Z}_\ell$.

Proposition 4.5 Suppose that $\psi \in \mathrm{End}(E)$. Then $\det(\psi_\ell) = \deg(\psi)$ and $\mathrm{Tr}(\psi_\ell) = 1 - \deg(\psi) - \deg(1 - \psi)$. (So $\det(\psi_\ell), \mathrm{Tr}(\psi_\ell) \in \mathbb{Z}$ and are independent of ℓ.)

Proof. Choose a $\mathbb{Z}_\ell$-basis v_1, v_2 of $T_\ell(E)$. Write the matrix of ψ_ℓ with respect to this basis as

$$\psi_\ell = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

There is a nondegenerate, bilinear, alternating Weil pairing $e : T_\ell(E) \times T_\ell(E) \to$

$T_\ell(\boldsymbol{\mu}) = \mathbb{Z}_\ell(1)$. So we have

$$\begin{aligned} e(v_1, v_2)^{\deg(\psi)} &= e([\deg \psi]v_1, v_2) \\ &= e(\hat{\psi}_\ell \circ \psi_\ell(v_1), v_2) \\ &= e(\psi_\ell(v_1), \psi_\ell(v_2)) \\ &= e(av_1 + cv_2, bv_1 + dv_2) \\ &= e(v_1, v_2)^{ad-bc} \\ &= e(v_1, v_2)^{\det \psi_\ell}. \end{aligned}$$

Hence $\deg \psi = \det \psi_\ell$, since e is nondegenerate. For any 2×2 matrix A, say, we have $\mathrm{Tr}(A) = 1 + \det(A) - \det(1 - A)$.

Let $\phi : E \to E$ be the q^{th} power Frobenius morphism. Then $\#E(K) = \deg(1 - \phi)$, $\#E(K_n) = \deg(1 - \phi^n)$. The characteristic polynomial of ϕ_ℓ has coefficients in $\mathbb{Z}$ and so may be factored over $\mathbb{C}$:

$$\det(T - \phi_\ell) = T^2 - \mathrm{Tr}(\phi_\ell)T - \det(\phi_\ell) = (T - \alpha)(T - \beta),$$

say. Next observe that for each $m/n \in \mathbb{Q}$, we have

$$\det\left(\frac{m}{n} - \phi_\ell\right) = \frac{\det(m - n\phi_\ell)}{n^2} = \frac{\deg(m - n\phi)}{n^2} \geq 0,$$

and so $\det(T - \phi_\ell)$ has complex conjugate roots. Hence $|\alpha| = |\beta|$, and so, since $\alpha\beta = \det \phi_\ell = \deg \phi = q$, we have $|\alpha| = |\beta| = \sqrt{q}$.

Now the characteristic polynomial of ϕ_ℓ^n is given by $\det(T - \phi_\ell^n) = (T - \alpha^n)(T - \beta^n)$, so

$$\#E(K_n) = \deg(1 - \phi^n) = \det(1 - \phi_\ell^n) = 1 - \alpha^n - \beta^n + q^n.$$

Theorem 4.6 Let $K = \mathbb{F}_q$, and let E/K be an elliptic curve. Then there is an $a \in \mathbb{Z}$ such that

$$Z(E/K; T) = \frac{1 - aT + qT^2}{(1 - T)(1 - qT)}.$$

Also

$$Z\left(E/K; \frac{1}{qT}\right) = Z(E/K; T),$$

and $1 - aT + qT^2 = (1 - \alpha T)(1 - \beta T)$, with $|\alpha| = |\beta| = \sqrt{q}$.

Proof. We have

$$\begin{aligned}\log Z(E/K; T) &= \sum_{n=1}^{\infty} (\#E(K_n)) \frac{T^n}{n} \\ &= \sum_{n=1}^{\infty} \frac{(1 - \alpha^n - \beta^n + q^n)}{n} T^n \\ &= -\log(1 - T) + \log(1 - \alpha T) + \log(1 - \beta T) - \log(1 - qT),\end{aligned}$$

so

$$Z(E/K; T) = \frac{(1 - \alpha T)(1 - \beta T)}{(1 - T)(1 - qT)}.$$

Thus

$$a = \alpha + \beta = \mathrm{Tr}(\phi_\ell) = 1 + q - \deg(1 - \phi) \in \mathbb{Z}_\ell.$$

[This is $\varepsilon = 0$ in the functional equation

$$Z\left(V/K; \frac{1}{q^n T}\right) = \pm q^{n\varepsilon/2} T^{\varepsilon} Z(V/K; T),$$

where $\dim V = n$.]

Remark. Suppose we make a change of variable $T = q^{-s}$. Then

$$\zeta_{E/K}(s) := Z(E/K; q^{-s}) = \frac{1 - \alpha q^{-s} + q^{1-2s}}{(1 - q^{-s})(1 - q^{1-s})}.$$

The functional equation becomes $\zeta_{E/K}(1 - s) = \zeta_{E/K}(s)$, and $\zeta_{E/K}(s) = 0$ implies that $|q^s| = \sqrt{q}$, so $\Re(s) = 1/2$.

Question. Suppose $E/\mathbb{Q}$ is an elliptic curve $y^2 = ax^3 + bx + c$, with $a, b \in \mathbb{Z}$. We can look at $E/\mathbb{F}_p$. This is an elliptic curve for all but finitely many p. Let $\phi : E/\mathbb{F}_p \to E/\mathbb{F}_p$ be the Frobenius morphism. For any $\ell \neq p$, we can look at $\phi_\ell : T_\ell(E/\mathbb{F}_p) \to T_\ell(E/\mathbb{F}_p)$. ϕ_ℓ has complex conjugate eigenvalues α_p and β_p, say (independent of ℓ). We've just shown that $|\alpha_p| = |\beta_p| = p^{1/2}$. So

$$\alpha_p = p^{1/2} e^{i\theta_p}, \qquad \beta_p = p^{1/2} e^{-i\theta_p}.$$

How do the angles θ_p vary with p?

4.2 Equidistribution

Suppose $E/\mathbb{Q}$ is an elliptic curve without complex multiplication, and let p be a prime such that $\tilde{E}/\mathbb{F}_p$ (the reduction of E modulo p) is nonsingular. Theorem 4.2 (Hasse) implies that $|\#\tilde{E}(\mathbb{F}_p) - p - 1| \leq 2\sqrt{p}$, i.e.

$$p + 1 - 2\sqrt{p} \leq \tilde{E}(\mathbb{F}_p) \leq p + 1 + 2\sqrt{p}.$$

So we may write $\tilde{E}(\mathbb{F}_p) = p + 1 - a_p$, with $|a_p| \leq 2\sqrt{p}$. $p+1$ is the "main term," and a_p is the "error term." We may write $a_p = 2\sqrt{p}\cos\theta$, with $\theta \in [0, \pi]$.

Question. How does θ_p vary with p?

Suppose we are given a sequence $\{x_n\}_{n\geq 1}$ in a compact space X with probability measure μ.

Definition 4.7 Say that $\{x_n\}$ is **equidistributed** with respect to μ if for all continuous functions $f : X \to \mathbb{C}$, we have

$$\int_X f \, d\mu = \lim_{N\to\infty} \sum_{i=1}^{N} f(x_n).$$

[It suffices to check this on a set of test functions $\{f_i\}$ whose $\mathbb{C}$-span is uniformly dense.]

Suppose that G is a compact group equipped with a Haar measure (so G has total mass 1). Let $X = \{\text{conjugacy classes in } G\}$, and write μ for the Haar measure on X induced from the Haar measure on G. There is a bijection between continuous functions on X and continuous central (class) functions on G given by $\int_X f \, d\mu = \int_G f \, dg$.

We can take our uniformly dense set of functions $\{f_i\}$ to be functions of the form $g \mapsto \operatorname{Tr}\Lambda(g)$, for Λ an irreducible representation of G (Peter-Weyl Theorem). We have

$$\int_X \mathbb{1} \, d\mu = 1, \qquad \int_X \operatorname{Tr}(\Lambda) \, d\mu = 0$$

if $\Lambda \neq 1$ is irreducible (via orthogonality relations for characters).

Weyl Criterion for Equidistribution. For all irreducible nontrivial representations Λ,

$$\sum_{i=1}^{N} \operatorname{Tr}(\Lambda(x_i)) = o(N).$$

4.3 The L-Function Method

Suppose that G is a compact group, and let $N \geq 1$ be an integer. Suppose that for each prime p with $p \nmid N$, we are given a conjugacy class θ_p of G. When is $\{\theta_p\}_{p \nmid N}$ equidistributed in X? See Serre's book *Abelian ℓ-adic Representations and Elliptic Curves* (1968).

For each nontrivial irreducible representation Λ of G, form the L-function

$$L(s, \Lambda) := \prod_{p \nmid N} \frac{1}{\det(1 - \Lambda(\theta_p) p^{-s})}.$$

This converges for $\Re(s) > 1$.

Theorem 4.8 (Serre's book). For Λ as above, suppose

(1) $L(s, \Lambda)$ has an analytic representation on an open set contained in $\Re(s) \geq 1$, and

(2) $L(s, \Lambda)$ is nowhere zero on $\Re(s) = 1$.

Then $\{\theta_p\}_{p \nmid N}$ is equidistributed in X.

Theorem 4.9 (Deligne, Weil II). In the Serre setup, (1) implies (2) with at most one exception. This exception, if it exists, is a 1-dimensional character $\Lambda : G \to \{\pm 1\}$.

Corollary 4.10 There are no exceptions if either G is connected or if for all $\Lambda : G \to \{\pm 1\}$, the map $p \mapsto \Lambda(\theta_p)$ is a Dirichlet character (i.e. a character of $(\mathbb{Z}/N\mathbb{Z})^\times$).

Examples.

(a) Dirichlet (1837). There exist infinitely many primes in arithmetic progressions unless there clearly aren't. Dirichlet introduces Dirichlet L-functions $L(s, \chi)$ and prove that $L(1, \chi) \neq 0$ if $\chi \neq 1$.

(b) Chebotarev (1915) Let $K/\mathbb{Q}$ be Galois, with $G = \mathrm{Gal}(K/\mathbb{Q})$. Consider the map sending p to the conjugacy class of Frob_p in G. Then $\{\mathrm{Frob}_p\}_{p \nmid \mathrm{disc}(K/\mathbb{Q})}$ is equidistributed in X.

(c) Early 1960's: Back to our original elliptic curve example. Salo does computer experiments. In 1963 Tate writes down the Sato-Tate Conjecture.

Sato-Tate Conjecture. Let $E/\mathbb{Q}$ be an elliptic curve without complex multiplication. For almost all p, we know that $E(\mathbb{F}_p) = p + 1 - a_p$, where $a_p = 2\sqrt{p}\cos\theta_p$ for some $\theta_p \in [0, \pi]$. Then $\{\theta_p\}$ is equidistributed in $[0, \pi]$ with respect to the (Sato-Tate) measure $\frac{2}{\pi}\sin^2\theta \, d\theta$.

Each conjugacy class in $G := \mathrm{SU}(2, \mathbb{C})$ contains a unique element of the form

$$\begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

The Haar measure on the set X of conjugacy classes is $\frac{2}{\pi}\sin^2\theta \, d\theta$ (cf e.g. Bröcker and tom Dieck, *Representations of Compact Lie Groups*).

(d) (2006) Clozel, Harris, Shepherd-Barron, Taylor: The Sato-Tate Conjecture holds for $E/\mathbb{Q}$ with $j(E) \notin \mathbb{Z}$. They prove this via the L-function method.

4.4 The Hasse Invariant and the Endomorphism Ring

Theorem 4.11 Suppose that K is a perfect field with characteristic $p > 0$, and let E/K be an elliptic curve. Let $\phi_r : E \to E^{(p^r)}$ and $\hat{\phi}_r : E^{(p^r)} \to E$ be the Frobenius map and its dual.

(a) The following conditions are equivalent:

   (i) $E_{p^r} = O$ for one (and therefore all) $r \geq 1$.

   (ii) $\hat{\phi}_r$ is purely inseparable for one (and therefore all) $r \geq 1$.

   (iii) The map $[p] : E \to E$ is purely inseparable, and $j(E) \in \mathbb{F}_{p^2}$.

   (iv) $\mathrm{End}(E)$ is an order in a quaternion algebra.

   In this case, we say that E is **supersingular** or has Hasse invariant 0.

(b) If (a) does not hold, then $E_{p^r} \simeq \mathbb{Z}/p^r\mathbb{Z}$ for all $r \geq 1$. In this case, if $j(E) \in \bar{F}_p$, then $\mathrm{End}(E)$ is an order in an imaginary quadratic field. If $j(E) \notin \bar{F}_p$, then $\mathrm{End}(E) \simeq \mathbb{Z}$. In this case, we say that E is **ordinary** or has Hasse invariant 1.

Proof.

(a) We first show (i) iff (ii). Recall that ϕ_r is purely inseparable (Theorem 2.5). So

$$\deg_s(\hat{\phi}_r) = \deg_s[p^r] = (\deg_s[p])^r = (\deg_s \hat{\phi})^r.$$

Thus

$$\#E_{p^r} = \deg_s(\hat{\phi}_r) = (\deg_s \hat{\phi})^r.$$

Thus $\#E_{p^r} = 1$ iff $\deg_s \hat{\phi}_r = 1$, as required.

We now show (ii) implies (iii). Since ϕ is purely inseparable, and (ii) implies that $\hat{\phi}$ is purely inseparable, we see that $[p] = \hat{\phi} \circ \phi$ is also purely inseparable. Now recall (Theorem 2.5) that every map $\tau : C_1 \to C_2$ between smooth curves over a field of characteristic p factors as

$$\begin{tikzcd}
C_1 \arrow[rr, "\tau"] \arrow[dr, "\phi^{(q)}"'] & & C_2 \\
& C_1^{(q)} \arrow[ur, "\lambda"'] &
\end{tikzcd}$$

where $q := \deg_s(\tau)$, $\phi^{(q)}$ is the q^{th} power Frobenius map, and λ is separable. Applying this to $\hat{\phi} : E^{(p)} \to E$, we see that we have a diagram

$$\begin{tikzcd}
E^{(p)} \arrow[rr, "\hat{\phi}"] \arrow[dr, "\Phi"'] & & E \\
& E^{(p^2)} \arrow[ur, "\lambda"'] &
\end{tikzcd}$$

where Φ is the p^{th} power Frobenius map on $E^{(p)}$, and λ is of degree 1. Hence λ is an isomorphism, and so $j(E) = j(E^{(p^2)}) = j(E)^{p^2}$, so $j(E) \in \mathbb{F}_{p^2}$.

We now show that (iii) implies (iv). The proof proceeds via contradiction. We first observe that if $\mathrm{End}(E)$ is not an order in a quaternion algebra and $K := \mathrm{End}(E) \otimes_{\mathbb{Z}} \mathbb{Q}$, then $K = \mathbb{Q}$ or K is an imaginary quadratic field. Suppose that E' is isogenous to E, with $\psi : E \to E'$. We have $\psi \circ [p] = [p] \circ \psi$, and since $[p]$ is purely inseparable on E, $[p]$ is also purely inseparable on E'. (Compare inseparable degrees of both sides.) This in turn implies that $j(E') \in \mathbb{F}_{p^2}$, and so there are only finitely many possibilities for E'. As there are only finitely many $\mathrm{End}(E')$'s, we may choose a prime $\ell \in \mathbb{Z}$ such that $\ell \neq p$ and ℓ remains prime in $\mathrm{End}(E')$ for every E' isogenous to E (exercise). Now $E[\ell^i] \simeq \mathbb{Z}/\ell^i\mathbb{Z} \times \mathbb{Z}/\ell^i\mathbb{Z}$, so there exists a sequence of subgroups $\Phi_1 \subset \Phi_2 \subset \cdots \subset E$ with $\Phi_i \simeq \mathbb{Z}/\ell^i\mathbb{Z}$ for each $i \geq 1$. Set $E_i := E/\Phi_i$, so E_i is isogenous to E. Then there exist integer m and n such that $E_{m+n} \simeq E_m$ with $\tau : E_{m+n} \xrightarrow{\sim} E_m$, say. Then we have

$$\begin{tikzcd}
E_m \arrow[rr, "\lambda"] \arrow[dr, "proj"'] & & E_m \\
& E_{m+n} \arrow[ur, "\sim", "\tau"'] &
\end{tikzcd}$$

where $\ker(\lambda) \simeq \mathbb{Z}/\ell^n\mathbb{Z}$ (i.e. $\ker(\lambda) \simeq \Phi_{m+n}/\Phi_m$). Since ℓ is prime in $\mathrm{End}(E_m)$, it follows (by looking at degrees) that $\lambda = u \circ [\ell^{n/2}]$, $u \in \mathrm{Aut}(E_m)$, and n is even. This is a contradiction, because $\ker([\ell^{n/2}])$ is never cyclic for any $n > 0$. Hence (iii) implies (iv), as claimed.

We now show that (iv) implies (ii). Our strategy is to show that if (ii) is false (so $\hat{\phi}_r$ is separable for all $r \geq 1$), then $\mathrm{End}(E)$ is commutative (which contradicts (iv)). Suppose therefore that $\hat{\phi}_r$ is separable for all $r \geq 1$. Then $E_{p^r} \simeq \mathbb{Z}/p^r\mathbb{Z}$ for all $r \geq 1$ (since (i) iff (ii)) and $T_p(E) \simeq \mathbb{Z}_p$. We claim that the natural map $\mathrm{End}(E) \to \mathrm{End}(T_p(E))$ is injective. For suppose that $\psi \in \mathrm{End}(E)$ lies in the kernel of this map. Then $\psi(E_{p^r}) = 0$ for all $r \geq 1$, so $\# \ker(\psi) \geq p^r$

for all $r \geq 1$, so $\psi = 0$. Since $\operatorname{End}(T_p(E)) \simeq \operatorname{End}(\mathbb{Z}_p) \simeq \mathbb{Z}_p$, we deduce that $\operatorname{End}(E)$ is commutative, as desired.

(b) If (a) does not hold, then (i) above implies that $E_{p^r} \simeq \mathbb{Z}/p^r\mathbb{Z}$ for all $r \geq 1$. Suppose that $j(E) \in \bar{F}_p$, and that (a) does not hold. Then $j(E) \in K$, K is a finite field, and there exists an elliptic curve E'/K with $E' \simeq E$ (cf Silverman III, Proposition 1.4). Suppose $\#K = p^r$; then $\phi_r \in \operatorname{End}(E') \simeq \operatorname{End}(E)$. If $\Phi_r \in \mathbb{Z} \subset \operatorname{End}(E') \simeq \operatorname{End}(E)$, then $\phi_r = [\pm p^{r/2}]$, and r is even (compare degrees!). Then $\#E'_{p^{r/2}} = \deg_s \phi_r = 1$, which is a contradiction. Hence $\phi_r \notin \mathbb{Z}$, and so $\operatorname{End}(E')$ is strictly larger than $\mathbb{Z}$. Therefore $\operatorname{End}(E')$ is an order in an imaginary quadratic field (since by assumption, it is not an order in a quaternion algebra).

4.5 Interlude: Legendre Normal Form

Definition 4.12 We say that a Weierstraß equation is in **Legendre form** if it can be written as $y^2 = x(x-1)(x-\lambda)$.

Theorem 4.13 Let K be any field with $\operatorname{char}(K) \neq 2$.

(a) Every elliptic curve E/K is isomorphic over $\bar{K}$ to an elliptic curve $E_\lambda : y^2 = x(x-1)(x-\lambda)$ for some $\lambda \in \bar{K}$, with $\lambda \neq 0, 1$.

(b) $j(E_\lambda) = \frac{2^8(\lambda^2-\lambda+1)}{\lambda^2(\lambda-1)^2}$.

(c) The map $\bar{K} \setminus \{0,1\} \to \bar{K}$ given by $\lambda \mapsto j(E_\lambda)$ is surjective. It is

$$\begin{aligned} &\text{six-to-one} && \text{if } j \neq 0 \text{ or } 1728, \\ &\text{two-to-one} && \text{if } j = 0, \\ &\text{three-to-one} && \text{if } j = 1728. \end{aligned}$$

Proof.

(a) If $\mathrm{char}(K) \neq 2$, then E has a Weierstraß equation of the form $y^2 = 4x^3 + b_2x^2 + 2b_4x + b_6$. The transformation $x \mapsto x$, $y \mapsto 2y$ yields $y^2 = (x-e_1)(x-e_2)(x-e_3)$, $e_1, e_2, e_3 \in \bar{K}$. The e_i's are distinct since $\Delta = 16(e_1-e_2)^2(e_1-e_3)^2(e_2-e_3)^2 \neq 0$. Now apply the substitution $x \mapsto (e_1 - e_2)x' + e_1$, $y \mapsto (e_2 - e_1)^{3/2}y'$ to obtain an equation in Legendre form with $\lambda = \frac{e_3-e_1}{e_2-e_1} \in \bar{K}$, $\lambda \neq 0$ or 1.

(b) This follows from a calculation.

(c) Suppose $j(E_\lambda) = j(E_\mu)$, say. Then $E_\lambda \simeq_{\bar{K}} E_\mu$, and so the Weierstraß equations of these curves in Legendre form are related by $x \mapsto u^2x' + r$, $y \mapsto u^3 + y'$. Equating yields

$$x(x-1)(x-\mu) = \left(x + \frac{r}{u^2}\right)\left(x + \frac{r-1}{u^2}\right)\left(x + \frac{r-\lambda}{u^2}\right).$$

There are six ways of assigning the linear terms. These yield the possibilities

$$\mu \in \left\{\lambda, \frac{1}{\lambda}, 1-\lambda, \frac{1}{1-\lambda}, \frac{\lambda}{\lambda-1}, \frac{\lambda-1}{\lambda}\right\}.$$

Thus $\lambda \mapsto j(E_\lambda)$ is six-to-one unless two or more of the values for μ coincide. The only possibilities are $\lambda = -1, 2, \frac{1}{2}$, in which case $j(E_\lambda) = 1728$, and the set has three elements, or $\lambda^2 - \lambda + 1 = 0$, in which case $j(E_\lambda) = 0$, and the set has two elements.

Question. How can we tell when an elliptic curve is supersingular?

Theorem 4.14 Suppose K is a finite field with $\mathrm{char}(K) > 2$.

(a) Let E/K be an elliptic curve with Weierstraß equation $E : y^2 = f(x)$, where $f(x) \in K[x]$ is a cubic with distinct roots. Then E is supersingular iff the coefficient of x^{p-1} in $f(x)^{(p-1)/2}$ is zero.

(b) Let $m = \frac{1}{2}(p-1)$, and set

$$H_p(t) = \sum_{i=0}^{m} \binom{m}{i}^2 t^i.$$

Suppose $\lambda \in \bar{K}$ with $\lambda \neq 0$ or 1. Then the elliptic curve $E_\lambda : y^2 = x(x-1)(x-\lambda)$ is supersingular iff $H_p(\lambda) = 0$.

(c) $H_p(t)$ has distinct roots in $\bar{K}$. There are (up to isomorphism) exactly $\lfloor \frac{p}{12} \rfloor + \varepsilon_p$ supersingular elliptic curves in characteristic p, where

$$\varepsilon_p = \begin{cases} 0 & \text{if } p \equiv 1 \pmod{12}, \\ 1 & \text{if } p \equiv 5 \text{ or } 7 \pmod{12}, \\ 2 & \text{if } p \equiv 11 \pmod{12}. \end{cases}$$

Proof.

(a) Set $q = \#K$. If $\chi : K^\times \to \{\pm 1\}$ is the unique nontrivial character of order 2, then, setting $\chi(0) = 0$, we have

$$\#E(K) = 1 + q + \sum_{x \in K} \chi(f(x)) = 1 + \sum_{x \in K} f(x)^{(q-1)/2}$$

in K. (Since $K^\times$ is cyclic of order $q-1$, $\chi(z) = z^{(q-1)/2}$ for all $z \in K^\times$.) Since $K^\times$ is cyclic of order $q-1$, we have

$$\sum_{x \in K} x^i = \begin{cases} -1 & \text{if } (q-1) \mid i, \\ 0 & \text{if } (q-1) \nmid i. \end{cases}$$

Now $f(x)$ has degree 3, so the only nonzero term in $\sum_{x \in K} f(x)^{(q-1)/2}$ comes from x^{q-1}. So if A_q is the coefficient of x^{q-1} in $f(x)^{(q-1)/2}$, then $\#E(K) = 1 - A_q$ in K. Now if $\phi : E \to E$ is the q^{th} power Frobenius endomorphism, we have

$$\#E(K) = \deg(1 - \phi) = (1 - \phi)(1 - \hat{\phi}) = 1 - (\phi + \hat{\phi}) + q = 1 - a + q,$$

say, whence $a = A_q$ in K. So $A_q = 0$ in K if and only if $a \equiv 0 \pmod{p}$ (since a is an integer). Now $\hat{\phi} = [a] - \phi$, so $a \equiv 0 \pmod{p}$ iff $\hat{\phi}$ is separable iff E is supersingular. Hence $A_q = 0$ in K iff E is supersingular. We claim that $A_p = 0$ in K iff $A_q = 0$ in K. For we have

$$f(x)^{(p^{r+1}-1)/2} = f(x)^{(p^r-1)/2} \left(f(x)^{(p-1)/2} \right)^{p^r}.$$

Equating coefficients (using the fact that $f(x)$ is a cubic!) yields $A_{p^{r+1}} = A_{p^r} \cdot A_p^{p^r}$, and this implies the claim via induction on r.

(b) We apply (a). Recall that $m = \frac{1}{2}(p-1)$. We have to calculate the coefficient of x^{p-1} in $[x(x-1)(x-\lambda)]^{(p-1)/2}$, which is the coefficient of $x^{(p-1)/2}$ in $(x-1)^{(p-1)/2}(x-\lambda)^{(p-1)/2}$, which is

$$\sum_{i=0}^{m} \binom{m}{i} (-\lambda)^i \binom{m}{m-i} (-1)^{m-i} = (-1)^m \sum_{i=0}^{m} \binom{m}{i}^2 \lambda^i = (-1)^m H_p(\lambda),$$

which implies the result.

(c) In order to show that $H_p(t)$ has simple roots, we introduce the differential operator

$$\mathscr{D} = 4t(1-t)\frac{d^2}{dt^2} + 4(1-2t)\frac{d}{dt} - 1.$$

A routine calculation yields

$$\mathscr{D}H_p(t) = p\sum_{i=0}^{m}(p-2-4i)\binom{m}{i}^2 t^i,$$

so

$$\mathscr{D}H_p(t) = 0 \qquad \text{in } K[t]. \tag{†}$$

Suppose $H_p(t) = (t-\alpha)^n f(t)$, say, with $2 \le n \le m$ and $f(\alpha) \neq 0$ in K. Substituting this expression into (†) and simplifying yields $4\alpha(\alpha-1) = 0$, so $\alpha = 0$ or 1. We have $H_p(0) = 1$ and

$$H_p(1) = \sum_{i=0}^{m} \binom{m}{i}^2 = \binom{2m}{m} = \frac{(2m)!}{(m!)^2} \not\equiv 0 \pmod{p}.$$

Hence the roots of $H_p(t)$ are simple, as claimed. Each root λ of $H_p(t)$ yields an elliptic curve $E_\lambda : y^2 = x(x-1)(x-\lambda)$.

If $p = 3$, then $H_p(t) = 1 + t$, so there is exactly one supersingular curve in this case, with j-invariant $j(E_{-1}) = 1728$. Suppose therefore that $p \ge 5$. Recall that the map $\lambda \mapsto j(E_\lambda)$ is six-to-one if $j \neq 0$ or 1728, two-to-one if $j = 0$, and three-to-one if $j = 1728$. Furthermore, if $H_p(\lambda) = 0$ and $j(E_\lambda) = j(E_{\lambda'})$, then $H_p(\lambda') = 0$ also, since $E_\lambda \simeq E_{\lambda'}$ and the roots of $H_p(t)$ consist precisely of all values of λ for which E_λ is supersingular. For each number β, say, define

$$\varepsilon_p(\beta) = \begin{cases} 1 & \text{if } \beta \text{ is a supersingular } j\text{-invariant over } \mathbb{F}_p, \\ 0 & \text{if } \beta \text{ is an ordinary } j\text{-invariant over } \mathbb{F}_p. \end{cases}$$

Then the number of supersingular elliptic curves in characteristic $p \geq 5$ is

$$\frac{1}{6}\left(\frac{p-1}{2} - 2\varepsilon_p(0) - 3\varepsilon_p(1728)\right) + \varepsilon_p(0) + \varepsilon_p(1728) = \frac{p-1}{2} + \frac{2}{3}\varepsilon_p(0) + \frac{1}{2}\varepsilon_p(1728).$$

We have to determine for which primes $p \geq 5$ the curve $E : y^2 = x^3 + 1$ (with j-invariant 0) is supersingular. Apply part (a): the coefficient of x^{p-1} in $(x^3+1)^{(p-1)/2}$ is

$$\begin{cases} 0 & \text{if } p \equiv 2 \pmod 3 \text{ — supersingular,} \\ \binom{(p-1)/2}{(p-1)/3} \not\equiv 0 \pmod p & \text{if } p \equiv 1 \pmod 3 \text{ — ordinary.} \end{cases}$$

We now have to determine for which primes $p \geq 5$ the curve $E : y^2 = x^3 + x$ (with j-invariant 1728) is supersingular. The coefficient of x^{p-1} in $(x^3+x)^{(p-1)/2}$ is equal to the coefficient of $x^{(p-1)/2}$ in $(x^2+1)^{(p-1)/2}$, which is

$$\begin{cases} 0 & \text{if } p \equiv 2 \pmod 3 \text{ — supersingular,} \\ \binom{(p-1)/2}{(p-1)/4} \not\equiv 0 \pmod p & \text{if } p \equiv 1 \pmod 4 \text{ — ordinary.} \end{cases}$$

Chapter 5

Elliptic Curves over $\mathbb{C}$

Basic Facts. We have $E(\mathbb{C}) \xrightarrow{\sim} \mathbb{C}/\Lambda$, a Riemann surface of genus 1. Over $\mathbb{C}$, every lattice gives rise to an elliptic curve. (In higher dimensions, it's possible to have lattices that give rise to abelian varieties that are not algebraic.)

Definition 5.1 Fix a lattice $\Lambda \subset \mathbb{C}$.

(a) Elliptic functions (relative to Λ) are meromorphic functions on $\mathbb{C}/\Lambda$, or meromorphic functions on $\mathbb{C}$, periodic with respect to Λ. The set of elliptic functions is denoted $\mathbb{C}(\Lambda)$. This is a field.

(b) A **fundamental parallelogram** for Λ is a set $P = \{a + t_1\omega_1 + t_2\omega_2 : 0 \leq t_1, t_2 \leq 1\}$, where $a \in \mathbb{C}$, and ω_1 and ω_2 are a basis of Λ.

Theorem 5.2 Suppose that $f \in \mathbb{C}(\Lambda)$.

(1) If f has no zeros or poles, then f is constant.

(2) $\sum_{w \in \mathbb{C}/\Lambda} \operatorname{ord}_w(f) = 0$.

(3) $\sum_{w \in \mathbb{C}/\Lambda} \operatorname{res}_w(f) = 0$.

(4) $\sum_{w \in \mathbb{C}/\Lambda} w \operatorname{ord}_w(f) \in \Lambda$.

Proof.

(1) If f has no poles, then it is bounded on the fundamental parallelogram. Hence f is a bounded entire function, and so is constant. If f has no zeros, then $1/f$ has no poles, and so we just argue as above.

The proofs of the remaining assertions follow via applying the residue theorem to suitable functions on P.

(2)

$$\sum_{w \in \mathbb{C}/\Lambda} \mathrm{ord}_w(f) = \frac{1}{2\pi i} \int_{\partial P} \frac{f'(z)}{f(z)}\, dz = 0.$$

(3)

$$\sum_{w \in \mathbb{C}/\Lambda} \mathrm{res}_w(f) = \frac{1}{2\pi i} \int_{\partial P} f(z)\, dz = 0.$$

(4)

$$\begin{aligned}\sum_{w \in \mathbb{C}/\Lambda} w\, \mathrm{ord}_w(f) &= \frac{1}{2\pi i} \int_{\partial P} z \frac{f'(z)}{f(z)}\, dz \\ &= \frac{1}{2\pi i} \left(\int_0^{\omega_1} + \int_{\omega_1}^{\omega_1+\omega_2} + \int_{\omega_1+\omega_2}^{\omega_2} + \int_{\omega_2}^{0} \right) z \frac{f'(z)}{f(z)}\, dz \\ &= \frac{-\omega_2}{2\pi i} \int_0^{\omega_1} \frac{f'(z)}{f(z)}\, dz + \frac{\omega_1}{2\pi i} \int_0^{\omega_2} \frac{f'(z)}{f(z)}\, dz.\end{aligned}$$

Now use the fact that e.g. $\frac{1}{2\pi i} \int_0^{\omega_1} \frac{f'(z)}{f(z)}\, dz$ is the winding number around 0 of the path $[0, 1] \to \mathbb{C}$ given by $t \mapsto f(t\omega_1)$, which is an integer, since $f(0) = f(\omega_1)$.

Definition 5.3 The **order** of an elliptic function f is the number of poles (counted with multiplicity) inside any fundamental parallelogram.

Corollary 5.4 Any nonconstant elliptic function f has order at least 2.

Proof. Suppose that f has a single simple pole. Then Theorem 5.2(3) implies that the residue of f at this pole is zero, so $f(z)$ is holomorphic. This implies that $f(z)$ is constant.

Definition 5.5 Let Λ be a lattice. The Weierstraß $\wp$-function relative to Λ is defined by the series

$$\wp(z;\Lambda) = \frac{1}{z^2} + \sum_{\substack{\omega\in\Lambda \\ \omega\neq 0}} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right).$$

This is periodic with respect to Λ. It has double poles at the lattice points and no other poles. $\wp'(z;\Lambda) = -2\sum_{\omega\in\Lambda} \frac{1}{(z-\omega)^3}$.

The **Eisenstein series** of weight $2k$ is

$$G_{2k}(\Omega) = \sum_{\substack{\omega\in\Lambda \\ \omega\neq 0}} \omega^{-2k}.$$

Lemma 5.6

$$\wp(z) = \frac{1}{z^2} + \sum_{\substack{n \text{ even} \\ n>0}} (n+1)G_{n+2}z^n.$$

Proof. We have

$$\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} = \frac{1}{\omega^2}\left(\frac{1}{\left(1-\frac{z}{\omega}\right)^2} - 1 \right) = \frac{1}{\omega^2}\left(\sum_{n>1} \left(\frac{z}{\omega}\right)^{n-1} \right).$$

Thus

$$\wp(z) = \frac{1}{z^2} + \sum_{n>1} \left(\sum \frac{1}{\omega^{n+1}} \right) nz^{n-1} = \frac{1}{z^2} + \sum_{\substack{n \text{ even} \\ n>0}} (n+1)G_{n+1}z^n.$$

Next, we observe that

$$\begin{aligned}
\wp(z) &= z^{-2} + 3G_4 z^2 + \cdots\ , \\
\wp(z)^2 &= z^{-4} + \text{constant} + \cdots\ . \\
\wp(z)^3 &= z^{-6} + * \cdot z + \cdots\ . \\
\wp'(z) &= -2z^3 + * \cdot z + \cdots\ . \\
\wp'(z)^2 &= 4z^{-6} + * \cdot z^{-2} + \cdots\ .
\end{aligned}$$

Look at

$$f(z) = \wp'(z)^2 - 4\wp(z)^3 + 60G_4\wp(z) + 140G_6.$$

This function $f(z)$ is holomorphic in a neighborhood of $z = 0$, and $f(0) = 0$. Since f is elliptic and holomorphic away from Λ, it follows that f is a holomorphic elliptic function and is therefore identically zero. So

$$\wp'(z)^2 = 4\wp(z)^3 - 60G_4\wp(z) - 140G_6.$$

In the future we will write g_2 for $60G_4$ and g_3 for $140G_6$.

Proposition 5.7 The equation $4x^3 - g_2x - g_3$ has only simple zeros.

Proof. Observe that $\wp'(z)$ is an odd function. So if $\omega = \frac{1}{2}\Lambda$, $\omega \notin \Lambda$, then $\wp'(\omega) = -\wp'(-\omega) = -\wp'(\omega)$, and so $\wp'(\omega) = 0$. It therefore follows that $4x^3 - g_2x - g_3$ has zeros at $x = \wp\left(\frac{\omega_1}{2}\right)$, $x = \wp\left(\frac{\omega_2}{2}\right)$, and $x = \wp\left(\frac{\omega_1+\omega_2}{2}\right)$. We now show that these three values of x are distinct. The function $f(z) := \wp(z) - \wp\left(\frac{\omega_1}{2}\right)$ has a double pole at $z = 0$, and a double zero at $z = \frac{\omega_1}{2}$. Hence $(f) = 2\left(\frac{\omega_1}{2}\right) - 2(0)$, so $f\left(\frac{\omega_2}{2}\right) \neq 0$ and $f\left(\frac{\omega_1+\omega_2}{2}\right) \neq 0$, i.e. $\wp\left(\frac{\omega_1}{2}\right) \neq \wp\left(\frac{\omega_2}{2}\right)$ and $\wp\left(\frac{\omega_1}{2}\right) \neq \wp\left(\frac{\omega_1+\omega_2}{2}\right)$. A similar argument shows that $\wp\left(\frac{\omega_2}{2}\right) \neq \wp\left(\frac{\omega_1+\omega_2}{2}\right)$.

Consequence. The equation $E : y^2 = 4x^3 - g_2x - g_3$ defines an elliptic curve over $\mathbb{C}$.

Theorem 5.8 $\mathbb{C}(\Lambda) = \mathbb{C}(\wp(z), \wp'(z))$.

Proof. Suppose that $f(z) \in \mathbb{C}(\Lambda)$. Then $f(z) = \frac{1}{2}(f(z) + f(-z)) + \frac{1}{2}(f(z) - f(-z))$, where the first term is even and the second term is odd. Observe that if $g(z) \in \mathbb{C}(\Lambda)$

is odd, then $\wp'(z)g(z)$ is even, and so we are reduced to considering even functions.

We claim that if $2\omega \in \Lambda$, then $\mathrm{ord}_\omega(f)$ is even. For $f(z) = f(-z)$, so $f^{(i)}(z) = (-1)^i f^{(i)}(z)$ for all $i \geq 0$. Now if $2\omega \in \Lambda$, then $f^{(i)}(\omega) = f^{(i)}(-\omega)$ for all i, and so we deduce that $f^{(i)}(\omega) = 0$ for all odd i. Hence $\mathrm{ord}_\omega(f)$ is even, as claimed.

We therefore see that if f is an even function, then $(f) = \sum_w n_w((w)+(-w))$, $n_w \in \mathbb{Z}$ for all w. Now

$$\mathrm{div}\left(\prod_w (\wp(z) - \wp(w))^{n_w}\right) = \sum_w n_w(-2(0) + (w) + (-w)) = (g(z)),$$

say. Hence $f(z)$ and $g(z)$ have exactly the same zeros and poles except possibly at $z = 0$. But now Theorem 5.2(2) implies that $\mathrm{ord}_0 f(z) = \mathrm{ord}_0 g(z)$ also. Therefore $(f) = (g)$, and now the result follows.

The map $\varphi : \mathbb{C}/\Lambda \to E(\mathbb{C}) \subseteq \mathbb{P}^2(\mathbb{C})$ given by $z \mapsto [\wp(z), \wp'(z), 1]$ is an analytic map.

φ is surjective. For any $x \in \mathbb{C}$, the function $\wp(z) - x$ has zeros (since $\wp(z)$ has a double pole). Thus there exists a z with $\wp(z) = x$. Then $(\wp(z), \wp'(z)) = (x, \pm y)$, and $(\wp(-z), \wp'(-z)) = (x, \mp y)$.

φ is injective. Suppose $\varphi(z_1) = \varphi(z_2)$. If $2z_1 \notin \Lambda$, then $\wp(z) - \wp(z_1)$ has order 2 and has zeros at z_1, $-z_1$, and z_2. Hence $z_1 \equiv \pm z_2 \pmod{\Lambda}$. Therefore $\wp'(z_1) = \wp'(z_2) = \wp'(\pm z_1) = \pm\wp'(z_1)$, so $z_1 \equiv z_2 \pmod{\Lambda}$ (since $\wp'(z_1) \neq 0$ from the proof of Theorem 5.8). If $2z_1 \in \Lambda$, then $\wp(z) - \wp(z_1)$ has a double zero at z_1, and vanishes at z_2. So $z_2 \equiv z_1 \pmod{\Lambda}$.

Chapter 6

Elliptic Curves over Local Fields

6.1 Formal Groups

The formal group of an elliptic curve (motivating example). Consider the Weierstraß equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$. Make the change of variables $z = -\frac{x}{y}$, $w = -\frac{1}{y}$. (So $y = -\frac{1}{w}$, $x = \frac{z}{w}$.) This yields

$$w = z^3 + a_1zw + a_2z^2w + a_3w^2 + a_4zw^2 + a_6w^3 =: f(z, w).$$

Now substitute this equation for w into itself repeatedly to obtain a formal power series. We obtain $w = z^3(1 + A_1z + A_2z^2 + \cdots)$, where $A_n \in \mathbb{Z}[a_1, \ldots, a_6]$. By the above procedure (assuming everything makes sense!), we have constructed $w(z)$ satisfying $w(z) = f(z, w(z))$. We may do this more precisely by using Hensel's Lemma:

Lemma 6.1 (Hensel's Lemma) Suppose that R is a ring which is complete with respect to an ideal I. Let $F(w) \in R[w]$ be a polynomial, and suppose that $a \in R$ satisfies $F(a) \in I^n$, $F'(a) \in R^\times$ (for some $n \geq 1$). Then for any $\alpha \in R$ satisfying $\alpha \equiv F'(a) \pmod{I}$, the sequence $w_0 = a$, $w_{m+1} = w_m - \frac{F(w_m)}{\alpha}$ converges to an element $b \in R$ satisfying $F(b) = 0$ and $b \equiv a \pmod{I^n}$. (b is uniquely determined if R is an integral domain.)

Proof. See Silverman IV, Lemma 1.2 or Fröhlich-Taylor, page 84.

Now define a sequence of polynomials $f_m(z, w)$ by $f_1(z, w) = f(z, w)$ and $f_{m+1}(z, w) =$

$f_m(z, f(z,w))$. Set

$$w(z) = \lim_{m\to\infty} f_m(z,0) \in \mathbb{Z}[a_1,\dots,a_6][[z]]$$

(assuming that this makes sense — see below).

Proposition 6.2

(a) The above procedure yields a power series

$$w(z) = z^3(1 + A_1 z + A_2 z^2 + \cdots) \in \mathbb{Z}[a_1,\dots,a_6][[z]].$$

(b) $w(z)$ is the unique power series satisfying $w(z) = f(z, w(z))$.

(c) Suppose that $\mathbb{Z}[a_1,\dots,a_6]$ is made into a graded ring by assigning weights $\mathrm{wt}(a_i) = i$. Then A_n is a homogeneous polynomial of weight n.

Proof.

(a) and (b) Apply Hensel's Lemma with $R = \mathbb{Z}[a_1,\dots,a_6][[z]]$, $I = (z)$, $F(w) = f(z,w) - w$, $a = 0$, and $\alpha = 1$.

(c) Use induction, starting with the fact that $f(z,w)$ is homogeneous of weight -3.

Now we may write down Laurent series for x and y:

$$x(z) = \frac{z}{w(z)} = \frac{1}{z^2} - \frac{a_1}{z} - a_3 z + (a_4 + a_1 a_3)z^2 + \cdots ,$$
$$y(z) = \frac{-1}{w(z)} = \frac{-1}{z^3} + \frac{a_1}{z^2} + \frac{a_2}{z} + a_3 + (a_4 + a_1 a_3)z + \cdots .$$

The coefficients of $x(z)$ and $y(z)$ lie in $\mathbb{Z}[a_1,\dots,a_6]$. For the invariant differential, we have

$$\frac{\omega(z)}{dz} = \frac{dx(z)/dz}{2y + a_1 x + a_3} = \frac{-2z^{-3} + \cdots}{-2z^{-3} + \cdots} \in \mathbb{Z}\left[\frac{1}{2}, a_1,\dots,a_6\right][[z]],$$

and

$$\frac{\omega(z)}{dz} = \frac{dy(z)/dz}{3x^2 + 2a_2 x + a_4 - a_1 y} = \frac{3z^{-4} + \cdots}{3z^{-4} + \cdots} \in \mathbb{Z}\left[\frac{1}{3}, a_1,\dots,a_6\right][[z]].$$

Hence $\frac{\omega(z)}{dz} \in \mathbb{Z}[a_1, \dots, a_6][[z]]$ also.

Now suppose that $a_1, \dots, a_6 \in \mathbb{Z}_p$. Then, for all $z \in p\mathbb{Z}_p$, we have $(x(z), y(z)) \in E(\mathbb{Q}_p)$. So we have a map $p\mathbb{Z}_p \hookrightarrow E(\mathbb{Q}_p)$. (This is *not* a group homomorphism.) We would now like to define an addition:

$$(z_1, w(z_1)) + (z_2, w(z_2)) = (z_3(z_1, z_2), w(z_3)).$$

(For brevity, we write $w_1 = w(z_1)$ and $w_2 = w(z_2)$. We will allow $z_3 \in R[[z_1, z_2]]$ for some ring R.) (Think of all of these as points on the curve $E(R[[z]])$. This is what the addition actually means.) We apply the chord-tangent method: The slope of the line joining $(z_1, w(z_1))$ and $(z_2, w(z_2))$ is

$$\lambda = \frac{w_2 - w_1}{z_2 - z_1} = \sum_{n \geq 3} A_n \left(\frac{z_2^n - z_1^n}{z_2 - z_1} \right).$$

Substituting into the Weierstraß equation gives a cubic in z whose third root is

$$z_3' = z_3'(z_1, z_2) \in \mathbb{Z}[a_1, \dots, a_6][[z_1, z_2]].$$

The inverse of a point (z, w) will have z-coordinate given by (recall $z = -x/y$)

$$i(z) = \frac{x(z)}{y(z) + a_1 x(z) + a_3} \in \mathbb{Z}[a_1, \dots, a_6][[z]].$$

Finally, we obtain

$$z_3 = F(z_1, z_2) = i(z_3'(z_1, z_2)) \in \mathbb{Z}[a_1, \dots, a_6][[z_1, z_2]].$$

From the properties of addition on E, it follows that $F(z_1, z_2)$ satisfies the following:

- $F(z_1, z_2) = F(z_2, z_1)$ (commutativity)
- $F(z_1, F(z_2, z_3)) = F(F(z_1, z_2), z_3)$ (associativity)
- $F(z, i(z)) = 0$ (inverse).

So $F(z_1, z_2)$ is a "group law without any elements."

Let us now pass to the general case:

Definition 6.3 A one-dimensional formal group over a ring R is a power series $F \in R[[x, y]]$ satisfying

1. $F(X, Y) = X + Y +$ higher order terms.
2. $F(X, Y) = F(Y, X)$.
3. $F(F(X, Y), Z) = F(X, F(Y, Z))$.
4. $F(X, 0) = F(0, X) = X$.
5. There exists $i(X) \in R[[X]]$ such that $F(X, i(X)) = 0$.

Examples.

- $\mathbb{G}_a : F(X, Y) = X + Y$.
- $\mathbb{G}_m : F(X, Y) = (1 + X)(1 + Y) - 1 = X + Y + XY$.
- $\hat{E}$: the formal group of an elliptic curve E.

Definition 6.4 A homomorphism between two formal groups F and G is a power series $\varphi \in R[[T]]$ (with no constant term) satisfying

$$G(\varphi(X), \varphi(Y)) = \varphi(F(X, Y)).$$

We say that F and G are isomorphic over R if there are homomorphisms $f : F \to G$ and $g : G \to F$ defined over R satisfying $f(g(T)) = g(f(T)) = T$.

Example. Define $[m](X) : F \to F$ by $[m](X) = F(X, [m-1](X))$ for $m \geq 0$ and $[-m](X) = i([m](X))$.

Proposition 6.5 Let F be a formal group over R, and suppose that $m \in \mathbb{Z}$. Then

(a) $[m](T) = mT +$ higher order terms.

(b) If $m \in R^\times$, then $[m] : F \to F$ is an isomorphism.

Proof.

(a) This follows by induction.

(b) This follows from the following fact: If $f(X) \in R[[X]]$ with $f(0) = 0$ and $f'(0) \in R^\times$, then there exists $g(X) \in R[[X]]$ such that $g(f(X)) = X$. To show existence, we inductively construct a sequence of polynomials $g_n(X) \in R[X]$ such that $f(g_n(X)) = X \pmod{X^{n+1}}$ and $g_{n+1}(X) \equiv g_n(X) \pmod{X^{n+1}}$. Then $g(X) := \lim_{n\to\infty} g_n(X)$ exists and satisfies $f(g(X)) = X$. Set $a = f'(0) \in R^\times$, and take $g_1(X) = a^{-1}X$. Suppose we've constructed $g_{n-1}(X)$. We seek $\lambda \in R$ such that $g_n(X) = g_{n-1}(X) + \lambda X^n$ satisfies the desired property:

$$\begin{aligned} f(g_n(X)) &= f(g_{n-1}(X) + \lambda X^n) \\ &\equiv f(g_{n-1}(X)) + a\lambda X^n \pmod{X^{n+1}} \\ &\equiv X + \alpha X^n + a\lambda X^n \pmod{X^{n+1}} \end{aligned}$$

for some $\alpha \in R$, via our inductive hypothesis. So we can take $\lambda = -\alpha a^{-1} \in R$ (remember that $a \in R^\times$!). It now follows that $g(X)$ exists. Now $f(g(X)) = X$, so $g(f(g(X))) = g(X)$ in $R[[g(X)]]$, so $g(f(X)) = X$. To show uniqueness, note that if $f(h(X)) = X$, then $g(X) = g(f(h(X))) = (g \circ f)(h(X)) = h(X)$. So $g(X)$ is unique.

Suppose now that R is a complete local ring with maximal ideal $\mathfrak{m}$ and residue field k. Let F be a formal group over R. We may endow $\mathfrak{m}$ with a new group structure via F as follows:

Definition 6.6 The group law associated to F is the set $\mathfrak{m}$ endowed with the following operations: addition $x \oplus_F y = F(x, y)$ for $x, y \in \mathfrak{m}$, and inverses $\ominus_F x = i(x)$ for $x \in \mathfrak{m}$. The power series $F(x, y)$ and $i(x)$ converge for $x, y \in \mathfrak{m}$ (since R is complete). Hence $\mathfrak{m}$ endowed with this structure is a group. (We often write $F(\mathfrak{m})$ for this group.)

Examples.

(a) $\hat{\mathbb{G}}_a(\mathfrak{m})$ is $\mathfrak{m}$ with the usual addition law. There is an exact sequence

$$0 \to \hat{\mathbb{G}}_a(\mathfrak{m}) \to R \to k \to 0.$$

(b) $\hat{\mathbb{G}}_m(\mathfrak{m})$ is the group of 1-units of R with the usual multiplication law. There is an exact sequence

$$1 \to \hat{\mathbb{G}}_m(\mathfrak{m}) \to R^\times \to k^\times \to 1.$$

(c) Let K be the field of fractions of R, and let $\hat{E}$ be the formal group of an elliptic curve E/K. Then there is a natural map $\mathfrak{m} \to E(K)$ given by $z \mapsto (x(z), y(z))$. This yields a homomorphism $\hat{E}(\mathfrak{m}) \to E(k)$. There is often (but not always!) an exact sequence

$$0 \to \hat{E}(\mathfrak{m}) \to E(K) \to \tilde{E}(k) \to 0.$$

Proposition 6.7

(a) Suppose $n \geq 1$. Then the natural map

$$\frac{F(\mathfrak{m}^n)}{F(\mathfrak{m}^{n+1})} \to \frac{\mathfrak{m}^n}{\mathfrak{m}^{n+1}}$$

induced by the identity on sets is an isomorphism.

(b) Suppose that $\text{char}(k) = p$. Then if $p \nmid m$, $F(\mathfrak{m})$ has no nontrivial m-torsion.

Proof.

(a) For $x, y \in \mathfrak{m}^n$, we have $x \oplus_F y = F(x, y) = x + y+$ higher order terms $\equiv x + y \pmod{\mathfrak{m}^{2n}}$.

(b) Suppose that $x \in F(\mathfrak{m})$ satisfies $[m](x) = 0$. Since m is prime to p, we have $m \notin \mathfrak{m}$, and so $[m] : F(\mathfrak{m}) \to F(\mathfrak{m})$ is an isomorphism. Hence $x = 0$.

Definition 6.8 A **differential** on a formal group F is an expression of the form

$$P(T)\, dT = \omega(T) \in R[[T]]\, dT.$$

An **invariant differential** is one satisfying $\omega \circ F(T, S) = \omega(T)$, i.e.

$$P(F(T, S))F_X(T, S)\, dT = P(T)\, dT.$$

($F_X(T, S)$ is the partial derivative with respect to the first variable.) We say that $\omega(T)$ is **normalized** if $P(0) = 1$.

Example. Suppose that E/R is an elliptic curve, and let $\omega = \frac{dx}{2y+a_1x+a_3}$. Then $\omega(z) = 1 + \cdots \in R[[z]]$. This translates into an invariant differential for the formal group $\hat{E}$.

Lemma 6.9

1. $F_X(0,T)^{-1}\,dT$ is an invariant differential on F.
2. If $P(T)\,dT$ is an invariant differential on F, then $P(T) = P(0)F_X(0,T)^{-1}\,dT$.

Proof.

1. From the associative law, we have $F(F(U,T),S) = F(U,F(T,S))$. Taking $\frac{\partial}{\partial U}$ gives
$$F_X(F(U,T),S)F_X(U,T) = F_X(U,F(T,S)).$$
Now setting $U = 0$ yields
$$F_X(T,S)F_X(0,T) = F_X(0,F(T,S)).$$
We set $P(T)^{-1} = F_X(0,T)$ and $P(F(T,S))^{-1} = F_X(0,F(T,S))$. This just says that $F(0,T)^{-1}\,dT$ is an invariant differential.
2. We have $P(F(T,S))F_X(T,S) = P(T)$. Setting $T = 0$ gives $P(S)F_X(0,S) = P(0)$, i.e. $P(S) = P(0)F_X(0,S)^{-1}$.

Corollary 6.10 Suppose that F and G are formal groups over R, with normalized invariant differentials $\omega_F(T)$ and $\omega_G(T)$, respectively. Let $f : F \to G$ be a homomorphism. Then $\omega_G \circ f = f'(0)\omega_F$.

Proof. We first observe that $\omega_G \circ f$ is an invariant differential on F:

$$\omega_G \circ f(F(T,S)) = \omega_G(F(f(T),f(S))) = \omega_G \circ f(T).$$

Lemma 6.9 implies that $\omega_G \circ f = \alpha\omega_F$ for some $\alpha \in R$, so $\alpha = f'(0)$ (compare initial terms).

Corollary 6.11 Suppose that F is a formal group over R, and let p be a rational prime. Then there exist $f(T), g(T) \in R[[T]]$ with $f(0) = g(0) = 0$ such that $[p](T) = pf(T) + g(T^p)$.

Proof. Let $\omega(T)$ be the normalized invariant differential on F. Proposition 6.7(a) implies that $[p]'(0) = p$. Thus Corollary 6.10 implies that

$$p\omega(T) = \omega \circ [p](T) = (1 + \cdots)[p]'(T),$$

so $[p]'(T) \in pR[[T]]$ since $1 + \cdots \in R[[T]]^\times$. Hence, for any term aT^n of the power series $[p](T)$, we have either $a \in pR$ or $p \mid n$.

Definition 6.12 Suppose that F is a formal group over R, and let K be the field of fractions of R, with characteristic 0. Let

$$\lambda_F(T) := \int F_X(0,T)^{-1}\, dT$$

(formal integral), i.e. if $\omega_F(T)$ is the normalized invariant differential on F, then $\lambda_F(T) := \int \omega_F(T)$.

Proposition 6.13 $\lambda_F(F(S,T)) = \lambda_F(S) + \lambda_F(T)$.

Proof. Let ω_F be the normalized invariant differential on F. Then $\omega_F(F(T,S)) = \omega_F(T)$, so integrating with respect to T, we have $\lambda_F(F(S,T)) = \lambda_F(T) + f(S)$, where $f(S) \in K[[S]]$. Setting $T = 0$ gives $\lambda_F(S) = \lambda_F(0) + f(S) = f(S)$.

λ_F is called the formal logarithm of F. Note that Proposition 6.13 implies that $\lambda_F : F \to \hat{\mathbb{G}}_a$ is a homomorphism of formal groups over K, since $\lambda_F \in K[[T]]$.

Definition 6.14 Observe that $\lambda_F(T) = T + \cdots$, so λ_F is a formal group isomorphism over K. We write $\exp_F$ for the inverse of λ_F, so $\exp_F$ is the unique power series satisfying $\exp_F \circ \lambda_F = \lambda_F \circ \exp_F = 1$ (cf the proof of Proposition 6.7(b)).

Theorem 6.15 $\lambda_F(T)$ converges on $\mathfrak{m}$ and $\exp_F(T)$ converges on $\mathfrak{m}^n$, where $n > \frac{v(p)}{p-1}$. (Here $v(p)$ denotes the largest integer such that $p \in \mathfrak{m}^{v(p)}$.) Also $\exp_F(T), \lambda_F(T) : \mathfrak{m}^n \to \mathfrak{m}^n$ if $n > \frac{v(p)}{p-1}$.

Corollary 6.16 $F(\mathfrak{m}^n) \xrightarrow{\sim} \mathfrak{m}^n$ if $n > \frac{v(p)}{p-1}$.

Let K be a finite extension of $\mathbb{Q}_p$, and v a valuation on K. Let R be the ring of integers of K, $\mathfrak{m}$ the maximal ideal of R, π a uniformizer in $\mathfrak{m}$, and $k = R/\mathfrak{m}$ the residue field.

Let E/K be an elliptic curve with Weierstraß model $y^2+a_1xy+a_3y = x^3+a_2x^2+a_4x+a_6$. Since $\text{char}(K) \neq 2$ or 3, we may put this equation in the form $E : y^2 = x^3+Ax+B$, with distriminant $\Delta = -16(4A^3 + 27B^2)$. E is nonsingular iff $\Delta \neq 0$.

Definition 6.17 A **minimal model** of E/R is a model of E (with all coefficients in R) such that $v(\Delta)$ is minimal.

Proposition 6.18 A minimal model of E/R is unique up to isomorphism over R.

Proof. Suppose E_1 and E_2 are minimal with $E_1 \overset{\sim/K}{\to} E_2$. Then the isomorphism must be of the form $x \mapsto u^2x + r$, $y \mapsto u^3x + sx + t$, $u \in K^\times$ (Corollary 3.4). Now under this transformation, $\Delta \mapsto u^{\pm 12}\Delta$. So if E_1/R and E_2/R are both minimal, then $v(\Delta_1) = v(\Delta_2)$, and $u \in R^\times$. This implies that $r, s, t \in R$ (see transformation formulae in Silverman III 1.2). Hence $E_1 \overset{\sim/R}{\to} E_2$.

6.2 Reduction

Suppose that E/K is an elliptic curve with a given minimal Weierstraß equation. Then we may reduce the coefficients of this equation $\pmod{\pi}$; this gives us a (possibly singular) curve over the residue field k via

$$\tilde{E} : y^2 + \tilde{a}_1 xy + \tilde{a}_3 y = x^3 + \tilde{a}_2 x^2 + \tilde{a}_4 x + \tilde{a}_6.$$

Suppose $P \in E(K)$. Then we may write $P = [x_0, y_0, z_0]$ with $x_0, y_0, z_0 \in R$ (and at least one of $x_0, y_0, z_0 \in R^\times$). So we have a reduction map $E(K) \to \tilde{E}(k)$ given by $P = [x_0, y_0, z_0] \mapsto \tilde{P} = [\tilde{x}_0, \tilde{y}_0, \tilde{z}_0]$. Let $\tilde{E}_{ns}(k)$ be the set of nonsingular points of $\tilde{E}(k)$. This is a group (direct check; see e.g. Silverman III, Proposition 2.5). We define $E_0(K) = \{P \in E(K) : \tilde{P} \in \tilde{E}_{ns}(k)\}$ and $E_1(K) = \{P \in E(K) : \tilde{P} = \tilde{O}\}$ (the kernel of reduction).

Proposition 6.19 There is an exact sequence of abelian groups

$$0 \to E_1(K) \to E_0(K) \to \tilde{E}_{ns}(k) \to 0.$$

Proof. First observe that reduction yields a group homomorphism, since if $P, Q \in \tilde{E}_{ns}(k)$, the line ℓ through P and Q intersects the curve again in $R \in \tilde{E}_{ns}(k)$. Now we show surjectivity on the right. Suppose $(\bar{x}, \bar{y}) \in \tilde{E}_{ns}(k)$, and let $f(x, y) = y^2 + a_1 xy + a_3 y - x^3 - a_2 x^2 - a_4 x - a_6 = 0$ be a minimal Weierstraß equation for E. Then $\frac{\partial f}{\partial x}(\bar{x}, \bar{y})$ and $\frac{\partial f}{\partial y}(\bar{x}, \bar{y})$ are not both zero, since the point $(\bar{x}, \bar{y})$ is nonsingular. Suppose that $\frac{\partial f}{\partial x}(\bar{x}, \bar{y}) \neq 0$, and take any $y_0 \in R$, reducing to $\bar{y} \pmod{\mathfrak{m}}$. Then $f(x, y_0) \in R[x]$. Suppose that $x_0 \in R$ reduces to $\bar{x}$. Then $f(x_0, y_0) \in \mathfrak{m}$ and $\frac{\partial f}{\partial x}(x_0, y_0) \notin \mathfrak{m}$. Hence there exists $x' \in R$ with $x' \equiv x_0 \pmod{\mathfrak{m}}$ such that $f(x', y_0) = 0$ and $(x', y_0) \mapsto (\bar{x}, \bar{y})$. (Hensel's Lemma — Lemma 6.1.) So we have surjectivity on the right.

Consequence. Suppose that $v(\Delta) = 0$. Then $\tilde{E}$ is nonsingular, and $\tilde{E}_{ns} = \tilde{E}$. So we have an exact sequence

$$\begin{tikzcd}
0 \arrow[r] & E_1(K) \arrow[r] & E_0(K) \arrow[r] \arrow[d, equal] & \tilde{E}(k) \arrow[r] & 0. \\
& & E(K) & &
\end{tikzcd}$$

We now analyze $E_1(K)$.

Proposition 6.20 The map $\hat{E}(m) \to E_1(K)$ given by $z \mapsto \left(\frac{z}{w(z)}, \frac{-1}{w(z)}\right)$ is an isomorphism.

Proof. We know that $w(z)$ converges for $z \in \mathfrak{m}$, and $\left(\frac{z}{w(z)}, \frac{-1}{w(z)}\right)$ satisfies the Weierstraß equation of E. So $\left(\frac{z}{w(z)}, \frac{-1}{w(z)}\right) \in E(K)$. Recall

$$w(z) = z^3(1 + A_1 z + A_2 z^2 + \cdots),$$

$A_n \in \mathbb{Z}[a_1, \ldots, a_6]$. So $v\left(\frac{-1}{w(z)}\right) = -3v(z)$, so $\left(\frac{z}{w(z)}, \frac{-1}{w(z)}\right) \in E_1(K)$. $w(z) = 0$ only if $z = 0$, so the map is injective. Now suppose $x, y \in E_1(K)$. Then $y^2 + \cdots = x^3 + \cdots$, so $3v(x) = 2v(y) = -6r$ (some $r \geq 1$). So $\frac{x}{y} \in \mathfrak{m}$, and so we have an injective homomorphism (Exercise!) $E_1(K) \to \hat{E}(\mathfrak{m})$ given by $(x, y) \mapsto -\frac{x}{y}$. Hence we have injections

$$\hat{E}(\mathfrak{m}) \to E_1(K) \to \hat{E}(\mathfrak{m}),$$

and so these must be isomorphisms.

Now we can look at points of finite order.

Proposition 6.21 Suppose that E/K is an elliptic curve, and $m \geq 1$ is an integer coprime to $\mathrm{char}(k)$.

(a) $E_1(K)$ has no nontrivial points of order m.

(b) Suppose that the reduced curve $\tilde{E}(k)$ is nonsingular, and let $E(K)[m]$ denote the set of points of order m in $E(K)$. Then the reduction map $E(K)[m] \to \tilde{E}(k)$ is injective.

Proof. Consider the exact sequence

$$0 \to E_1(K) \to E_0(K) \to \tilde{E}_{ns}(k) \to 0.$$

(a) $E_1(K) \simeq \hat{E}(\mathfrak{m})$, and so $E_1(K)$ contains no nontrivial points of order m (since this is true of $\hat{E}(\mathfrak{m})$) (Proposition 6.7(b)).

(b) If $\tilde{E}/k$ is nonsingular, then $E_0(K) = E(K)$ and $\tilde{E}_{ns}(k) = \tilde{E}(k)$. Hence the m-torsion in $E(K)$ injects into $\tilde{E}(k)$.

Corollary 6.22 If E has good reduction and $p \nmid m$, then $(x, y) \in E(K)[m]$ implies $x, y \in R$.

Proof. $E_1(K) = \{(x, y) : x, y \notin R\}$.

Example. Finally all torsion on $E : y^2 = x^3 - x$ over $\mathbb{Q}$. Observe that $\Delta = -64 = -2^6$. Consider $E \pmod 3$. $\tilde{E}$ is nonsingular. So we have (prime-to-3 torsion) $\hookrightarrow \tilde{E}(\mathbb{F}_3)$.

x	$x^3 - x$	y
0	0	0
1	0	0
-1	0	0

and the point at infinity. So $\#\tilde{E}(\mathbb{F}_3) = 4$, and this bounds the prime-to-3 torsion. Now consider $E \pmod 5$.

x	$x^3 - x$	y
0	0	0
1	1	0
2	2	± 1
-2	-2	± 2
-1	-1	0

and the point at infinity. So $\#\tilde{E}(\mathbb{F}_5) = 8$, and this implies that there is no 3-torsion. Hence $\#E(\mathbb{Q})_{\text{tors}} \leq 4$, and in fact

$$E(\mathbb{Q})_{\text{tors}} = \{(0, 0), (1, 0), (-1, 0), \infty\}$$

— all killed by 2.

Theorem 6.23 Suppose that K is a local field, and that E/K is an elliptic curve with good reduction. Let $p = \mathrm{char}(k)$, and suppose $m \in \mathbb{Z}$ with $p \nmid m$. Then $K(E_m)/K$ is unramified.

Proof. Suppose σ is in the inertia subgroup of $\mathrm{Gal}(K(E_m)/K)$. If $P \in E_m$, then

$$\widetilde{\sigma(P) - P} = \widetilde{\sigma(P)} - \tilde{P} = \tilde{P} - \tilde{P} = \tilde{O}.$$

Hence $\sigma(P) = P$ for all $P \in E_m$, and so $\sigma = 1$, as required.

Remark.

(a) Theorem 6.23 is false without the good reduction hypothesis.

(b) Suppose that F is a number field. Then $F(E_m)/F$ is ramified only at primes dividing m and primes of bad reduction.

Theorem 6.24 Suppose that K is a local field, and that E/K has good reduction. Let $P \in E(\bar{K})$ with $mP \in E(K)$ and $p \nmid m$. Then $K(P)/K$ is unramified.

Proof. If $\sigma \in \mathrm{Gal}(\bar{K}/K)$, then $m(\sigma P - P) = \sigma(mP) - mP = O$. Now, as before, if σ is in the inertia subgroup of $\mathrm{Gal}(K(P)/K)$, then $\widetilde{\sigma P - P} = \tilde{O}$, so $\sigma P - P = O$, and the result follows as previously.

The previous two theorems may be formulated in terms of Galois action: Let K^{nr} be the maximal unramified extension of K, and let I_v be the inertia subgroup of $\mathrm{Gal}(\bar{K}/K)$. Then there is an exact sequence

$$\begin{tikzcd}
1 \arrow[r] & \mathrm{Gal}(\bar{K}/K^{nr}) \arrow[r] \arrow[d, equal] & \mathrm{Gal}(\bar{K}/K) \arrow[r] & \mathrm{Gal}(K^{nr}/K) \arrow[r] \arrow[d, "\simeq"] & 1. \\
 & I_v & & \mathrm{Gal}(\bar{k}/k) &
\end{tikzcd}$$

Definition 6.25 Let Σ be a set on which $\mathrm{Gal}(\bar{K}/K)$ acts. Then Σ is said to be **unramified** at v if the action of I_v upon Σ is trivial.

Theorem 6.26 Let E/K be an elliptic curve with $\tilde{E}/k$ nonsingular.

(i) Suppose $m \geq 1$, with $p \nmid m$ ($p = \mathrm{char}(k)$). Then E_m is unramified at v.

(ii) If $\ell \neq p$, then $T_\ell(E)$ is unramified at v.

Proof. See above.

Definition 6.27 Suppose that E/K is an elliptic curve and that $\tilde{E}/k$ is the reduced curve for a minimal Weierstraß equation.

(i) E has **good** (or **stable**) reduction over K if $\tilde{E}/k$ is nonsingular.

(ii) E has **multiplicative** (or **semistable**) reduction over K if $\tilde{E}$ has a node.

(iii) E has **additive** (or **unstable**) reduction if $\tilde{E}$ has a cusp.

In case (ii) above, E is said to have split (respectively non-split) multiplicative reduction if the slopes of the tangent lines at the node are in k (respectively not in k).

The reasons for (some of) the above terminology are summarized by the following proposition.

Proposition 6.28 Let E/K be an elliptic curve with minimal Weierstraß equation $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$, and of discriminant Δ. Set $c_r = (a_1^2 + 4a_2)^2 - 24(2a_4 + a_1a_3)$.

(a) E has good reduction iff $v(\Delta) = 0$.

(b) E has multiplicative reduction iff $v(\Delta) > 0$ and $v(c_4) = 0$. In this case, $\tilde{E}_{ns}(\bar{k}) \simeq \bar{k}^\times$.

(c) E has additive reduction iff $v(\Delta) > 0$ and $v(c_4) > 0$. In this case, $\tilde{E}_{ns}(\bar{k}) \simeq \bar{k}^+$.

Proof. Tedious case-by-case analysis. See Silverman III 1.4 and III 2.5.

Definition 6.29 An elliptic curve E/K is said to have **potential good reduction** over K if there is a finite extension K'/K such that E/K' has good reduction.

Exercise. If E/K has complex multiplication, then E/K has potential good reduction.

Theorem 6.30 (Semistable reduction theorem) Let E/K be an elliptic curve.

(a) Suppose that K'/K is an unramified extension. Then the reduction type of E over K is the same as that of E over K'.

(b) Suppose that K'/K is a finite extension, and that E has either good or multiplicative reduction over K. Then it has the same type of reduction over K'.

(c) There exists a finite extension K'/K such that E/K' has either good or split multiplicative reduction.

Proof. We suppose that we have v'/v, K'/K, R'/R, Δ'/Δ, and c_4'/c_4.

(a) For simplicity, assume $\mathrm{char}(k) \geq 5$, and let $y^2 = x^3 + Ax + B$ be a minimal Weierstraß equation of E over K. Let $x \mapsto (u')^2 x'$, $y \mapsto (u')^3 y'$ be a change of coordinates giving a minimal equation for E over K'. Then K'/K is unramified implies that there exists $u \in K$ such that $u/u' \in (R')^\times$. So we see that the substitution $x \mapsto u^2 x'$, $y \mapsto u^3 y'$ also gives a minimal equation for E/K' because $v'(u^{-12}\Delta) = v'((u')^{-12}\Delta)$. Since this new equation also has coefficients in R, we have $v(u) = 0$, as the original equation was minimal over K. Thus the original equation is also minimal over K'. Now $v(\Delta) = v'(\Delta)$ and $v(c_4) = v'(c_4')$, so by Proposition 6.28, E has the same reduction type over K and K'.

(b) Let Δ and c_4 be the quantities associated to a minimal Weierstraß equation for E/K, and let $x \mapsto u^2 x' + r$, $y \mapsto u^3 y' + s u^2 x' + t$ be a change of coordinates giving a minimal equation over K' with associated quantities Δ' and c_4'. Then $0 \leq v'(\Delta') = v'(u^{-12}\Delta)$ and $0 \leq v'(c_4') = v'(u^{-4}c_4)$. Since $v'(\Delta')$ is minimal, $u \in R'$, and so

$$0 \leq v'(u) \leq \min\left\{\frac{1}{12}v'(\Delta), \frac{1}{4}v'(c_4)\right\}.$$

Now good reduction implies $v(\Delta) = 0$, and multiplicative reduction implies $v(c_4) = 0$, so $v'(u) = 0$ in the case of either good or multiplicative reduction. So we have $v'(\Delta') = v'(\Delta)$ and $v'(c_4') = v'(c_4)$, and now the result follows from Proposition 6.28.

(c) Assume for simplicity that $\mathrm{char}(k) \neq 2$, and that (possibly over a finite extension of K) E has a Weierstraß equation in Legendre normal form

$$E : y^2 = x(x-1)(x-\lambda), \qquad \lambda \neq 0, q.$$

Then $c_4 = 16(\lambda^2 - \lambda + 1)$ and $\Delta = 16\lambda^2(\lambda - 1)^2$. There are three cases to consider:

- (i) $\lambda \in R$, $\lambda \not\equiv 0$ or $1 \pmod{\mathfrak{m}}$. Then $\Delta \in R^\times$, so E has good reduction over K.
- (ii) $\lambda \in R$, $\lambda \equiv 0$ or $1 \pmod{\mathfrak{m}}$. Then $\Delta \in \mathfrak{m}$ and $c_4 \in R^\times$, so E has split multiplicative reduction.
- (iii) $\lambda \notin R$. Choose $r \geq 1$ so that $\pi^r \lambda \in R^\times$, and make the substitution $x \mapsto \pi^{-r}x'$, $y \mapsto \pi^{-3/2}y'$ (passing to $K(\pi^{1/2})$ if necessary). This yields a Weierstraß equation $(y')^2 = x'(x' - \pi^r)(x' - \pi^r\lambda)$ with integral coefficients, $\Delta' \in \mathfrak{m}$, $c_4' \in R^\times$, so E has split multiplicative reduction.

Proposition 6.31 Let E/K be an elliptic curve. Then E has potential good reduction iff $j(E) \in R$.

Proof. Assume that $\mathrm{char}(k) \neq 2$ and that $E : y^2 = x(x-1)(x-\lambda)$, $\lambda \neq 0$ or 1. Then we have

$$2^8(1-(1-\lambda))^3 - j\lambda^2(1-\lambda)^2 = 0.$$

Thus $j(E) \in R$ implies that λ is an integer and $\lambda \equiv 0$ or $1 \pmod{\mathfrak{m}}$. So the Legendre model has integral coefficients and good reduction. Suppose conversely that E has potential good reduction, and let K'/K be a finite extension such that E has good reduction over K'. Then

$$j(E) = \frac{(c_4')^3}{\Delta'} \in R',$$

so $j(E) \in R$ since $j(E) \in K$.

To study questions involving reduction, we introduce the notion of the Néron minimal model.

Definition 6.32 Let X be a scheme with a morphism to another scheme, $X \to S$. We say that X is a **group scheme** over S if there are

- A section $e : S \to X$ (the identity).
- A morphism $\rho : X \to X$ over S (the inverse).
- A morphism $\mu : X \times X \to X$ over S (group multiplication) such that
  - (a) The composition $\mu \circ (id \times \rho) : X \to X$ is equal to the projection $X \to S$ followed by e.
  - (b) The two morphisms $\mu \circ (\mu \times id)$ and $\mu \circ (id \times \mu)$ from $X \times X \times X \to X$ are the same.

Now let K be a local field as before, and let E/K be an elliptic curve.

Definition 6.33 A **Néron model** $\mathcal{E}/R$ for E/K is a smooth group scheme over R whose generic fiber is E/K and which satisfies the following universal property: Let X/R be a smooth scheme, and let $\phi_K : X \times_R K \to \mathcal{E} \times_R K$ be a rational map. Then ϕ_K extends uniquely to a morphism $\phi : X/R \to \mathcal{E}/K$. This universal property characterizes the Néron model.

By analyzing the special fiber of $\mathcal{E}/R$ (there are only finitely many possibilities), it is possible to prove the following result:

Theorem 6.34 Let E/K be an elliptic curve. If E has split multiplicative reduction over K, then $E(K)/E_0(K)$ is a cyclic group of order $v(\Delta)$. In all other cases, $E(K)/E_0(K)$ has order at most 4.

Corollary 6.35 $E_0(K)$ is of finite index in $E(K)$.

This fact can be used to give further insight into $E(K)$:

Proposition 6.36 Suppose K is a finite extension of $\mathbb{Q}_p$. Then $E(K)$ contains a subgroup of finite index which is isomorphic to the additive group R^+.

Proof. We know that $E(K)/E_0(K)$ is finite, and that $E_0(K)/E_1(K) \simeq \tilde{E}_{ns}(k)$, which is also finite. So it suffices to show that $E_1(K)$ has a subgroup of finite index which is isomorphic to R^+. We have $E_1(K) \simeq \hat{E}(\mathfrak{m})$. Now $\hat{E}(\mathfrak{m})$ has a filtration

$$\hat{E}(\mathfrak{m}) \supset \hat{E}(\mathfrak{m}^2) \supset \hat{E}(\mathfrak{m}^3) \supset \cdots,$$

and

$$\hat{E}(\mathfrak{m}^i)/\hat{E}(\mathfrak{m}^{i+1}) \simeq \mathfrak{m}^i/\mathfrak{m}^{i+1}$$

(Proposition 6.7(a)), which is finite. For $r > \frac{v(p)}{p-1}$ (where v is the valuation on K), we have (Corollary 6.16) $\hat{E}(\mathfrak{m}^r) \simeq \mathfrak{m}^r$ (via the formal logarithm), which in turn is isomorphic to $\pi^r R$, where π is a local uniformizer of K.

Theorem 6.37 (Criterion of Néron-Ogg-Shafarevich) Let E/K be an elliptic curve. The following statements are equivalent:

(a) E has good reduction over K.

(b) $E[m]$ is unramified at v for all integers $m \geq 1$ coprime to char(k).

(c) The Tate module $T_\ell(E)$ is unramified for some (or all) $\ell \neq$ char(k).

(d) $E[m]$ is unramified for infinitely many integers $m \geq 1$ coprime to char(k).

Proof. It suffices to show (d) implies (a). Let m be an integer satisfying:

(i) m is coprime to char(k).

(ii) $m > \#E(K^{nr})/E_0(K^{nr})$.

(iii) $E[m]$ is unramified at v.

Look at the two exact sequences

$$0 \to E_0(K^{nr}) \to E(K^{nr}) \to \frac{E(K^{nr})}{E_0(K^{nr})} \to 0 \tag{$*$}$$

and

$$0 \to E_1(K^{nr}) \to E_0(K^{nr}) \to \tilde{E}_{ns}(\bar{k}) \to 0. \tag{$**$}$$

Now $E[m] \subset E(K^{nr})$, and so $E(K^{nr})$ has a subgroup isomorphic to $(\mathbb{Z}/m\mathbb{Z})^2$. Since $m > \#E(K^{nr})/E_0(K^{nr})$, there exists a prime $\ell \mid m$ such that $E_0(K^{nr})$ contains a subgroup isomorphic to $(\mathbb{Z}/\ell\mathbb{Z})^2$. Now $(**)$ implies that $\tilde{E}_{ns}(\bar{k})$ contains a subgroup isomorphic to $(\mathbb{Z}/\ell\mathbb{Z})^2$, since $E_1(K^{nr})$ contains no ℓ-torsion. This can only happen if E has good reduction over K^{nr}, which in turn implies that E has good reduction over K.

Corollary 6.38 Suppose E_1/K and $E_2(K)$ are elliptic curves which are isogenous over K. Then either they both have good reduction or they both do not.

Proof. Let $\varphi : E_1 \to E_2$ be a nonzero isogeny defined over K. Suppose $m \geq 2$ is coprime to both $\mathrm{char}(k)$ and $\deg(\varphi)$. Then $\varphi : E_1[m] \to E_2[m]$ is an isomorphism of $\mathrm{Gal}(\bar{K}/K)$-modules. Hence both modules are ramified or both are not.

Corollary 6.39 Let E/K be an elliptic curve. Then E has potential good reduction iff the inertia group I_v acts on $T_\ell(E)$ through a finite quotient for some (or all) primes $\ell \neq \mathrm{char}(k)$.

Proof. Suppose that E/K has potential good reduction. Then there exists a finite extension K'/K such that E has good reduction over K'; we may assume that K'/K is Galois. Theorem 6.37 implies that $I_{v'}$ acts trivially on $T_\ell(E)$ for all $\ell \neq \mathrm{char}(k)$. Hence the action of I_v on $T_\ell(E)$ factors though the finite quotient $I_v/I_{v'}$, as desired.

Suppose conversely that for some $\ell \neq \mathrm{char}(k)$, the action of I_v on $T_\ell(E)$ factors though a finite quotient I_v/H. Then $\bar{K}^H$ is a finite extension of $\bar{K}^{I_v} = K^{nr}$. Hence there exists a finite extension K'/K such that $\bar{K}^H = K' \cdot K^{nr}$. Then $I_{v'} = H$, which acts trivially on $T_\ell(E)$ by hypothesis. Theorem 6.37 now implies that E has good reduction over K'.

Chapter 7

A Cohomological Interlude

7.1 Cohomology of Finite Groups

Suppose that G is a finite group, and M is a G-module.

Definition 7.1 We define $H^0(G, M) = M^G = \{m \in M \mid \sigma(m) = m \text{ for all } \sigma \in G\}$.

Definition 7.2 A **(1)-cocycle** or **crossed homomorphism** is a map $f : G \to M$ such that $f(\sigma\tau) = f(\sigma) + \sigma f(\tau)$ for all $\sigma, \tau \in G$. So

$$f(1) = f(1 \cdot 1) = f(1) + f(1),$$

and so $f(1) = 0$. For any fixed $m \in M$, the map $\sigma \mapsto \sigma(m) - m$ is a cocycle. We say that such a cocycle is a **coboundary** (or that such a crossed homomorphism is **principal**). The sets of cocycles and coboundaries are closed under addition and subtraction.

Definition 7.3 We define $H^1(G, m) = \frac{\{\text{cocycles}\}}{\{\text{coboundaries}\}}$.

Remark. If G acts trivially on M, then a cocycle is a homomorphism, and every coboundary is zero. So $H'(G, M) = \mathrm{Hom}(G, M)$ (and $H^0(G, M) = M^G = M$).

Theorem 7.4 (Hilbert's Theorem 90) Suppose L/K is a finite Galois extension with $G = \mathrm{Gal}(L/K)$. Then $H^1(G, L^\times) = 0$.

Proof. Suppose $f : G \to L^\times$ is a cocycle. So $f(\sigma\tau) = f(\sigma)f(\tau)^\sigma$. We seek $\gamma \in L^\times$ such that $f(\sigma) = \frac{\sigma(\gamma)}{\gamma}$ for all $\sigma \in G$. Now since f is not the zero map, it follows via linear independence of characters that the map $L \to L$ given by

$$x \mapsto \sum_{\tau \in G} f(\tau)\tau(x)$$

is not the zero map, i.e. there exists $\alpha \in L$ such that

$$\beta := \sum_{\tau \in G} f(\tau)\tau(\alpha) \neq 0.$$

Then

$$\begin{aligned}\sigma(\beta) &= \sum_{\tau \in G} \sigma(f(\tau))\sigma\tau(\alpha) \\ &= \sum_{\tau \in G} f(\sigma)^{-1} f(\sigma\tau)\sigma\tau(\alpha) \\ &= f(\sigma)^{-1} \sum_{\tau \in G} f(\sigma\tau)\sigma\tau(\alpha) \\ &= f(\sigma)^{-1}\beta.\end{aligned}$$

Thus

$$f(\sigma) = \frac{\beta}{\sigma(\beta)} = \frac{\sigma(\beta^{-1})}{\beta^{-1}},$$

as desired.

Corollary 7.5 A point $P = (x_0 : \cdots : x_n) \in \mathbb{P}^n(L)$ is fixed by G iff it is represented by an $(n+1)$-tuple in K.

Proof. Suppose that $\sigma(P) = P$ for all $\sigma \in G$. Then we have $\sigma(x_0, \ldots, x_n) = c(\sigma)(x_0, \ldots, x_n)$ for some $c(\sigma) \in L^\times$. Check that $\sigma \mapsto c(\sigma)$ is a cocycle. Then Theorem 7.4 implies that $c(\sigma) = \frac{\alpha}{\sigma(\alpha)}$ for some $\alpha \in L^\times$. Thus $\sigma(\alpha x_0, \ldots, \alpha x_n) = (\alpha x_0, \ldots, \alpha x_n)$, and so $\alpha x_i \in K$ for $i = 0, \ldots, n$.

Proposition 7.6 For any exact sequence of G-modules

$$0 \to M \xrightarrow{g} M \xrightarrow{f} P \to 0,$$

there is a natural exact sequence

$$0 \to H^0(G,M) \to H^0(G,N) \to H^0(G,P) \xrightarrow{\delta} H^1(G,M) \to H^1(G,N) \to H^1(G,P).$$

Proof. Here is the definition of the connecting homomorphism δ: Suppose $p \in H^0(G,P) = P^G$. Then there exists $n \in N$ with $f(n) = p$. For any $\sigma \in G$, $f(\sigma(n) - n) = \sigma(p) - p = 0$, and so $\sigma(n) - n \in M$. Then $G \to M$ given by $\sigma \mapsto \sigma(n) - n$ is a cocycle. Check that this is well-defined, etc.

Definition 7.7 Suppose $H \leq G$. Then the restriction map $f \mapsto f \mid_H$ on cocycles induces a restriction homomorphism Res : $H^1(G,M) \to H^1(H,M)$ on cohomology groups.

Remark. Suppose that $H \lhd G$, and that M is a G-module. Then M^H is a G/H-module. A cocycle $f : G/H \to M^H$ induces a cocycle $\tilde{f} : G \to M$

$$\begin{tikzcd}
G \arrow[r, dashed, "\tilde{f}"] \arrow[d] & M \\
G/H \arrow[r, "f"'] & M^H \arrow[u, hook]
\end{tikzcd}$$

and so we obtain an inflation homomorphism Inf : $H^1(G/H, M^H) \to H^1(G,M)$. Then the following sequence is exact (exercise):

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

7.2 Cohomology of Infinite Galois Groups

Suppose K is a perfect field, and set $G = \mathrm{Aut}(\bar{K}/K)$. We define the Krull topology on G as follows: $H \leq G$ is open iff $\mathrm{Fix}(H)/K$ is a finite extension. We write $\mathrm{Gal}(\bar{K}/K)$ for G endowed with the Krull topology. We have the Galois correspondence

$$\{\text{finite extensions of } K\} \leftrightarrow \{\text{open subgroups of } G\}.$$

Definition 7.8

- WE say that a G-module M is **discrete** if the map $G \times M \to M$ is continuous relative to the discrete topology on M and the Krull topology on G. This is equivalent to $M = \bigcup_H M^H$, where H runs over open subgroups of G (i.e. every element of M is fixed by a subgroup of G fixing a finite extension of K).
- Suppose that M is discrete. Then a cocycle $f : G \to M$ is continuous iff f is constant on the cosets of some open normal subgroup H of G. (Then f arises via inflation from a cocycle $G/H \to M$.) Every coboundary is continuous.

Definition 7.9 $H^1(G, M) = \frac{\{\text{continuous cocycles}\}}{\{\text{coboundaries}\}}$. So

$$H^1(G, M) = \varinjlim_{H} H^1(G/H, M^H),$$

where H runs over open normal subgroups of G.

Example. (Kummer Theory) We have

$$H^1(\mathrm{Gal}(\bar{K}/K), \bar{K}^\times) = \varinjlim_{L} H^1(\mathrm{Gal}(L/K), L^\times) = 0$$

(via Hilbert's Theorem 90). Now consider the exact sequence

$$\begin{tikzcd}
1 \arrow[r] & \mu_n(\bar{K}) \arrow[r] & \bar{K}^\times \arrow[r] & \bar{K}^\times \arrow[r] & 1. \\
 & & x \arrow[r, mapsto] & x^n &
\end{tikzcd}$$

This yields the following exact sequence of cohomology groups:

$$\begin{tikzcd}
1 \arrow[r] & \mu_n(K) \arrow[r] & K^\times \arrow[r] & K^\times \arrow[r, "\delta"] & H^1(\mathrm{Gal}(\bar{K}/K), \mu(\bar{K})) \arrow[r] & H^1(\mathrm{Gal}(\bar{K}/K), \bar{K}^\times) = 1. \\
 & & x \arrow[r, mapsto] & x^n & &
\end{tikzcd}$$

So we have

$$H^1(\mathrm{Gal}(\bar{K}/K), \mu_n(\bar{K})) = \frac{K^\times}{(K^\times)^n}.$$

Notice that if $\mu_n(\bar{K}) \subseteq K^\times$, then

$$H^1(\mathrm{Gal}(\bar{K}/K), \mu_n(\bar{K})) = \mathrm{Hom}(\mathrm{Gal}(\bar{K}/K), \mu_n(\bar{K})),$$

and so

$$\frac{K^\times}{(K^\times)^n} \simeq \mathrm{Hom}(\mathrm{Gal}(\bar{K}/K), \mu_n(\bar{K})).$$

If $x \in K^\times$, then $\delta(x)$ is the cocycle given by $\sigma \mapsto \frac{\sigma(x^{1/n})}{x^{1/n}}$.

Chapter 8

Elliptic Curves over Global Fields

Mordell-Weil Theorem. If K is a number field and E/K is an elliptic curve, then $E(K)$ is finitely generated.

Weak Mordell-Weil Theorem. Suppose in addition that $n \in \mathbb{N}$. Then $E(K)/nE(K)$ is finite.

Notation. We write $H^i(\mathrm{Gal}(\bar{K}/K), -) = H^i(K, -)$.

Proposition 8.1

(a) If K is a number field or a local field, then there is an exact sequence

$$0 \to \frac{E(K)}{nE(K)} \to H^1(K, E_n) \to H^1(K, E)_n \to 0.$$

(b) IF K is a number field and v is any place of K, then the following diagram commutes:

$$\begin{tikzcd}
0 \arrow[r] & \frac{E(K)}{nE(K)} \arrow[r] \arrow[d] & H^1(K, E_n) \arrow[r] \arrow[d] \arrow[dr, dashed, "\partial_v"] & H^1(K, E)n \arrow[r] \arrow[d, "\mathrm{loc}_v"] & 0 \\
0 \arrow[r] & \frac{E(K_v)}{nE(K_v)} \arrow[r] & H^1(K_v, E_n) \arrow[r] & H^1(K_v, E)_n \arrow[r] & 0
\end{tikzcd}$$

Proof.

(a) There is an exact sequence

$$0 \to E_n \to E(\bar{K}) \overset{[n]}{\to} E(\bar{K}) \to 0.$$

Taking $\mathrm{Gal}(\bar{K}/K)$-cohomology of this sequence yields

$$0 \to E_n(K) \to E(K) \overset{[n]}{\to} E(K) \to H^1(K, E_n) \to H^1(K, E) \overset{[n]}{\to} H^1(K, E),$$

whence we obtain

$$0 \to \frac{E(K)}{nE(K)} \to H^1(K, E_n) \to H^1(K, E)_n \to 0.$$

If K is a number field, then $H^1(K, E_n)$ is usually infinite: e.g. suppose $E_n \subseteq E(K)$. Then also $\mu_n \subset K$ (via the existence of the Weil pairing). So

$$H^1(K, E_n) \simeq H^1(K, \mu_n \times \mu_n) \simeq \left(\frac{K^\times}{(K^\times)^n} \right)^2.$$

This motivates the following definitions:

Definition 8.2

(i) The n-**Selmer group** $S^{(n)}(E/K)$ is defined by

$$S^{(n)}(E/K) = \ker \left\{ H^1(K, E_n) \overset{\prod_v \partial v}{\longrightarrow} \prod_v H^1(K_v, E) \right\}.$$

(ii) The **Tate-Shafarevich group** of E/K is defined by

$$Ш(E/K) = \ker \left\{ H^1(K, E) \overset{\prod_v \mathrm{loc}_v}{\longrightarrow} H^1(K_v, E) \right\}.$$

Theorem 8.3 There is an exact sequence

$$0 \to \frac{E(K)}{nE(K)} \to S^{(n)}(E/K) \to Ш(E/K)_n \to 0.$$

Proof. Follows directly from the definitions. Alternatively, use the following:

Kernel-Cokernel Exact Sequence. Suppose A, B, and C are abelian groups with

$$A \xrightarrow{\alpha} B \xrightarrow{\beta} C.$$

Then there is an exact sequence

$$0 \to \ker(\alpha) \to \ker(\beta\alpha) \xrightarrow{\alpha} \ker(\beta) \to \operatorname{coker}(\alpha) \xrightarrow{\beta} \operatorname{coker}(\beta\alpha) \to \operatorname{coker}(\beta) \to 0.$$

To see this, apply the snake lemma to the following diagram:

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r, "id"] \arrow[d, "\alpha"'] & A \arrow[r, "\alpha"] \arrow[d, "\beta\alpha"] & B \arrow[d, "\beta"] \\
 & B \arrow[r, "\beta"] & C \arrow[r, "id"] & C
\end{tikzcd}$$

Then we have

$$H^1(K, E_n) \xrightarrow{\alpha} H^1(K, E)_n \xrightarrow{\beta} \prod_v H^1(K_v, E)_n,$$

so

$$0 \to \frac{E(K)}{nE(K)} \to S^{(n)}(E/K) \to Ш(E/K)_n \to 0$$

is exact.

Goal. We show that $S^{(n)}(E/K)$ is finite.

The essential idea is to show that each element of $S^{(n)}(E/K)$ becomes trivial over an extension of bounded degree which is unramified away from a set of primes depending only on n, E, and K. We shall then appeal to the classical finiteness theorems of algebraic number theory to complete the proof.

Lemma 8.4 Let v be a finite place of K, and suppose that E/K_v has good reduction. Suppose also that $\mathrm{char}(k_v) \nmid n$ (k_v is the residue field of K_v). Then for any $P \in E(K_V)$, there exists a finite unramified extension $M(v;P)/K_v$ such that $P \in nE(M(v;P))$.

Proof. Follows immediately from the fact that $K_v\left(\frac{1}{n}P\right)/K_v$ is unramified (see Theorem 6.24).

Proposition 8.5 Let T be the set of infinite places of K, together with the finite set of finite places of K dividing $2n\Delta_E$. Then, for any $\gamma \in S^{(n)}(E/K)$ and any $v \notin T$, there exists a finite unramified extension $K_v(\gamma)$ of K_v such that γ maps to zero under the following sequence of maps

$$\begin{tikzcd}
H^1(K, E_n) \arrow[r, "\mathrm{loc}_v"] \arrow[d, "\subseteq"'] & H^1(K_v, E_n) \arrow[r, "\mathrm{Res}"] & H^1(K_v(\gamma), E_n). \\
S^{(n)}(E/K) & &
\end{tikzcd}$$

Proof. For any place v of K, there exists $P_v \in E(K_v)$ mapping to the image $\gamma_v \in H^1(K_v, E_n)$ of $\gamma \in S^{(n)}(E/K)$. If $v \notin T$, then E/K_v has good reduction. The result now follows via considering the following diagram (cf Lemma 8.4):

$$\begin{tikzcd}
E(K) \arrow[r, "{[n]}"] \arrow[d] & E(K) \arrow[r] \arrow[d] & H^1(K, E_n) \arrow[d] \\
E(K_v) \arrow[r, "{[n]}"] \arrow[d] & E(K_v) \arrow[r] \arrow[d] & H^1(K_v, E_n) \arrow[d] \\
E(M(v;P)) \arrow[r] & E(M(v;P_v)) \arrow[r] & H^1(M_v(v;P_v))
\end{tikzcd}$$

Lemma 8.6 For any finite extension L/K, the kernel of the restriction map $S^{(n)}(E/K) \to S^{(n)}(E/L)$ is finite.

Proof. Observe that the kernel of $H^1(K, E_n) \to H^1(L, E_n)$ is $H^1(\mathrm{Gal}(L/K), E_n)$, which is finite.

Consequence. In order to show that $S^{(n)}(E/K)$ is finite, we may assume that $E_n \subseteq E(K)$. Then

$$H^1(K, E_n) \simeq H^1(K, \mu_n) \times H^1(K, \mu_n) \simeq (K^\times / K^{\times n})^2.$$

We make this assumption from now on.

Observe that for any finite place v of K, we have a natural homomorphism $(K_v^\times, K_v^{\times n})^2 \to (\mathbb{Z}/n\mathbb{Z})^2$ given by $(\alpha, \beta) \mapsto (\mathrm{ord}_v(\alpha), \mathrm{ord}_v(\beta))$.

Proposition 8.7 Suppose that $\gamma \in S^{(n)}(E/K)$ and $v \notin T$. Then the image of γ under the sequence of maps

$$H^1(K, E_n) \to H^1(K_v, E_n) \xrightarrow{\sim} (K_v^\times / K_v^{\times n})^2 \xrightarrow{\text{``ord''}} (\mathbb{Z}/n\mathbb{Z})^2$$

is equal to zero.

Proof. Proposition 8.5 implies that there exists a finite unramified extension $K_v(\gamma)$ of K_v such that the image of $\gamma_v \in H^1(K_v, E_n)$ in $H^1(K_v(\gamma), E_n)$ is zero. The result now follows from the following diagram:

$$\begin{tikzcd}
H^1(K_v, E_n) \arrow[r, "\sim"] \arrow[d, "\mathrm{Res}"'] & (K_v^\times / K_v^{\times n})^2 \arrow[r] \arrow[d] & (\mathbb{Z}/n\mathbb{Z})^2 \arrow[d, equal] \\
H^1(K_v(\gamma), E_n) \arrow[r, "\sim"] & (K_v(\gamma)^\times / K_v(\gamma)^{\times n})^2 \arrow[r] & (\mathbb{Z}/n\mathbb{Z})^2
\end{tikzcd}$$

where the map on the right is the identity map because $K_v(\gamma)/K_v$ is unramified.

Theorem 8.8

(a) The ideal class group $\mathrm{Cl}(\mathfrak{o}_K)$ is finite.

(b) The unit group $\mathfrak{o}_K^\times$ of $\mathfrak{o}_K$ is finitely generated. Recall (to orient yourself) that there is an exact sequence

$$\begin{tikzcd}
1 \arrow[r] & \mathfrak{o}_K^\times \arrow[r] & \bigoplus_{v \text{ finite}} \mathbb{Z} \arrow[r] & \mathrm{Cl}(\mathfrak{o}_K) \arrow[r] & 0. \\
 & & \alpha \arrow[r, mapsto] & (\mathrm{ord}_v(\alpha)) &
\end{tikzcd}$$

(c) Let $\mathfrak{o}_{K,T}^\times$ and $\mathrm{Cl}(\mathfrak{o}_{K,T})$ be defined via exactness of the sequence

$$1 \to \mathfrak{o}_{K,T}^\times \to K^\times \to \bigoplus_{v \notin T} \mathbb{Z} \to \mathrm{Cl}(\mathfrak{o}_{K,T}) \to 0.$$

Then $\mathfrak{o}_{K,T}^\times$ is finitely generated, and $\mathrm{Cl}(\mathfrak{o}_{K,T})$ is finite.

Proof.

(c) This follows from the fact that (from the definitions) there is an exact sequence

$$1 \to \mathfrak{o}_K^\times \to \mathfrak{o}_{K,T}^\times \to \bigoplus_{v \in T} \mathbb{Z} \to \mathrm{Cl}(\mathfrak{o}_K) \to \mathrm{Cl}(\mathfrak{o}_{K,T}) \to 0. \tag{$*$}$$

Aliter: Apply the kernel-cokernel exact sequence to

$$K^\times \xrightarrow{\alpha} \bigoplus_{\text{all } v} \mathbb{Z} \xrightarrow{\beta} \bigoplus_{v \notin T} \mathbb{Z}$$

to obtain $(*)$.

Lemma 8.9 For *any* finite subset T of places of K which contains the infinite places of K, write N_T for the kernel of the map

$$K^\times / K^{\times n} \to \bigoplus_{v \notin T} \mathbb{Z}/n\mathbb{Z}$$

given by $\alpha \mapsto (\mathrm{ord}_v(\alpha))_{v \notin T}$. Then there is an exact sequence

$$1 \to \frac{\mathfrak{o}_{K,T}^\times}{\mathfrak{o}_{K,T}^{\times n}} \to N_T \to \mathrm{Cl}(\mathfrak{o}_{K,T})_n.$$

Proof. Consider the following diagram:

$$\begin{tikzcd}
1 \arrow[r] & \mathfrak{o}_{K,T}^{\times} \arrow[r] \arrow[d, "n"] & K^{\times} \arrow[r] \arrow[d, "n"] & \bigoplus_{v\notin T} \mathbb{Z} \arrow[r] \arrow[d, "n"] & \mathrm{Cl}(\mathfrak{o}_{K,T}) \arrow[r] \arrow[d, "n"] & 0 \\
1 \arrow[r] & \mathfrak{o}_{K,T}^{\times} \arrow[r] & K^{\times} \arrow[r] \arrow[d] & \bigoplus_{v\notin T} \mathbb{Z} \arrow[r] \arrow[d] & \mathrm{Cl}(\mathfrak{o}_{K,T}) \arrow[r] & 0 \\
 & & \frac{K^{\times}}{K^{\times n}} \arrow[r] & \bigoplus_{v\notin T} \frac{\mathbb{Z}}{n\mathbb{Z}} & &
\end{tikzcd}$$

Suppose $\alpha \in K^{\times}$ represents an element of N_T. Then $n \mid \mathrm{ord}_v(\alpha)$ for all $v \notin T$, so we can map α to the class of

$$c = \left(\frac{\mathrm{ord}_v(\alpha)}{n}\right)_{v\notin T} \in \mathrm{Cl}(\mathfrak{o}_{K,T}).$$

Clearly we have $nc = 0$. Suppose $c = 0$. Then there exists $\beta \in K^{\times}$ such that $\mathrm{ord}_v(\beta) = \frac{\mathrm{ord}_v(\alpha)}{n}$ for all $v \notin T$. Then $\frac{\alpha}{\beta^n} \in \mathfrak{o}_{K,T}^{\times}$ and is well-defined up to an element of $\mathfrak{o}_{K,T}^{\times}$.

Corollary 8.10 $S^{(n)}(E/K)$ is finite.

Proof. Follows from Proposition 8.7 and Lemma 8.9.

Theorem 8.11 (Descent Theorem.) Suppose that A is an abelian group and that there is a function $h : A \to \mathbb{R}$ (a **height function**) satisfying the following properties:

(i) Suppose $Q \in A$. Then there exists a constant C_Q (depending only on A and Q) such that for all $P \in A$, $h(P + Q) \leq 2h(P) + C_Q$.

(ii) There exists an integer $m \geq 2$ and a constant C_2 (depending only on A) such that for all $P \in A$, $h(mP) \geq m^2 h(P) - C_2$.

(iii) For every constant C_3, $\#\{P \in A : h(P) \leq C_3\} < \infty$.

Then, if A/mA is finite (m as in (ii)), the group A is finitely generated.

Proof. Let $Q_1, \dots, Q_r \in A$ be a set of representatives of the cosets in A/mA, and suppose that $P \in A$. Then we may write

$$\begin{aligned} P &= mP_1 + Q_{i_1} \qquad (1 \le i_1 \le r), \\ P_1 &= mP_2 + Q_{i_2}, \\ \vdots\ &\ \ \vdots \\ P_{n-1} &= mP_n + Q_{i_n}. \end{aligned}$$

Then, for any $j \ge 1$,

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

(using (i) above), where $C_1' = \max_{1 \le i \le r}\{C_{Q_i}\}$. Note that C_1' and C_2 are both independent of P.

Now apply (†) repeatedly starting from P_n and working backward to P. We obtain

$$h(P_n) \le \left(\frac{2}{m^2}\right) h(P_{n-1}) + \frac{1}{m^2}(C_1' + C_2),$$

so

$$\begin{aligned} h(P_n) &\le \left(\frac{2}{m^2}\right)^n h(P) + \left(\frac{1}{m^2} + \frac{2}{m^4} + \frac{4}{m^6} + \cdots + \frac{2^{n-1}}{m^{2n}}\right)(C_1' + C_2) \\ &< \left(\frac{2}{m^2}\right)^n h(P) + \frac{m^2}{m^2 - 2}(C_1' + C_2) \\ &\le 2^{-n} h(P) + 2(C_1' + C_2) \end{aligned}$$

(since $m \ge 2$). So by taking n sufficiently large, we may ensure that $h(P_n) < 1 + 2(C_1' + C_2)$. Now

$$P = m^n P_n + \sum_{j=1}^{n} m^{j-1} Q_{i_j}$$

(from the definitions). Hence it follows that each element $P \in A$ is a linear combination of points in the set

$$\{Q_1, \ldots, Q_r\} \cup \{Q \in A \mid h(Q) \leq 1 + 2(C_1' + C_2)\},$$

and (iii) implies that this set is finite.

Definition 8.12 Suppose that K is a number field, and let $P = [x_0 : \cdots : x_N] \in \mathbb{P}^N(K)$, $x_i \in K$, $0 \leq i \leq N$. The **height** $H_K(P)$ of P relative to K is defined by

$$H_K(P) = \prod_{v \in M_K} \max\{|x_0|_v, \ldots, |x_N|_v\}^{[K_v:\mathbb{Q}_v]} = \prod_{v \in M_K} \{|x_0|_v, \ldots, |x_N|_v\}^{n_v}.$$

Proposition 8.13

(a) $H_K(P)$ is independent of the choice of homogeneous coordinates of P.

(b) $H_K(P) \geq 1$ for all $P \in \mathbb{P}^N$.

(c) If L/K is any finite extension, then

$$H_L(P) = H_K(P)^{[L:K]}.$$

Proof.

(a) For any $\lambda \in K^\times$, we have

$$\begin{aligned} \prod_v \max_i \{|\lambda x_i|_v\}^{n_v} &= \left(\prod_v |\lambda|_v^{n_v}\right) \prod_v \max_i \{|x_i|_v\}^{n_v} \\ &= \prod_v \max_i \{|x_i|_v\}^{n_v} \end{aligned}$$

(via the product formula).

(b) For any point $P \in \mathbb{P}^N(x)$, we may choose coordinates $[x_0 : \cdots : x_N]$ such that at least one $x_i = 1$. Then every factor in

$$\prod_v \max_i \{|x_i|_v\}^{n_v}$$

is at least 1.

(c) [Recall that

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

for $v \in M_K$.] We have

$$\begin{aligned} H_L(P) &= \prod_{w \in M_L} \max\{|x_i|_w\}^{n_w} \\ &= \prod_{v \in M_K} \prod_{\substack{w \in M_L \\ w|v}} \max\{|x_i|_v|\}^{n_w} \\ &= \prod_{v \in M_K} \max\{|x_i|_v\}^{[L:K]n_v} \\ &= H_K(P)^{[L:K]}. \end{aligned}$$

Definition 8.14 Suppose that $P \in \mathbb{P}^N(\bar{\mathbb{Q}})$. The **absolute height** $H(P)$ of P is defined by

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

where K is any number field such that $P \in \mathbb{P}^N(K)$.

Proposition 8.15 Suppose $P = [x_0 : \cdots : x_N] \in \mathbb{P}^N(\bar{\mathbb{Q}})$ and $\sigma \in \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$. Then $H(\sigma(P)) = H(P)$.

Proof. Choose a number field K such that $P \in \mathbb{P}^N(K)$. Write M_K and $M_{\sigma(K)}$ for the set of absolute values on K and $\sigma(K)$, respectively. Then we have isomorphisms $\sigma : K \xrightarrow{\sim} \sigma(K)$ and $\sigma : M_K \xrightarrow{\sim} M_{\sigma(K)}$ given by $v \mapsto \sigma(v)$ and $|\sigma(x)|_{\sigma(v)} = |x|_v$ for all

$x \in K$. Also $\sigma : K_v \xrightarrow{\sim} \sigma(K)_{\sigma(V)}$ (isomorphism on completions), and so $n_v = n_{\sigma(v)}$ (equality of local degrees). Thus

$$\begin{aligned} H_{\sigma(K)}(\sigma(P)) &= \prod_{w \in M_{\sigma(K)}} \max\{|\sigma(x_i)|_w\}^{n_w} \\ &= \prod_{v \in M_K} \max\{|\sigma(x_i)|_{\sigma(v)}\}^{n_{\sigma(v)}} \\ &= \prod_{v \in M_K} \max\{|x_i|_v\}^{n_v} \\ &= H_K(P), \end{aligned}$$

whence the result follows.

Theorem 8.16 For any numbers $B, D \geq 0$, the set

$$\{P \in \mathbb{P}^N(\bar{\mathbb{Q}}) \mid H(P) \leq B \text{ and } [\mathbb{Q}(P) : \mathbb{Q}] \leq D\}$$

is finite. So, for any fixed number field K, the set $\{P \in \mathbb{P}^N(K) \mid H_K(P) \leq B\}$ is finite.

Proof. Suppose that $P = [x_0 : \cdots : x_N]$ with some $x_i = 1$. Then, for any v and for any i, we have

$$\max\{|x_1|_v, \ldots, |x_N|_v\}^{n_v} \geq \max\{|x_i|_v, 1\}^{n_v}.$$

So, multiplying over all v and taking an appropriate root gives $H(P) \geq H(x_i)$ for $0 \leq i \leq N$. Plainly $\mathbb{Q}(x_i) \subseteq \mathbb{Q}(P)$.

It suffices to prove that for each $1 \leq d \leq D$, the set

$$\{x \in Q\bar{Q} \mid H(x) \leq B \text{ and } [\mathbb{Q}(x) : \mathbb{Q}] = d\}$$

is finite (i.e. we've reduced to the case $N = 1$). Set $K = \mathbb{Q}(x)$, and suppose $[K : \mathbb{Q}] = d$. Let $x_1, \ldots, x_d$ denote the Galois conjugates of x over $\mathbb{Q}$. Set

$$F_x(T) = \prod_{j=1}^{d}(T - x_j) = \sum_{r=0}^{d}(-1)^r S_r(x) T^{d-r}$$

— the minimal polynomial of x over $\mathbb{Q}$. Then

$$\begin{aligned}
|S_r(x)|_v &= \left| \sum_{1 \le i_1 < \cdots < i_r \le d} x_{i_1} \cdots x_{i_r} \right|_v \\
&\le c(v,r,d) \max_{1 \le i_1 < \cdots < i_r \le d} |x_{i_1} \cdots x_{i_r}|_v \\
&\le c(v,r,d) \max_{1 \le i \le d} |x_i|_v^r,
\end{aligned}$$

where

$$c(v,r,d) = \begin{cases} \binom{d}{r} \le 2^d & \text{if } v \text{ is archimedean,} \\ 1 & \text{if } v \text{ is nonarchimedean.} \end{cases}$$

Hence

$$\max\{|S_0(x)|_v, \ldots, |S_d(x)|_v\} \le c(v,d) \prod_{i=0}^{d} \max\{|x_i|_v, 1\}^d,$$

where

$$c(v,d) = \begin{cases} 2^d & \text{if } v \text{ is archimedean,} \\ 1 & \text{if } v \text{ is nonarchimedean.} \end{cases}$$

Multiplying over all $v \in M_K$ and taking $[K : \mathbb{Q}]^{\text{th}}$ roots yields

$$H(S_0(x), \ldots, S_d(x)) \le 2^d \prod_{i=0}^{d} H(x_i)^d = 2^d H(x)^{d^2}$$

(via Proposition 8.15). So if x lies in the set

$$\{x \in \bar{\mathbb{Q}} \mid H(x) \le B \text{ and } [\mathbb{Q}(x) : \mathbb{Q}] = d\},$$

then x is a root of a polynomial $F_x(T) \in \mathbb{Q}[T]$ whose coefficients $S_0, \ldots, S_d$ satisfy

$$H(S_0, \ldots, S_d) \le s^3 B^{d^2}.$$

There are only finitely many possibilities for such an $F_x(T)$, and hence there are only finitely many possibilities for x.

Corollary 8.17 (Kronecker) Let K be a number field and $P = [x_0 : \cdots : x_N] \in \mathbb{P}^N(K)$. Fix any i with $x_i \neq 0$. Then $H(P) = 1$ iff x_j/x_i is a root of unity or zero for every $0 \le j \le n$.

Proof. Without loss of generality we may divide the coordinates of $\mathbb{P}$ and then reorder them so that $P = (1, x_1, \ldots, x_N)$. Suppose that every x_j is zero or a root of unity. Then $\max\{1, |x_j|_v\} = 1$ for every v, and so $H(P) = 1$. Suppose conversely that $H(P) = 1$. Set

$$P^r = (1, x_1^r, x_2^r, \ldots, x_N^r),$$

$r = 1, 2, \ldots$ Then $H(P^r) = H(P)^r = 1$ for each r. Theorem 8.16 implies that the sequence $P, P^2, P^3, \ldots$ contains only finitely many distinct points, and so we may choose $r > s \geq 1$ such that $P^s = P^r$. Then $x_j^s = x_j^r$ $(1 \leq j \leq n)$ (since we've dehomogenized with $x_0 = 1$), so each x_j is a root of unity or zero.

8.1 Heights on Elliptic Curves

Recall. If $f \in \bar{K}(E)$, then we have a max $f : E \to \mathbb{P}^1$ given by

$$P \mapsto \begin{cases} [1, 0] & \text{if } P \text{ is a pole of } f, \\ [f(P), 1] & \text{otherwise.} \end{cases}$$

Definition 8.18 The (absolute logarithmic) height on $\mathbb{P}^N(\bar{\mathbb{Q}})$ is defined by $h : \mathbb{P}^N(\bar{\mathbb{Q}}) \to \mathbb{R}$, given by $P \mapsto \log H(P)$. (So $H(P) \geq 1$ implies $h(P) \geq 0$.)

Definition 8.19 Suppose that E/K is an elliptic curve and that $f \in \bar{K}(E)$ is a nonconstant function. The **height** on E relative to f is defined by $h_f : E(\bar{K}) \to \mathbb{R}$, given by $P \mapsto h(f(P))$.

Proposition 8.20 Suppose E/K is an elliptic curve and $f \in K(E)$ a nonconstant function. Then, for any constant C,

$$\#\{P \in E(K) \mid h_f(P) \leq C\} < \infty.$$

Proof. The function f gives a map from $\{P \in E(K) \mid h_f(P) \leq C\}$ to $\{Q \in \mathbb{P}^1(K) \mid H(Q) \leq e^C\}$, and this map is finite-to-one. Now apply Theorem 8.16.

Definition 8.21 A morphism of degree d between projective spaces is a map $F : \mathbb{P}^N \to \mathbb{P}^M$ given by $P \mapsto [f_0(P), \ldots, f_M(P)]$, where $f_0, \ldots, f_M \in \bar{\mathbb{Q}}[X_0, \ldots, X_N]$ are homogeneous polynomials of degree d with no common zero in $\bar{\mathbb{Q}}$ except $X_0 = \cdots = X_N = 0$.

Theorem 8.22 Suppose $F : \mathbb{P}^N \to \mathbb{P}^M$ is a morphism of degree d. Then there exist constants C_1 and C_2, depending only upon F, such that for all $P \in \mathbb{P}^N(\bar{\mathbb{Q}})$,

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

Proof. Set $F = [f_0, \ldots, f_M]$, f_i homogeneous for all i, and let $P = [x_0, \ldots, x_N] \in \mathbb{P}^N(\bar{\mathbb{Q}})$. Let K be a field containing all x_i's and all of the coefficients of all of the f_j's. Define

$$\begin{aligned} |P|_v &= \max_i \{|x_i|_v\}, \\ |F(P)|_v &= \max_j \{|f_j(P)|_v\}, \\ |F|_v &= \max\{|a|_v : a \text{ is a coefficient of some } f_j\}. \end{aligned}$$

Then

$$H_K(P) = \prod_v |P|_v^{n_v}$$

and

$$H_K(F(P)) = \prod_v |F(P)|_v^{n_v}.$$

So we define

$$H_K(F) = \prod_v |F|_v^{n_v}.$$

(This means that $H_K(F) := H_K([a_0, a_1, \ldots])$, where the a_i's are all of the coefficients of the f_j's.) Set

$$\varepsilon(v) = \begin{cases} 1 & \text{if } v \mid \infty, \\ 0 & \text{if } v \nmid \infty. \end{cases}$$

So, for example,

$$|t_1 + \cdots + t_n|_v \leq n^{\varepsilon(n)} \max\{|t_1|_v, \ldots, |t_n|_v\}$$

(triangle inequality).

We now show the upper bound. Each f_i is homogeneous of degree d. So, for each i, we have

$$|f_i(P)|_v \leq C_1^{\varepsilon(v)} |F|_v |P|_v^d$$

(via the triangle inequality) (e.g. take $C_1 = \binom{N+d}{d}$, the number of monomials of degree d in $N+1$ variables). Also

$$|F(P)|_v \leq C_1^{\varepsilon(v)} |F|_v |P|_v^d.$$

Now raising to the n_v^{th} power, multiplying over all v, and taking $[K : \mathbb{Q}]^{\text{th}}$ roots yields

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

We now show the lower bound. Recall Hilbert's Nullstellensatz: Suppose that $\mathfrak{a}$ is an ideal of $K[X_0, \ldots, X_N]$, and let f be any polynomial in $K[X_0, \ldots, X_N]$ such that $f(\alpha_0, \ldots, \alpha_N) = 0$ for every zero of $\mathfrak{a}$ in $\bar{\mathbb{Q}}$. Then there exists an integer $m > 0$ such that $f^m \in \mathfrak{a}$.

Suppose

$$\{Q \in \mathbb{A}^{N+1}(\bar{\mathbb{Q}}) : f_0(Q) = \cdots = f_N(Q) = 0\} = \{(0, \ldots, 0)\}.$$

Then by the Nullstellensatz, the ideal $(f_0, \ldots, f_M) \in \bar{\mathbb{Q}}[X_0, \ldots, X_N]$ contains some power of each of $X_0, \ldots, X_N$. Thus for some $e \geq 1$, there exist polynomials $g_{ij} \in \bar{\mathbb{Q}}[X_0, \ldots, X_N]$ such that

$$X_i^e = \sum_{j=0}^{M} g_{ij} f_j, \qquad 0 \leq i \leq N. \tag{$\dagger$}$$

Without loss of generality, we may assume:

- Each $g_{ij} \in K[X_0, \ldots, X_N]$.
- Each g_{ij} is homogeneous of degree $e - d$.

Set

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

and

$$H_K(G) := \prod_v |G|_v^{n_v}.$$

Now $P = [X_0, \ldots, X_N]$, and so (†) implies

$$|x_i|_v^e = \left| \sum_{j=0}^{M} g_{ij}(P) f_j(P) \right|_v \leq C_2^{\varepsilon(v)} \max_j \{|g_{ij}(P) f_j(P)|_v\},$$

so

$$|P|_v^e \leq C_2^{\varepsilon(v)} \max_{i,j}\{|g_{ij}(P)|_v\} |F(P)|_v \tag{$*$}$$

(taking the maximum over i). Now $\deg g_{ij} = e - d$, so

$$|g_{ij}(P)|_v \leq C_3^{\varepsilon(v)} |G|_v |P|_v^{e-d},$$

whence $(*)$ gives

$$|P_v|^d \leq C_4^{\varepsilon(v)} |G|_v |F(P)|_v,$$

and now the lower bound follows.

Theorem 8.23 Suppose E/K is an elliptic curve, and that $f \in K(E)$ is even (i.e. $f \circ [-1] = f$). Then for all $P, Q \in E(\bar{K})$, we have

$$h_f(P+Q) + h_f(P-Q) = 2h_f(P) + 2h_f(Q) + O_{E,f}(1).$$

Proof. Let $E : y^2 = x^3 + Ax + B$, say. We first consider the case of $f = x$. Then $h_x(O) = 0$ and $h_x(-P) = h_x(P)$, so the result holds if $P = O$ or $Q = O$. Thus suppose $P \neq O$ and $Q \neq O$. Let $x(P) = [x_1, 1]$, $x(Q) = [x_2, 1]$, $x(P+Q) = [x_3, 1]$, and $x(P-Q) = [x_4, 1]$. Then (by the addition formulae and algebra)

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

Define $g : \mathbb{P}^2 \to \mathbb{P}^2$ by

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

We claim that there is a commutative diagram

$$\begin{tikzcd}
 & (P,Q) \arrow[r, mapsto] & (P+Q, P-Q) & \\
(P,Q) \arrow[d, mapsto] & E \times E \arrow[r, "G"] \arrow[d] \arrow[dd, bend left, "\sigma"] & E \times E \arrow[d] \arrow[dd, bend right, "\sigma"'] & \\
(x(P), x(Q)) & \mathbb{P}^1 \times \mathbb{P}^1 \arrow[d] & \mathbb{P}^1 \times \mathbb{P}^1 \arrow[d] & ([\alpha_1, \beta_1], [\alpha_2, \beta_2]) \arrow[d, mapsto] \\
 & \mathbb{P}^2 \arrow[r, "g"] & \mathbb{P}^2 & [\beta_1\beta_2, \alpha_1\beta_2 + \alpha_2\beta_1, \alpha_1\alpha_2]
\end{tikzcd}$$

[The idea is to treat t, u, and v as 1, $x_1 + x_2$, and x_1x_2, respectively.] This follows from formulae for x_3 and x_4.

We claim that g is a morphism. We are required to prove that the three homogeneous polynomials defining g have no common zero except $t = u = v = 0$. So suppose $t = 0$. Then $u^2 - 4tv = 0$ and $(v - At)^2 - 4Btu = 0$ imply $u = v = 0$. Thus we may assume $t \neq 0$ and define $x := u/2t$. Then $u^2 - 4tv = 0$ becomes $x^2 = v/t$, $2u(At+v)+4Bt^2 = 0$ become $\psi(x) := 4x^3+4Ax+4B = 0$, and $(v-At)^2-4Btu = 0$ becomes $\phi(x) := x^4 - 2Ax^2 - 8Bx + A^2 = 0$. Observe that

$$(12x^2 - 16A)\phi(x) - (3x^2 - 5Ax + 27B)\psi(x) = 4(4A^3 + 27B^2) \neq 0$$

(since E is nonsingular), so $\psi(x)$ and $\phi(x)$ have no common zeros, so g is a morphism. Hence, from the diagram, we have

$$\begin{aligned}
h(\sigma(P + Q, P - Q)) &= h(\sigma \circ G(P, Q)) \\
&= h(g \circ \sigma(P, Q)) \\
&= 2h(\sigma(P, Q)) + O(1)
\end{aligned}$$

(via Theorem 8.22, since g is a morphism of degree 2).

We claim that for all $R_1, R_2 \in E(\bar{K})$, we have $h(\sigma(R_1, R_2)) = h_x(R_1)+h_x(R_2)+O(1)$. [Then

$$h(\sigma(P + Q, P - Q)) = 2h(\sigma(P, Q)) + O(1),$$

so

$$h_x(P+Q) + h_x(P-Q) = h_x(P) + h_x(Q) + O(1),$$

as desired.] First observe that if $R_1 = O$ or $R_2 = O$, then $h(\sigma(R_1 + R_2)) = h_x(R_1) + h_x(R_2)$. Thus assume $R_1 \neq O$ and $R_2 \neq O$, and set $x(R_1) = [\alpha_1, 1]$ and $x(R_2) = [\alpha_2, 1]$. Then $h(\sigma(R_1, R_2)) = h([1, \alpha_1 + \alpha_2, \alpha_1\alpha_2])$ and $h_x(R_1) + h_x(R_2) = h(\alpha_1) + h(\alpha_2)$. Now, just as in the proof of Theorem 8.16 (using the polynomial $(T - \alpha_1)(T - \alpha_2)$), we have

$$h([1, \alpha_1 + \alpha_2, \alpha_1\alpha_2]) \leq h(\alpha_1) + h(\alpha_2) + \log 2.$$

This establishes the claim. So we have now proven the theorem when $f = x$.

For an arbitrary even function f, we argue as follows: Suppose that $f, g \in \bar{K}(E)$ are even functions. We claim that $(\deg g)h_f = (\deg f)h_g + O(1)$. $K(x)$ is the subfield of even functions in $K(E)$ (see Silverman III, §2.3.1). Thus there exists $\rho(x)$ in $K(x)$ such that the following diagram commutes:

$$\begin{tikzcd}
E \arrow[d, "x"'] \arrow[dr, "f"] & \\
\mathbb{P}^1 \arrow[r, "\rho(x)"'] & \mathbb{P}^1
\end{tikzcd}$$

Then $h_f = h_{x \circ \rho} = (\deg \rho)h_x + O(1)$ (via Theorem 8.22). The diagram implies that $\deg(f) = \deg(x)\deg(\rho) = 2\deg(\rho)$. So

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

Similarly,

$$2h_g = (\deg g)h_x + O(1), \tag{$**$}$$

and now the claim follows from $(*)$ and $(**)$.

From this claim, we have $h_f = \frac{1}{2}(\deg f)h_x + O(1)$, and now the theorem follows for f because we've already shown that

$$h_x(P+Q) + h_x(P-Q) = 2h_x(P) + 2h_x(Q) + O(1).$$

Corollary 8.24 Suppose that E/K is an elliptic curve, and $f \in K(E)$ is an even function.

(a) Let $Q \in E(K)$. Then for all $P \in E(\bar{K})$, we have

$$h_f(P+Q) \leq 2h_f(P) + O_{E,f,Q}(1).$$

(b) Suppose $m \in \mathbb{Z}$. Then for all $P \in E(\bar{K})$,

$$h_f([m]P) = m^2 h_f(P) + O_{E,f,m}(1).$$

Proof.

(a) Theorem 8.23 implies that

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

since $h_f(P-Q) \geq 0$.

(b) It suffices to prove the result for $m \geq 0$ since f is even. It is true for $m = 0$ and 1 plainly! So assume that the result holds for m and $m-1$. Applying Theorem 8.23 with P replaced by $[m]P$ and Q by P gives

$$\begin{aligned} h_f([m+1]P) &= -h_f([m-1]P) + 2h_f([m]P) + 2hf(P) + O(1) \\ &= -((m-1)^2 + 2m^2 + 2)h_f(P) + O(1) \\ &= (m+1)^2 h_f(P) + O(1), \end{aligned}$$

as desired.

Theorem 8.25 (Mordell-Weil Theorem.) Let K be a number field, and let E/K be an elliptic curve. Then E/K is finitely generated.

Proof. We apply Theorem 8.11 (the Descent Theorem) with $m = 2$. Let $f \in K(E)$ be any nonconstant even function, and consider $h_f : E(\bar{K}) \to \mathbb{R}$. Then h_f satisfies the following properties:

(i) Suppose $Q \in E(K)$. Then there exists a constant C_1 (depending only on E, f, and Q) such that for all $P \in E(K)$, $h_f(P+Q) \leq 2h_f(P) + C_1$. (This follows from Corollary 8.24(a).)

(ii) There exists a constant C_2 (depending only upon E and f) such that $h_f([2]P) \geq 4h_f(P) - C_2$. (This follows from Corollary 8.24(b) with $m = 2$.)

(iii) For every constant C_3, $\#\{P \in E(K) \mid h_f(P) \leq C_3\} < \infty$. (This follows from Proposition 8.20.)

The goal is to construct an actual quadratic form on $E(K)$ that differs from h_f by a bounded quantity.

Proposition 8.26 Suppose E/K is an elliptic curve. Let $f \in K(E)$ be a nonconstant even function, and let $P \in E(\bar{K})$. Then

$$\frac{1}{\deg f} \lim_{N \to \infty} 4^{-N} h_f([2^N]P)$$

exists and is independent of f.

Proof. The strategy is to show that $4^{-N}h_f([2^N]P)$ is a Cauchy sequence. Corollary 8.24(b) (with $m = 2$) implies that there exists a constant C such that for all $Q \in E(\bar{K})$, we have

$$|h_f([2]Q) - 4h_f(Q)| \leq C. \tag{†}$$

Hence if $N \geq M \geq 0$, then

$$\begin{aligned} |4^{-N}h_f([2^N]P) - 4^{-M}h_f([2^M]P)| &= \left| \sum_{n=M}^{N-1} \left\{ 4^{-n-1}h_f([2^{n+1}]P) - 4^{-n}h_f([2^n]P) \right\} \right| \\ &\leq \sum_{n=M}^{N-1} 4^{-n-1} |h_f([2^{n+1}]P) - 4h_f([2^n]P)| \\ &\leq \sum_{n=M}^{N-1} 4^{-n-1} C \qquad \text{(by (†) with } Q = [2^n]P) \\ &\leq \frac{C}{4^M}. \end{aligned}$$

Hence the sequence is Cauchy and so converges.

Suppose now that $g \in K(E)$ is another nonconstant even function. Then, from the proof of Theorem 8.23, we have

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

Hence

$$(\deg g)4^{-N}h_f([2^N]P) - (\deg f)4^{-N}h_g([2^N]P) = 4^{-N}O(1) \to 0$$

as $N \to \infty$. Therefore the limit is independent of the choice of f, as claimed.

Definition 8.27 The canonical (or Néron-Tate) height $\hat{h} : E(\bar{K}) \to \mathbb{R}$ is defined by

$$\hat{h}(P) = \frac{1}{\deg f} \lim_{N\to\infty} 4^{-N}h_f([2^N]P).$$

Theorem 8.28 (Néron-Tate.) Let E/K be an elliptic curve.

(a) For all $P, Q \in E(\bar{K})$,

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

(the parallelogram law).

(b) For all $P \in E(\bar{K})$ and for all $m \in \mathbb{Z}$, $\hat{h}([m]P) = m^2\hat{h}(P)$.

(c) $\hat{h}$ is a quadratic form on E, i.e. $\hat{h}$ is even, and the pairing $\langle\ ,\ \rangle : E(\bar{K}) \times E(\bar{K}) \to \mathbb{R}$ given by $\langle P, Q\rangle = \hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)$ is bilinear.

(d) Suppose that $P \in E(\bar{K})$. Then $\hat{h}(P) \geq 0$, and $\hat{h}(P) = 0$ iff $P \in E(\bar{K})_{\text{tors}}$.

(e) Suppose that $f \in K(E)$ is a nonconstant even function. Then $(\deg f)\hat{h} = h_f + O_{E,f}(1)$.

(f) If $\hat{h}' : E(\bar{K}) \to \mathbb{R}$ is any other function which satisfies (e) for some nonconstant function f and (b) for any one integer $m \geq 2$, then $\hat{h}' = \hat{h}$.

Proof.

(e) From the proof of Proposition 8.26, there exists a constant C such that $N \geq M \geq 0$, and for all $P \in E(\bar{K})$ we have

$$|4^{-N} h_f([2^N]P) - 4^{-M} h_f([2^M]P)| \leq \frac{C}{4^M}.$$

Set $M = 0$ and let $N \to \infty$ to obtain

$$|(\deg f)\hat{h}(P) - h_f(P)| \leq C,$$

as desired.

(a) Theorem 8.23 implies that

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

Replace P by $[2^n]P$ and Q by $[2^n]Q$; multiply through by $\frac{1}{\deg f} 4^{-N}$, and let $N \to \infty$. This yields

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

(the $O(1)$ disappears).

(b) Corollary 8.24(b) implies that $h_f([m]P) = m^2 hf(P) + O(1)$. Replace P by $[2^N]P$, multiply through by $\frac{1}{\deg f} 4^{-N}$, and let $N \to \infty$. This gives $\hat{h}([m]P) = m^2 \hat{h}(P)$.

(c) (Linear algebra: *Any* function satisfying the parallelogram law is quadratic.) Setting $P = 0$ in the parallelogram law yields $\hat{h}(Q) = \hat{h}(-Q)$, so $\hat{h}$ is even. It suffices to prove that $\langle P + Q, R\rangle = \langle P, R\rangle + \langle Q, R\rangle$. Now we have (using the parallelogram law and the fact that $\hat{h}$ is even):

$$\hat{h}(P + Q + R) + \hat{h}(P + R - Q) - 2\hat{h}(P + R) - 2\hat{h}(Q) = 0, \tag{1}$$
$$\hat{h}(P - R + Q) + \hat{h}(P + R - Q) - 2\hat{h}(P) - 2\hat{h}(R - Q) = 0, \tag{2}$$
$$\hat{h}(P - R + Q) + \hat{h}(P + R + Q) - 2\hat{h}(P + Q) - 2\hat{h}(R) = 0, \tag{3}$$
$$2\hat{h}(R + Q) + 2\hat{h}(R - Q) - 4\hat{h}(R) - 4\hat{h}(Q) = 0. \tag{4}$$

Then (1) − (2) + (3) − (4) implies the result.

(d) Plainly $h_f(P) \geq 0$, so $\hat{h}(P) \geq 0$ for all $P \in E(\bar{K})$. Suppose that $P \in E(\bar{K})_{\text{tors}}$. Then $[m]P = 0$ for some $m \geq 1$, and now (b) gives

$$\hat{h}(P) = m^{-2}\hat{h}([m]P) = m^{-2}\hat{h}(O) = 0.$$

Suppose conversely that $P \in E(K')$ (K'/K a finite extension) with $\hat{h}(P) = 0$. Then, for every $m \in \mathbb{Z}$, we have (from (b)) $\hat{h}([m]P) = m^{-2}\hat{h}(P) = 0$. Now (e) implies that there exists a constant C such that for each $m \in \mathbb{Z}$, we have

$$h_f([m]P) = |\deg(f)\hat{h}([m]P) - h_f([m]P)| \leq C.$$

So $\{P, 2P, 3P, \ldots\} \subseteq \{Q \in E(K') \mid h_f(Q) \leq C\}$, and therefore P has finite order since this last set is finite.

(f) Suppose that $\hat{h}'$ satisfies $\hat{h}' \circ [m] = m^2\hat{h}'$ and $(\deg f)\hat{h}' = h_f + O(1)$ for some $m \geq 2$. Then $\hat{h}' \circ [m^N] = m^{2N}\hat{h}'$ *and*

$$\begin{aligned} \hat{h}' &= m^{-2N}\hat{h}' \circ [m^N] \\ &= m^{-2N}(\hat{h} \circ [m^N] + O(1)) \\ &= \hat{h} + m^{-2N}O(1) \end{aligned}$$

(since $\hat{h}$ satisfies (b)). Now let $N \to \infty$ to obtain $\hat{h}' = \hat{h}$.

Lemma 8.29 Suppose that V is a finite-dimensional $\mathbb{R}$-vector space, and let $L \subset V$ be a lattice. Let $q : V \to \mathbb{R}$ be a positive definite quadratic form satisfying:

(i) If $P \in L$, then $q(P) = 0$ iff $P = 0$.

(ii) For every constant C, $\#\{P \in L \mid q(P) \leq C\} < \infty$. Then q is positive definite on V.

Proof. We may choose a basis of V such that for any $X = (x_1, \ldots, x_n) \in V$, we have

$$q(X) = \sum_{i=1}^{s} x_i^2 - \sum_{i=1}^{t} x_{s+i}^2,$$

where $s + t \leq n = \dim(V)$. We may view $V \simeq \mathbb{R}^n$ via this choice of basis. Suppose that $s \neq n$. Let λ be the length of the shortest vector in L, i.e.

$$\lambda = \inf\{q(P) \mid P \in L, P \neq 0\}.$$

Then (i) and (ii) imply that $\lambda > 0$. Now consider the set

$$B(\delta) := \left\{ (x_1, \ldots, x_n) \in \mathbb{R}^n \;\middle|\; x_1^2 + \cdots + x_s^2 \leq \frac{\lambda}{2},\ x_{s+1}^2 + \cdots + x_t^2 \leq \delta \right\}.$$

Then length (using q!) of any vector in $B(\delta)$ is at most $\lambda/2$, and so $B(\delta) \cap L = \{0\}$. Now $B(\delta)$ is compact, convex, and symmetric about the origin, and $\mathrm{Vol}(B(\delta)) \to \infty$ as $\delta \to \infty$. This contradicts Minkowski's convex body theorem.

Theorem 8.30 (Minkowski.) Let L be a lattice in $\mathbb{R}^n$ with fundamental parallelepiped D, and suppose that $B \subseteq \mathbb{R}^n$ is compact, convex, and symmetric about the origin. If $\mathrm{Vol}(B) \geq 2^n \,\mathrm{Vol}(D)$, then B contains a nonzero point of L.

Proof. We claim that if S is a measurable set in $\mathbb{R}^n$ with $\mathrm{Vol}(S) > \mathrm{Vol}(D)$, then S contains distinct points α and β with $\alpha, \beta \in L$.

Note that

$$\mathrm{Vol}(S) = \sum_{\ell \in L} \mathrm{Vol}(S \cap (D + \ell)),$$

D will contain a unique translate (by an element of L) of each set $S \cap (D + \ell)$. Since $\mathrm{Vol}(S) > \mathrm{Vol}(D)$, at least two of these sets will overlap, so there exist $\alpha, \beta \in S$ such that $\alpha - \lambda = \beta - \lambda'$ for distinct $\lambda, \lambda' \in L$, so $\alpha - \beta = \lambda - \lambda' \in L \setminus \{0\}$, as claimed.

Now take $S = \frac{1}{2}B = \left\{ \frac{x}{2} \;\middle|\; x \in B \right\}$. Then $\mathrm{Vol}(S) = \frac{1}{2^n}\mathrm{Vol}(B) > \mathrm{Vol}(D)$, so there exist $\alpha, \beta \in B$ such that $\frac{\alpha}{2} - \frac{\beta}{2} \in L$. Since B is symmetric about the origin, $-\beta \in B$. Since B is convex, $\frac{1}{2}(\alpha + (-\beta)) \in B$.

Theorem 8.31 The Néron-Tate height is a positive definite quadratic form on $R \otimes E(K)$.

Proof. Apply Lemma 8.29 to the lattice $E(K)/E(K)_{\mathrm{tors}}$ in $E(K) \otimes \mathbb{R}$.

Definition 8.32 The **Néron-Tate height pairing** on E/K is defined by $\langle\ ,\ \rangle : E(\bar{K}) \times E(\bar{K}) \to \mathbb{R}$, given by $\langle P, Q \rangle = \hat{h}(P + Q) - \hat{h}(P) - \hat{h}(Q)$.

Definition 8.33 The **elliptic regulator** $R_{E/K}$ is the volume of the fundamental domain of $E(K)/E(K)_{\text{tors}}$ with respect to $\hat{h}$, i.e. if $P_1, \dots, P_r \in E(K)$ form a basis of $E(K)/E(K)_{\text{tors}}$, then $R_{E/K} := \det(\langle P_i, P_j \rangle)$. (If $r = 0$, set $R_{E/K} = 1$.)

Corollary 8.34 $R_{E/K} > 0$.

So now we have: $E(K) \simeq E(K)_{\text{tors}} \times \mathbb{Z}^r$.

Conjecture 8.35 For any fixed K, r can be arbitrarily large.

Conjecture 8.36 Suppose K is a number field, and E/K is a number field. Then there exists a constant $c([K:\mathbb{Q}])$ such that for any point $P \in E(K)$ of infinite order, we have

$$\hat{h}(P) \geq c([K:\mathbb{Q}]) \max\{1, h(j_E), \log |N_{K/\mathbb{Q}}(\mathscr{D}_{E/K})|\},$$

where $\mathscr{D}_{E/K}$ is the minimal discriminant of E/K.

Theorem 8.37 (Cassels.) Suppose K is a local field with $\operatorname{char}(K) = 0$, $\operatorname{char}(k) = p > 0$, and let E/K be an elliptic curve with Weierstraß equation

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6, \qquad a_i \in \mathfrak{o}_K. \tag{$\dagger$}$$

Let $P \in E(K)$ have exact order $m \geq 2$.

(a) If $m \neq p^n$ for some n, then $x(P), y(P) \in \mathfrak{o}_K$.

(b) If $m = p^n$, then $\pi^{2r}x(P), \pi^{3r}y(P) \in \mathfrak{o}_K$, with

$$r = \left\lfloor \frac{v(p)}{p^n - p^{n-1}} \right\rfloor.$$

Proof. First observe that $x(P) \in \mathfrak{o}_K$ implies $y(P) \in \mathfrak{o}_K$, so in this case, there is nothing to prove. Thus $v(x(P)) < 0$. Without loss of generality, we may assume that the Weierstraß equation for E is minimal, for if (x', y') are coordinates for a minimal Weierstraß equation, then $v(x(P)) \geq v(x'(P))$ and $v(y(P)) \geq v(y'(P))$.

(a) (†) implies that $3v(x(P)) = 2v(y(P)) = -6s$ (some integer $s \geq 2$). Also $v(x(P)) > 0$ implies that $P \in E_1(K)$ (the kernel of reduction), so $P \leftrightarrow -x(P)/y(P) \in \hat{E}(\mathfrak{m})$. But $\hat{E}(\mathfrak{m})$ contains no prime-to-p torsion, so (a) follows.

(b) This follows from a general property of formal groups (see Silverman, Ch. IV, Theorem 6.1): if $-x(P)/y(P)$ has exact order p^n in $\hat{E}(\mathfrak{m})$, then

$$s = v\left(\frac{-x(P)}{y(P)}\right) \leq \frac{v(P)}{p^n - p^{n-1}}.$$

Thus $\pi^{2s}x(P), \pi^{3s}y(P) \in \mathfrak{o}_K$, implying the result.

Theorem 8.38 Suppose that K is a number field, and let E/K be an elliptic curve with Weierstraß equation

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6,\ a_i \in \mathfrak{o}_K \text{ for all } i.$$

Let $P \in E(K)$ be a point of exact order $m \geq 2$.

(a) If m is not a prime power, then $x(P), y(P) \in \mathfrak{o}_K$.

(b) Suppose $m = p^n$ for some prime p. For each finite place v of K, set

$$r_v := \left\lfloor \frac{\mathrm{ord}_v(p)}{p^n - p^{n-1}} \right\rfloor.$$

Then $\mathrm{ord}_v(x(P)) \geq -2r_v$ and $\mathrm{ord}_v(y(P)) \geq -3r_v$, and so if $\mathrm{ord}_v(p) = 0$, then $x(P)$ and $y(P)$ are v-integral.

Theorem 8.39 (Nagell-Lutz.) Suppose $E/\mathbb{Q}$ is an elliptic curve with Weierstraß equation

$$E : y^2 = x^3 + Ax + B, \qquad A, B \in \mathbb{Z}.$$

Let $P \in E(\mathbb{Q})_{\mathrm{tors}}$, $P \neq O$.

(a) We have $x(P), y(P) \in \mathbb{Z}$.

(b) Either $2P = O$ or $y(P)^2 \mid (4A^3 + 27B^2)$.

Proof.

(a) Set m to be the exact order of P. If $m = 2$, then $y(P) = 0$ (chord-tangent method!), so $x(P)$ is integral, so $x(P) \in \mathbb{Z}$ since $P \in E(\mathbb{Q})$. If $m > 2$, the result follows from Theorem 8.38 ($r_v = 0$ for all v).

(b) Assume $2P \neq O$; then $y(P) \neq 0$. So applying (a) to P and $2P$, we have $x(P), y(P), x(2P) \in \mathbb{Z}$. Set $\phi(x) := x^4 - 2Ax^2 - 8Bx + A^2$ and $\psi(x) := x^3 + Ax + B$ so that $x(2P) = \frac{\phi(x(P))}{4\psi(x(P))}$ (duplication formula — see Silverman III, §2.3(d)) and $f(x)\phi(x) - g(x)\psi(x) = 4A^3 + 27B^2$, where $f(x) = 3x^2 + 4A$ and $g(x) = 3x^2 - 5Ax - 27B$. Then

$$y(P)^2[4f(x(P))x(2P) - g(x(P))] = 4A^3 + 27B^2 \tag{$*$}$$

(via the duplication formula). The result follows since all quantities in $(*)$ lie in $\mathbb{Z}$.

Chapter 9

Diophantine Approximation on Elliptic Curves

Proposition 9.1 (Dirichlet.) Suppose $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Then there exist infinitely many $p/q \in \mathbb{Q}$ such that

$$\left|\frac{p}{q} - \alpha\right| \leq \frac{1}{q^2}.$$

Proof. Let $Q \in \mathbb{Z}$ be large, and consider $\{q\alpha - \lfloor q\alpha \rfloor \mid q = 0, 1, 2, \ldots, Q\}$. Since α is irrational, the numbers in this set are distinct. There are $Q+1$ of them, and so the pigeonhole principle implies that there exist $0 \leq q_1 < q_2 \leq Q$ such that

$$|(q_1\alpha - \lfloor q_1\alpha \rfloor) - (q_2\alpha - \lfloor q_2\alpha \rfloor)| \leq \frac{1}{Q}.$$

Thus

$$\left|\frac{\lfloor q_2\alpha \rfloor - \lfloor q_1\alpha \rfloor}{q_2 - q_1} - \alpha\right| \leq \frac{1}{(q_2 - q_1)Q} \leq \frac{1}{(q_2 - q_1)^2}.$$

The result now follows, since Q may be taken to be arbitrarily large.

[**Hurwitz**: $\frac{1}{q^2}$ may be replaced by $\frac{1}{\sqrt{5}q^2}$, and this is the best possible.]

Proposition 9.2 (Liouville.) Suppose $\alpha \in \bar{\mathbb{Q}}$ with $[\mathbb{Q}(\alpha) : \mathbb{Q}] = d \geq 2$. Then there exists a constant $C = C(\alpha) > 0$ such that for all $\frac{p}{q} \in \mathbb{Q}$, we have

$$\left|\frac{p}{q} - \alpha\right| \geq \frac{C}{q^d}.$$

Proof. Let $f(T) = a_d T^d + \cdots + a_0 \in \mathbb{Z}[T]$ be the minimal polynomial for α, and set $C_1 = \sup\{f'(t) \mid \alpha - 1 \leq t \leq \alpha + 1\}$. Suppose $\left|\frac{p}{q} - \alpha\right| \leq 1$. Then

$$\left|f\left(\frac{p}{q}\right)\right| = \left|f\left(\frac{p}{q}\right) - f(\alpha)\right| \leq C_1 \left|\frac{p}{q} - \alpha\right| \tag{$*$}$$

via the mean value theorem. Also $q^d f\left(\frac{p}{q}\right) \in \mathbb{Z}$, and certainly $f\left(\frac{p}{q}\right) \neq 0$ (since f can't have any rational roots). Thus

$$\left|q^d f\left(\frac{p}{q}\right)\right| \geq 1. \tag{$**$}$$

Thus $(*)$ and $(**)$ give

$$\left|\frac{p}{q} - \alpha\right| \geq \frac{C}{q^d},$$

where $C = \min\left\{\frac{1}{C_1}, 1\right\}$.

Question. If $d = 3$, what is the best possible exponent?

Definition 9.3 Suppose we are given a function $\tau : \mathbb{N} \to \mathbb{R}_{>0}$. We say that a number field K has **approximation exponent** τ if the following holds: Suppose $\alpha \in \bar{K}$ with $[K(\alpha) : K] = d$. Let v be an absolute value on K, extended to $K(\alpha)$. Then, for any constant C, there exist only finitely many $x \in K$ such that $|x - \alpha|_v < C \cdot H_K(x)^{-\tau(d)}$, where $H_K(x) = \prod_v \max\{1, |x|_v\}^{n_v}$ and $n_v := [K_v : \mathbb{Q}_v]$.

Liouville: $\mathbb{Q}$ has approximation exponent $d + \varepsilon$ for any $\varepsilon > 0$.

Theorem 9.4 (Roth.) For every $\varepsilon > 0$, every number field has approximation exponent $2+\varepsilon$.

Example. $x^3 - 5y^3 = \alpha$ for some fixed α. Suppose $x, y \in \mathbb{Z}$ is a solution, with $y \neq 0$. Then if $\zeta^3 = 1$, $\zeta \neq 1$, we have

$$\left(\frac{x}{y} - \sqrt[3]{5}\right)\left(\frac{x}{y} - \zeta\sqrt[3]{5}\right)\left(\frac{x}{y} - \zeta^2\sqrt[3]{5}\right) = \frac{\alpha}{y^3},$$

so

$$\left(\frac{x}{y} - \sqrt[3]{5}\right) = \frac{\alpha}{y^3}\left(\frac{x}{y} - \zeta\sqrt[3]{5}\right)^{-1}\left(\frac{x}{y} - \zeta^2\sqrt[3]{5}\right)^{-1},$$

so

$$\left|\frac{x}{y} - \sqrt[3]{5}\right| \leq \frac{C}{|y|^3},$$

for some constant C independent of x and y. Then Theorem 9.4 implies that there exist only finitely many possibilities for x and y. Thus the Diophantine equation $x^3 - 5y^3 = \alpha$ has only finitely many solutions with $x, y \in \mathbb{Z}$.

Definition 9.5 Suppose C/K is a curve with $P, Q \in C(K_v)$. Let $t_Q \in K_v(C)$ be a function with a zero of order $e \geq 1$ at Q. The v-adic **distance** $d_v(P, Q)$ from P to Q is $d_v(P, Q) = \min\{|t_Q(P)|_v^{1/e}, 1\}$. We sometimes write $d_v(P, t_Q)$ instead.

Proposition 9.6 Suppose $Q \in C(K_v)$, and let $t_Q, t'_Q \in K_v(C)$ be functions vanishing at Q. Then

$$\lim_{\substack{P \in C(K_1) \\ P \to Q}} \frac{\log d_v(P, t'_Q)}{\log d_v(P, t_Q)} = 1.$$

Proof. Suppose $\operatorname{ord}_Q(t_Q) = p$ and $\operatorname{ord}_Q(t'_Q) = e'$. Then $\phi := \frac{(t'_Q)^e}{(t_Q)^{e'}}$ has neither a zero nor a pole at Q, and so $|\phi(P)|_v$ is bounded away from 0 and ∞ as $P \to Q$. Hence as $P \to Q$, we have

$$\frac{\log d_v(P, t'_Q)}{\log d_v(P, t_Q)} = 1 + \frac{\log |\phi(P)|^{1/ee'}}{\log d_v(P, t_Q)} \to 1$$

as $P \to Q$.

Proposition 9.7 Suppose C_1/K and C_2/K are curves, and that $f : C_1 \to C_2$ is a finite map defined over K. Let $Q \in C_1(K_v)$, and set $e_f(Q)$ to be the ramification index of f at Q. Then

$$\lim_{\substack{P \in C_1(K_v)\\ P \to Q}} \frac{\log d_v(f(P), f(Q))}{\log d_v(P, Q)} = e_f(Q).$$

Proof. Let $t_Q \in K_v(C_1)$ and $t_{f(Q)} \in K_v(C_2)$ be uniformizers at Q and $f(Q)$, respectively. Then $t_{f(Q)} \circ f = t_Q^{e_f(Q)}\phi$, where $\phi \in K_v(C_1)$ has neither a zero nor a pole at Q. Thus $|\phi(P)|_v$ is bounded away from 0 and ∞ as $P \to Q$. Thus

$$\begin{aligned} \frac{\log d_v(f(P), f(Q))}{\log d_v(P, Q)} &= \frac{\log |t_{f(Q)}(f(P))|_v}{\log |t_Q(P)|_v} \\ &= \frac{e_f(Q) \log |t_Q(P)|_v + \log |\phi(P)|_v}{\log |t_Q(P)|_v} \\ &\to e_f(Q) \end{aligned}$$

as $P \to Q$.

Theorem 9.8 Suppose C/K is a curve, $f \in K(C)$ is a nonconstant function, and $Q \in C(\bar{K})$. Then

$$\varliminf_{\substack{P \in C(K)\\ P \to Q}} \frac{\log d_v(P, Q)}{\log H_K(f(P))} \geq -2.$$

["$P \to Q$" means $P \to Q$ with respect to the v-adic topology. If Q is not a v-adic limit point of $C(K)$, we define $\varliminf = 0$.]

Proof. Without loss of generality we may assume $f(Q) \neq \infty$. (Replace f by $1/f$ if necessary, and observe that $H_K\left(\frac{1}{f(P)}\right) = H_K(f(P))$.) Thus

$$d_v(P, Q) = \min\{|f(P) - f(Q)|_v^{1/e}, 1\},$$

where $e = \operatorname{ord}_Q(f - f(Q)) \geq 1$. Thus

$$\varliminf_{P \to Q} \frac{\log d_v(P,Q)}{\log H_K(f(P))} = \varliminf_{P \to Q} \frac{\log |f(P) - f(Q)|}{e \log H_K(f(P))}$$
$$= \frac{1}{e} \varliminf_{P \to Q} \left\{ \frac{\log\{H_K(f(P))^2 \cdot |f(P) - f(Q)|_v\}}{\log H_K(f(P))} - \tau \right\}.$$

Roth's Theorem (Theorem 9.4) implies that if $\tau = 2 + \varepsilon$, then we have $H_K(f(P))^T \cdot |f(P) - f(Q)|_v \geq 1$ for all but finitely many $P \in C(K)$. Hence

$$\varliminf_{P \to Q} \frac{\log d_v(P,Q)}{\log H_K(f(P))} \geq \frac{-\tau}{e} = \frac{-2 + \varepsilon}{e}.$$

This implies the desired result since $\varepsilon > 0$ is arbitrary and $e \geq 1$.

Theorem 9.9 (Siegel.) Let E/K be an elliptic curve with $E(K)$ infinite. Suppose $f \in E(K)$ is a nonconstant even function, $Q \in E(\bar{K})$, and $v \in M_K$. Then

$$\lim_{\substack{P \in E(K) \\ h_f(P) \to \infty}} \frac{\log d_v(P,Q)}{h_f(P)} = 0.$$

Proof. Let

$$L = \varliminf_{\substack{P \in E(K) \\ h_f(P) \to \infty}} \frac{\log d_v(P,Q)}{h_f(P)}.$$

Now $L \leq 0$ since $h_f(P) \geq 0$ and $d_v(P,Q) \leq 1$ for all P. Thus it suffices to prove that $L \geq 0$ to deduce that $L = 0$. Choose a sequence $\{P_i\} \subseteq E(K)$ (P_i's distinct) such that $\lim_{i \to \infty} \frac{\log d_v(P_i,Q)}{h_f(P_i)} = L$. Choose $m \in \mathbb{N}$ large. Then since $E(K)/mE(K)$ is finite, some coset of $E(K)/mE(K)$ contains infinitely many P_i. Thus passing to a subsequence and relabeling, we have $P_i = mP_i' + R$, where $R \in E(K)$ is independent of i and $P_i' \in E(K)$. Now

$$m^2 h_f(P_i') = h_f(mP_i') + O(1) = h_f(P_i - R) + O(1) \leq 2h_f(P_i) + O(1), \qquad (\dagger)$$

where the $O(1)$ term is independent of i, and may be taken to be positive.

First observe that if P_i is bounded away from Q (with respect to the v-adic topology), then $\log d_v(P_i, Q)$ is bounded, and so $L = 0$, and we're done.

Otherwise, by passing to a subsequence, we may assume that $P_i \in Q$ as $i \to \infty$. Then $mP'_i \to Q - R$, so $\{P'_i\}$ has an m^{th} root $Q' \in E(\bar{K})$, say, of $Q - R$ as a limit point. So, by passing to a subsequence again, we may assume that $P'_i \to Q'$, with $Q = mQ' + R$.

Now the map $E \to E$ given by $P \mapsto mP + R$ is unramified (Proposition 3.9(3)) everywhere. Thus Proposition 9.7 implies that

$$\lim_{i\to\infty} \frac{\log d_v(P_i, Q)}{\log d_v(P'_i, Q')} = 1. \tag{‡}$$

Combining (†) and (‡) gives

$$L = \lim_{i\to\infty} \frac{\log d_v(P_i, Q)}{h_f(P_i)} \geq \lim_{i\to\infty} \frac{\log d_v(P'_i, Q'_i)}{\frac{1}{2}m^2 h_f(P'_i) + O(1)}. \tag{§}$$

Now Theorem 9.8 implies that

$$\varliminf_{i\to\infty} \frac{\log d_v(P'_i, Q')}{\log H_K(f(P))} \geq -2,$$

i.e.

$$\varliminf_{i\to\infty} \frac{\log d_v(P'_i, Q')}{[K:\mathbb{Q}]h_f(P)} \geq -2. \tag{$*$}$$

Combining (§) and ($*$) gives

$$L \geq \frac{-4[K:\mathbb{Q}]}{m^2 + O(1)} \geq \frac{-4[K:\mathbb{Q}]}{m^2}.$$

Since m is arbitrary, it follows that $L \geq 0$.

Theorem 9.10 Suppose E/K is an elliptic curve with Weierstraß coordinate functions x and y. Let S be a finite set of places of K containing the infinite places of K. Set $\mathfrak{o}_{K,S} := \{x \in K \mid v(x) \geq 0 \text{ for all } v \notin S\}$. Then $\#\{P \in E(K) \mid x(P) \in \mathfrak{o}_{K,S}\} < \infty$.

Proof. We apply Theorem 9.9 with $f = x$. Suppose if possible that $\{P_i\}$ is an infinite sequence of distinct points in $E(K)$ with $x(P_i) \in \mathfrak{o}_{K,S}$ for all i. Then

$$h_x(P_i) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in S} \log(\max\{1, |x(P_i)|_v^{n_v}\})$$

(since $v \notin S$ implies that $|x(P_i)|_v \leq 1$). Hence, by passing to a subsequence if necessary, we may assume that $h_x(P_i) \leq |S| \cdot \log |x(P_i)|_v$ for all i (note $n_v \leq [K:\mathbb{Q}]$), where v is a *fixed* absolute value. So $|x(P_i)|_v \to \infty$ as $i \to \infty$ (there exist only finitely many points of bounded height). The only pole of x is O, so $d_v(P_i, O) \to 0$. x has a pole of order 2 at O, so we can take $d_v(P_i, O) = \min\{|x(P_i)|_v^{-1/2}, 1\}$ as our distance function. Thus for all $i \gg 0$, we have

$$\frac{-\log d_v(P_i, Q)}{h_x(P_i)} \geq \frac{1}{2|S|},$$

which is a contradiction since Theorem 9.9 implies that the left side tends to 0 as $i \to \infty$.

Corollary 9.11 Suppose that C/K is a curve of genus 1, and $f \in K(C)$ is any nonconstant function. Then $\{\{P \in C(K) \mid f(P) \in \mathfrak{o}_{K,S}\} < \infty$.

Proof. Without loss of generality we may extend K and enlarge S. Thus we may assume C contains a pole Q of f. So (C, Q) is an elliptic curve over K. Let x and y be Weierstraß coordinates for (C, Q) with $y^2 = x^3 + Ax + B$. Now $[K(x,y) : K(x)] = 2$, and if $f \in K(x,y) = K(C)$, then

$$f(x,y) = \frac{\phi(x) + \psi(x)y}{\eta(x)},$$

where $\phi, \psi, \eta \in K[x]$. Also $\mathrm{ord}_Q(x) = -2$, $\mathrm{ord}_Q(y) = -3$, and $\mathrm{ord}_Q(f) < 0$, so

$$2\deg(\eta) < \max\{2\deg\phi, 2\deg\psi + 3\}. \tag{$*$}$$

We claim that x satisfies a monic polynomial over $K[f]$:

$$(f\eta(x) - \phi(x))^2 = (\psi(x) - y)^2 = \psi(x)^2(x^3 + Ax + B).$$

Viewed as a polynomial in f, the highest power of x will come from one of the terms $f^2\eta(x)^2$, $\phi(x)^2$, or $\psi(x)^2x^3$. Now $(*)$ implies that $\deg(f^2\eta(x)^2) < \deg(\phi(x)^2)$ or $\deg(\psi(x)^2x^3)$, and $\deg(\phi(x)^2) \neq \deg(\psi(x)^2x^3)$. This implies that the leading terms of $\phi(x)^2$ and $\psi(x)^2x^3$ cannot cancel. So, clearing denominators, we have

$$a_nx^n + a_{n-1}(f)x^{n-1} + \cdots + a_1(f)x + a_0(f) = 0,$$

with $a_n \in \mathfrak{o}_{K,S}$ and $a_i(f) \in \mathfrak{o}_{K,S}[f]$ for $o \leq i \leq n-1$. Without loss of generality we may assume $a_n \in \mathfrak{o}_{K,S}^\times$ (by enlarging S if necessary). Suppose $P \in C(K)$ satisfies $f(P) \in \mathfrak{o}_{K,S}$. Then P is *not* a pole of f, and

$$a_nx(P)^n + a_{n-1}(f(P))x(P)^{n-1} + \cdots + a_1(f(P))x(P) + a_0(f(P)) = 0,$$

so $x(P)$ is integral over $\mathfrak{o}_{K,S}$. Thus $x(P) \in K$ and $\mathfrak{o}_{K,S}$ is integrally closed in K, so $x(P) \in \mathfrak{o}_{K,S}$. Thus

$$\{P \in C(K) \mid f(P) \in \mathfrak{o}_{K,S}\} \subseteq \{P \in C(K) \mid x(P) \in \mathfrak{o}_{K,S}\},$$

and now the result follows from Theorem 9.10.

Example. Consider $C : y^2 = x^3 + Ax + B$, $A, B \in \mathbb{Z}$, $4A^3 + 27B^2 \neq 0$. Theorem 9.10 implies that this equation has only finitely many solutions with $x, y \in \mathbb{Z}$. What does Theorem 9.9 (i.e. the strong form of Siegel's Theorem) give us? In Theorem 9.9, we take $Q = O$, $f = x$, and v the infinite place of $\mathbb{Q}$. Suppose that $C(\mathbb{Q})$ is infinite, with $\{P_i\} \subseteq C(\mathbb{Q})$ with $h(P_i) \leq h(P_{i+1})$. Write $x(P_i) = \frac{a_i}{b_i} \in \mathbb{Q}$, fractions in lowest terms. Recall that x has a pole of order 2 at O, so $1/x$ has a zero of order 2 at O. Thus

$$d_v(P_i, O) = \frac{1}{2}\log\min\left\{\left|\frac{b_i}{a_i}\right|, 1\right\}$$

and

$$h_x(P_i) = \log\max\{|a_i|, |b_i|\}.$$

Now Theorem 9.9 implies that

$$\lim_{i\to\infty}\frac{\min\left\{\log\left|\frac{b_i}{a_i}\right|, 0\right\}}{\max\{\log|a_i|, \log|b_i|} = 0. \tag{$*$}$$

Now let $Q_1 \in C(Q)$ be any point with $x(Q_1) = 0$. Then

$$\log d_v(P_i, Q_1) = \log\min\left\{\left|\frac{a_i}{b_i}\right|, 1\right\},$$

and now Theorem 9.9 gives

$$\lim_{i\to\infty} \frac{\min\left\{\log\left|\frac{a_i}{b_i}\right|, 0\right\}}{\max\{\log|a_i|, \log|b_i|} = 0. \tag{$**$}$$

$(*)$ and $(**)$ imply that

$$\lim_{i\to\infty} \frac{|\log|a_i| - \log|b_i||}{\max\{\log|a_i|, \log|b_i|\}} = 0,$$

so

$$\lim_{i\to\infty} \frac{\log|a_i|}{\log|b_i|} = 1.$$

The upshot of all this is that the numerators and denominators of the points P_i tend to have about the same number of digits as $i \to \infty$.

Theorem 9.12 Let S be a finite set of places of K, and suppose $a, b \in K^\times$. Then the equation

$$ax + by = 1 \tag{$\dagger$}$$

has only finitely many solutions x, y with $x, y \in \mathfrak{o}_{K,S}^\times$.

Proof. Choose $m \in \mathbb{N}$ to be large. Dirichlet's unit theorem implies that $\mathfrak{o}_{K,S}^\times/\mathfrak{o}_{K,S}^{\times m}$ is finite. Let $c_1, \ldots, c_r \in \mathfrak{o}_{K,S}^\times$ be a set of coset representatives. Then if $x, y \in \mathfrak{o}_{K,S}^\times$ is a solution to $(\dagger)$, we may write $x = c_i X^m$, $y = c_j Y^m$ (some $X, Y \in \mathfrak{o}_{K,S}^\times$), and so (X, Y) is a solution of $ac_i X^m + bc_j Y^m = 1$. Thus it suffices to prove that for any $\alpha, \beta \in K^\times$, the equation

$$\alpha X^m + \beta Y^m = 1 \tag{\S}$$

admits only finitely many solutions with $X, Y \in \mathfrak{o}_{K,S}^\times$. Now appeal to the fact that when m is large, the curve defined by (§) has genus greater than 1, and use Siegel's Theorem for curves of large genus.

This proof is ineffective.

Theorem 9.13 (Siegel.) Suppose $f(x) \in K[x]$ is of degree $d \geq 3$ and that $f(x)$ has distinct roots in $\bar{K}$. Then the equation $y^2 = f(x)$ has only finitely many solutions in $\mathfrak{o}_{K,S}$.

Proof. We are certainly at liberty to enlarge S and make a finite extension of K. So we may assume $f(x) = a(x - \alpha_1) \cdots (x - \alpha_d)$, $x_i \in K$, with

(i) $a \in \mathfrak{o}_{K,S}^{\times}$.

(ii) $\alpha_i - \alpha_j \in \mathfrak{o}_{K,S}^{\times}$ for $i \neq j$.

(iii) $\mathfrak{o}_{K,S}$ is a PID.

Thus suppose that $x, y \in \mathfrak{o}_{K,S}$ are such that $y^2 = f(x)$, and let $\mathfrak{p}$ be a prime ideal of $\mathfrak{o}_{K,S}$. (ii) implies that $\mathfrak{p}$ divides at most one $x - \alpha_i$ (since if $\mathfrak{p}$ divides $x - \alpha_i$ and $x - \alpha_j$, then $\mathfrak{p} \mid (\alpha_i - \alpha_j)$, which is a contradiction). (i) implies that $\mathfrak{p} \nmid a$. Thus $y^2 = a(x - \alpha_1) \cdots (x - \alpha_i)$ implies that $\operatorname{ord}_{\mathfrak{p}}(x - \alpha_i)$ is even. So $(x - \alpha_i)\mathfrak{o}_{K,S} = \mathfrak{a}^2$, say, but since $\mathfrak{o}_{K,S}$ is a PID, it follows that there exists $z_i \in \mathfrak{o}_{K,S}$ and $b_i \in \mathfrak{o}_{K,S}^{\times}$ such that $x - \alpha_i = b_i z_i^2$. Set $L := K(\sqrt{\mathfrak{o}_{K,S}^{\times}})$.

$$\begin{tikzcd}
L \arrow[d, dash] & T \arrow[d, dash] \\
K & S
\end{tikzcd}$$

Then $b_i = \beta_i^2$, $\beta_i \in \mathfrak{o}_{L,T}$, whence $x - \alpha_i = (\beta_i z_i)^2$. Therefore, taking the difference of any two of these equations gives:

$$(x - \alpha_i) - (x - \alpha_j) = \alpha_j - \alpha_i = (\beta_i z_i - \beta_j z_j)(\beta_i z_i + \beta_j z_j).$$

Now $\alpha_j - \alpha_i \in \mathfrak{o}_{L,T}^{\times}$ and $\beta_i z_i \pm \beta_j z_j \in \mathfrak{o}_{L,T}$. It therefore follows that in fact $\beta_i z_i \pm \beta_j z_j \in \mathfrak{o}_{L,T}^{\times}$ for all $i \neq j$.

We appeal to Siegel's Identity:

$$\frac{\beta_1 z_1 \pm \beta_2 z_2}{\beta_1 z_1 - \beta_3 z_3} \mp \frac{\beta_2 z_2 \pm \beta_3 z_3}{\beta_1 z_1 - \beta_3 z_3} = 1.$$

Thus Theorem 9.12 implies that there exist only finitely many possibilities for $\frac{\beta_1 z_1 + \beta_2 z_2}{\beta_1 z_1 - \beta_3 z_3}$ and $\frac{\beta_1 z_1 - \beta_2 z_2}{\beta_1 z_1 - \beta_3 z_3}$, so there exist only finitely many possibilities for $\frac{\alpha_2 - \alpha_1}{(\beta_1 z_1 - \beta_3 z_3)^2}$ (multiplying the above numbers together), so there exist only finite many possibilities for $\beta_1 z_1 - \beta_3 z_3$, so there exist only finitely many possibilities for

$$\beta_1 z_1 = \frac{1}{2}\left[(\beta_1 z_1 - \beta_3 z_3) + \frac{\alpha_3 - \alpha_1}{\beta_1 z_1 - \beta_3 z_3}\right],$$

so there exist only finitely many possibilities for $x = \alpha_1 + (\beta_1 z_1)^2$, so there exist only finitely many possibilities for y.

9.1 Effectivity

Theorem 9.14 (Gelfond-Schneider.) Suppose $\alpha, \beta \in \bar{\mathbb{Q}}$ with $\alpha \neq 0, 1$ and $\beta \notin \mathbb{Q}$. Then α^β is transcendental.

Aliter. If $\alpha_1, \alpha_2 \in \bar{\mathbb{Q}}^\times$ and if $\log \alpha_1$ and $\log \alpha_2$ are linearly independent over $\mathbb{Q}$, then they are linearly independent over $\bar{\mathbb{Q}}$, i.e. $\frac{\log \alpha_1}{\log \alpha_2}$ is either rational or transcendental.

Theorem 9.15 (Baker.) Suppose that $\alpha_1, \ldots, \alpha_n \in K^\times$ and $\beta_1, \ldots, \beta_n \in K$. For any constant κ, set

$$\tau(\kappa) := \tau(\kappa; \alpha_1, \ldots, \alpha_n; \beta_1, \ldots, \beta_n) = h([1, \beta_1, \ldots, \beta_n]) h([1, \alpha_1, \ldots, \alpha_n])^\kappa.$$

Suppose that $\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n \neq 0$. Then there exist effectively computable constants $C(n, [K : \mathbb{Q}])$ and $\kappa(n, [K : \mathbb{Q}]) > 0$ such that

$$|\beta_1 \log \alpha_1 + \cdots + \beta_n \log \alpha_n| > C^{-\tau(\kappa)}.$$

($K \hookrightarrow \mathbb{C}$ with absolute value $|\cdot|$.)

Lemma 9.16 Let V be a finite dimensional $\mathbb{R}$-vector space. Suppose $\mathbf{e} = (e_1, \ldots, e_n)$ is a basis of V, and define

$$\|x\|_{\mathbf{e}} = \left\|\sum_{i=1}^{n} x_i e_i\right\|_{\mathbf{e}}$$

(sup norm). Suppose $\mathbf{f} = (f_1, \ldots, f_n)$ is another basis of V. Then there exist constants $c_1, c_2 > 0$ (depending on $\mathbf{e}$ and $\mathbf{f}$) such that for all $x \in V$,

$$c_1 \|x\|_{\mathbf{e}} \leq \|x\|_{\mathbf{f}} \leq c_2 \|x\|_{\mathbf{e}}.$$

Proof. Let $A = (a_{ij})$ be such that $e_i = \sum_{j=1}^{n} a_{ij} f_j$ (change of basis matrix), and set $\|A\| = \max_{i,j} |a_{ij}|$. Then if $x = \sum_{i=1}^{n} x_i e_i \in V$, we have $x = \sum_{i,j=1}^{n} x_i a_{ij} f_j$, whence

$$\|x\|_{\mathbf{f}} = \max_j \left\{ \left| \sum_i x_i a_{ij} \right| \right\} \leq n \max_{i,j} \{|a_{ij}|\} \cdot \max_i \{|x_i|\} = n\|A\| \cdot \|x\|_{\mathbf{e}},$$

and the other equality follows by symmetry.

Application. Let S be a finite set of places of K. Assume S contains the infinite places, $s := |S|$, $\alpha_1, \ldots, \alpha_{s-1}$ form a basis of $\mathfrak{o}_{K,S}^{\times}/(\mathfrak{o}_{K,S}^{\times})_{\text{tors}}$. So, if $\alpha \in \mathfrak{o}_{K,S}^{\times}$, then $\alpha = \zeta \alpha_1^{m_1} \cdots \alpha_{s-1}^{m_{s-1}}$, where ζ is a root of unity. Define $m(\alpha) := \max_i \{|m_i|\}$.

Lemma 9.17 There exist constants $c_1, c_2 > 0$ (depending only on K and S) such that for all $\alpha \in \mathfrak{o}_{K,S}^{\times}$, we have $c_1 h(\alpha) \leq m(\alpha) \leq c_2 h(\alpha)$.

Proof. Suppose $S = \{v_1, \ldots, v_s\}$, and set $n_i := n_{v_i} = [K_{v_i} : \mathbb{Q}_{v_i}]$. Define $\rho_s : \mathfrak{o}_{K,S}^{\times} \to \mathbb{R}^s$ by $\alpha \mapsto (n_1 v_1(\alpha), \ldots, n_s v_s(\alpha))$. Then $\operatorname{Im}(\rho_s) \subseteq H = \{x_1 + \cdots + x_s = 0\}$ and $\operatorname{Im}(\rho_s)$ spans H. Let $\|\cdot\|_1$ be the sup norm on $\mathbb{R}^s$ with respect to the standard basis and $\|\cdot\|_2$ the sup norm on $\mathbb{R}^s$ with respect to the basis $\{\rho_s(\alpha_1), \ldots, \rho_s(\alpha_{s-1}), (1, \ldots, 1)\}$. Lemma 9.16 implies that there exist constants $c_1, c_2 > 0$ such that

$$c_1 \|x\|_1 \leq \|x\|_2 \leq c_2 \|x\|_1 \qquad \text{for all } x \in \mathbb{R}^s. \tag{$*$}$$

Now if $\alpha \in \mathfrak{o}_{K,S}^{\times}$ with

$$\rho_s(\alpha) = \sum_{i=1}^{n} m_i \rho_s(\alpha_i),$$

then

$$\|\rho_s(\alpha)\|_2 = \max\{|m_i|\} = m(\alpha),$$
$$\|\rho_s(\alpha)\|_1 = \max\{n_i |v_i(\alpha)|\},$$

and

$$h_x(\alpha) = \sum_i \max\{0, -n_i v_i(\alpha)\}.$$

- If $x = (x_1, \dots, x_s) \in H$, then

$$h(x) = \sum_i \max\{0, -x_i\} \leq \sum_i |x_i| \leq s\|x\|_1.$$

- $x_i = \max\{0, x_i\} - \max\{0, -x_i\}$.

Thus summing, and using $\sum x_i = 0$, gives $0 = h(-x) - h(x)$. Thus $h(x) = h(-x)$. So

$$\begin{aligned} 2h(x) &= h(x) + h(-x) \\ &= \sum_i (\max\{0, -x_i\} + \max\{0, x_i\}) \\ &= \sum_i |x_i| \\ &\geq \max\{|x_i|\} \\ &= \|x\|_1. \end{aligned}$$

Thus

$$\frac{1}{2}\|x\|_1 \leq h(x) \leq s\|x\|_1. \qquad (**)$$

Now combining $(*)$ and $(**)$ gives us what we want.

Theorem 9.18 Suppose $a, b \in K^\times$. Then there exists an effectively computable constant $C = C(K, S, a, b)$ such that any solution $\alpha, \beta \in \mathfrak{o}_{K,S}^\times$ of the S-unit equation $a\alpha + b\beta = 1$ satisfies $H(\alpha) < C$.

Proof. Set $s = |S|$. Suppose α, β is a solution, and let $v \in S$ be such that $|\alpha|_v$ is maximal. Then

$$|\alpha|_v^{[K:\mathbb{Q}]} \geq \prod_{w \in S} \max\{1, |\alpha|_w^{n_w}\} = H_K(\alpha), \qquad |\alpha_v \geq H_K(\alpha)^{1/s}. \qquad (1)$$

We make the simplifying assumption that v is archimedean. Now apply the Mean Value Theorem to $\log x$ to obtain

$$\left| \frac{\log x - \log y}{x - y} \right|_v \leq \frac{1}{\min\{|x|_v, |y|_v\}}.$$

Set $x = a\alpha$ and $y = -b\beta$. Then $x - y = 1$, and so

$$|\log(a\alpha) - \log(b\beta)|_v \le \min\{|a\alpha|_v, |a\alpha - 1|_v\}^{-1} \le 2(|\alpha| \cdot H(\alpha)^{1/s})^{-1} \tag{2}$$

from (1), and assuming $|\alpha| > 2|a|$ (otherwise we'd have a good bound on $H(\alpha)$). Choose a basis $\alpha_1, \ldots, \alpha_{s-1}$ of $\mathfrak{o}_{K,S}^\times/(\mathfrak{o}_{K,S}^\times)_{\text{tors}}$ and write $\alpha = \zeta\alpha_1^{m_1} \cdots \alpha_{s-1}^{m_{s-1}}$ and $\beta = \zeta'\alpha_1^{m_1'} \cdots \alpha_{s-1}^{m_{s-1}'}$ (where ζ and ζ' are roots of unity). Substituting into (2) yields

$$\left|\sum_i (m_i - m_i') \log \alpha_i + \log\left(\frac{a\zeta}{b\zeta'}\right)\right| \le c_1 H(\alpha)^{-1/s}, \tag{3}$$

where c_1 is an effectively computable constant depending only upon K, S, a, and b. Next observe that since $a\alpha + b\beta = 1$,

$$h(\alpha) = h\left(\frac{1}{a} - \frac{b}{a}\beta\right) \le h(\beta) + C,$$

so $|h(\alpha) - h(\beta)| \le c_2$, and we apply Lemma 9.17 to both α and β to obtain

$$c_3 m(\alpha) \le m(\beta) \le c_4 m(\alpha).$$

This in turn implies

$$|m_i - m_i'| \le m(\alpha) + m(\beta) \le c_5 h(\alpha). \tag{§}$$

Set $q_i := m_i - m_i'$ and $\gamma := a\zeta/b\zeta'$. Then (3) gives

$$|q_1 \log \alpha_1 + \cdots + q_{s-1} \log \alpha_{s-1} + \log \gamma| < c_1 H(\alpha)^{-1/s}, \tag{4}$$

where $\alpha_1, \ldots, \alpha_{s-1}$ and γ are fixed, and $q_i \in \mathbb{N}$ satisfies $|q_i| \le c_5 h(\alpha)$. Now apply Theorem 9.15 (i.e. Baker's Theorem). This implies that

$$|q_1 \log \alpha_1 + \cdots + q_{s-1} \log \alpha_{s-1} + \log \gamma| \ge c_6^{-\tau}, \tag{5}$$

where $\tau = h([1, q_1, \ldots, q_{s-1}])h([1, \alpha_1, \ldots, \alpha_{s-1}, \gamma])^\kappa$, where κ is a constant depending only upon K and s. Now (§) implies that

$$h([1, q_1, \ldots, q_{s-1}]) = \log \max\{1, |q_1|, \ldots, |q_{s-1}|\} \le \log(c_5 - h(\alpha)). \tag{6}$$

Thus (4), (5), and (6) give

$$c_7^{-\log(c_5 h(\alpha))} \le c_1 H(\alpha)^{-1/s},$$

so $H(\alpha) \leq c_8 h(\alpha)^{c_9}$, i.e. $H(\alpha) \leq c_{10} \log H(\alpha)$, so we have a bound on $H(\alpha)$.

Theorem 9.19 For any $a, b \in K^\times$, the equation

$$aX^m + bY^m = 1 \tag{§}$$

has only finitely many solutions $X, Y \in \mathfrak{o}_{K,S}^\times$ if m is large.

Proof. Suppose (§) has infinitely many solutions $X, Y \in \mathfrak{o}_{K,S}^\times$. The idea is to show that X/Y is too good an approximation to $(-b/a)^{1/m}$. Since S is finite, there exists some $w \in S$ such that (§) has infinitely many solutions $X, Y \in \mathfrak{o}_{K,S}^\times$ such that

$$|Y|_w^{n_w} = \max\{|Y|_v^{n_v} \mid v \in S\}.$$

Fix an m^{th} root α of $-b/a$. Then

$$\frac{1}{aY^m} = \frac{X^m}{Y^m} + \frac{b}{a} = \frac{X^m}{Y^m} - \alpha^m = \prod_{\zeta \in \mu_m} \left(\frac{X}{Y} - \zeta\alpha \right).$$

If Y is "large," then at least one of $\frac{X}{Y} - \zeta\alpha$ is "small."

We claim that "only one of $\frac{X}{Y} - \zeta\alpha$ can be small." For suppose $\zeta, \zeta' \in \mu_m$ are distinct. Then

$$\left| \frac{X}{Y} - \zeta\alpha \right|_w + \left| \frac{X}{Y} - \zeta'\alpha \right|_w \geq \|\zeta'\alpha - \zeta\alpha|_w \geq C_1(K, S, m). \tag{†}$$

Therefore

$$\frac{1}{|aY^m|_w} = \prod_{\zeta \in \mu_m} \left| \frac{X}{Y} - \zeta\alpha \right|_w \geq \left(\min_{\zeta \in \mu_m} \left| \frac{X}{Y} - \zeta\alpha \right|_w \right) \cdot \left(\frac{C_1}{2} \right)^{m-1}$$

(since (†) implies that all but one of the terms in the product must be at least $C_1/2$).

A consequence is that

$$\frac{1}{|Y^m|_w^{n_w}} \geq C_2(K, S, m) \min_{\zeta \in \mu_m} \left| \frac{X}{Y} - \zeta\alpha \right|_w^{n_w}.$$

Since μ_m is finite, there is some $\xi \in \mu_m$ such that there exist infinitely many solutions $X, Y \in \mathfrak{o}_{K,S}^{\times}$ of (§) such that

$$\frac{1}{|Y^m|_w^{n_w}} \geq C_2 \left| \frac{X}{Y} - \xi\alpha \right|_w^{n_2} \tag{‡}$$

(i.e. X/Y is a good approximation to $\xi\alpha$).

Recall that w was chosen to maximize $|Y|_w^{n_w}$. Hence (since $|Y|_v = 1$ for all $v \notin S$)

$$\begin{aligned} |Y|_w^{n_w} &= \max_{v \in S} |Y|_v^{n_v} \\ &\geq \left(\prod_{v \in S} |Y|_v^{n_v} \right)^{1/s} \qquad (s := \#S) \\ &= \left(\prod_{\text{all } v} |Y|_v^{n_v} \right)^{1/s} \\ &= H_K(Y)^{1/s} \end{aligned} \tag{§§}$$

Thus we can compute

$$\begin{aligned} H_K\left(\frac{X^m}{Y^m}\right) &= H_K\left(\frac{1}{aY^m} - \frac{b}{a}\right) \\ &\leq 2^{[K:\mathbb{Q}]} H_K\left(\frac{1}{aY^m}\right) H_K\left(\frac{b}{a}\right) \\ &\leq 2^{[K:\mathbb{Q}]} H_K\left(\frac{1}{Y^m}\right) H_K\left(\frac{1}{a}\right) H_K\left(\frac{b}{a}\right). \end{aligned}$$

Taking m^{th} roots yields

$$H_K\left(\frac{X}{Y}\right) \leq C_3(K,S,m) \cdot H_K\left(\frac{1}{Y}\right) = C_3(K,S,m) H_K(Y).$$

Now applying (§§) gives

$$|Y|_w^{n_w} \geq C_4(K,S,m) H_K\left(\frac{X}{Y}\right)^{1/s}.$$

Substituting this into (‡) yields

$$\frac{C_5}{H_K(X/Y)^{m/s}} \geq \left| \frac{X}{Y} - \xi\alpha \right|_w^{n_w},$$

and Roth's Theorem implies that this is only satisfied by finitely many X, Y if m is large.

Theorem 9.20 (Shafarevich.) Let K be a number field, and let S be a finite set of places of K, with $S_\infty \subseteq S$. Then, up to isomorphism over K, there are only finitely many elliptic curves E/K that have good reduction away from S (i.e. good reduction at all primes not in S).

Proof. Without loss of generality, we may assume

- S contains all primes above 2 and 3.
- $\mathrm{Cl}(\mathfrak{o}_{K,S}) = 1$.

Then we may write $E : y^2 = x^3 + Ax + B$ with $A, B \in \mathfrak{o}_{K,S}$ and $\Delta = -16(4A^3 + 27B^2)$. $\Delta\mathfrak{o}_{K,S} = \mathscr{D}_{E/K}\mathfrak{o}_{K,S}$, where $\mathscr{D}_{E/K}$ is the minimal discriminant of E/K, so $\Delta \in \mathfrak{o}_{K,S}^\times$ since E has good reduction away from S. Now suppose $E_1/K, E_2/K, \ldots$ is a sequence of elliptic curves, and that E_i/K has good reduction away from S. Let

$$E_i : y^2 = x^3 + A_i x + B_i, \qquad A_i, B_i \in \mathfrak{o}_{K,S}, \qquad \Delta_i = -16(4A_i^3 + 27B_i^2),$$
$$\Delta_i\mathfrak{o}_{K,S} = \mathscr{D}_{E_i/K}\mathfrak{o}_{K,S}, \qquad \Delta_i \in \mathfrak{o}_{K,S}^\times. \quad (\dagger)$$

By passing to a subsequence if necessary, we may assume that all of the Δ_i have the same image in the (finite!) group $\mathfrak{o}_{K,S}^\times/(\mathfrak{o}_{K,S}^\times)^{12}$, i.e. we may write

$$\Delta_i = CD_i^{12}, \qquad C \text{ fixed}, \qquad D_i \in \mathfrak{o}_{K,S}^\times. \qquad (\ddagger)$$

Now ($\dagger$) and ($\ddagger$) imply $CD_i^{12} = -16(4A_i^3 + 27B_i^2)$, so

$$C = \left(\frac{-4A_i}{D_i^4}\right)^3 - 3\left(\frac{12B_i}{D^6}\right)^2,$$

so

$$27C = \left(\frac{-12A_i}{D_i^4}\right)^3 - \left(\frac{108B_i}{D_i^6}\right)^2 = X^3 - Y^2,$$

so for each i, the point $\left(\frac{-12A_i}{D_i^4}, \frac{108B_i}{D_i^6}\right)$ is an S-integral point on the curve $Y^2 = X^3 - 27C$. Theorem 9.13 (Siegel's Theorem) implies that there are only finitely many such points, so there are only finitely many possibilities for A_i/D_i^4 and B_i/D_i^6. But if $\frac{A_i}{D_i^4} = \frac{A_j}{D_j^4}$ and $\frac{B_i}{D_i^6} = \frac{B_j}{D_j^6}$, then we have $E_i \xrightarrow{\sim} E_j$ given by $x \mapsto \left(\frac{D_i}{D_j}\right)^2 x'$, $y \mapsto \left(\frac{D_i}{D_j}\right)^3 y'$.

So the sequence contains only finitely many K-isomorphism classes of elliptic curves.

Corollary 9.21 Let E/K be a fixed elliptic curve. Then there are only finitely many elliptic curves E'/K that are K-isogenous to E.

Proof. Corollary 6.38 implies that if E and E' are K-isogenous, then they have the same set of primes of bad reduction. The result now follows from Theorem 9.20.

Corollary 9.22 (Serre.) Suppose that E/K is an elliptic curve without complex multiplication. Then for all but finitely many primes ℓ, the group $E[\ell]$ has no nontrivial $\mathrm{Gal}(\bar{K}/K)$-invariant subgroups, i.e. the representation $\rho_\ell : \mathrm{Gal}(\bar{K}/K) \to \mathrm{Aut}(E[\ell]) \simeq \mathrm{GL}_2(\mathbb{F}_\ell)$ is irreducible.

Proof. If $\Phi_\ell \subset E[\ell]$ is a nontrivial $\mathrm{Gal}(\bar{K}/K)$-invariant subgroup, then $\Phi_\ell \simeq \mathbb{Z}/\ell\mathbb{Z}$, since $E[\ell] \simeq (\mathbb{Z}/\ell\mathbb{Z})^2$. Theorem 3.11 implies that there exists an elliptic curve E_ℓ/K and a K-isogeny $\varphi_\ell : E \to E_\ell$ such that $\ker(\varphi_\ell) = \Phi_\ell$. Corollary 9.21 implies that the curves E_ℓ fall into finitely many isomorphism classes since each E_ℓ is isomorphic to E. Suppose then that $E_\ell \simeq E_{\ell'}$, and consider the following sequence of maps:

$$E \overset{\varphi_\ell}{\to} E_\ell \simeq E_{\ell'} \overset{\hat{\varphi}_{\ell'}}{E} .$$

This is an element of $\mathrm{End}(E)$ of degree $(\deg \varphi_\ell)(\deg \hat{\varphi}_{\ell'}) = \ell\ell'$. Since E does not have complex multiplication, $\mathrm{End}(E) \simeq \mathbb{Z}$, every element of $\mathrm{End}(E)$ has degree n^2, and so it follows that $\ell = \ell'$. So if $\ell \neq \ell'$, $E_\ell \not\simeq E_{\ell'}$, and therefore there are only finitely many primes for which Φ_ℓ can exist.

Conjecture 9.23 (Frey.) Let E/K be an elliptic curve. Then there are only finitely many pairs (E_i, p_i) consisting of

- An elliptic curve E_i/K which is *not* isogenous to E.
- A prime $p_i > 5$ such that $E[p_i] \simeq E_i[p_i]$ as $\mathrm{Gal}(\bar{K}/K)$-modules.

Definition 9.24 (Darmon.) Say that an integer n has the **isogeny property** (relative to a number field K) if the implication

$$E[n] \simeq E'[n] \text{ as } \mathrm{Gal}(\bar{K}/K)\text{-modules} \qquad \Rightarrow \qquad E \text{ is isogenous to } E' \qquad (*)$$

holds for all elliptic curves $E, E'/K$.

Conjecture 9.25 (Darmon.) Given any global field K, there exists a constant M_K such that all $n \geq M_K$ have the isogeny property.

Say that n satisfies the **weak isogeny property** if $(*)$ holds with at most finitely many exceptions.

Conjecture 9.26 (Darmon.) There exists an absolute constant M such that all $n \geq M$ have the weak isogeny property over all number fields K.

Chapter 10

Geometric Interpretation of Cohomology Groups

Basic Idea. Let E/K be an elliptic curve (K a number field, say). We have an exact sequence

$$0 \to \frac{E(K)}{nE(K)} \to H^1(K, E_n) \to H^1(K, E)_n \to 0.$$

We will try to understand these cohomology groups geometrically.

To a genus one curve C/K, we associate an elliptic curve $E/K = \mathrm{Jac}(C)$, the Jacobian of C.

There is a bijection $Ш(E/K) \leftrightarrow \{$curves C/K of genus 1 such that $\mathrm{Jac}(C) = E$, and the Hasse principle fails for $C/K\}$.

General Principle. $H^1(K, ?) \leftrightarrow$ 'Objects over K that become isomorphic over $\bar{K}$ to a fixed object with automorphism group "?".'

Definition 10.1 Let G be an abelian group. A (right) G-set P is called a G-**torsor** (or a **principal homogeneous space** for G) if $P \neq \varnothing$ and the map $P \times G \to P \times P$ given by $(p, g) \mapsto (p, p + g)$ is bijective (i.e. for every pair $P_1, P_2 \in P$, there exists a unique $g \in G$ such that $P_1 + g = P_2$).

Example. The addition map $G \times G \to G$ makes G a G-torsor (the trivial G-torsor).

Definition 10.2 A morphism $\varphi : P \to P'$ of G-torsors is just a map of G-sets.

Some basic properties:

(a) For any points $\pi \in P, \pi' \in P'$, there is a unique morphism $P \to P'$ such that $\varphi(\pi) = \pi'$.

(b) Every morphism $P \to P'$ is an isomorphism.

(c) For any point $\pi \in P$, there is a unique morphism $G \to P$ (of G-torsors) such that $0 \mapsto \pi$.

(d) Any element $g \in G$ defines an automorphism $\pi \mapsto \pi + g$ of P. Every automorphism of P is one of this form, for some $g \in G$.

Consequence. $\mathrm{Aut}(P) = G$ for any G-torsor P.

Definition 10.3 Let E/K be an elliptic curve. An E-torsor is a curve C/K together with a right action of E given by a regular map $C \times E \to C$ given by $(w, Q) \mapsto w + Q$ such that the map $C \times E \to C \times C$ given by $(w, Q) \mapsto (w, w + Q)$ is an isomorphism of algebraic varieties.

Consequence. For any extension L/K, $C(L) = \varnothing$ or $C(L)$ is an $E(L)$-torsor (as sets).

A **morphism** of E-torsors is a regular map $\varphi : C \to C'$ such that the following diagram commutes:

$$\begin{tikzcd}
C \times E \arrow[r] \arrow[d, "\varphi \times id_E"'] & C \arrow[d, "\varphi"] \\
C' \times E \arrow[r] & C'
\end{tikzcd}$$

All the statements made following Definition 10.2 hold in this setting.

Remark. If C is an E-torsor and $w \in C(K)$ is any point, then there is a unique morphism $E \to C$ (of E-torsors) such that $O \mapsto w$, and this morphism is an isomorphism. So C is trivial iff $C(K) \neq \varnothing$.

10.1 Classifying E-Torsors

Suppose that C is an E-torsor over K, and choose a point $w_0 \in C(\bar{K})$. For any $\sigma \in \mathrm{Gal}(\bar{K}/K)$, we have $\sigma(w_0) = w_0 + f(\sigma)$ ($f(\sigma) \in E(\bar{K})$ unique). Then

$$(\sigma\tau)(w_0) = \sigma(\tau(w_0)) = \sigma(w_0 + f(\tau)) = w_0 + f(\sigma) + \sigma f(\tau),$$

and $(\sigma\tau)(w_0) = w_0 + f(\sigma\tau)$ (from the definition of f). So

$$f(\sigma\tau) = f(\sigma) + \sigma f(\tau),$$

i.e. $f : \mathrm{Gal}(\bar{K}/K) \to E(\bar{K})$ is a 1-cocycle. w_0 has coordinates in a finite extension of K, so f is continuous.

Suppose we choose $w_1 \in C(\bar{K})$. Then $w_1 = w_0 + P$ for some $P \in E(\bar{K})$. Thus

$$\sigma(w_1) = \sigma(w_0 + P) = w_0 + f(\sigma) + \sigma(P) = w_1 + f(\sigma) + \sigma(P) - P,$$

so f and f_1 differ by a coboundary, so the cohomology class of f depends only upon C.

Suppose $[f] \in H^1(K, E)$ is zero. Then $f(\sigma) = \sigma(P) - P$ for some $P \in E(\bar{K})$. Then

$$\sigma(w_0 - P) = \sigma(w_0) - \sigma(P) = w_0 + \sigma(P) - P - \sigma(P) = w_0 - P,$$

so $w_0 - P \in C(K)$, so C is a trivial E-torsor.

Theorem 10.4 The map $\frac{\{E\text{-torsors}\}}{\simeq} \to H^1(K, E)$ given by $C \mapsto [f]$ is a bijection, sending the trivial E-torsor to the zero element.

Proof. We'll come back to this later, if ever.

Remark. Set $WC(E/K)$ to be equal to the set of isomorphism classes of E-torsors over K. The group structure on $WC(E/K)$ may be described concretely as follows: Suppose $C, C' \in WC(E/K)$. Define $C \wedge C'$ to be equal to the quotient by the diagonal action of E. So

$$(C \wedge C')(\bar{K}) = \frac{C(\bar{K}) \times C'(\bar{K})}{\sim},$$

where $(w, w') \sim (w + Q, w' + Q)$, $Q \in E(\bar{K})$. Then $C \wedge C'$ represents $C + C'$ in $WC(E/K)$.

10.2 Geometric Interpretation of $H^1(K, E_n)$

Definition 10.5 An n-**covering** is a pair (C, α) consisting of

- An E-torsor C.
- A regular map $\alpha : C \to E$ defined over K such that for some $w_1 \in C(\bar{K})$ we have $\alpha(w_1 + P) = [n]P$ for all $P \in E(\bar{K})$.

A **morphism** $(C, \alpha) \to (C', \alpha')$ of n-coverings is a morphism $\varphi : C \to C'$ of E-torsors such that $\alpha = \alpha'\varphi$.

For $\sigma \in \mathrm{Gal}(\bar{K}/K)$, we have $\sigma(w_1) = w_1 + f(\sigma)$, $f(\sigma) \in E(\bar{K})$.

Check that f is an $E(\bar{K})$-valued 1-cocycle.

We have $\alpha(\sigma(w_1)) = \alpha(w_1 + f(\sigma)) = [n]f(\sigma)$ and $\alpha(\sigma(w_1)) = \sigma(\alpha(w_1)) = \sigma(\alpha(w_1 + O)) = O$, so $[n]f(\sigma) = O$, i.e. $f(\sigma) \in E_n$. w_1 is unique up to translation by $Q \in E_n$, so $[f] \in H^1(K, E_n)$ is independent of w_1.

Theorem 10.6 The map $\{n\text{-coverings}\}/\simeq \to H^1(K, E_n)$ given by $(C, \alpha) \mapsto [f]$ is a bijection.

Proof. Write $WC(E_n/K)$ for the set of n-coverings of E modulo isomorphism, and consider the forgetful map $WC(E_n/K) \to WC(E/K)$ given by $(C, \alpha) \mapsto C$.

Exercise.

(a) Show that this map defines a surjection

$$WC(E_n/K) \to WC(E/K)_n. \tag{$\dagger$}$$

(b) Show that the fibers of ($\dagger$) are $E(K)/nE(K)$-torsors.

For example, if C is trivial, then there exists $w_0 \in C(K)$ with $\alpha(w_0) \in E(K)$. If $w_1 \in C(K)$, then $w_1 = w_0 + P$ for some $P \in E(K)$, so $\alpha(w_1) = \alpha(w_0 + P) = \alpha(w_0) + [n]P$, so $\alpha(w_0) \in E(K)/nE(K)$ is well-defined.

Now consider the following diagram:

$$\begin{tikzcd}
 & & WC(E_n/K) \arrow[r, "\beta"] \arrow[d, "\alpha"'] & WC(E/K)_n \arrow[d, "\sim"] & \\
0 \arrow[r] & \frac{E(K)}{nE(K)} \arrow[r] & H^1(K, E_n) \arrow[r, "\gamma"] & H^1(K, E)_n \arrow[r] & 0.
\end{tikzcd}$$

The diagram commutes, so α maps the fibers of β into the fibers of γ. These fibers are $E(K)/nE(K)$-torsors, so α is bijective on each fiber, so α is bijective on the entire set.

10.3 Twisting

Problem. Given an elliptic curve E/K, find all elliptic curves E'/K that become isomorphic to E over $\bar{K}$. (E' is called a **twist** of E.)

Example. Consider the elliptic curves $E : y^2 = f(x)$, $E_d : dy^2 = f(x)$. The change of variables $x \mapsto x$, $y \mapsto y\sqrt{d}$ show that $E \simeq E_d$ over $K(\sqrt{d})$.

In order to apply cohomology, we need to understand $\mathrm{Aut}(E, O)$.

Proposition 10.7 We have

$$\mathrm{Aut}_K(E, O) = \begin{cases} \mu_6(K) & \text{if } j(E) = 0, \\ \mu_4(K) & \text{if } j(E) = 1728, \\ \mu_2(K) & \text{if } j(E) \neq 0 \text{ or } 1728. \end{cases}$$

Proof. See Silverman III, §10.

Fix E/K, and let E'/K be an elliptic curve such that there is an isomorphism $\varphi : E \xrightarrow{\sim} E'$ over $\bar{K}$ If $\sigma \in \mathrm{Gal}(\bar{K}/K)$, then $\sigma\varphi := \sigma\varphi\sigma^{-1} : E \xrightarrow{\sim} E'$ is also an isomorphism over $\bar{K}$. We have $\sigma\varphi = \varphi \circ \alpha(\sigma)$, $\alpha(\sigma) \in \mathrm{Aut}_{\bar{K}}(E, O)$. Observe that

$$(\sigma\tau)\varphi = \sigma(\tau\varphi) = \sigma(\varphi \circ \alpha(\tau)) = \varphi \circ \alpha(\sigma) \cdot \sigma(\alpha\tau),$$

whence $\alpha(\sigma\tau) = \alpha(\sigma)\sigma(\alpha(\tau))$, i.e. $\alpha : \mathrm{Gal}(\bar{K}/K) \to \mathrm{Aut}_{\bar{K}}(E, O)$ is a 1-cocycle.

Check that choosing a different φ replaces α by its composite with a coboundary.

Theorem 10.8 The map

$$\frac{\{E'/K \text{ such that } E \simeq_{\bar{K}} E'\}}{\simeq} \to H^1(K, \mathrm{Aut}_{\bar{K}}(E, O))$$

is a bijection.

Corollary 10.9 If $j(E) \neq 0$ or 1728, then every twist of E is of the form E_d as in the example above.

Proof. $\mathrm{Aut}_{\bar{K}}(E, O) = \{\pm 1\} = \mu_2$, and $H^1(K, \mu_2) \simeq K^\times/(K^\times)^2$ under the correspondence in Theorem 10.8 given by $E_d \mapsto d \pmod{K^{\times 2}}$.

Remark. Set $\mathrm{Aut}(E)$ to be the group of *all* automorphisms of E (not necessarily preserving O). Then $E(K) \hookrightarrow \mathrm{Aut}(E)$, $Q \mapsto \tau_Q$ (translation by Q).

Claim. $\mathrm{Aut}(E) = E(K) \rtimes \mathrm{Aut}(E, O)$, i.e.

(a) $E(K) \lhd \mathrm{Aut}(E)$,

(b) $E(K) \cap \mathrm{Aut}(E, O) = \{0\}$,

(c) $\mathrm{Aut}(E) = E(K) \cdot \mathrm{Aut}(E, O)$.

Proof.

(a) Suppose $Q \in E(K)$ and $\gamma \in \mathrm{Aut}(E, O)$. Then for any $P \in E(\bar{K})$,

$$(\alpha \circ \tau_Q \circ \alpha^{-1})(P) = \alpha(\alpha^{-1}(P) + Q) = P + \alpha(Q) = \tau_{\alpha(Q)}(P),$$

and so $E(K) \lhd \mathrm{Aut}(E)$.

(b) Clear.

(c) Let $\gamma \in \mathrm{Aut}(E)$, and set $\gamma(O) = Q$. Then we have $\gamma \circ \tau_Q \circ (\tau_{-Q} \circ \gamma) = \gamma$, and $\tau_{-Q} \circ \gamma \in \mathrm{Aut}(E, O)$.

Theorem 10.10 Let C/K be a nonsingular projective curve of genus 1. Then there exists an elliptic curve E_0/K such that C is an E_0-torsor. The curve E_0 is unique up to K-isomorphism.

Proof. (Sketch.) There exists an isomorphism $\varphi : C \xrightarrow{\sim} E$ over $\bar{K}$, where $E/\bar{K}$ is an elliptic curve $E : y^2 = x^3 + ax + b$, $a, b \in \bar{K}$, $\Delta = 4a^3 + 27b^2 \neq 0$. For any $\sigma \in \mathrm{Gal}(\bar{K}/K)$, we have $\sigma\varphi : \sigma C = C \xrightarrow{\sim} \sigma E$, so $E \simeq C = \sigma(C) \simeq \sigma(E)$, so $j(E) = j(\sigma(E)) = \sigma(j(E))$, so $j(E) \in K$. Choose a curve E_0/K such that $j(E_0) = j(E)$. (Such a curve certainly exists — see Theorem 4.13.)

The problem is that E_0 might be the wrong curve. We fix this by twisting. Choose an isomorphism $\psi : E_0 \xrightarrow{\sim} C$ over $\bar{K}$. For $\sigma \in \mathrm{Gal}(\bar{K}/K)$, let

$$\begin{tikzcd} E_0 \arrow[r, "\psi", "\sim"'] & C \arrow[r, "\sigma", "\approx"'] & \sigma(C) = C \end{tikzcd}$$

be $\psi \circ \alpha(\sigma)$, where $\alpha(\sigma) \in \mathrm{Aut}_{\bar{K}}(E_0)$. Then $\sigma \mapsto \alpha(\sigma)$ is an $\mathrm{Aut}_{\bar{K}}(E_0)$-valued 1-cocycle of $\mathrm{Gal}(\bar{K}/K)$. This gives us some $[\alpha] \in H^1(K, \mathrm{Aut}_{\bar{K}}(E_0))$. The Remark implies that there is an exact sequence

$$1 \to E_0(\bar{K}) \to \mathrm{Aut}_{\bar{K}}(E_0) \to \mathrm{Aut}_{\bar{K}}(E, O) \to 1,$$

so

$$H^1(K, E_0) \to H^1(K, \mathrm{Aut}_{\bar{K}}(E_0)) \to H^1(K, \mathrm{Aut}_{\bar{K}}(E, O))$$

is exact, where the last map is defined by $[\alpha] \mapsto \widetilde{[\alpha]}$. If $\widetilde{[\alpha]} = 0$, then $[\alpha] \in H^1(K, E_0)$, and C is an E_0-torsor. If $\widetilde{[\alpha]} \neq 0$, we can twist E_0 by $\widetilde{[\alpha]}$ to obtain a new curve E_1. Check that $[\alpha] \in H^1(K, E_1)$, so C is an E_1 torsor.

Remark. $\mathrm{Aut}_{\bar{K}}(E)$ is noncommutative in general.

10.4 $H^1(G, M)$ for M Noncommutative

A 1-cocycle is a map $f : G \to M$ such that $f(\sigma\tau) = f(\sigma) \cdot \sigma(f(\tau))$ for all $\sigma, \tau \in G$. Say that two 1-cocycles f and g are equivalent if there exists an $m \in M$ such that $g(\sigma) = m^{-1} \cdot f(\sigma) \cdot \sigma(m)$. Define $H^1(G, M)$ to be the set of equivalence classes of 1-cocycles. This is a **pointed set**, with distinguished element $\sigma \mapsto 1$.