Rational points on curves

Henri Darmon

April 11, 2008

Contents

Introduction

Algebraic number theory is first and foremost the study of diophantine equations. Such a definition is arguably too narrow for a subject whose scope has expanded over the years to encompass an ever-growing list of fundamental notions: number fields and their class groups, abelian varieties, moduli spaces, Galois representations, p-divisible groups, modular forms, Shimura varieties, and L-functions, to name just a few. All of these subjects will be broached (sometimes too briefly, for reasons having less to do with their relative importance than with limitations of time, space, and the author's grasp of the subject) in this survey, which is devoted to the first nontrivial class of diophantine equations: those associated to varieties of dimension one, or *algebraic curves.*

The term *diophantine equation* refers to a system of polynomial equations

$$X : \begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ \qquad \vdots \\ f_m(x_1, \ldots, x_n) = 0 \end{cases} \quad (\text{with } f_i \in \mathbf{Z}[x_1, \ldots, x_n]). \qquad (1)$$

Given such a system, one wishes to understand (and, if possible, determine completely) its set of integer or rational solutions.

Little of the essential features of the question are lost, and much flexibility is gained, if one replaces the base ring $\mathbf{Z}$ by a more general ring $\mathcal{O}$. The prototypical examples are the ring of integers $\mathcal{O}_K$ of a number field K, or the ring $\mathcal{O}_{K,S}$ of its S-integers, for a suitable finite set S of primes of $\mathcal{O}_K$.

Fix such a base ring $\mathcal{O} = \mathcal{O}_{K,S}$ from now on, and assume that the polynomials in (1) have coefficients in $\mathcal{O}$.

If R is any $\mathcal{O}$-algebra, the set of solutions of (1) with coordinates in R is denoted $X(R)$:

$$X(R) := \{(x_1, \ldots, x_n) \in R^n \text{ satisfying } (1)\}.$$

The functor $R \mapsto X(R)$ from the category of $\mathcal{O}$-algebras to the category of sets is *representable*,

$$X(R) = \mathrm{Hom}_{\mathcal{O}}(A_X, R), \quad \text{where } A_X = \mathcal{O}[x_1, \ldots, x_n]/(f_1, \ldots, f_m). \quad (2)$$

In this way the system (1) determines the *affine scheme* $X := \mathrm{Spec}(A_X)$ over $\mathrm{Spec}(\mathcal{O})$.

When the polynomials in (1) are homogenous, it is customary to view X as giving rise to a *projective scheme* over $\mathcal{O}$. When R is a principal ideal domain, the set $X(R)$ is a subset of the set $\mathbb{P}_{n-1}(R)$ of n-tuples $(x_1, \ldots, x_n) \in R^n$ satisfying $Rx_1 + \cdots + Rx_n = R$, taken modulo the equivalence relation defined by

$$(x_1, \ldots, x_n) \sim (x'_1, \ldots, x'_n) \quad \text{if } x_i x'_j - x_j x'_i = 0, \qquad \forall \quad 1 \leq i, j \leq n.$$

Specifically,

$$X(R) := \{(x_1, \ldots, x_n) \in \mathbb{P}_{n-1}(R) \text{ satisfying } (1)\}.$$

In the projective setting, replacing the base ring $\mathcal{O}$ by its fraction field K, and X by its *generic fiber* X_K—a projective variety over K—does not change the diophantine problem. For instance, the natural map $X(\mathcal{O}) \longrightarrow X_K(K)$ is a bijection. So there is no distinction between the study of integral and rational points on a scheme whose generic fiber is a projective variety.

Here are some of the basic questions that can be asked about the behaviour of $X(\mathcal{O})$.

Question 1. *What is the cardinality of $X(\mathcal{O})$? Is it finite, or infinite?*

Question 2. *If $X(\mathcal{O})$ is finite, can its cardinality be bounded by a quantity depending in a simple way on X and $\mathcal{O}$?*

Question 3. *Can $X(\mathcal{O})$ be effectively determined?*

The arithmetic complexity of a point $P \in X(\mathcal{O})$—roughly speaking, the amount of space that would be required to store the coordinates of P on a computer—is measured by a (logarithmic) *height function*

$$h : X(\mathcal{O}) \longrightarrow \mathbf{R}.$$

The precise definitions and basic properties of heights are discussed elsewhere in this volume. Let us just mention that for any real $B > 0$, the number $N(X; B)$ of $P \in X(\mathcal{O})$ with $h(P) \leq B$ is finite, in any reasonable definition of h.

Question 4. *When $X(\mathcal{O})$ is infinite, what can be said about the asymptotics of the function $N(X; B)$ as $B \longrightarrow \infty$?*

A related question is concerned with the *equidistribution* properties of the points in $X(\mathcal{O})$ (ordered by increasing height), relative to some natural measure on $X(\mathbf{R})$ or $X(\mathbf{C})$.

An *algebraic curve* over $\mathcal{O}$ is a scheme X (either affine or projective) of relative dimension one over $\mathrm{Spec}(\mathcal{O})$. If its generic fiber is smooth, the set $X(\mathbf{C})$ (relative to a chosen embedding of $\mathcal{O}$ into $\mathbf{C}$, through which $\mathbf{C}$ becomes an $\mathcal{O}$-algebra) is a one-dimensional complex manifold. While a curve is often described by equations like (1), it is to be viewed up to isomorphism, as an equivalence class of such equations modulo suitable changes of variables. The main objects we will study are curves X over $\mathrm{Spec}(\mathcal{O})$, and the behaviour of the sets $X(R)$ as R ranges over different $\mathcal{O}$-algebras.

Remark. The term "integral points on elliptic curves" is often used (particularly by number theorists) to refer to to the integral solutions of an affine Weierstrass equation:
$$E_0 : y^2 = x^3 + ax + b$$

which describes an *affine curve* over the base ring $\mathbf{Z}[a, b]$. This is an abuse of terminology, since elliptic curves are always defined as projective varieties by passing to the projective equation

$$E : y^2 z = x^3 + axz^2 + bz^3,$$

resulting in the addition of the "point at infinity" $O := (0, 1, 0)$ to E_0. This passage is crucial. Note, for instance, that E has the structure of an algebraic group, while E_0 does not. It should be kept in mind that the common usage "integral points on E" refers to the integral points on the affine curve $E_0 = E - \{O\}$, which is not an elliptic curve at all, and that, according to the definitions in standard usage, $E(\mathcal{O})$ is equal to $E(K)$ because E is projective.

The fundamental trichotomy for curves
Suppose that the curve X is *generically smooth*, i.e., its generic fiber is a non-singular curve over K, so that $X(\mathbf{C})$ has the structure of a smooth Riemann surface. The set $X(\mathbf{C})$ is (topologically and analytically) identified with

$$X(\mathbf{C}) \simeq S - \{P_1, \ldots, P_s\},$$

where S is a compact Riemann surface (of genus g, say) and $P_1, \ldots, P_s$ are distinct points. The invariants g and s, which completely determine the topological isomorphism class of $X(\mathbf{C})$, can be packaged into the *Euler characteristic*

$$\chi(X) = 2 - 2g - s.$$

The answers to Questions 1—4 above depend on the sign of $\chi(X)$ in an essential way.

I. Positive Euler characteristic. If $\chi(X) > 0$, then $g = 0$ and $s = 0$ or 1. Therefore X is isomorphic over $\bar{K}$ either to the projective line $\mathbb{P}_1$ or the affine line $\mathbb{A}^1$. Forms of $\mathbb{P}_1$ over K correspond to conics, for which one has the following basic result.

Theorem 5. *Let X be a smooth conic over K. The following are equivalent.*

(a) *The curve X is isomorphic over K to $\mathbb{P}_1$.*

(b) *The set $X(K)$ is nonempty.*

(c) *The set $X(K_v)$ is nonempty, for all completions K_v of K.*

The equivalence between (a) and (b) follows from the Riemann–Roch theorem: given a rational point $\infty \in X(K)$, there is a rational function with only a simple pole at ∞; such a function gives an isomorphism between X and $\mathbb{P}_1$ over K. The equivalence between (a) and (c) is the Hasse–Minkowski theorem, one of the most basic instances of the so-called *local-global principle* which is discussed at greater length elsewhere in this volume.

Remark 6. The proof of the Hasse–Minkowski theorem, which relies on geometry of numbers, leads to an upper bound on the smallest height of a point on $X(K)$, and thus is effective. Attempts to generalise Theorem 5 to higher dimensional varieties have led to a rich theory which forms the basis for some of the articles in this volume.

The case of positive Euler characteristic, for which the basic questions 1—4 are in some sense well-understood thanks to Theorem 5, will not be treated any further in these notes.

II. Euler characteristic zero. There are two types of curve with Euler characteristic zero:

- The affine case: $g = 0$ and $s = 2$.

- The projective case: $g = 1$ and $s = 0$.

The prototypical example of the affine case is when

$$X = \mathbb{P}_1 - \{0, \infty\} = \mathbb{G}_m.$$

The set $X(\mathcal{O}) = \mathcal{O}^\times$ is an abelian group under multiplication, and X is naturally equipped with the structure of a commutative group scheme over $\mathcal{O}$. Something similar happens in the projective case: since X is a curve of genus one, it is isomorphic over K either to an elliptic curve, if $X(K) \neq \emptyset$, or to a principal homogeneous space over such a curve. For the following theorem, suppose that $X(\mathcal{O}) \neq \emptyset$, and that X can be equipped with the structure of a group scheme over $\mathcal{O}$.

Theorem 7. *The group $X(\mathcal{O})$ is finitely generated.*

In the affine case, Theorem 7 is essentially Dirichlet's S-unit theorem, while in the projective case it corresponds to the Mordell–Weil Theorem that the group of rational points on an elliptic curve over a number field is finitely generated.

III. Negative Euler characteristic. The theory of curves with negative Euler characteristic is dominated by the following basic finiteness result.

Theorem 8. *If $\chi(X) < 0$, then $X(\mathcal{O})$ is finite.*

In the affine case this is a theorem of Siegel proved in 1929, and can be reduced to the most interesting special case where $X = \mathbb{P}_1 - \{0, 1, \infty\}$. The points in $X(\mathcal{O})$ then correspond to solutions to the so-called *S-unit equation*

$$u + v = 1, \quad \text{with } u, v \in \mathcal{O}^{\times}.$$

In the projective case Theorem 8 used to be known as the *Mordell Conjecture*. Its proof by Faltings in 1983 represents a significant achievement in the diophantine theory of curves.

We now describe the contents of these notes.

Chapter 1 recalls some preliminary results that are used heavily in later chapters: the main finiteness results of algebraic number theory, and the method of descent based on unramified coverings and the Chevalley–Weil theorem. Hugo Chapdelaine's article [Chap] in these proceedings further develops these themes by describing a relatively elementary application of Falting's theorem to a diophantine equation—the *generalised Fermat equation $x^p + y^q + z^r = 0$*—that appears to fall somewhat beyond the scope of the study of algebraic curves, but to which, it turns out, the "fundamental trichotomy" described in this introduction can still be applied.

The main goal of Chapter 2 is to give a survey of Faltings' proof of the Mordell Conjecture. In many ways, this chapter forms the heart of these notes. The ideas in Chapter 2 are used to motivate the startlingly diverse array of techniques that arise in the diophantine study of curves. These techniques are deployed in subsequent chapters to study several important and illustrative classes of algebraic curves—specifically, modular curves, Fermat curves, and elliptic curves.

Chapter 3 focusses on what may appear at first glance to be a rather special collection of algebraic curves, the so-called *modular curves* over $\mathbf{Q}$ classifying isomorphism classes of elliptic curves with extra level structure. Singling out modular curves for careful study can be justified on (at least) two grounds.

1. They are the simplest examples of *moduli spaces*. Classifying the rational points on modular curves translates into "uniform boundedness" statements for the size of torsion subgroups of elliptic curves over $\mathbf{Q}$, and therefore leads to nontrivial results concerning rational points on curves of genus one.

2. Modular curves are also the simplest examples of *Shimura varieties*, and their jacobians and ℓ-adic cohomology are closely tied to spaces of modular forms. (It is from this connection that they derive their name.) This makes it feasible to address finer questions about the rational points on modular curves, following a line of attack that was initiated by Mazur [Ma1] in his landmark paper on the Eisenstein ideal.

Chapter 3 attempts to convey some of the flavour of Mazur's approach by describing a simple but illustrative special case of his general results: namely, his proof of the conjecture, originally due to Ogg, that the size of the torsion subgroup of elliptic curves over $\mathbf{Q}$ is uniformly bounded, by 14. The approach we describe incorporates an important strengthening due to Merel exploiting progress on the Birch and Swinnerton-Dyer conjecture that grew out of later work of Gross–Zagier and Kolyvagin–Logachev. Marusia Rebolledo's article [Reb] in these proceedings takes this development one step further by describing Merel's proof of the *strong uniform boundedness conjecture* over number fields: given $d \geq 1$, the modular curves $Y_1(p)$ contain no points of degree d when p is large enough (relative to d).

Chapter 4 describes the approach initiated by Frey, Serre, and Ribet for reducing Fermat's Last Theorem to deep questions about the relationship between elliptic curves and modular forms. This subject is only lightly touched upon in these notes. Pierre Charollois's article in this volume [Char] describes a technique of Halberstadt and Kraus that strengthens the "modular approach" to prove a result on the generalised Fermat equation $ax^p + by^p + cz^p = 0$ that is notable for its generality. This result also suggests that it might be profitable to view the modular approach as part of a general method, rather than just a serendipitous "trick" for proving Fermat's Last Theorem.

Chapter 5 gives a rapid summary of the author's second week of lectures at the Göttingen summer school, devoted largely to curves of genus 1, particularly elliptic curves. This chapter is less detailed than the others, partly because it covers topics that have already been treated elsewhere, notably in [Da2]. The main topics that are touched upon (albeit briefly) in Chapter 5 are:

1. The collection of Heegner points on a modular elliptic curve, and Kolyvagin's use of them to prove essentially all of the Birch and Swinnerton-Dyer conjecture for elliptic curves with analytic rank ≤ 1. Kolyvagin's techniques also supply a crucial ingredient in Merel's proof of the uniform boundedness conjecture, further justifying its inclusion as a topic

8

in the present notes. The article by Samit Dasgupta and John Voight [DV] in these proceedings describes an application of the theory of Heegner points to Sylvester's conjecture on the primes that can be expressed as a sum of two rational cubes.

2. Variants of the modular parametrisation which can be used to produce more general systems of algebraic points on elliptic curves over $\mathbf{Q}$. Such systems are likely to continue to play an important role in further progress on the Birch and Swinnerton-Dyer conjecture. A key example is the fact that many elliptic curves defined over totally real fields are expected to occur as factors of the jacobians of Shimura curves attached to certain quaternion algebras. The articles by John Voight [Vo] and Matthew Greenberg [Gre1] in these proceedings discuss the problem of calculating with Shimura curves and their associated parametrisations from two different angles: from the point of view of producing explicit equations in [Vo], and relying on the Cerednik–Drinfeld p-adic uniformisation in [Gre1].

3. The theory of Stark–Heegner points which is meant to generalise classical Heegner points. Matthew Greenberg's second article [Gre2] in these proceedings discusses Stark–Heegner points attached to elliptic curves over imaginary quadratic fields. Proving the existence and basic algebraicity properties of the points that Greenberg describes how to calculate numerically would lead to significant progress on the Birch and Swinnerton-Dyer conjecture—at present, there is no elliptic curve that is "genuinely" defined over a quadratic imaginary field for which this conjecture is proved in even its weakest form.

1 Preliminaries

1.1 Zero-dimensional varieties

In order to get a good understanding of algebraic varieties of dimension $d+1$, it is useful to understand the *totality* of algebraic varieties of dimension d. Such a principle is hardly surprising, since a $(d + 1)$-dimensional variety can be expressed as a family of d-dimensional varieties, parametrized by a one-dimensional base. Any discussion of the diophantine properties of curves must therefore necessarily begin with a mention of the zero-dimensional case.

A zero-dimensional variety (of finite type) over a field K is an affine scheme of the form $X = \mathrm{Spec}(R)$, where R is a finite-dimensional commutative K-algebra without nilpotent elements. Let

$$n := \#X(\bar{K}) = \#\mathrm{Hom}(R, \bar{K}) = \dim_K(R),$$

where $\bar{K}$ denotes as usual an algebraic closure of the field K. Finding the rational points on X amounts to solving a degree n polynomial in one variable over K.

An *integral model* of X over $\mathcal{O}$ is an affine scheme of the form $\mathrm{Spec}(R_\mathcal{O})$, where $R_\mathcal{O} \subset R$ is an $\mathcal{O}$-algebra satisfying $R_\mathcal{O} \otimes_\mathcal{O} K = R$. Such a model is said to be *smooth* if $R_\mathcal{O}$ is finitely generated as an $\mathcal{O}$-module and $R_\mathcal{O}/\mathfrak{p}$ is a ring without nilpotent elements for all $\mathfrak{p} \in \mathrm{Spec}(\mathcal{O})$. The reader can check that X has a smooth model over $\mathrm{Spec}(\mathcal{O})$ if and only if $R = \prod_i L_i$ is a product of field extensions L_i/K which are unramified outside of S.

It is of interest to consider the collection of zero-dimensional varieties of fixed cardinality n which possesses a smooth model over $\mathrm{Spec}(\mathcal{O})$. The following classical finiteness result is extremely useful in the study of curves.

Theorem 1.1 (Hermite–Minkowski). *Given n and $\mathcal{O} = \mathcal{O}_{K,S}$, there are finitely many isomorphism classes of varieties of cardinality n over K which possess a smooth model over $\mathrm{Spec}(\mathcal{O})$. Equivalently, there are finitely many field extensions of K of degree at most n which are unramified outside of S.*

The proof is explained, for example, in [Sz], p. 91. In the simplest special case where $K = \mathbf{Q}$ and $S = \emptyset$, we mention the following more precise statement:

Theorem 1.2 (Minkowski). *Any zero–dimensional variety over $\mathbf{Q}$ which has a smooth model over $\mathrm{Spec}(\mathbf{Z})$ is isomorphic to $\mathrm{Spec}(\mathbf{Q}^n)$ for some $n \geq 1$. Equivalently, there are no nontrivial everywhere unramified field extensions of $\mathbf{Q}$.*

1.2   Etale morphisms and the Chevalley–Weil theorem

If $\pi : X \longrightarrow Y$ is a nonconstant, finite morphism of projective curves defined over K (or of affine curves over $\mathcal{O} = \mathcal{O}_{K,S}$), then π induces finite-to-one maps $\pi_K : X(K) \longrightarrow Y(K)$ and $\pi_\mathcal{O} : X(\mathcal{O}) \longrightarrow Y(\mathcal{O})$. In particular, if $Y(K)$ is finite, then so is $X(K)$. This simple principle reduces the study of rational

points on a curve X to the often simpler study of points on the image curve Y. (For instance, the genus of Y is less than or equal to the genus of X, by the Riemann–Hurwitz formula.) As an historical illustration, Fermat proved that the equation $x^4 + y^4 = z^4$ (which corresponds to a projective curve of genus 3 over $\mathbf{Q}$) has no nontrivial rational points by studying the integer solutions of the auxiliary equation $x^4 + y^4 = z^2$ which are *primitive* in the sense of [Chap]. These primitive solutions correspond to rational points on a curve of genus one (in line with the principles explained in [Chap]), and Fermat is able to dispose of these rational points by his method of descent.

In contrast, the finiteness of $X(K)$ does not imply the finiteness of $Y(K)$ in general, because the maps π_K or $\pi_{\mathcal{O}}$ need not be surjective, and in fact are usually far from being so. The following weakening of the notion of surjectivity is frequently useful in practice.

Definition 1.3. The map $\pi : X \longrightarrow Y$ of curves over $\mathrm{Spec}(\mathcal{O}_{K,S})$ is said to be *almost surjective* if there is a finite extension L of K and a finite set T of primes of L containing the primes above those in S, such that $Y(\mathcal{O}_{K,S})$ is contained in the image of $\pi_{\mathcal{O}_{L,T}}$.

Definition 1.4. A morphism $\pi : X \longrightarrow Y$ of curves over $\mathrm{Spec}(\mathcal{O}_{K,S})$ is said to be *generically étale* if it satisfies any of the following equivalent conditions:

(a) The induced map $\pi_{\mathbf{C}} : X(\mathbf{C}) \longrightarrow Y(\mathbf{C})$ is an unramified covering of Riemann surfaces;

(b) The map $\pi_K : X_K \longrightarrow Y_K$ is an étale morphism of K-varieties on the generic fibers;

(c) There exists a finite set $S' \supset S$ of primes of K such that the map $\pi_{\mathcal{O}_{K,S'}} : X_{\mathcal{O}_{K,S'}} \to Y_{\mathcal{O}_{K,S'}}$ is a finite étale morphism of schemes over $\mathrm{Spec}(\mathcal{O}_{K,S'})$.

The following result, known as the *Chevalley–Weil theorem*, gives a criterion for a map π to be almost surjective.

Theorem 1.5 (Chevalley–Weil). *If the morphism π is generically étale, then it is almost surjective.*

Proof. Suppose that π is generically étale. By Property (c) in the definition, we may suitably enlarge S so that the map π becomes étale over $\mathrm{Spec}(\mathcal{O}_{K,S})$.

If P belongs to $Y(\mathcal{O}) = \text{Hom}(\text{Spec}(\mathcal{O}), Y)$, let $P^*(X) = \pi^{-1}(P)$ denote the fiber of π above P. This fiber can be described as a scheme over $\text{Spec}(\mathcal{O})$ by viewing P as a morphism $\text{Spec}(\mathcal{O}) \longrightarrow Y$, and $\pi^{-1}(P)$ as the scheme-theoretic pull-back of π to $\text{Spec}(\mathcal{O})$ via P, for which the following diagram is cartesian

$$
\begin{array}{ccc}
P^*(X) & \longrightarrow & X \\
\downarrow & & \downarrow \\
\text{Spec}(\mathcal{O}) & \xrightarrow{P} & Y.
\end{array}
$$

Note that $\pi^{-1}(P)$ is a zero-dimensional scheme over $\text{Spec}(\mathcal{O})$ of cardinality $n = \deg(\pi)$, which is smooth because π is étale. By the Hermite–Minkowski theorem (Theorem 1.1) there are finitely many possibilities for $\pi^{-1}(P)$, as P ranges over $Y(\mathcal{O})$. Hence the compositum L of their fraction fields is a finite extension of K. Let T denote the set of primes of L above those in S. Then, by construction, $Y(\mathcal{O})$ is contained in $\pi(X(\mathcal{O}_{L,T}))$. It follows that π is almost surjective. $\qquad\qquad\square$

Example 1.6. *The Klein and Fermat curves.* The quartic curve

$$Y : x^3 y + y^3 z + z^3 x = 0 \tag{3}$$

studied by Felix Klein is a curve of genus 3 having an automorphism group $G = \mathbf{PSL}_2(\mathbb{F}_7)$ of order 168. By the Hurwitz bound, this is the largest number of automorphisms a curve of genus 3 may have. (A curve with this property is in fact unique up to $\bar{\mathbf{Q}}$-isomorphism.) The curve Y is also a model for the *modular curve* $X(7)$. (Cf. Section 4.1 for a brief discussion of $X(n)$.) The automorphism group $\mathbf{PSL}_2(7)$ arises from the transformations that preserve the fibers of the natural projection of $Y(7)$ onto the j-line. In [Hur], Hurwitz proved that Y has no nontrivial rational points, as follows: let (x, y, z) be a point on the Klein quartic with integer coordinates, satisfying $\gcd(x, y, z) = 1$. Although x, y and z have no common factor, they need not be pairwise coprime; setting

$$u = \gcd(x, y), \quad v = \gcd(y, z), \quad w = \gcd(z, x),$$

one sees (after changing the signs of u, v, and/or w if necessary) that

$$(x, y, z) = (u^3 w, v^3 u, w^3 v) =: \pi(u, v, w). \tag{4}$$

Substituting back into the original equation (3) and dividing by $u^3v^3w^3$, one finds that (u, v, w) is a rational point on the Fermat curve of degree 7:

$$X : u^7 + v^7 + w^7 = 0.$$

Through this argument, Hurwitz shows that the map $\pi : X \longrightarrow Y$ given by (4), a generically étale map of degree 7, is almost surjective (in fact, surjective) on rational points. This is a simple special case of Theorem 1.5. Hurwitz then applies Lamé's result for the Fermat equation of degree 7 to conclude that the Klein quartic has no integer solutions except the trivial ones.

Note that this example gives a nontrivial diophantine relation between modular curves and Fermat curves. More sophisticated connections between these two classes of curves are discussed in Section 4.

Example 1.7. *Algebraic groups.* Recall that $\mathcal{O}$ is the ring of S-integers of a number field K. Let G be any commutative group scheme of finite type over $\mathrm{Spec}(\mathcal{O})$. Then for any integer $n \geq 1$, the morphism $[n]$ given by $g \mapsto g^n$ is generically étale (more precisely, étale over $\mathrm{Spec}(\mathcal{O}[1/n])$). Therefore, the Chevalley-Weil theorem implies that there is a *finite* extension L of K for which $G(\mathcal{O})/nG(\mathcal{O})$ maps to the kernel of the natural map $G(K)/nG(K) \longrightarrow G(L)/nG(L)$. A standard construction shows that this kernel injects into the finite group $H^1(\mathrm{Gal}(L/K), G[n](L))$, where $G[n](L)$ is the finite group of n-torsion-points on $G(L)$. It follows that $G(\mathcal{O})/nG(\mathcal{O})$ is finite. (When $G = \mathbb{G}_m$, this statement is a weak form of Dirichlet's S-unit theorem, while when $G = A$ is an elliptic curve or an abelian variety, it is the *weak Mordell–Weil theorem* asserting that $A(K)/nA(K)$ is finite.)

Example 1.8. It is not hard to exhibit a projective curve X of genus greater than 1 eqipped with a map $\pi : X \longrightarrow \mathbb{P}_1$ which is unramified outside $\{0, 1, \infty\}$. Examples include

(a) The Fermat curve $x^n + y^n = z^n$ with $\pi(x, y, z) = x^n/z^n$;

(b) The modular curves $X_0(n)$ and $X_1(n)$ introduced in Section 3.1, with their natural maps to the j-line.

One can use the map π to show that Theorem 8 for projective curves (Faltings' Theorem) implies the case $X = \mathbb{P}_1 - \{0, 1, \infty\}$ over $\mathrm{Spec}(\mathcal{O})$ of Theorem 8 (Siegel's Theorem).

More generally, a celebrated theorem of Belyi asserts that *any* projective curve X/K can be equipped with a morphism $\pi : X \longrightarrow \mathbb{P}_1$ which is unramified outside $\{0, 1, \infty\}$. (See Hugo Chapdelaine's article in these proceedings.) This fact has been exploited by Elkies [El] to prove that the abc conjecture implies Faltings' theorem.

Further topic: Hugo Chapdelaine's article in these proceedings explains how the discussion of unramified coverings and the Chevalley–Weil Theorem can be adapted to treat the primitive solutions of the *generalised Fermat equation* $x^p + y^q + z^r = 0$. The reader who has mastered the ideas in Section 1 may skip directly to this article if so inclined.

2 Faltings' theorem

This chapter is devoted to explaining the main ideas in Faltings' proof of the Mordell Conjecture (Theorem 8 for projective curves over number fields).

Theorem 2.1 (Faltings). *Let X be a smooth projective curve of genus ≥ 2 defined over a number field K. Then $X(K)$ is finite.*

The proof will be presented as a series of reductions.

2.1 Prelude: the Shafarevich problem

The first of these reductions, explained in Section 2.2, reduces Theorem 2.1 to a finiteness conjecture of Shafarevich. The Shafarevich problem is concerned with the collection of all arithmetic objects sharing certain common features and having "good reduction" over the ring $\mathcal{O}$ of S-integers of a number field K, taken, of course, up to isomorphism over K. Some key examples are:

1. the set $\mathcal{F}_d(\mathcal{O})$ of smooth zero-dimensional schemes over $\mathrm{Spec}(\mathcal{O})$ of cardinality d;

2. the set $\mathcal{M}_g(\mathcal{O})$ of smooth curves of genus g over $\mathrm{Spec}(\mathcal{O})$;

3. the set $\mathcal{A}_g(\mathcal{O})$ of abelian schemes of dimension g over $\mathrm{Spec}(\mathcal{O})$;

4. the set $\mathcal{I}_g(\mathcal{O})$ of K-isogeny classes of abelian varieties of dimension g over $\mathrm{Spec}(\mathcal{O})$.

The following question is known as the Shafarevich problem:

Question 2.2. *How large are the sets above? Are they finite?*

One can also ask what happens for specific values of K and S, the most interesting special case being $\mathcal{O} = \mathbf{Z}$ (i.e., $K = \mathbf{Q}$ and $S = \emptyset$).

We now discuss these questions for the various cases listed above:

1. The set $\mathcal{F}_d(\mathcal{O})$ corresponds to the set of étale K-algebras (i.e., products of seperable field extensions) of rank d over K which are unramified outside S. The finiteness of $\mathcal{F}_d(\mathcal{O})$ is just a restatement of the Hermite–Minkowski Theorem (Theorem 1.1).

2. The set $\mathcal{M}_0(\mathcal{O})$ consists of the set of smooth conics over K which have good reduction outside of S. It admits a cohomological interpretation, via the exact sequence

$$0 \longrightarrow \mathcal{M}_0(\mathcal{O}) \longrightarrow H^2(K, \pm 1) \longrightarrow \oplus_{v \notin S} H^2(K_v, \pm 1).$$

The fundamental results of local and global class field theory imply that $\mathcal{M}_0(\mathcal{O})$ is finite, and in fact, its order can be evaluated precisely:

$$\#\mathcal{M}_0(\mathcal{O}) = 2^{\#S + r - 1},$$

where r is the number of real places of K. In particular, when $K = \mathbf{Q}$ and $S = \emptyset$, then $\mathcal{M}_0(\mathbf{Z})$ consists of one element, corresponding to the projective line $\mathbb{P}_1$ over $\mathbf{Q}$.

3. The set $\mathcal{M}_1(\mathcal{O})$ can be infinite; in fact, an infinite set of curves of genus 1 which are all isomorphic over $\bar{K}$ and have good reduction outside of S can sometimes be found, even if S consists of just one prime of K. (See [Ma3], p. 241.) On the other hand, a deep conjecture of Shafarevich and Tate implies that $\mathcal{M}_1(\mathcal{O})$ is finite if S is empty. Also, if one replaces $\mathcal{M}_1$ by the set $\mathcal{E}$ of K-isomorphism classes of elliptic curves, i.e., curves of genus 1 equipped with a K-rational point, then Shafarevich [Sha] showed that $\mathcal{E}(\mathcal{O})$ is always finite.

When $g > 1$, the following conjecture of Shafarevich can be viewed as a one-dimensional analogue of the Hermite–Minkowski theorem (Theorem 1.1):

Conjecture 2.3. *Let $g \geq 2$ be an integer, and let $\mathcal{O}$ be the ring of S-integers of a number field K, for a finite set S of primes of K.*

(a) *(Shafarevich conjecture for curves). The set $\mathcal{M}_g(\mathcal{O})$ is finite, i.e., there are only finitely many K-isomorphism classes of curves of genus g defined over K and having good reduction outside of S.*

(b) *(Shafarevich conjecture for abelian varieties). The set $\mathcal{A}_g(\mathcal{O})$ is finite, i.e., there are only finitely many isomorphism classes of abelian varieties of dimension g defined over K and having good reduction outside of S.*

(c) *(Shafarevich conjecture for isogeny classes). The set $\mathcal{I}_g(\mathcal{O})$ is finite, i.e., there are only finitely many K-isogeny classes of abelian varieties of dimension g with good reduction outside of S.*

Remark 2.4. It is a deep theorem of Fontaine [Fo] that the sets $\mathcal{A}(\mathbf{Z})$ and $\mathcal{M}_g(\mathbf{Z})$ are *empty* for $g \geq 2$, i.e., there are no abelian varieties, or smooth curves of genus ≥ 2, over $\mathrm{Spec}(\mathbf{Z})$.

2.2   First reduction: the Kodaira–Parshin trick

In [Pa1], Parshin showed that part 2.3 of Conjecture 2.3 implies Theorem 2.1.

Theorem 2.5. *(Kodaira–Parshin). The Shafarevich conjecture for curves implies Mordell's conjecture.*

Sketch of proof. Let X be a curve of genus $g > 1$ defined over a number field K. To each point $P \in X(K)$ one associates a curve X_P and a covering map $\phi_P : X_P \longrightarrow X$ with the following properties:

1. The curve X_P and the map ϕ_P can be defined over a finite extension K' of K which does not depend on P.

2. The genus g' of X_P (and the degree of ϕ_P) is fixed and in particular does not depend on P.

3. The map ϕ_P is ramified only over the point P.

4. The curve X_P has good reduction outside a finite set of primes S' of K' which does not depend on P.

For a description of this assignment, see [Ma3], p. 243-244, [FW], p. 191-197, or [Pa1]. The reader should note that one has some leeway in constructing it, and that different versions appear in the literature.

We will describe one approach here, which consists in considering the embedding $X \longrightarrow J$ of X into its jacobian that sends P to the origin of J, and letting $\tilde{X}$ be the pullback to X of the multiplication-by-2 map $[2] : J \longrightarrow J$. This map induces an unramified covering $\pi : \tilde{X} \longrightarrow X$ of degree 2^{2g}, and hence the genus of $\tilde{X}$ can be calculated explicitly using the Riemann–Hurwitz formula. The fiber $\pi^{-1}(P)$ can be written as

$$\pi^{-1}(P) = \tilde{P} + D,$$

where $\tilde{P}$ corresponds to the identity element of J, and hence belongs to $\tilde{X}(K)$, and D is an effective divisor of degree $2^{2g} - 1$ defined over K with support disjoint from $\tilde{P}$. Let J_D be the *generalised jacobian* attached to $\tilde{X}$ and D: the group $J_D(\bar{K})$ is identified with the group of degree zero divisors on $\tilde{X}$ with support outside D, modulo the subgroup of principal divisors of the form $\mathrm{div}(f)$, as f ranges over the functions satisfying $f(D_0) = 1$, for all degree zero divisors D_0 supported on D. The functor $L \mapsto J_D(\bar{K})^{G_L}$ (where $G_L := \mathrm{Gal}(\bar{L}/L)$) on finite extensions of K is representable by the algebraic group over K denoted J_D, which is an extension of J by a torus T over K of rank $(2^{2g} - 2)$. In other words, there is a natural exact sequence

$$1 \longrightarrow T \longrightarrow J_D \longrightarrow J \longrightarrow 1$$

of commutative algebraic groups over K.

One can embed $\tilde{X} - D$ into J_D by sending a point Q to the equivalence class of the divisor $(Q) - (\tilde{P})$. The multiplication-by-2 map $[2]$ on J_D induces a map $X_P^0 \longrightarrow \tilde{X} - D$, as summarised by the following diagram with cartesian squares in which the vertical maps are induced by multiplication by 2:

$$
\begin{array}{ccc}
J_D & \longleftarrow & X_P^0 \\
\downarrow & & \downarrow \\
J_D & \longleftarrow & \tilde{X} - D & \longrightarrow & J \\
& & \downarrow & & \downarrow \\
& & X & \longrightarrow & J.
\end{array}
\tag{5}
$$

The closure X_P of X_P^0 has the desired properties 1-4: it is defined over K, and it follows directly from the Riemann–Hurwitz formula that its genus g' does not depend on P. Furthermore, the map $X_P^0 \longrightarrow \tilde{X} - D$ is unramified, and hence $X_P \longrightarrow X$ is ramified only over the point P. Finally, if X is smooth over $\mathrm{Spec}(\mathcal{O})$, the curve X_P has a smooth model over $\mathcal{O}' := \mathcal{O}[1/2]$.

The assignment $P \mapsto X_P$ therefore gives rise to a well-defined map

$$R_1 : X(K) \longrightarrow \mathcal{M}_{g'}(\mathcal{O}').$$

But this assignment is finite-to-one; for otherwise there would be a curve Y and infinitely many (by property 3) distinct maps $\phi_P : Y \longrightarrow X$. This would contradict the following geometric finiteness result of De Franchis (cf. [Ma3], p. 227).

Theorem 2.6. *If X and Y are curves over any field K, and Y has genus $g \geq 2$, then the set $\mathrm{Mor}_K(X, Y)$ of K-morphisms from X to Y is finite.*

18

The Shafarevich conjecture for curves, which asserts the finiteness of $\mathcal{M}_{g'}(\mathcal{O}')$, therefore implies the finiteness of $X(K)$. This completes the proof of Theorem 2.5. $\qquad\square$

Remark 2.7. The reader will note that the proof of Theorem 2.5 breaks down (as it should!) when $g = 1$, because the set $\mathrm{Mor}_K(Y, X)$ can be (and in fact, frequently is) infinite when X has genus 1.

2.3 Second reduction: passing to the jacobian

The second step in the proof of the Mordell conjecture consists in observing that the Shafarevich conjecture for curves would follow from the corresponding statement for abelian varieties.

Proposition 2.8. *The Shafarevich conjecture for curves follows from the Shafarevich conjecture for abelian varieties.*

To prove Proposition 2.8, one studies the map R_2 which associates to a curve X its jacobian J. If X is smooth over $\mathrm{Spec}(\mathcal{O})$, the same is true of J, and hence R_2 defines a map $\mathcal{M}_g(\mathcal{O}) \longrightarrow \mathcal{A}_g(\mathcal{O})$. Key to Proposition 2.8 is the following corollary of Torelli's theorem:

Theorem 2.9. *If $g \geq 2$, then the map R_2 is finite-to-one.*

Proof. Torelli's theorem asserts that a curve X of genus ≥ 2 can be recovered by the data of its jacobian J together with the principal polarisation associated to the Riemann theta-divisor. But a given abelian variety can carry only finitely many principal polarisations. (See [CS] for a more detailed exposition of the Torelli Theorem and surrounding concepts.) $\qquad\square$

2.4 Third reduction: passing to isogeny classes

The third, crucial and more difficult reduction was carried out by Faltings himself.

Theorem 2.10. *(Faltings). The Shafarevich conjecture for abelian varieties follows from the Shafarevich conjecture for isogeny classes.*

As one would expect, the proof is based on showing that the natural map $R_3 : \mathcal{A}_g(\mathcal{O}) \longrightarrow \mathcal{I}_g(\mathcal{O})$ has finite fibers. This is a consequence of the following key result:

Theorem 2.11. *(Faltings) There are finitely many isomorphism classes of abelian varieties over K in a given K-isogeny class.*

This result is the technical heart of Faltings' proof, and rests on his theory of heights on moduli spaces of abelian varieties. Things become somewhat simpler if we assume that the abelian varieties in the isogeny class are semistable. This can be assumed without loss of generality because of Grothendieck's semistable reduction theorem which asserts that every abelian variety becomes semistable after a finite extension of the ground field (for instance, one over which the points of order 3 become rational). For a finite extension K'/K, there are finitely many K-isomorphism classes of abelian varieties that are K'-isomorphic to a given abelian variety over K', and hence the finiteness of the K-isogeny class follows from that of any K'-isogeny class.

Faltings defines a height function (now called the Faltings height) of an abelian variety. We will not dwell on the definition, but will content ourselves with stating two of its main finiteness properties:

Theorem 2.12. *Let K be a number field and H be a positive constant. There are finitely many isomorphism classes of g-dimensional abelian varieties over K with height less than H.*

The second finiteness property concerns the behaviour of the Faltings height on a K-isogeny class. Given a prime ℓ, the ℓ-isogeny class of an abelian variety A is the set of abelian varieties which are isogenous to A via an isogeny of ℓ-power degree. More generally, if M is any finite set of rational primes, two abelian varieties are said to be M-isogenous if they are related by a K-isogeny whose degree is a product of primes in M.

Theorem 2.13. *If A is a semistable abelian variety over a number field K, then:*

1. *There exists a finite set M of rational primes, depending only on the isogeny class of A, such that if $A \longrightarrow B$ is a K-isogeny of degree not divisible by the primes in M, then*

$$h(A) = h(B).$$

2. *For any finite set S of rational primes, the Faltings height is bounded on S-isogeny classes.*

The proof of this theorem relies on deep results of Tate and Raynaud on group schemes and p-divisible groups; cf. Theorems 2.4 and 2.6 of [De].

For more details on the proof of theorems 2.12 and 2.13 see the expositions [CS], [FW], [Sz], [De], or [ZP]. Note that these two theorems together imply:

Proposition 2.14. *Let A be a semistable abelian variety over K, and let M be as in part 1 of Theorem 2.13.*

1. *Up to K-isomorphism, there are finitely many abelian varieties that are K-isogenous to A via an isogeny of degree not divisible by the primes in M.*

2. *Given any abelian variety B over K and any finite set S of rational primes, there are finitely many abelian varieties in the S-isogeny class of B.*

Proof of Theorem 2.11: Let $\phi : A \longrightarrow B$ be a K-isogeny. We can write ϕ as a composition of isogenies

$$A \xrightarrow{\phi_o} B_0 \xrightarrow{\phi_1} B_1,$$

where ϕ_0 is of degree not divisible by the primes in M, and ϕ_1 is an M-isogeny. By part 1 of Proposition 2.14, there are finitely many possibilities for ϕ_0 and for B_0. By part 2 of this proposition, for each B_0 there are finitely many possibilities for B_1. Theorem 2.11 follows.

2.5 Fourth reduction: from isogeny classes to ℓ-adic representations

To an abelian variety A over K of dimension g and a prime ℓ, one can associate the ℓ-adic Tate module and ℓ-adic representations

$$T_\ell(A) := \varprojlim A[\ell^n], \qquad V_\ell(A) := T_\ell(A) \otimes \mathbf{Q}_\ell,$$

where the inverse limit is taken with respect to the multiplication-by-ℓ maps. The $\mathbf{Q}_\ell$ vector space $V_\ell(A)$ is $2g$-dimensional and is equipped with a $\mathbf{Q}_\ell$-linear action by two *commuting* $\mathbf{Q}_\ell$-algebras E and Π_K defined by

$$E = \mathrm{End}_K(A) \otimes \mathbf{Q}_\ell, \qquad \Pi_K := \mathbf{Z}_\ell[[G_K]] \otimes \mathbf{Q}_\ell.$$

Here $\mathbf{Z}_\ell[[G_K]]$ denotes the profinite group ring $\varprojlim \mathbf{Z}_\ell[\mathrm{Gal}(L/K)]$, where the projective limit is taken over all finite Galois extensions $L \subset \bar{K}$ of K.

If A and B are K-isogenous abelian varieties, they give rise to ℓ-adic representations that are isomorphic as Π_K-modules. In other words, the assignment $A \mapsto V_\ell(A)$ yields a map

$$R_4 : \mathcal{I}_g(\mathcal{O}) \longrightarrow \left\{ \begin{array}{l} \text{Isomorphism classes of} \\ 2g\text{-dimensional } \ell\text{-adic} \\ \text{representations of } \Pi_K \end{array} \right\}.$$

The strategy will now consist in showing that R_4 has finite fibers, and finally in describing the image R_4 precisely enough to show that it is finite.

We begin by introducing some further notations and recalling some background. Given a prime v of K, let $I_v \subset G_v \subset G_K$ be inertia and decomposition subgroups of G_K attached to v. Note that the groups G_v and I_v are only well-defined up to conjugation in G_K, since they depend on a choice of a prime of $\bar{K}$ above v. The quotient G_v/I_v is procyclic with a canonical generator Frob_v called the *Frobenius element* at v, which induces the automorphism $x \mapsto x^{\mathbf{N}v}$ on the residue field, where $\mathbf{N}v$ denotes the norm of v (the cardinality of the associated residue field).

If V is any finite-dimensional $\mathbf{Q}_\ell$-vector space equipped with a continuous Π_K-action, we say that V is *unramified* at v if I_v acts trivially on V. When this happens, the Frobenius element $\mathrm{Frob}_v \in G_v/I_v$ gives an element of $\mathbf{GL}(V)$ which is well-defined up to conjugation in this group.

The following theorem lists some of the basic properties of $V_\ell(A)$.

Theorem 2.15. *Let A be an abelian scheme over $\mathrm{Spec}(\mathcal{O}_{K,S})$. The ℓ-adic Galois representation $V_\ell(A)$ satisfies the following properties:*

1. *It is semisimple as a representation of E.*

2. *It is unramified at all $v \notin S' := S \cup \{\lambda | \ell\}$.*

3. *(Rationality) If $v \notin S'$, then the characteristic polynomial of Frob_v has rational integer coefficients. The complex roots of this polynomial have absolute value $\mathbf{N}v^{1/2}$.*

4. *(Tate conjecture) The representation $V_\ell(A)$ is semisimple as a representation of Π_K.*

Property 1 follows from the basic theory of duality for abelian varieties, and properties 2 and 3 were shown by Weil (cf. [We]). Property 4, a particular case of the Tate conjecture, is one of Faltings' important contributions. We now explain how Faltings proved the semisimplicity of $V_\ell(A)$ over Π_K, adapting an idea used by Tate to prove the corresponding statement over finite fields.

Lemma 2.16. *For every Π_K-invariant subspace W in $V_\ell(A)$, there is an element $u \in E$ such that*
$$uV_\ell(A) = W.$$

Proof. The $\mathbf{Z}_\ell$-module $W_\infty = W \cap T_\ell(A)$ gives rise to a collection of groups $W_n = W_\infty/\ell^n W_\infty \subset A[\ell^n]$ which are defined over K and compatible under the natural maps $A[\ell^{n+1}] \longrightarrow A[\ell^n]$. Let

$$\alpha_n : A \longrightarrow A_n := A/W_n,$$

be the natural isogeny with kernel W_n, and let β_n denote the isogeny characterised by

$$\alpha_n\beta_n = \ell^n, \qquad \beta_n\alpha_n = \ell^n.$$

Note that $\beta_n(A_n[\ell^n]) = W_n$ by construction, in light of the first identity above. By Faltings' finiteness theorem 2.11, there exists an infinite set $I = \{n_0, n_1, \ldots\} \subset \mathbf{Z}^{>0}$ for which there exist isomorphisms

$$\nu_i : A_{n_0} \simeq A_i$$

for all $i \in I$. Now define a sequence of K-endomorphisms of A by the rule

$$u_i := \beta_i\nu_i\alpha_{n_0}.$$

Since $\mathrm{End}_K(A) \otimes \mathbf{Z}_\ell$ is compact in the ℓ-adic topology, the sequence (u_i) has a convergent subsequence $(u_i)_{i\in J}$ in this topology. Let u denote the limit of such a subsequence. After eventually refining J further, we can assume that for each $i \in J$, we have natural maps

$$u(A[\ell^i]) = u_i(A[\ell^i]) \longrightarrow \beta_i(A_i[\ell^i]) = W_i,$$

with kernel and cokernel bounded independently of i, because they arise from α_{n_0}. It follows that

$$u(V_\ell(A)) = W,$$

as was to be shown. $\qquad\qquad\square$

Corollary 2.17. *The representation $V_\ell(A)$ is a semisimple Π_K-module.*

Proof. Let W be a Π_K-stable subspace of $V_\ell(A)$, and let $u \in E$ be an element constructed in Lemma 2.16, satisfying $u(V_\ell(A)) = W$. Consider the right ideal uE in the algebra E. Because E is semisimple, this ideal is generated by an idempotent u_0. Note that $u_0(V_\ell(A)) = W$. The subspace $\ker(u_0)$ is therefore a Π_K-stable complement of W in $V_\ell(A)$. Hence $V_\ell(A)$ is semisimple over Π_K. $\qquad\square$

In conclusion, let $\mathrm{Rep}_S(G_K, 2g)$ be the set of isomorphism classes of rational semisimple ℓ-adic representations of G_K of dimension $2g$ which are unramified outside of S. We have shown that R_4 maps $\mathcal{I}_g(\mathcal{O})$ to $\mathrm{Rep}_S(G_K, 2g)$. To complete the proof of the Mordell conjecture, it remains to show:

1. The map R_4 is finite-to-one.

2. The set $\mathrm{Rep}_S(G_K, 2g)$ is finite.

We will prove the first in the next section, and the second in Section 2.7.

2.6   The isogeny conjecture

The proof of the following deep conjecture of Tate is a cornerstone of Faltings' strategy for proving the Mordell conjecture.

Theorem 2.18. *(Isogeny conjecture). Let A and B be abelian varieties defined over a number field K. If $V_\ell(A)$ is isomorphic to $V_\ell(B)$ as a Π_K-module, then the abelian varieties A and B are isogenous.*

In other words, the map R_4 is *injective.*
We first note that Theorem 2.18 can be reduced to the following statement, known as the *Tate conjecture* for abelian varieties.

Theorem 2.19. *(Tate conjecture). Let A and B be abelian varieties defined over K. Then the natural map*

$$\mathrm{Hom}_K(A, B) \otimes \mathbf{Q}_\ell \longrightarrow \mathrm{Hom}_{\Pi_K}(V_\ell(A), V_\ell(B))$$

is surjective.

To see that Theorem 2.19 implies Theorem 2.18, let $j : V_\ell(A) \simeq V_\ell(B)$ be a Π_K-equivariant isomorphism. By Theorem 2.19, this isomorphism comes from an element $u \in \operatorname{Hom}_K(A, B) \otimes \mathbf{Q}_\ell$. After multiplying u by some power of ℓ, we can assume that u belongs to $\operatorname{Hom}_K(A, B) \otimes \mathbf{Z}_\ell$. Note that $\operatorname{Hom}_K(A, B)$ is dense in $\operatorname{Hom}_K(A, B) \otimes \mathbf{Z}_\ell$. Any good enough ℓ-adic approximation to u in $\operatorname{Hom}_K(A, B)$ gives the desired K-isogeny between A and B. Theorem 2.18 follows.

We next observe that Theorem 2.19 can be reduced to the following special case:

Theorem 2.20. *Let A be an abelian variety over K. The natural map*

$$\operatorname{End}_K(A) \otimes \mathbf{Q}_\ell \longrightarrow \operatorname{End}_{\Pi_K}(V_\ell(A))$$

is surjective.

The fact that Theorem 2.20 implies Theorem 2.19 can be seen by applying Theorem 2.20 to the abelian variety $A \times B$, since

$$\operatorname{End}_K(A \times B) = \operatorname{End}_K(A) \oplus \operatorname{Hom}_K(A, B) \oplus \operatorname{Hom}_K(B, A) \oplus \operatorname{End}_K(B)$$

and likewise for $\operatorname{End}(V_\ell(A \times B)) = \operatorname{End}(V_\ell(A) \times V_\ell(B))$.

Proof of Theorem 2.20: Let ϕ be an element of $\operatorname{End}_{\Pi_K}(V_\ell(A))$, and let

$$W = \{(x, \phi(x)) \in V_\ell(A) \times V_\ell(A)\} \subset V_\ell(A \times A)$$

be the graph of ϕ. Note that W is Π_K-stable. Hence there is an endomorphism $u \in \operatorname{End}_K(A \times A) \otimes \mathbf{Q}_\ell = M_2(E)$ associated to W by Lemma 2.16, satisfying $u(V_\ell(A \times A)) = W$.

Let $E^0 = \operatorname{End}_E(V_\ell(A))$ denote the commutant of E in $\operatorname{End}(V_\ell(A))$. For any $\alpha \in E^0$, the matrix $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix}$ with entries in $\operatorname{End}(V_\ell(A))$ commutes with $u \in M_2(E) \subset M_2(\operatorname{End}(V_\ell(A))$. It follows that this matrix preserves $W = \operatorname{image}(u)$, and hence α commutes with ϕ. Since this argument is valid for any $\alpha \in E^0$, the endomorphism ϕ belongs to the double commutant E^{00} which is equal to E by the semisimplicity of $V_\ell(A)$ as a module over E.

2.7 The finiteness principle for rational ℓ-adic representations

Now that R_4 has been shown to be injective, it remains to prove that the target $\mathrm{Rep}_S(G_K, 2g)$ is finite. The main theorem of this section is:

Theorem 2.21. *(Finiteness principle for rational semisimple ℓ-adic representations). Let K be a number field and S a finite set of primes of K. Then there are finitely many isomorphism classes of rational, semisimple ℓ-adic representations of G_K of dimension d which are unramified outside of S.*

Remark. The reader will observe that this finiteness principle is close in spirit to the Hermite–Minkowski theorem: it asserts that there are only finitely many extensions of K (albeit, of infinite degree) of a certain special kind with bounded ramification. The proof of Theorem 2.21 will in fact rely crucially on the Hermite–Minkowski theorem, as well as on the Chebotarev density theorem.

We begin by establishing the following key lemma.

Lemma 2.22. *There exists a finite set T of primes of K (depending on S and d) satisfying the following two properties:*

1. *T is disjoint from $S_\ell := S \cup \{v|\ell\}$.*

2. *Two representations $\rho_1, \rho_2 \in \mathrm{Rep}_S(G_K, d)$ are isomorphic if and only if*
$$trace(\rho_1(\mathrm{Frob}_v)) = trace(\rho_2(\mathrm{Frob}_v)), \quad \text{for all } v \in T.$$

Proof. Consider the set of all extensions of K of degree $\leq l^{2d^2}$ which are unramified outside S_ℓ. By Theorem 1.1 (Hermite–Minkowki), there are finitely many such extensions, and hence their compositum L is a finite extension of K. Let $T = \{v_1, \dots v_N\}$ be a set of primes of K which are not in S and such that the Frobenius conjugacy classes Frob_{v_i} generate $\mathrm{Gal}(L/K)$. The existence of such a finite set follows from the Chebotarev density theorem. We claim that this set T satisfies the conclusion of Lemma 2.22. Given $\rho_1, \rho_2 \in \mathrm{Rep}_S(G_K, d)$, a choice of G_K-stable $\mathbf{Z}_\ell$-lattices in the underlying representation spaces makes it possible to view each ρ_i as a homomorphism from $\mathbf{Z}_\ell[[G_K]]$ to $M_d(\mathbf{Z}_\ell)$. Let

$$j = \rho_1 \oplus \rho_2 : \mathbf{Z}_\ell[[G_K]] \longrightarrow M_d(\mathbf{Z}_\ell) \times M_d(\mathbf{Z}_\ell),$$

and let M denote the image of j. The induced homomorphism

$$\bar{j} : G_K \longrightarrow (M/\ell M)^\times$$

factors through $\mathrm{Gal}(L/K)$, since the cardinality of $M/\ell M$ is at most ℓ^{2d^2} and $\bar{j}$ is unramified outside of S_ℓ. It follows that the elements

$$\bar{j}(\mathrm{Frob}_{v_1}), \ldots, \bar{j}(\mathrm{Frob}_{v_N})$$

generate $M/\ell M$. By Nakayama's lemma, the elements

$$j(\mathrm{Frob}_{v_1}), \ldots, j(\mathrm{Frob}_{v_N})$$

generate M as a $\mathbf{Z}_\ell$-module.

In particular, if

$$\mathrm{trace}(\rho_1(\mathrm{Frob}_{v_j})) = \mathrm{trace}(\rho_2(\mathrm{Frob}_{v_j})), \quad \text{for } j = 1, \ldots, N,$$

then

$$M \subseteq \Delta \subset M_d(\mathbf{Z}_\ell) \times M_d(\mathbf{Z}_\ell),$$

where $\Delta = M_d(\mathbf{Z}_\ell)$ is embedded diagonally. Therefore one has

$$\mathrm{trace}(\rho_1(\sigma)) = \mathrm{trace}(\rho_2(\sigma)) \quad \text{for all } \sigma \in \Pi_K.$$

Hence ρ_1 and ρ_2 have the same traces. Since they are semisimple, it follows that they are isomorphic as Π_K-representations. $\qquad\square$

Proof of Theorem 2.21. Let $T = \{v_1, \ldots, v_N\}$ be as in the statement of Lemma 2.22. The assignment

$$\rho \mapsto (\mathrm{Tr}\,(\rho(\mathrm{Frob}_{v_1})), \ldots, \mathrm{Tr}\,(\rho(\mathrm{Frob}_{v_N})))$$

is injective on $\mathrm{Rep}_S(G_K, d)$, and can only assume finitely many values, by the rationality of ρ. (More precisely, each $\mathrm{Tr}\,(\mathrm{Frob}_{v_i})$ is a rational integer of absolute value $\leq d\mathbf{N}v_i^{1/2}$.) Theorem 2.21 follows.

2.8 A summary of Faltings' proof

Faltings' proof of Mordell's conjecture is based on a sequence of maps (here X is a curve of genus g defined over K and having good reduction outside of the finite set S of primes of K):

$$
\left\{
\begin{array}{l}
K\text{-rational} \\
\text{points on } X
\end{array}
\right\}
\xrightarrow{R_1}
\left\{
\begin{array}{l}
\text{Curves of genus } g' \text{ over } K' \\
\text{with good reduction outside } S'
\end{array}
\right\}
$$

$$
\xrightarrow{R_2}
\left\{
\begin{array}{l}
\text{Isomorphism classes of semistable} \\
\text{abelian varieties of dimension } g' \\
\text{with good reduction outside } S'
\end{array}
\right\}
$$

$$
\xrightarrow{R_3}
\left\{
\begin{array}{l}
\text{Isogeny classes of abelian varieties} \\
\text{of dimension } g' \\
\text{with good reduction outside } S'
\end{array}
\right\}
$$

$$
\xrightarrow{R_4}
\left\{
\begin{array}{l}
\text{Rational semisimple } \ell\text{-adic representations} \\
\text{of dimension } 2g' \text{ unramified outside } S'_\ell
\end{array}
\right\}
$$

1. The map R_1 is given by Parshin's construction, and is finite-to-one, by the geometric theorem of De Franchis.

2. The map R_2 is defined by passing to the jacobian of a curve, and is finite-to-one by Torelli's theorem.

3. The map R_3 is the obvious one, and is finite-to-one, by Falting's fundamental Theorem 2.11 on finiteness of abelian varieties in a given isogeny class.

4. The map R_4 is defined by passing to the Tate module, and is one-to-one, thanks to the Tate conjectures proved by Faltings. The proof of the Tate conjectures is obtained by combining a strategy of Tate with the finiteness Theorem 2.11. These ideas are also used to show that the Galois representations arising in the image of R_4 are *semisimple.*

5. The last set in this sequence of maps is finite by the finiteness principle for rational semisimple ℓ-adic representations, which is itself a consequence of the Chebotarev density theorem and the Hermite–Minkowski theorem.

3 Modular curves and Mazur's theorem

The first step in the proof of the Mordell conjecture (the Kodaira–Parshin reduction) consists in transforming a question about rational points on a given curve into the Shafarevich conjecture. This new diophantine question is concerned with the moduli space of curves themselves, to which an array of techniques (notably, jacobians, ℓ-adic representations, etc.) can be applied. It is therefore apparent that the extra structures afforded by moduli spaces are of great help in studying the diophantine questions that are associated to them. So it is natural to examine more closely the simplest class of moduli spaces, which are also curves in their own right: the *modular curves* classifying elliptic curves with extra level structure.

3.1 Modular curves

Let p be a prime ≥ 5, and write Z for the ring $\mathbf{Z}[1/p]$. The functor $Y_1(p)$ which to any Z-algebra R associates the set of R-isomorphism classes of pairs (E, P) where E is an elliptic curve over $\mathrm{Spec}(R)$ and P is a point of order p on E_R is representable by a smooth affine scheme over $\mathrm{Spec}(Z)$ of relative dimension one, denoted $Y_1(p)$.

The group $(\mathbf{Z}/p\mathbf{Z})^\times$ acts on $Y_1(p)$ by the rule $t \cdot (E, P) := (E, tP)$, and the quotient of $Y_1(p)$ by this action is an affine scheme $Y_0(p)$ over $\mathrm{Spec}(Z)$ which is a coarse moduli scheme classifying pairs (E, C) consisting of an elliptic curve over R and a cyclic subgroup scheme $C \subset E$ of order p defined over R.

These curves admit analytic descriptions as quotients of the Poincaré upper half-plane
$$\mathcal{H} = \{\tau \in \mathbf{C}, \qquad \mathrm{Im}(\tau) > 0\}$$
by the action of the following discrete subgroups of $\mathbf{SL}_2(\mathbf{Z})$:

$$\Gamma_1(p) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{with } a - 1 \equiv c \equiv d - 1 \equiv 0 \pmod{p} \right\},$$

$$\Gamma_0(p) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{with } c \equiv 0 \pmod{p} \right\}.$$

For example, the curve $Y_0(1)$ is identified with $\mathrm{Spec}(Z[j])$, and a birational (singular, even on the generic fiber) model for $Y_0(p)$ over $\mathrm{Spec}(Z)$ is given by

$$\mathrm{Spec}(Z[j, j']/\Phi_p(j, j')),$$

where $\Phi_p(x, y) \in \mathbf{Z}[x, y]$ is the canonical *modular polynomial* of bidegree $p+1$ satisfying $\Phi_p(j(\tau), j(p\tau)) = 0$, for all $\tau \in \mathcal{H}$.

A rational point on $Y_1(p)$ (resp. on $Y_0(p)$) determines an elliptic curve over $\mathbf{Q}$ with a $\mathbf{Q}$-rational point of order p (resp. a rational subgroup of order p). The main goal of this chapter is to explain the proof of the following theorem of Mazur.

Theorem 3.1. *If $p > 13$, then $Y_1(p)(\mathbf{Q}) = \emptyset$.*

Remark 3.2. Note that Theorem 3.1 can be viewed as a theorem about curves in two different ways. Firstly, it asserts that the collection of modular curves $Y_1(p)$, whose genera grow with p, have no rational points once p is large enough—a type of statement that is similar in flavour to Fermat's Last Theorem. Secondly, it leads to the *uniform boundedness* of the size of the torsion subgroups $E(\mathbf{Q})_{\mathrm{tors}}$ as E ranges over all elliptic curves over $\mathbf{Q}$, and is therefore also a theorem about curves of genus one.

3.2 Mazur's criterion

An important role is played in Mazur's argument by the compactification $X_0(p)$ of the affine curve $Y_0(p)$. As a Riemann surface, $X_0(p)(\mathbf{C})$ is obtained by adjoining to $Y_0(p)$ a finite set of cusps which are in bijection with the orbits of $\Gamma_0(p)$ acting on $\mathbb{P}_1(\mathbf{Q})$ by Mobius transformations. More precisely, letting $\mathcal{H}^* := \mathcal{H} \cup \mathbb{P}_1(\mathbf{Q})$, we have

$$X_0(p)(\mathbf{C}) = \Gamma_0(p)\backslash\mathcal{H}^* = (\Gamma_0(p)\backslash\mathcal{H}) \cup \{0, \infty\} = Y_0(p)(\mathbf{C}) \cup \{0, \infty\}.$$

The complex structure in a neighbourhood of ∞ is defined by letting $q = e^{2\pi i \tau}$ be a local parameter at ∞.

The equation for the universal elliptic curve in a formal punctured neighbourhood of ∞ is given by the Tate curve

$$E_q = Z[[q]]^\times / q^{\mathbf{Z}} : y^2 + xy = x^3 + a(q)x + b(q) \qquad \text{over } Z((q)),$$

where

$$a(q) = -5 \sum_{n=1}^{\infty} \sigma_3(n)q^n, \quad b(q) = -\frac{1}{12} \sum_{n=1}^{\infty} (7\sigma_5(n) + 5\sigma_3(n))q^n.$$

(Recall that $\sigma_k(n) = \sum_{d|n} d^k$.) The discriminant of E_q is equal to

$$\Delta(E_q) = q \prod_{n \geq 1}(1 - q^n)^{24},$$

and therefore E_q defines an elliptic curve over $Z((q))$.

The important *q-expansion principle* asserts that the parameter q is also a local parameter for the scheme $X_0(p)_Z$ in a neighborhood of ∞. Thanks to the q-expansion principle, the completion of the local ring of $X_0(p)_Z$ at ∞ is identified with the power series ring $Z[[q]]$:

$$\hat{\mathcal{O}}_{X_0(p),\infty} = Z[[q]].$$

A basic technique in Mazur's proof is to study the behaviour of certain maps on modular curves, via their behaviour in a formal neighbourhood of ∞. The following definition will be useful.

Definition 3.3. A morphism $j : X \longrightarrow Y$ of schemes over Z is a *formal immersion* at $x \in X(Z)$ if the induced map on completed local rings

$$j^* : \hat{\mathcal{O}}_{Y,j(x)} \longrightarrow \hat{\mathcal{O}}_{X,x}$$

is surjective.

Let $J_0(p)$ denote the jacobian of $X_0(p)$. It is an abelian variety over Z and is equipped with an embedding

$$\Phi : X_0(p) \longrightarrow J_0(p)$$

defined by letting $\Phi(x)$ be the class of the degree zero divisor $(x) - (\infty)$.

If $J_\sharp(p)$ is any quotient of $J_0(p)$, let $j_\sharp : X_0(p) \longrightarrow J_\sharp(p)$ be the map obtained by composing Φ with the projection to $J_\sharp(p)$. The following criterion of Mazur for $Y_1(p)(\mathbf{Q}) = \emptyset$ is the main result of this section.

Theorem 3.4. *Assume that $p > 7$. Suppose that there is an abelian variety quotient $J_\sharp(p)$ of $J_0(p)$ satisfying the following conditions:*

(a) *The map $j_\sharp : X_0(p) \longrightarrow J_\sharp(p)$ is a formal immersion at ∞.*

(b) *$J_\sharp(p)(\mathbf{Q})$ is finite.*

Then $Y_1(p)(\mathbf{Q}) = \emptyset$.

Sketch of proof. Let $\tilde{x}$ be a point in $Y_1(p)(\mathbf{Q})$ corresponding to the pair (E, P), where E is an elliptic curve over $\mathbf{Q}$ and $P \in E(\mathbf{Q})$ is of order p. Let $\mathcal{E}$ be the minimal Weierstrass model of E over Z.

The proof is divided into four steps.

Step 1. If E has potentially good reduction at the prime 3, then the special fiber $\mathcal{E}_{\mathbb{F}_3}$ is either an elliptic curve, or an extension of a finite group of connected components of cardinality $2^a 3^b$ by the additive group $\mathbb{G}_{a/\mathbb{F}_3}$. Such a group cannot contain a point of order $p > 7$, by the Hasse bound. Hence E has potentially mutiplicative reduction at 3.

Step 2. Let $x \in X_0(p)(\mathbf{Q})$ be the image of $\tilde{x}$ under the natural map: it corresponds to the pair $(E, \langle P \rangle)$ consisting of the curve E and the cyclic subgroup generated by P. By Step 1, the point x reduces to one of the cusps 0 or ∞ of $X_0(p)$ modulo 3. It can be asssumed without loss of generality that x reduces to ∞, by replacing $(E, \langle P \rangle)$ by $(E/\langle P \rangle, E[p]/\langle P \rangle)$ otherwise.

Step 3. Consider the element $j_\sharp(x) \in J_\sharp(p)(\mathbf{Q})$. By step 2 this element belongs to the formal group $J_\sharp^1(p)(\mathbf{Q}_3)$, which is torsion-free because $\mathbf{Q}_3$ is absolutely unramified. It also belongs to $J_\sharp(p)(\mathbf{Q})$, which is torsion by assumption. It follows that $j_\sharp(x) = 0$.

Step 4. We now use the fact that $j_\sharp$ is a formal immersion to deduce that $x = \infty$. To see this, let $\mathrm{Spec}(R)$ be an affine neighborhood of ∞ containing x. The point x gives rise to a ring homomorphism $x : R \longrightarrow \mathbf{Z}_3$, which factors through the local ring $\hat{\mathcal{O}}_{X_0(p), \infty} = Z[[q]]$, so that x can be viewed as a map $Z[[q]] \longrightarrow \mathbf{Z}_3$. By step 3, we have

$$x \circ j_\sharp^* = \infty \circ j_\sharp^*.$$

It follows that $x = \infty$, since $j_\sharp^*$ was assumed to be surjective, contradicting the initial assumption that x belongs to $Y_0(p)$. $\square$

Mazur's criterion reduces Theorem 3.1 to the problem of exhibiting a quotient $J_\sharp(p)$ of $J_0(p)$ satisfying the conditions of Theorem 3.4.

3.3 The jacobian $J_0(p)$

The fact that makes it possible to analyse the jacobian $J_0(p)$ precisely, and exhibit a nontrivial quotient of it with finite Mordell–Weil group, arises from two related ingredients.

(a) *Hecke operators.* If n is an integer that is not divisible by p, the modular curve $X_0(np)$ is equipped with two maps π_1, π_2 to $X_0(p)$, defined by

$$\pi_1(E, C) = (E, C[p]), \qquad \pi_2(E, C) = (E/C[n], C/C[n]).$$

The pair (π_1, π_2) gives rise to an embedding of $X_0(pn)$ in the product $X_0(p) \times X_0(p)$. The image in this product, denoted T_n, is an algebraic correspondence on $X_0(p)$ defined over $\mathbf{Q}$, which gives rise to an endomorphism of $J_0(p)$ defined over $\mathbf{Q}$. On the level of divisors, T_n is described by

$$T_n(E, C) = \sum_{E \longrightarrow E'} (E', C'), \tag{6}$$

where the sum is taken over the cyclic isogenies $\varphi : E \longrightarrow E'$ of degree n, and $C' = \varphi(C)$. Let $\mathbf{T}$ denote the subring of $\mathrm{End}_\mathbf{Q}(J_0(p))$ generated by the Hecke operators T_n. It is finitely generated (as a ring, and even as a module) over $\mathbf{Z}$. Our basic approach to constructing $J_\sharp(p)$ is to use the endomorphisms in $\mathbf{T}$ to decompose the abelian variety $J_0(p)$ (up to $\mathbf{Q}$-isogeny) into smaller pieces which can then be analysed individually. If R is any ring, let $\mathbf{T}_R$ denote the R-algebra $\mathbf{T} \otimes R$.

(b) *Modular forms.* If R is any Z-algebra, let $S_2(p, R)$ denote the space of regular differentials on $X_0(p)_R$. Restriction to the formal neighborhood $\mathrm{Spec}(R[[q]])$ of $\infty \in X_0(p)$ gives rise to a map (called the *q-expansion map*)

$$q\text{-exp} : S_2(p, R) \longrightarrow R[[q]]dq.$$

When $R = \mathbf{C}$, the space $S_2(p, \mathbf{C})$ is identified with the vector space of homomorphic functions $f : \mathcal{H} \longrightarrow \mathbf{C}$ for which

(i) the differential $2\pi i f(\tau) d\tau$ is invariant under $\Gamma_0(p)$, i.e.,

$$f\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^2 f(\tau), \qquad \text{for all } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(p).$$

(ii) $2\pi i f(\tau) d\tau$ extends to a holomorphic differential on the compactified modular curve $X_0(p)$. In particular, it admits a Fourier expansion of the form

$$f(\tau) = \sum_{n=1}^{\infty} a_n e^{2\pi i n \tau},$$

so that $q\text{-exp}(2\pi i f(\tau) d\tau) = \sum_{n=1}^{\infty} a_n q^n \frac{dq}{q}$.

The action of the Hecke operators T_n on $J_0(p)_R$ induces an action on the cotangent space $S_2(p, R)$, which can be described explicitly on the level of the q-expansions. For example, if $\ell \neq p$ is prime,

$$T_\ell \left(\sum_{n=1}^\infty a_n q^n \frac{dq}{q} \right) = \left(\sum_{\ell|n} a_n q^{n/\ell} + \ell \sum_{n=1}^\infty a_n q^{n\ell} \right) \frac{dq}{q}. \qquad (7)$$

There is an extra Hecke operator T_p defined via an algebraic correspondence $X_0(p^2) \subset X_0(p) \times X_0(p)$ which admits the following simpler formula for its action on q-expansions:

$$T_p \left(\sum_{n=1}^\infty a_n q^n \frac{dq}{q} \right) = \left(\sum_{p|n} a_n q^{n/p} \right) \frac{dq}{q}. \qquad (8)$$

The definition of T_ℓ for ℓ prime can then be extended to all integers n by the multiplicativity relations implicit in the following identity of formal Dirichlet series:

$$\sum_{n \geq 1} T_n n^{-s} = (1 - T_p p^{-s})^{-1} \prod_{\ell \neq p} (1 - T_\ell \ell^{-s} + \ell^{1-2s})^{-1}. \qquad (9)$$

In other words,

$$T_{mn} = T_m T_n \text{ if } \gcd(m, n) = 1, \qquad T_{\ell^{n+1}} = a_\ell T_{\ell^n} - \ell T_{\ell^{n-1}}.$$

Proposition 3.5. *The algebra* $\mathbf{T_Q}$ *is a commutative semisimple algebra of dimension* $g := \dim_{\mathbf{Q}} S_2(p, \mathbf{Q}) = \mathrm{genus}(X_0(p))$.

Sketch. The fact that $\mathbf{T_Q}$ is commutative follows from the explicit description of the operators T_n as correspondences given in (6) (or, if one prefers, from equation (7) describing its effect on q-expansions). The semisimplicity arises from the fact that the operators T_n are self-adjoint with respect to the Hermitian pairing on $S_2(p, \mathbf{C})$ (Petersson scalar product) defined by

$$\langle \omega_1, \omega_2 \rangle = \frac{1}{2i} \int_{\Gamma_0(p) \backslash \mathcal{H}} \omega_1 \wedge \bar{\omega}_2.$$

(We mention in passing that in general, the operator T_ℓ acting on $S_2(N, \mathbf{C})$ need not be self-adjoint when $\ell | N$, but it is self-adjoint when restricted to the space of so-called *newforms*. We are using implicitly the fact that $S_2(p, \mathbf{C})$

is equal to its subspace of newforms.) One computes the dimension of $\mathbf{T_Q}$ by showing that the bilinear pairing

$$\mathbf{T_Q} \times S_2(p, \mathbf{Q}) \longrightarrow \mathbf{Q}, \qquad (T, f) := a_1(Tf)$$

is left and right nondegenerate, and in fact positive definite. The details are left to the reader (who may also consult Section 2.2 of [Da2] for more details). □

As a consequence of Proposition 3.5 and its proof, one has the decomposition

$$\mathbf{T_Q} = K_1 \times \cdots \times K_t$$

of $\mathbf{T_Q}$ into a product of totally real fields, with $\sum_{j=1}^{t}[K_j : \mathbf{Q}] = n$. The factors K_j are indexed by:

(a) The points $\phi_1, \ldots, \phi_t$ of $\mathrm{Spec}(\mathbf{T_Q})$, viewed as algebra homomorphisms $\phi_j : \mathbf{T_Q} \longrightarrow \bar{\mathbf{Q}}$ (taken modulo the natural action of $G_{\mathbf{Q}} = \mathrm{Gal}(\bar{\mathbf{Q}}/\mathbf{Q})$).

(b) The distinct $G_{\mathbf{Q}}$-equivalences classes $f_1, \ldots, f_t$ of eigenforms for $\mathbf{T}$, normalised so that $a_1(f_j) = 1$. The q-expansions of these eigenforms are described by

$$f_j = \sum_{n=1}^{\infty} \phi_j(T_n)q^n.$$

The quotient A_f attached to f is defined by letting

$$A_f := J_0(p)/I_f, \quad \text{where } I_f := \ker(\mathbf{T} \longrightarrow K_f).$$

With these notations, the main result of this section is the following *Eichler-Shimura decomposition*, which asserts that $J_0(p)$ is isogenous to a product of $\mathbf{Q}$-simple factors indexed by the ($G_{\mathbf{Q}}$-orbits of) normalised eigenforms f_j ($j = 1, \ldots, t$).

Theorem 3.6. *The abelian variety $J_0(p)$ is $\mathbf{Q}$-isogenous to the product*

$$\prod_{i=1}^{t} A_{f_i},$$

of $\mathbf{Q}$-simple abelian varieties A_{f_i}. The varieties A_f that occur in this decomposition have the following properties:

(a) $\dim(A_f) = [K_f : \mathbf{Q}]$;

(b) *The natural image of* $\mathbf{T_Q}$ *in* $\mathrm{End}_{\mathbf{Q}}(A_f) \otimes \mathbf{Q}$ *is isomorphic to* K_f.

For more details on this decomposition see Chapter 2 of [Da2].

Thanks to Theorem 3.6, we are reduced to the following question:

Question 3.7. *Find a criterion involving the normalised eigenform f for the quotient A_f to have finite Mordell–Weil group.*

3.4 The Birch and Swinnerton-Dyer conjecture

The key to bounding the rank of $A_f(\mathbf{Q})$ (and showing that this rank is zero, for a sufficiently large collection of normalised eigenforms f) lies in studying the so-called *Hasse–Weil L-series* attached to A_f.

Let A be an abelian variety (of dimension d, say) defined over $\mathbf{Q}$. The Hasse–Weil L-series of A is most conveniently defined in terms of the ℓ-adic representation $V_\ell(A)$ that was introduced in Section 2.5. If $p \neq \ell$ is a prime, the Frobenius element acts naturally on on the space $V_\ell(A)^{I_p}$ of vectors in $V_\ell(A)$ that are fixed under the action of the inertia group at p. (Recall that $V_\ell(A)^{I_p} = V_\ell(A)$ if A has good reduction at $p \neq \ell$.) By the rationality of the representation $V_\ell(A)$, the characteristic polynomial

$$F_p(T) := \det(1 - \mathrm{Frob}_p|_{V_\ell(A)^{I_p}} T)$$

has integer coefficients. Furthermore, it does not depend on the choice of ℓ, and can therefore be defined for all p. This makes it possible to define the Hasse–Weil L-series as a function of the complex variable s, by the infinite product

$$L(A, s) = \prod_p F_p(p^{-s})^{-1}.$$

Using the rationality of the Galois representation $V_\ell(f)$ in the sense of Theorem 2.15, one can show that the infinite product defining $L(A, s)$ converges uniformly on compact subsets of $\{s \in \mathbf{C} | \Re(s) > 3/2\}$, and hence defines an analytic function in this region.

Concerning the behaviour of $L(A, s)$ and its connection to the arithmetic of A over $\mathbf{Q}$, there are the following two fundamental conjectures:

Conjecture 3.8. *The L-series $L(A, s)$ has an analytic continuation to the entire complex plane and a functional equation of the form*

$$\Lambda(A, s) := (2\pi)^{-ds}\Gamma(s)^d N^{s/2} L(A, s) = \pm\Lambda(A, 2 - s),$$

where N is the conductor *of A.*

In particular, if Conjecture 3.8 is true, the process of analytic continuation gives meaning to the behaviour of $L(A, s)$ in a neighborhood of the central critical point $s = 1$ for the functional equation, and in particular, the order of vanishing of $L(A, s)$ at $s = 1$ is defined. The *Birch and Swinnerton-Dyer conjecture* relates this order of vanishing to the arithmetic of A over $\mathbf{Q}$:

Conjecture 3.9. *If A is an abelian variety over $\mathbf{Q}$, then*

$$\text{rank}(A(\mathbf{Q})) = \text{ord}_{s=1}(L(A, s)).$$

In particular, $A(\mathbf{Q})$ is finite if $L(A, 1) \neq 0$.

Both Conjectures 3.8 and 3.9 are far from being proved in general. But much more is known when $A = A_f$ occurs in the Eichler–Shimura decomposition of the modular jacobian $J_0(N)$, as will be explained in the next section.

3.5   Hecke theory

A *newform* of level N is a normalised eigenform $f = \sum_{n\geq 1} a_n q^n \frac{dq}{q}$ on $\Gamma_0(N)$ whose associated sequence $(a_n)_{(n,N)=1}$ of Fourier coefficients is different from that of any eigenform g on $\Gamma_0(d)$ with $d|N$ and $d \neq N$.

To each newform $f = \sum_{n\geq 1} a_n q^n \frac{dq}{q} \in S_2(N, \mathbf{C})$, one can associate an *L*-series

$$L(f, s) := \sum_{n=1}^{\infty} a_n n^{-s}.$$

This *L*-series enjoys the following properties, which were established by Hecke:

(a) **Euler product**: It admits the Euler product factorisation given by

$$L(f, s) = \prod_{p\nmid N}(1 - a_p p^{-s} + p^{1-2s})^{-1} \prod_{p|N}(1 - a_p p^{-s})^{-1},$$

as can be seen by applying φ_f to the formal identity (9) expressing the Hecke operators T_n in terms of the operators T_ℓ for ℓ prime.

(b) **Integral representation**: The L-series $L(f, s)$ can be represented as an integral transform of the modular form f, by the formula:

$$\Lambda(f, s) := (2\pi)^{-s}\Gamma(s)N^{s/2}L(f, s) = N^{s/2}\int_0^\infty f(it)t^{s-1}dt, \qquad (10)$$

where $\Gamma(s) = \int_0^\infty e^{-t}t^{s-1}dt$ is the Γ-function. In particular, because f is of rapid decay at the cusps, this integral converges absolutely to an analytic function of $s \in \mathbf{C}$.

(c) **Functional equation**: The involution w defined on $S_2(N, \mathbf{C})$ by the rule

$$w(f)(\tau) = \frac{1}{N\tau^2}f\left(\frac{-1}{N\tau}\right) \qquad (11)$$

commutes with the Hecke operators and hence preserves its associated eigenspaces. It follows that for the eigenform f,

$$w(f) = \varepsilon f, \quad \text{where } \varepsilon = \pm 1. \qquad (12)$$

The L-series $L(f, s)$ satisfies the functional equation

$$\Lambda(f, s) = -\Lambda(w(f), 2 - s) = -\varepsilon\Lambda(f, 2 - s). \qquad (13)$$

It is a direct calculation to derive this functional equation from the integral representation of $\Lambda(f, s)$.

For the next result, we view f as an element of $S_2(N, K_f)$. (Recall that K_f is the totally real field generated by the Fourier coefficients of f.) Any complex embedding $\sigma : K_f \hookrightarrow \mathbf{C}$ yields an eigenform f^σ with complex coefficients, to which the Hecke L-function $L(f^\sigma, s)$ may be attached. The following result relates the L-series of Hasse–Weil and of Hecke.

Theorem 3.10. *Let A_f be the abelian variety associated to the newform $f \in S_2(N, \mathbf{C})$ by the Eichler–Shimura construction. Then*

$$L(A_f, s) = \prod_{\sigma:K_f \longrightarrow \mathbf{C}} L(f^\sigma, s).$$

In particular, Conjecture 3.8 holds for A_f.

38

The main ingredient in the proof of Theorem 3.10 is the Eichler–Shimura congruence which relates the Hecke correspondence $T_p \subset X_0(N)^2$ in characteristic p to the graph of the Frobenius morphism and its transpose. For more details and references see Chapter 2 of [Da2].

Theorem 3.10 reveals that one has better control of the arithmetic of the abelian varieties A_f—Conjecture 3.8 remains open for the general abelian variety A over $\mathbf{Q}$. In fact, one has the following strong evidence for Conjecture 3.9 for the abelian varieties A_f.

Theorem 3.11. *If $L(A_f, 1) \neq 0$, then $A_f(\mathbf{Q})$ is finite.*

The main ingredients that go into the proof of Theorem 3.11 are

1. The theory of Heegner points on modular curves;

2. The theorem of Gross–Zagier expressing the canonical heights of the images of these points in A_f in terms of special values of L-series closely related to $L(f, s)$;

3. A theorem of Kolyvagin which relates the system of Heegner points and the arithmetic of A_f over $\mathbf{Q}$.

These ingredients will be discussed in somewhat more detail in Section 5 devoted to elliptic curves and the Birch and Swinnerton-Dyer conjecture.

3.6 The winding quotient

The criterion for the finiteness of $A_f(\mathbf{Q})$ supplied by Theorem 3.11 allows us to construct a quotient $J_\sharp(p)$ which is in some sense the "largest possible" quotient with finite Mordell–Weil group.

We construct $J_\sharp(p)$, following Merel, by letting e_0 be the vertical path from 0 to $i\infty$ on $\mathcal{H}$; its image in $X_0(p)$ gives an element in the relative homology $H_1(X_0(p)(\mathbf{C}), \mathbf{Z}; \{\text{cusps}\})$. By a result of Manin–Drinfeld, the element e_0 gives rise to an element e in the rational homology $\mathbb{H} := H_1(X_0(p)(\mathbf{C}), \mathbf{Q})$. This element is referred to as the *winding element*.

The Hecke algebra $\mathbf{T} = \mathbf{T_Q}$ acts on $\mathbb{H}$ by functoriality of correspondences. Let e_f denote the image of e in $\mathbb{H}/I_f\mathbb{H}$. The integral formula (10) for $L(f, s)$ shows that

$$e_f \neq 0 \quad \text{if and only if } L(f, 1) \neq 0.$$

Hence it is natural to define

$$J_e(p) := J_0(p)/I_e, \quad \text{where } I_e := \mathrm{Ann}_{\mathbf{T}}(e).$$

Theorem 3.12. *The Mordell–Weil group $J_e(p)(\mathbf{Q})$ is finite.*

Proof. Up to isogeny, $J_e(p)$ decomposes as

$$J_e(p) \sim \prod_{e_f \neq 0} A_f = \prod_{L(f,1) \neq 0} A_f.$$

Theorem 3.11 implies that $A_f(\mathbf{Q})$ is finite for all the f that appear in this decomposition. The theorem follows. $\qquad\square$

In order to exploit Mazur's criterion with $J_\sharp(p) = J_e(p)$, and thereby prove Theorem 3.1, it remains to show that the natural map $j_e : X_0(p) \longrightarrow J_e(p)$ is a formal immersion at ∞. (So that in particular $J_e(p)$ is nontrivial, which is not clear *a priori* from its definition!) This is done in the article of Marusia Rebolledo in these proceedings (cf. Theorem 4 of Section 2.3 of [Reb]). Rebolledo's article goes significantly further by showing that the natural map from the d-th symmetric power $X_0(p)^{(d)}$ of $X_0(p)$ sending $(P_1, \ldots, P_d)$ to the image in $J_\sharp(p)$ of the divisor class $(P_1) + \cdots + (P_d) - d(\infty)$ is a formal immersion at $(\infty, \ldots, \infty)$, as soon as p is sufficiently large relative to d.

Remark 3.13. Our presentation of Mazur's argument incorporates an important simplification due to Merel, which consists in working with the winding quotient $J_e(p)$ whose finiteness is known thanks to Theorem 3.11. At the time of Mazur's original proof described in [Ma1], Theorems 3.11 and 3.12 were not available, and Mazur's approach worked with the so-called *Eisenstein quotient* $J_{\mathrm{eis}}(p)$. This quotient contains a rational torsion subgroup of order $n = \mathrm{numerator}(\frac{p-1}{12})$, and one of the key results in [Ma1] is to establish the finiteness of $J_{\mathrm{eis}}(\mathbf{Q})$ by an n-descent argument. In Merel's approach, Mazur's somewhat delicate "Eisenstein descent" is in effect replaced by Kolyvagin's descent based on Heegner points and the theorem of Gross–Zagier.

3.7 More results and questions

By various refinements of the techniques discussed above, Mazur was able to classify all possible rational torsion subgroups of elliptic curves over $\mathbf{Q}$ and obtained the following results:

Theorem 3.14. *Let T be the torsion subgroup of the Mordell–Weil group of an elliptic curve E over $\mathbf{Q}$. Then T is isomorphic to one of the following 15 groups:*

$$\mathbf{Z}/m\mathbf{Z} \qquad \text{for} \qquad 1 \leq m \leq 10 \text{ or } m = 12,$$
$$\mathbf{Z}/2m\mathbf{Z} \times \mathbf{Z}/2\mathbf{Z} \quad \text{for} \qquad 1 \leq m \leq 4.$$

For the proof, see [Ma1], p. 156. We note in passing that all possibilities for T that are not ruled out by Mazur's theorem do in fact occur infinitely often: the associated modular curves are of genus 0 and have a rational point.

Mazur's theorem implies that rational points of order p on elliptic curves cannot occur for $p > 7$. One can ask similar questions for rational subgroups. In this direction, Mazur proved the following result in [Ma2].

Theorem 3.15. *Suppose that there is an elliptic curve E over $\mathbf{Q}$ with a rational subgroup of prime order p. Then $p \leq 19$ or $p = 37, 43, 67,$ or 163.*

The four exceptional values of p in Theorem 3.15 correspond to discriminants of imaginary quadratic fields of class number one. The corresponding elliptic curves with complex multiplication can be defined over $\mathbf{Q}$ and have a rational subgroup of order p.

Theorem 3.15 implies that for large enough p, the Galois representation

$$\rho_{E,p} : G_{\mathbf{Q}} \longrightarrow \text{Aut}(E[p])$$

is always *irreducible*. One can also ask whether, for large enough p, this Galois representation is in fact necessarily *surjective*. The existence of elliptic curves with complex multiplication, for which $\rho_{E,p}$ is *never* surjective when $p \geq 3$, precludes an affirmative answer to this question. Discarding elliptic curves with complex multiplication, the following conjecture (which appears in [Se1], p. 299, §4.3, phrased more prudently as an open question) can be proposed:

Conjecture 3.16. *(Surjectivity conjecture) If E is an elliptic curve over $\mathbf{Q}$ without complex multiplication, and $p \geq 19$ is prime, then the Galois representation associated to $E[p]$ is surjective.*

The surjectivity conjecture remains open, more than 30 years after [Ma1]. The hypothetical cases that are the most difficult to dispose of are those where the image of $\rho_{E,p}$ is contained in the normalizer of a Cartan subgroup, particularly a nonsplit Cartan subgroup.

It is also natural to search for analogues of Theorem 3.14 over number fields other than $\mathbf{Q}$; a remarkable breakthrough was achieved on this problem by S. Kamienny and Merel around 1992 ([Ka], [Me]).

Theorem 3.17. *Let K be a number field. Then the size of $E(K)_{\mathrm{tors}}$ is bounded by a constant $B(K)$ which depends only on K. In fact, this constant can be made to depend only on the degree of K over $\mathbf{Q}$.*

The proof of this theorem is explained in the article by Marusia Rebolledo [Reb] in these proceedings.

We finish with a conjecture that can be viewed as a "mod p analogue" of Theorem 2.18 (Tate's isogeny conjecture).

Conjecture 3.18. *There exists an integer M such that, for all $p \geq M$, any two elliptic curves E_1 and E_2 over $\mathbf{Q}$ are isogenous if and only if $E_1[p] \simeq E_2[p]$ as $G_{\mathbf{Q}}$-modules.*

This conjecture appears to be difficult. It is not even clear what the best value M might be, assuming it exists. (Calculations of Cremona [Cr] based on his complete tables of elliptic curves over $\mathbf{Q}$ of conductor $\leq 30,000$ show that necessarily $M > 13$.) We mention Conjecture 3.18 here because it implies strong results about ternary diophantine equations analogous to Fermat's Last Theorem, thanks to the methods explained in Chapter 4.

4 Fermat curves

The purpose of this section is to discuss the Fermat curves

$$F_n : x^n + y^n = z^n,$$

and the proof of Fermat's Last Theorem, that these curves have no nontrivial rational points when $n \geq 3$. Fermat's Last Theorem has the same flavour as Mazur's Theorem 3.1, since it determines all of the rational points in a naturally arising infinite collection of algebraic curves. Although Fermat curves are simpler to write down as explicit equations, they do not admit a direct moduli interpretation, and therefore turn out to be harder to analyse than modular curves. In fact, the eventual solution of Fermat's Last Theorem is based on an elaborate *reduction* of the study of Fermat curves to diophantine questions about modular curves. In particular, Theorem 3.1—its statement, as well as some of the techniques used in its proof—play an essential role in the proof of Fermat's Last Theorem.

4.1 Motivation for the strategy

Hugo Chapdelaine's article in these proceedings discusses the more general problem of classifying the primitive integer solutions of the generalised Fermat equation

$$x^p + y^q + z^r = 0, \tag{14}$$

and sets up a "dictionary" relating

$$\left\{ \begin{array}{l} \text{Strategies for studying} \\ \text{primitive solutions of} \\ x^p + y^q + z^r = 0 \end{array} \right\} \quad \text{and} \quad \left\{ \begin{array}{l} \text{Unramified coverings} \\ \text{of } \mathbb{P}_1 - \{0, 1, \infty\} \\ \text{of signature } (p, q, r) \end{array} \right\}.$$

The idea explained in [Chap] is that, given an unramified covering

$$\pi : X \longrightarrow \mathbb{P}_1 - \{0, 1, \infty\},$$

one can study (14) by

1. Attempting to classify the possible fibers of π over the points in

$$\Sigma_{p,q,r} = \left\{ \frac{a^p}{c^r}, \quad \text{with } a^p + b^q = c^r \quad \text{and } (a, b, c) \text{ primitive} \right\} \subset \mathbb{P}_1(\mathbf{Q}).$$

   Since the ramification in these fibers is bounded, there can only be finitely many, by the Hermite–Minkowski theorem. In particular, the compositum of these extensions is a finite extension of $\mathbf{Q}$, denoted L.

2. Understanding the L-rational points on the curve X.

To apply these principles to the classical Fermat equation, one is led to consider unramified coverings of $\mathbb{P}_1 - \{0, 1, \infty\}$ of signature (p, p, p). Among such coverings, one finds:

1. the Fermat curve $F_p : x^p + y^p = z^p$ itself, equipped with the natural projection $\pi : (x, y, z) \mapsto t = \frac{x^p}{z^p}$ of degree p^2. For this π, it is clear that $\pi(F_p(\mathbf{Q})) \supset \Sigma_{p,p,p}$; but this merely leads to a tautological reformulation of the original question.

2. There are many coverings of signature (p, p, p) with solvable Galois groups, and studying these leads to classical attempts to prove Fermat's Last Theorem by factoring $x^p + y^p$ over the p-th cyclotomic fields. This circle of ideas led to many interesting questions on cyclotomic fields and their class groups, but has proved unsuccessful (so far) in settling Fermat's Last Theorem.

A third type of covering is obtained from modular curves. These coverings, which are nonsolvable, arises naturally in light of the strong results obtained in Section 3.

More precisely, let $Y(n)$ be the open modular curve that classifies elliptic curves with full level n structure, i.e., pairs

$$(E, \iota : \mathbf{Z}/n\mathbf{Z} \times \mu_n \longrightarrow E[n])$$

where ι is an identification which induces an isomorphism

$$\wedge^2 \iota : \wedge^2(\mathbf{Z}/n\mathbf{Z} \times \mu_n) = \mu_n \simeq \wedge^2(E[n]) = \mu_n.$$

Over the base $Z = \mathbf{Z}[1/2]$, the curve $Y(2)_Z$ is identified with

$$\mathrm{Spec}(Z[\lambda, 1/\lambda, 1/(\lambda - 1)]) = (\mathbb{P}_1 - \{0, 1, \infty\})_Z,$$

where λ is the parameter that occurs in the Legendre family

$$E_\lambda : y^2 = x(x - 1)(x - \lambda).$$

The natural covering map $\pi : Y(2p) \longrightarrow Y(2)$ is an unramified covering of signature (p, p, p), with Galois group $\mathbf{SL}_2(\mathbf{Z}/p\mathbf{Z})/\langle \pm 1 \rangle$. Given $\lambda = \frac{a^p}{c^p} \in \Sigma_{p,p,p}$, the fiber $\pi^{-1}(\lambda)$ is contained in the field of definition of the field of p-division points of the elliptic curve

$$y^2 = x(x - 1)(x - a^p/c^p). \tag{15}$$

In practice, it is more convenient to work with the closely related *Frey curve*,

$$E_{a,b,c} : y^2 = x(x - a^p)(x - c^p),$$

which differs from (15) by a quadratic twist, and replace the study of the fiber of π at λ with considerations involving the *mod p Galois representation*

$$\rho_{a,b,c} : G_\mathbf{Q} \longrightarrow \mathrm{Aut}(E_{a,b,c}[p]) \simeq \mathbf{GL}_2(\mathbf{Z}/p\mathbf{Z}).$$

We normalise (a, b, c) so that $a \equiv 3 \pmod 4$ and c is even. (This can always be done, by permuting a, b and c and changing their signs if necessary.) With this normalisation, the minimal discriminant, conductor, and j-invariant associated to $E_{a,b,c}$ are

$$\Delta = 2^{-8}(abc)^{2p}, \qquad N = \prod_{\ell | abc} \ell, \qquad j = \frac{2^8(b^{2p} + a^p c^p)^3}{(abc)^{2p}}. \tag{16}$$

In particular, the elliptic curve $E_{a,b,c}$ is *semistable*: it has either good or (split or nonsplit) multiplicative reduction at all primes. (The reader may wish to consult Section 2 of the article by Pierre Charollois in this proceedings volume, which discusses the local invariants of Frey curves in greater detail.)

4.2 Galois representations associated to Frey curves

The following theorem states the main *local properties* of the Galois representation $\rho_{a,b,c}$.

Theorem 4.1. *The representation $\rho = \rho_{a,b,c}$ has the following properties.*

(a) *It is unramified outside 2 and p;*

(b) *The restriction of ρ to a decomposition group D_2 at 2 is of the form*

$$\rho_{a,b,c}|_{D_2} = \begin{pmatrix} \chi_{\mathrm{cyc}}\psi & \kappa \\ 0 & \psi^{-1} \end{pmatrix},$$

where $\chi_{\mathrm{cyc}} : G_{\mathbf{Q}_2} \longrightarrow (\mathbf{Z}/p\mathbf{Z})^{\times}$ is the mod p cyclotomic character, and ψ is an unramified character of order 1 or 2.

(c) *The restriction of ρ to D_p comes from the Galois action on the points of a finite flat group scheme over $\mathbf{Z}_p$.*

Proof. (a) Let $\ell \neq 2, p$ be a prime. The analysis of the restriction of $\rho = \rho_{a,b,c}$ to D_ℓ can be divided into three cases:

Case 1: The prime ℓ does not divide abc. In that case, it is a prime of good reduction for $E_{a,b,c}$, and the action of D_ℓ on $E_{a,b,c}[p]$ is therefore unramified, by the criterion of Néron–Ogg–Shafarevich.

Case 2: The prime ℓ divides abc. It is therefore a prime of *multiplicative reduction* for $E_{a,b,c}$. Hence the curve $E_{a,b,c}$, or a twist of it over the unramified quadratic extension of $\mathbf{Q}_\ell$, is isomorphic to the Tate curve $\mathbb{G}_m/q_\ell^{\mathbf{Z}}$ over $\mathbf{Q}_\ell$. More precisely, replacing $E_{a,b,c}$ by its twist if necesary, we have an identification which respects the action of $G_{\mathbf{Q}_\ell}$ on both sides:

$$E(\bar{\mathbf{Q}}_\ell) \simeq \bar{\mathbf{Q}}_\ell^{\times}/\langle q_\ell \rangle, \tag{17}$$

where $q_\ell \in \mathbf{Q}_\ell^{\times}$ is the ℓ-adic Tate period, which is obtained by reverting the power series with integer coefficients

$$j = \frac{1}{q} + 744 + 196884q + \cdots$$

that expresses j in terms of q, to obtain a power series

$$q = \mathrm{Tate}(1/j) = 1/j + \cdots \quad \in (1/j)\mathbf{Z}[[1/j]]^{\times}.$$

In particular, note that, by (16),

$$\mathrm{ord}_\ell(q_\ell) = \mathrm{ord}_\ell(1/j) = \mathrm{ord}_\ell(\Delta) \equiv 0 \pmod{p}. \tag{18}$$

The explicit description of the $G_{\mathbf{Q}_\ell}$-module $E(\bar{\mathbf{Q}}_\ell)$ given by (17) implies that

$$E(\bar{\mathbf{Q}}_\ell)[p] \simeq \{\zeta_p^a q_\ell^{b/p}, \quad 0 \le a, b \le p-1\},$$

where ζ_p is a primitive pth root of unity in $\bar{\mathbf{Q}}_\ell^\times$. In the basis $(\zeta_p, q_\ell^{1/p})$ for $E[p]$, the restriction of $\rho = \rho_{a,b,c}$ to D_ℓ can be written as

$$\rho(\sigma) = \begin{pmatrix} \chi_{\mathrm{cyc}}(\sigma)\psi(\sigma) & \kappa(\sigma) \\ 0 & \psi^{-1}(\sigma) \end{pmatrix}, \tag{19}$$

where χ_{cyc} is the p-th cyclotomic character giving the action of D_ℓ on the p-th roots of unity, and ψ is an unramified character of order at most 2 (which is trivial precisely when E has split multiplicative reduction at ℓ.) Furthermore, the cocycle κ is unramified, by (18): this is because the extension $\mathbf{Q}_\ell(\zeta_p, q_\ell^{1/p})$ through which $\rho_{a,b,c}|_{G_{\mathbf{Q}_\ell}}$ factors is unramified. Part (a) of Theorem 4.1 follows.

(b) When $\ell = 2$, the elliptic curve $E_{a,b,c}$ has multiplicative reduction at 2, and hence is identified with a Tate curve over $\mathbf{Q}_2$. The result then follows from (19) with $\ell = 2$.

(c) When $\ell = p$ does not divide abc, the Galois representation $\rho_{a,b,c}$ arises from the p-torsion of an elliptic curve with good reduction at p, and is hence from a finite flat group scheme over $\mathbf{Z}_p$. In the case where $p | abc$ (which corresponds to what was known classically as the *second case* of Fermat's Last Theorem) one has a similar conclusion: essentially, the condition $\mathrm{ord}_p(q_p) \equiv 0 \pmod{p}$ limits the ramification of $\rho_{a,b,c}$ at p and implies that $E_{a,b,c}[p]$ extends to a finite flat group scheme over $\mathbf{Z}_p$, in spite of the fact that $E_{a,b,c}$ itself does not have a smooth model over $\mathbf{Z}_p$. $\qquad\square$

The following theorem gives a *global* property of the representation $\rho_{a,b,c}$.

Theorem 4.2 (Mazur). *The Galois representation $\rho_{a,b,c}$ is irreducible.*

Proof. This follows (at least when p is large enough) from Theorem 3.15. We will now give a self-contained proof which rests on the ideas developed in the proof of Theorem 3.4.

Suppose that $\rho_{a,b,c}$ is reducible. Then $E = E_{a,b,c}$ has a rational subgroup C of order p, and the pair (E, C) gives rise to a rational point x on the modular curve $X_0(p)$. Let $\ell \neq p$ be an odd prime that divides abc. Then E has multiplicative reduction at ℓ. Therefore, the point x reduces to one of the cusps 0 or ∞ of $X_0(p)$ modulo ℓ. It can be asssumed without loss of generality that x reduces to ∞, as in Step 2 of the proof of Theorem 3.4. Now recall the natural projection $\Phi_e : J_0(p) \longrightarrow J_e(p)$ of $J_0(p)$ to its winding quotient $J_e(p)$, and the resulting map $j_e : X_0(p) \longrightarrow J_e(p)$. The element $j_e(x)$ belongs to the formal group $J_e^1(p)(\mathbf{Q}_\ell)$, which is torsion-free, and to $J_e(p)(\mathbf{Q})$, which is torsion by Theorem 3.12. Hence $j_e(x) = 0$. We now use the fact that j_e is a formal immersion to deduce that $x = \infty$, as in Step 4 of the proof of Theorem 3.4 (with 3 replaced by ℓ). $\qquad\square$

Remark 4.3. The importance of the diophantine study of modular curves described in Chapter 3 in the proof of Fermat's Last Theorem, via Theorem 4.2, cannot be overemphasised. It is sometimes underplayed in expositions of Fermat's Last Theorem which tend to focus on the ingredients that were supplied later.

Thanks to Theorems 4.1 and 4.2, Fermat's Last Theorem is now reduced to the problem of "classifying" the irreducible two-dimensional mod p representations satisfying the strong restrictions on ramification imposed by Theorem 4.1—or in some sense, to make Theorem 1.1 precise for the class of extensions of $\mathbf{Q}$ arising from such representations. The control we have over questions of this type (which in general seem very hard) arises from the deep and largely conjectural connection that is predicted to exist between Galois representations and *modular forms*.

4.3   Modular forms and Galois representations

Let $f = \sum_n a_n q^n$ be a newform in $S_2(N, \mathbf{C})$. Let K_f denote as before the finite extension of $\mathbf{Q}$ generated by the Fourier coefficients of f, so that f belongs to $S_2(N, K_f)$. The Fourier coefficients of f belong to the ring $\mathcal{O}_f$ of integers of K_f. Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_f$ and let $K_{f,\mathfrak{p}}$ denote the completion of K_f at $\mathfrak{p}$.

Theorem 4.4. *There exists a Galois representation*

$$\rho_{f,\mathfrak{p}} : G_{\mathbf{Q}} \longrightarrow \mathbf{GL}_2(K_{f,\mathfrak{p}})$$

such that

1. *The representation $\rho_{f,\mathfrak{p}}$ is unramified ouside Np.*

2. *The characteristic polynomial of $\rho_{f,\mathfrak{p}}(\mathrm{Frob}_\ell)$ is equal to $x^2 - a_\ell x + \ell$, for all primes ℓ not dividing p.*

3. *The representation $\rho_{f,\mathfrak{p}}$ is odd, i.e., the image of complex conjugation has eigenvalues 1 and -1.*

Sketch of proof. Let A_f be the abelian variety quotient of $J_0(N)$ associated to f by the Eichler–Shimura construction (Theorem 3.6). Its endomorphism ring $\mathrm{End}_{\mathbf{Q}}(A_f)$ contains $\mathbf{T}/I_f$, which is an order in K_f. In this way, the Galois representation $V_p(A_f)$ is equipped with an action of $K_f \otimes \mathbf{Q}_p$ which commutes with the action of $G_{\mathbf{Q}}$. Let

$$V_{f,\mathfrak{p}} = V_p(A_f) \otimes_{K_f} K_{f,\mathfrak{p}}.$$

It is a two-dimensional $K_{f,\mathfrak{p}}$-vector-space, equipped with a continuous linear action of $G_{\mathbf{Q}}$. The fact that it has the desired properties, particularly property 2, is a consquence of the Eichler–Shimura congruence that was used to prove the equality of L-series given in Theorem 3.10. See Chapter 2 of [Da2] for further details and references. $\qquad\square$

4.4 Serre's conjecture

Modular forms can also be used to construct two-dimensional representations of $G_{\mathbf{Q}}$ over finite fields. More precisely, let $\mathcal{O}_{f,\mathfrak{p}}$ be the ring of integers of $K_{f,\mathfrak{p}}$. Since $G_{\mathbf{Q}}$ is compact and acts continuously on $V_{f,\mathfrak{p}}$, it preserves an $\mathcal{O}_{f,\mathfrak{p}}$-stable sublattice $V_{f,\mathfrak{p}}^0 \subset V_{f,\mathfrak{p}}$ of rank two over $\mathcal{O}_{\mathfrak{p},p}$. Let $\mathbb{F}_{\mathfrak{p}} := \mathcal{O}_{f,\mathfrak{p}}/\mathfrak{p}$ be the residue field of $\mathcal{O}_f$ at $\mathfrak{p}$. The action of $G_{\mathbf{Q}}$ on the two-dimensional $\mathbb{F}_{\mathfrak{p}}$-vector space $W_{f,\mathfrak{p}} := V_{f,\mathfrak{p}}^0/\mathfrak{p}V_{f,\mathfrak{p}}^0$ gives rise to a two-dimensional mod $\mathfrak{p}$ representation

$$\bar{\rho}_{f,\mathfrak{p}} : G_{\mathbf{Q}} \longrightarrow \mathbf{GL}_2(\mathbb{F}_{\mathfrak{p}}).$$

Like its $\mathfrak{p}$-adic counterpart, this representation is unramified outside of pN and also satisfies parts 2 and 3 of Theorem 4.4.

In [Se2], Serre associates to *any* two-dimensional Galois representation

$$\rho : G_{\mathbf{Q}} \longrightarrow \mathbf{GL}_2(\mathbb{F}) \tag{20}$$

with coefficients in a finite field $\mathbb{F}$ two invariants $N(\rho)$ and $k(\rho)$, called the *Serre conductor* and *Serre weight* of ρ, respectively. The Serre conductor

$N(\rho)$ is only divisible by primes distinct from the characteristic of $\mathbb{F}$ at which ρ is ramified. When $\rho = \bar{\rho}_{f,\mathfrak{p}}$ arises from a modular form, the Serre conductor $N(\rho)$ always divides (but is not necessarily equal to) the level N of f. In particular, using parts 1 and 2 of Theorem 4.1, one can show that

$$N(\rho_{a,b,c}) = 2. \tag{21}$$

The recipe for defining $k(\rho)$ is somewhat more involved, but depends only on the restriction of ρ to the decomposition group (in fact, the inertia group) at p. It will suffice, for the purposes of this survey, to note that when ρ arises from the p-division points of a finite flat group scheme over $\mathbf{Z}_p$, then Serre's recipe gives $k(\rho) = 2$. Hence, by part 3 of Theorem 4.1,

$$k(\rho_{a,b,c}) = 2. \tag{22}$$

In [Se2], Serre conjectures that *any* odd irreducible two-dimensional mod p Galois representation ρ as in (20) necessarily arises from an appropriate modular form mod p of weight $k(\rho)$ and level $N(\rho)$. This conjecture has recently been proved by Khare and Wintenberger (cf. Theorem 1.2 of [KW]) in the case where $N(\rho_{a,b,c})$ is odd, and follows in the general case from a similar method, using a result of Kisin [Ki].

Theorem 4.5. *Let ρ be an odd, irreducible two-dimensional mod p representation of $G_{\mathbf{Q}}$. Then there exists an eigenform f of weight $k(\rho)$ on $\Gamma_1(N(\rho))$, and a prime $\mathfrak{p}|p$ of the field K_f such that ρ is isomorphic to $\bar{\rho}_{f,\mathfrak{p}}$ as a representation of $G_{\mathbf{Q}}$.*

Proof of Fermat's Last Theorem. Let (a, b, c) be a primitive nontrivial solution of Fermat's equation $x^p + y^p = z^p$, and consider the Galois representation $\rho = \rho_{a,b,c}$ associated to the p-division points of the associated Frey curve. It follows from Theorem 4.2 that ρ is an odd, irreducible mod p representation of $G_{\mathbf{Q}}$. Its Serre conductor and weight are $N(\rho) = 2$ and $k(\rho) = 2$ by (21) and (22). Therefore Theorem 4.5 implies the existence of a nontrivial cusp form in $S_2(2, \mathbf{C})$. This leads to a contradiction, because there are no such cusp forms: the modular curve $X_0(2)$ has genus zero and hence has no regular differentials. This contradiction implies Fermat's Last Theorem.

4.5 The Shimura–Taniyama conjecture

Historically, the proof of Theorem 4.5 by Khare and Wintenberger came almost 10 years after Wiles proved Fermat's Last theorem. In essence, Wiles

proved enough of Theorem 4.5 to cover the Galois representations $\rho_{a,b,c}$ arising from hypothetical solutions of Fermat's equation.

More precisely, the articles [Wi] and [TW] proved the following result, known as the Shimura-Taniyama conjecture for semistable elliptic curves:

Theorem 4.6. *Let E be a semistable elliptic curve over $\mathbf{Q}$ of conductor N. Then there is a normalised eigenform f in $S_2(N, \mathbf{Z})$ such that $V_p(E)$ is isomorphic to $V_p(A_f)$.*

The proof of this theorem—or even, an outline of its main ideas—is beyond the scope of this survey. For details the reader is invited to consult [DDT] for example.

Theorem 4.6 implies that $\rho_{a,b,c}$ arises from a modular form in $S_2(N, \mathbf{C})$, where $N = \prod_{\ell | abc} \ell$. The Serre conjecture (Theorem 4.5) for $\rho_{a,b,c}$ then follows from an earlier theorem of Ribet (which also played an important role in Wiles' original approach to proving Theorem Theorem 4.6.)

Theorem 4.7. *Suppose that p is odd. Let ρ be an irreducible mod p Galois representation which arises from a modular form (of some weight and level). Then it also arises from an eigenform of weight $k(\rho)$ and level $N(\rho)$.*

Aside from the fact that it proves Fermat's Last Theorem, the importance of Theorem 4.6 can be justified on several other levels.

Firstly, the methods used to prove Theorem 4.6 were subsequently refined in [BCDT] to prove the full Shimura–Taniyama conjecture: all elliptic curves over $\mathbf{Q}$ are modular. This result is of great importance in understanding the arithmetic of elliptic curves over $\mathbf{Q}$, as will be explained in more detail in the next chapter.

Secondly—and this is a theme that we will not begin to do justice to, because it falls outside the scope of this survey—Wiles' method for proving Theorem 4.6 has led to a general, flexible method for establishing relationships between Galois representations and modular forms. It was by building on these techniques that Khare and Wintenberger proved Serre's conjecture (Theorem 4.5). Over the years, many other conjectures of this type have been proved building on the proof of Theorem 4.6: for instance, special cases of Artin's conjecture relating representations with finite image to modular forms of weight one (cf. for example [Tay2] and the references contained therein), and a proof of the Sato–Tate conjecture for elliptic curves over $\mathbf{Q}$ in [Tay1], [HST], and [CHT].

Closer to the themes that have been developed in this chapter, we mention a natural generalisation of Theorem 4.6 concerning abelian varieties *of* $\mathbf{GL}_2$-*type*. An abelian variety A over $\mathbf{Q}$ is said to be of $\mathbf{GL}_2$-type if $\mathrm{End}_{\mathbf{Q}}(A) \otimes \mathbf{Q}$ contains a field K with $[K : \mathbf{Q}] = \dim(A)$. The reason for this terminology is that such an A gives rise, for each prime ideal $\mathfrak{p}$ of K, to a two-dimensional Galois representation

$$\rho_{A,\mathfrak{p}} : G_{\mathbf{Q}} \longrightarrow \mathbf{GL}_2(K_{\mathfrak{p}})$$

arising from the action of $G_{\mathbf{Q}}$ on $V_p(A) \otimes_K K_{\mathfrak{p}}$. The abelian varieties A_f arising from the Eichler-Shimura construction are examples of abelian varieties of $\mathbf{GL}_2$-type. A conjecture of Fontaine and Mazur predicts that all abelian varieties of $\mathbf{GL}_2$-type arise as quotients of jacobians of modular curves. It can be shown that this generalisation of the Shimura-Taniyama conjecture follows from Theorem 4.5. (Cf. for example [Se2] or the introduction of [Ki].)

4.6   A summary of Wiles' proof

There are some enlightening parallels to be drawn between the proof of Fermat's Last Theorem and Faltings' proof of the Mordell conjecture as summarised in Section 2.8. Like Faltings' proof, the proof of Fermat's Last theorem is based on a sequence of maps, resulting in a sequence of transformations leading from the original problem to questions about other types of structures, such as Galois representations, and ultimately modular forms. These reductions are summarised in the diagram below.

$$
\left\{
\begin{array}{l}
\text{Integer solutions} \\
\text{(a,b,c) of} \\
x^p + y^p = z^p
\end{array}
\right\}
\xrightarrow{R_1}
\left\{
\begin{array}{l}
\text{Semistable elliptic curves} \\
\text{of conductor } N = abc \\
\text{and discriminant } 2^{-8}(abc)^{2p}
\end{array}
\right\}
$$

$$
\xrightarrow{R_4}
\left\{
\begin{array}{l}
\text{Irreducible galois representations} \\
\rho : G_{\mathbf{Q}} \longrightarrow \mathbf{GL}_2(\mathbb{F}_p) \\
\text{with } N(\rho) = 2 \text{ and } k(\rho) = 2.
\end{array}
\right\}
$$

$$
\xrightarrow{R_5}
\left\{
\begin{array}{l}
\text{Cusp forms in} \\
S_2(2, \mathbf{Z}/p\mathbf{Z}).
\end{array}
\right\}.
$$

1. The map R_1 is defined via the Frey curve, and is reminiscent of the Kodaira–Parshin construction of Section 2.2. An important difference is that the set of primes of bad reduction of the Frey curve associated to (a, b, c) is *not* bounded independently of (a, b, c). In fact, the set of

primes of bad reduction for E consists *exactly* of the primes that divide abc.

2. The map R_4 plays a role analogous to the passage to the ℓ-adic representations in Faltings' proof. An important difference here is that we consider mod p representations (with coefficients in a finite field) rather than p-adic representations. The justification for doing this is given by Theorem 4.1, which shows that the mod p representation ρ attached to $E_{a,b,c}$ has bounded ramification. Note that the corresponding p-adic representation would be ramified precisely at the primes dividing $pabc$. It is an exercise to show that the map R_4 is finite-to-one when $p \geq 7$. (Hint: use Faltings' theorem 2.1, and the fact that $X(p)$ has genus > 1 when $p \geq 7$.) It is even believed that R_4 is injective once p is large enough (cf. Conjecture 3.18), but this assertion is still unproved.

3. The map R_5 is a new ingredient that has no counterpart in Faltings' proof of Mordell's conjecture, and exploits the deep "dictionary" that is expected to exist between Galois representations and modular forms—in this case, the Serre conjecture proved by Khare and Wintenberger.

4. The final step in the argument exploits the fact that there are no modular forms of weight two and level two. This last point may seem like a "lucky accident" in the proof of Fermat's Last Theorem. Indeed the presence of modular forms of higher level presents an obstruction for the method based on Frey curves to yield results on more general ternary diophantine equations of Fermat type. However, see the article by Charollois in this volume [Char], where a refinement of the techniques described in this chapter leads to a strikingly general result on the generalised Fermat equation $ax^p + by^p + cz^p = 0$.

Remark 4.8. One of the consequences of Conjecture 3.18 is that the generalised Fermat equation $ax^n + by^n + cz^n = 0$ (with a, b, c fixed) has no primitive integer solutions (x, y, z) with $xyz \neq 0, \pm 1$, once n is large enough. (The reader who masters the ideas in the article by Pierre Charollois in this proceedings volume will be able to prove this assertion.)

5 Elliptic curves

After surveying curves of genus > 1, we turn our attention to curves of genus 1. A projective curve of genus 1 over a field K, equipped with a distinguished K-rational point over that field, is endowed with a natural structure of a commutative algebraic group over K for which the distinguished element becomes the identity. Such a curve is called an *elliptic curve.*

If E is an elliptic curve defined over a number field K, then the Mordell–Weil Theorem (cf. Theorem 7 of the introduction) asserts that the group $E(K)$ of K-rational points on E is *finitely generated.* Let $r(E, K)$ denote the rank of this finitely generated abelian group. Many of the important questions in the theory of elliptic curves revolves around calculating this invariant, and understanding its behaviour as E or K vary.

Question 5.1. *Is there an effective algorithm to calculate $r(E, K)$, given E and K?*

Showing that Fermat's method of descent yields such an effective algorithm is intimately connected to the Shafarevich–Tate conjecture asserting the finiteness of the Shafarevich–Tate group $Ш(E/K)$ of E/K.

One can also fix a base field (the most natural, and interesting, case being the case where $K = \mathbf{Q}$) and ask

Question 5.2. *Is the rank $r(E, K)$ unbounded, as E ranges over all elliptic curves defined over K?*

One can also fix an elliptic curve E and enquire about the variation of $r(E, K)$ as K ranges over different number fields.

The main tool available at present to study $r(E, K)$ is the relationship between the rank and the Hasse–Weil L-series predicted by the Birch and Swinnerton-Dyer conjecture (Conjecture 3.9).

Assume that E is an elliptic curve over $\mathbf{Q}$. Thanks to Theorem 4.6 (and its extension to all elliptic curves over $\mathbf{Q}$ given in [BCDT]), the Hasse–Weil L-series $L(E, s)$ is equal to $L(f, s)$ for some newform f of weight two. In particular, $L(E, s)$ has an analytic continuation to the entire complex plane, and a functional equation.

The main result we will discuss in this chapter is the following:

Theorem 5.3. *Let E be an elliptic curve over $\mathbf{Q}$, and let $L(E, s)$ be its Hasse–Weil L-series. If $r := \operatorname{ord}_{s=1} L(E, s) \leq 1$, then $r(E, \mathbf{Q}) = r$ and $Ш(E/\mathbf{Q})$ is finite.*

5.1 Modular parametrisations

Let E be an elliptic curve over $\mathbf{Q}$ of conductor N. Recall the modular curve $X_0(N)$ that was introduced in Section 3.1. The following theorem, which produces a dominant rational map from such a curve to E, plays a crucial role in the proof of Theorem 5.3.

Theorem 5.4. *There exists a nonconstant map of curves over* $\mathbf{Q}$

$$\varphi : X_0(N) \longrightarrow E.$$

Proof. By Theorem 4.6, there is a normalised eigenform f in $S_2(N, \mathbf{Z})$ satisfying $L(E, s) = L(f, s)$. Let A_f be the quotient of $J_0(N)$ associated to f via the Eichler–Shimura construction. By assumption, the Galois representations $V_p(E)$ and $V_p(A_f)$ are isomorphic. Hence the isogeny conjecture (Theorem 2.18) implies the existence of an isogeny $\alpha : A_f \longrightarrow E$ defined over $\mathbf{Q}$. Composing such an isogeny with the natural surjective morphism $J_0(N) \longrightarrow A_f$ gives a nonconstant map $\Phi : J_0(N) \longrightarrow E$. The modular parametrisation φ is defined by setting $\varphi(x) := \Phi((x) - (\infty))$. $\qquad\square$

It is useful to describe briefly how the modular parametrisation φ can be computed analytically. The pull-back $\varphi^*(\omega_E)$ is a nonzero rational multiple of the diferential form

$$\omega_f := 2\pi i f(\tau) d\tau = \sum_{n=1}^{\infty} a_n q^n \frac{dq}{q}.$$

Denote by Λ_f the collection of periods of ω_f (integrals of ω_f against smooth closed one-chains C in $X_0(N)(\mathbf{C})$):

$$\Lambda_f := \left\{ \int_C \omega_f, \text{ where } \partial C = 0 \right\}.$$

It is a lattice in $\mathbf{C}$, and $A_f(\mathbf{C}) \simeq \mathbf{C}/\Lambda_f$. Let us replace E by A_f, so that $\alpha = 1$. It is suggestive (for later generalisations) to view φ as a map

$$\varphi : \mathrm{Div}^0(X_0(N)) \longrightarrow E.$$

This map is defined on $\mathrm{Div}^0(X_0(N)(\mathbf{C}))$ by the rule

$$\varphi(\Delta) := \int_{\substack{C \text{ s.t.} \\ \partial(C) = \Delta}} \omega_f \quad (\mathrm{mod}\ \Lambda_f), \tag{23}$$

where the integral is taken over any smooth one-chain C whose boundary is Δ. The invariant $\varphi(\Delta) \in \mathbf{C}/\Lambda_f$ is viewed as a point on $E(\mathbf{C})$ via the Weierstrass uniformisation.

5.2   Heegner points

Perhaps the most important arithmetic application of the modular parametrisation arises from the fact that $X_0(N)$ is endowed with a systematic supply of algebraic points defined over abelian extensions of imaginary quadratic fields—the so-called CM-points. These points correspond, in the moduli interpretation of $X_0(N)$, to pairs (A, C) where A is an elliptic curve whose endomorphism ring $\mathcal{O} = \mathrm{End}(A)$ is an order in a quadratic imaginary field K. Such an elliptic curve is said to have *complex multiplication* by K, and the corresponding points on $X_0(N)$ are called *CM-points* attached to K. Let $\mathrm{CM}(K)$ denote the set of all CM points in $X_0(N)$ attached to K. It satisfies the following properties.

1. The set $\mathrm{CM}(K)$ is dense in $X_0(N)(\mathbf{C})$ (relative to the Zariski topology, and also the complex topology).

2. Let K^{ab} denote the maximal abelian extension of K. Then $\mathrm{CM}(K)$ is contained in $X_0(N)(K^{\mathrm{ab}})$.

3. Analytically, $\mathrm{CM}(K) = \Gamma_0(N)\backslash(\mathcal{H} \cap K)$.

Definition 5.5. The collection of points

$$HP(K) := \{\varphi(\Delta)\}_{\Delta \in \mathrm{Div}^0(\mathrm{CM}(K))} \subset E(K^{\mathrm{ab}})$$

is called the system of *Heegner points* on E attached to K.

The usefulness of Heegner points arises from two facts:

1. They can be related to L-series, thanks to the theorem of Gross–Zagier and its generalisations.

2. They can be used to bound Mordell–Weil groups and Shafarevich–Tate groups of elliptic curves, following a descent method that was discovered by Kolyvagin.

Heegner points and L-series.

For simplicity, suppose that the imaginary quadratic field K satisfies the following so-called *Heegner hypothesis*:

Hypothesis 5.6. *There exists a ideal $\mathcal{N}$ of norm N in $\mathcal{O}_K$ with cyclic quotient.*

This hypothesis is used to construct a distinguished element in HP(K). More precisely, let h denote the class number of K, and let H be its Hilbert class field. By the theory of complex multiplication, there are precisely h distinct (up to isomorphism over **C**) elliptic curves $A_1, \ldots, A_h$ having endomorphism ring equal to $\mathcal{O}_K$. The j-invariants of these curves belong to H, and are permuted simply transitively by the action of Gal(H/K). It is therefore possible to choose $A_1, \ldots, A_h$ in such a way that they are defined over H, and permuted by the action of Gal(H/K).

The pairs $(A_i, A_i[\mathcal{N}])$ (with $1 \leq i \leq h$) correspond to points P_i in $X_0(N)(H)$. Let

$$P_K := \varphi((P_1) + \cdots + (P_h) - h(\infty)) \in E(K). \tag{24}$$

The fact that the point P_K has an explicit moduli description makes it possible to establish some of its key properties. For example, let $\bar{P}_K$ denote the image of P_K under complex conjugation. Then it can be shown that

$$\bar{P}_K = wP_K \pmod{E(K)_{\text{tors}}}, \tag{25}$$

where $w \in \{\pm 1\}$ is the *negative* of the sign in the functional equation for $L(E, s) = L(f, s)$. (Cf. Chapter 3 of [Da2].) This provides a simple connection between the behaviour of P_K and the L-series $L(E, s)$.

We note that, in many cases where Hypothesis 5.6 is satisfied (for example, when all the primes dividing N are *split* in the quadratic imaginary field K), the sign in the functional equation for the Hasse–Weil L-series $L(E/K, s)$ is -1, so that $L(E/K, 1) = 0$. It then becomes natural to consider the first derivative $L'(E/K, 1)$ at the central critical point. The following theorem of Gross and Zagier establishes an explicit link between P_K and this quantity.

Theorem 5.7. *Let $\langle f, f \rangle$ denote the Petersson scalar product of f with itself, and let $h(P_K)$ denote the Néron–Tate canonical height of P_K on $E(K)$. There is an explicit nonzero rational number t such that*

$$L'(E/K, 1) = t \cdot \langle f, f \rangle \cdot h(P_K). \tag{26}$$

In particular, the point P_K is of infinite order if and only if $L'(E/K, 1) \neq 0$.

The proof of Theorem 5.7 given in [GZ] proceeds by a direct calculation in which both sides of (26) are computed explicitly, compared, and found to be equal.

Remark 5.8. Let P_n be a point in $\mathrm{CM}(K)$ corresponding to an elliptic curve with endomorphism ring equal to the order $\mathcal{O}_n$ of conductor n in K. Such a point can be defined over the ring class field H_n of K of conductor n, whose Galois group $G_n := \mathrm{Gal}(H_n/K)$ is canonically identified with the class group $\mathrm{Pic}(\mathcal{O}_n)$ by class field theory. If $\chi : G_n \longrightarrow \mathbf{C}^\times$ is a complex character, one can generalise (24) to define

$$P_\chi := \varphi \left(\sum_{\sigma \in G_n} \chi(\sigma) P_n^\sigma \right) \in E(H_n) \otimes \mathbf{C}. \tag{27}$$

A generalisation of Theorem 5.7 due to S. Zhang (cf. for example [Zh1], [Zh2], [Ho1] and [Ho2]) relates the height of P_χ to the derivative of the twisted L-series $L(E/K, \chi, s)$ at $s = 1$.

When $L'(E/K, 1) \neq 0$, the method of Heegner points gives an efficient method for producing a point of infinite order in $E(K)$. The following proposition asserts the existence of many K for which the L-series does not vanish.

Proposition 5.9. *Suppose that* $r := \mathrm{ord}_{s=1} L(E, s) \leq 1$. *Then there exist infinitely many quadratic imaginary fields K satisfying Hypothesis 5.6 for which*

$$\mathrm{ord}_{s=1}(L(E/K, s)) = 1.$$

The proof of this proposition is explained in [MM].

Heegner points and arithmetic: Kolyvagin's descent

Theorem 5.7 implies that if $L'(E/K, 1) \neq 0$, then P_K is of infinite order and hence $r(E, K) \geq 1$. The following theorem of Kolyvagin gives a bound in the other direction as well.

Theorem 5.10 (Kolyvagin). *Suppose that P_K is of infinite order in $E(K)$. Then $r(E, K) = 1$, and $Ш(E/K)$ is finite.*

For a proof of this theorem, see [Gr] or Chapter 10 of [Da2]. Let us just mention here that Kolyvagin's proof makes essential use of the fact that point P_K does not come alone, but rather is part of a norm-compatible system of points in $E(K^{\mathrm{ab}})$ arising from the (infinite) collection of points in $\mathrm{HP}(K)$. These points are used to construct global cohomology classes in $H^1(K, E[p])$ whose local behaviour can be controlled precisely and related to P_K. Under the assumption that P_K is of infinite order, this system of ramified cohomology classes is enough to bound the p-Selmer group of E/K and show that $r(E, K) = 1$.

5.3 Proof of Theorem 5.3

We will now explain how the properties of P_K and $\mathrm{HP}(K)$ described in the previous section can be combined to prove Theorem 5.3:

Proof of Theorem 5.3. Assume that $r \leq 1$. By Proposition 5.9, there is a quadratic imaginary field K satisfying Hypothesis 5.6, for which

$$\mathrm{ord}_{s=1}(L(E/K, s)) = 1.$$

Fix such a K, and consider the point P_K. Since $L'(E/K, 1) \neq 0$, Theorem 5.7 implies that P_K is of infinite order. Theorem 5.10 then shows that

$$r(E, K) = 1, \quad \text{and } Ш(E/K) \text{ is finite.}$$

Let E' denote the quadratic twist of E over K. We then have

$$1 = r(E, K) = r(E, \mathbf{Q}) + r(E', \mathbf{Q}).$$

To be able to ignore finer phenomena associated to torsion in $E(K)$, it is convenient to replace P_K by its image in $E(K) \otimes \mathbf{Q}$. Since $E(K) \otimes \mathbf{Q}$ is generated by P_K, it follows that

$$r(E, \mathbf{Q}) = \begin{cases} 0 & \text{if } \bar{P}_K = -P_K, \\ 1 & \text{if } \bar{P}_K = P_K. \end{cases}$$

Theorem 5.3 now follows from (25). $\qquad \square$

Remark 5.11. The proof of Theorem 5.3 carries over with only minor changes when E is replaced by the abelian variety quotient A_f attached to an arbitrary eigenform f of weight 2 on $\Gamma_0(N)$. This is how Theorem 3.11 is proved:

$$L(A_f, 1) \neq 0 \implies A_f(\mathbf{Q}) \text{ is finite.}$$

The reader will recall the key role played by this theorem in the proof of Theorem 3.1 and (even more importantly) in Merel's proof of the uniform boundedness conjecture for elliptic curves explained in Marusia Rebolledo's article in these proceedings.

5.4 Modularity of elliptic curves over totally real fields

Because of the crucial role played by the system $\mathrm{HP}(K)$ in the proof of Theorem 5.3, it is natural to ask whether such structures are present in more general settings. For example, we would like to prove analogues of Theorem 5.3 for elliptic curves defined over number fields other than $\mathbf{Q}$. The class of number fields for which this program is best understood is the class of *totally real fields.*

More precisely, let F be a totally real field of degree n, and let E be an elliptic curve over F, of conductor N. Assume, for simplicity, that F has narrow class number one, so that in particular the conductor can now be viewed as a totally positive element of $\mathcal{O}_F$ rather than just an ideal.

The group $\Gamma_0(N; \mathcal{O}_F) \subset \mathbf{SL}_2(\mathcal{O}_F)$ is defined as the group of matrices that are upper triangular modulo N. The n distinct real embeddings $v_1, \ldots, v_n : F \longrightarrow \mathbf{R}$ of F allow us to view $\Gamma_0(N; \mathcal{O}_F)$ as a subgroup of $\mathbf{SL}_2(\mathbf{R})^n$. This subgroup acts discretely on the product $\mathcal{H}^n$, and the analytic quotient $\Gamma_0(N; \mathcal{O}_F) \backslash \mathcal{H}^n$ represents the natural generalisation of modular curves to this setting:

1. This quotient is identified with the complex points of an n-dimensional algebraic variety $Y_0(N; \mathcal{O}_F)$ defined over F. This variety can be compactified by adjoining a finite set of cusps, much as in the setting $n = 1$ of classical modular curves. A suitable desingularisation of the resulting projective variety is denoted $X_0(N; \mathcal{O}_F)$, and is called a *Hilbert modular variety.* Hilbert modular varieties are basic examples of higher dimensional Shimura varieties.

2. The variety $X_0(N; \mathcal{O}_F)$ is equipped with natural Hecke correspondences T_λ indexed by the prime ideals of $\mathcal{O}_F$.

3. These correspondences induce linear actions on the n-th deRham cohomology $H^n_{dR}(X_0(N; \mathcal{O}_F))$, and the eigenvalues of the Hecke operators are expected to encode the same type of arithmetic information as in the case where $F = \mathbf{Q}$.

To amplify this last point and make it more precise, we state the following generalisation of the Shimura–Taniyama conjecture to elliptic curves over F:

Conjecture 5.12. *Let E be an elliptic curve over F of conductor N. There exists a closed (in fact, holomorphic) differential form $\omega \in H^n_{dR}(X_0(N; \mathcal{O}_F))$*

satisfying

$$T_\lambda(\omega) = a_\lambda(E)\omega,$$

for all primes $\lambda \nmid N$ of $\mathcal{O}_F$.

Remark 5.13. In some cases, the methods of Wiles for proving the modularity of elliptic curves over $\mathbf{Q}$ have been extended to the setting of elliptic curves over totally real fields, and many cases of Conjecture 5.12 can be made uncondtional.

5.5 Shimura curves

When $n > 1$, the holomorphic differential form ω whose existence is predicted by Conjecture 5.12 cannot be used to directly produce an analogue of the modular parametrisation. In this sense, there is no immediate generalisation of Theorem 5.4, which plays such a crucial role in the construction of $\mathrm{HP}(K)$ when $n = 1$.

To extend the notion of Heegner points, it is necessary to introduce another generalisation of modular curves: the so-called *Shimura curves* associated to certain quaternion algebras over F.

A quaternion algebra B over F is said to be *almost totally definite* if

$$B \otimes_{v_1} \mathbf{R} \simeq M_2(\mathbf{R}), \qquad B \otimes_{v_j} \mathbf{R} \simeq \mathbb{H}, \text{ for } 2 \leq j \leq n.$$

We can associate to any order R in B a discrete subgroup

$$\Gamma := v_1(R^\times) \subset \mathbf{SL}_2(\mathbf{R}),$$

which acts discretely on $\mathcal{H}$ by Mobius transformations. When $F = \mathbf{Q}$ and $B = M_2(\mathbf{Q})$ is the split quaternion algebra, one recovers the analytic description of the modular curves $X_0(N)$. Otherwise, the analytic quotient $\Gamma \subset \mathcal{H}$ is a *compact* Riemann surface which can be identified with the complex points of an algebraic curve X possessing a canonical model over F. The curve X can be related (following a construction of Shimura) to the solution of a moduli problem and is also equipped with a supply $\mathrm{CM}(K) \subset X(K^{\mathrm{ab}})$ of CM points, associated this time to any quadratic totally imaginary extension K of F.

An elliptic curve E over F is said to be *arithmetically uniformisable* if there is a nonconstant map defined over F, generalising Theorem 5.4,

$$\varphi : \mathrm{Div}^0(X) \longrightarrow E.$$

The theory of Jacquet–Langlands gives a precise (partly conjectural) understanding of the class of elliptic curves that should be arithmetically uniformisable:

Theorem 5.14. *Let E be an elliptic curve over F which is not isogenous to any of its Galois conjugates. Then E is arithmetically uniformisable if and only if*

1. *E is modular in the sense of Conjecture 5.12;*

2. *E has potentially semistable reduction at a prime of F, or can be defined over a field F of odd degree.*

The collection $\mathrm{HP}(K) := \varphi(\mathrm{Div}^0(CM(K))) \subset E(K^{\mathrm{ab}})$, for suitable totally complex quadratic extensions K/F, can be used to obtain results analogous to Theorem 5.3 for elliptic curves over totally real fields. See [Zh1] where general results in this direction are obtained.

The articles [Vo] and [Gre1] in this volume describe Shimura curves and the associated parametrisations in more detail, from a computational angle. The article [Vo] discusses explicit equations for Shimura curves of low degree, and [Gre1] explains how to approach the numerical calculation of the systems $\mathrm{HP}(K)$ of Heegner points via p-adic integration of the associated modular forms, exploiting the theory of p-adic uniformisation of these curves due to Cerednik and Drinfeld.

5.6 Stark–Heegner points

Heegner points arising from Shimura curve parametrisations do not completely dispell the mystery surrounding the Birch and Swinnerton-Dyer conjecture for (modular) elliptic curves over totally real fields, since (even assuming the modularity Conjecture 5.12) there remain elliptic curves over F that are *not* arithmetically uniformisable.

The simplest example of such an elliptic curve is one that has everywhere good reduction over a totally real field F of even degree, and is not isogenous to any of its Galois conjugates. (More generally, one can also consider any quadratic twist of such a curve.) For these elliptic curves, there is at present very little evidence for the Birch and Swinnerton-Dyer conjecture, and in particular the analogue of Theorem 5.3 is still unproved when $\mathrm{ord}_{s=1}L(E/F,s) = 1$. (In the case where $L(E/F,1) \neq 0$, see the work of Matteo Longo [Lo].)

The notion of Stark–Heegner points represents an attempt to remedy this situation (albeit conjecturally) by exploiting the holomorphic differential n-form ω whose existence is predicted by Conjecture 5.12 rather than resorting to a Shimura curve parametrisation. We note that the holomorphic form ω can be written

$$\omega = f(\tau_1, \ldots, \tau_n)d\tau_1 \cdots d\tau_n,$$

where f is a (holomorphic) Hilbert modular form of parallel weight 2 on $\Gamma_0(N)$, satisfying, for all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$,

$$f\left(\frac{a_1\tau_1 + b_1}{c_1\tau_1 + d_1}, \ldots, \frac{a_n\tau_n + b_n}{c_n\tau_n + d_n}\right) = (c_1\tau_1 + d_1)^2 \cdots (c_n\tau_n + d_n)^2 f(\tau_1, \ldots, \tau_n).$$

We let any unit $\epsilon \in \mathcal{O}_F^\times$ act on $\mathcal{H}^n$ by the rule:

$$\epsilon \star \tau_j = \begin{cases} \epsilon_j \tau_j & \text{if } \epsilon_j > 0; \\ \epsilon_j \bar{\tau}_j & \text{if } \epsilon_j < 0. \end{cases}$$

For any subset $S \subset \{2, \ldots, n\}$ of cardinality m, we can then define a closed differential n-form of type $(n-m, m)$ by choosing a unit ϵ of $\mathcal{O}_F^\times$ which is negative at the places of S, and positive at the other embeddings, and setting

$$\omega_S = f(\epsilon \star \tau_1, \ldots, \epsilon \star \tau_n)d(\epsilon \star \tau_1) \ldots d(\epsilon \star \tau_n).$$

Finally we set

$$\omega_E := \sum_{S \subset \{2, \ldots, n\}} \omega_S.$$

The following conjecture is due to Oda [Oda].

Conjecture 5.15. *The set of periods*

$$\Lambda_f := \left\{ \int_C \omega_E \quad \text{for } C \in H_n(X_0(N, F)(\mathbf{C}), \mathbf{Z}) \right\} \subset \mathbf{C}$$

is a lattice which is commensurable with the period lattice of $E_1 := v_1(E)$.

Conjecture 5.15 can be used to define a generalisation of the modular parametrisation of equation (23). This map is defined on homologically trivial $(n-1)$-cycles on $X_0(N; \mathcal{O}_F)(\mathbf{C})$ by the rule

$$\varphi(\Delta) := \int_{\substack{C \text{ s.t.} \\ \partial(C) = \Delta}} \omega_f \quad (\mathrm{mod}\ \Lambda_f). \tag{28}$$

The interest of this generalisation of (23) is that it is possible to define a collection of distinguished topological $(n-1)$-cycles on which φ is conjectured to take algebraic values.

These cycles, which play the same role that Heegner divisors of degree zero played in the case where $n = 1$, are defined in terms of certain quadratic extensions K of F. Such a quadratic extension is said to be *almost totally real* if
$$K \otimes_{v_1} \mathbf{R} \simeq \mathbf{C}, \qquad K \otimes_{v_j} \mathbf{R} \simeq \mathbf{R} \oplus \mathbf{R} \quad \text{for } 2 \le j \le n.$$
Let $\iota : K \longrightarrow M_2(F)$ be an F-algebra embedding, and let $K_1^\times$ be the group of elements whose norm to F is equal to 1. The torus $v_1(\iota(K_1^\times))$ acts on $\mathcal{H}$ with a unique fixed point τ_1, and $\iota(K_1^\times)$ acts on the region $\{\tau_1\} \times \mathcal{H}^{n-1}$ without fixed points. The orbit of any point in this region under the action of $\iota((K \otimes_F \mathbf{R})_1^\times)$ is a real $(n-1)$-dimensional manifold $Z_\iota \subset \{\tau_1\} \times \mathcal{H}^{n-1}$ which is homeomorphic to $\mathbf{R}^{n-1}$. The group $G_\iota := \iota(K^\times) \cap \Gamma_0(N, \mathcal{O}_F)$ is an abelian group of rank $n - 1$, corresponding to a finite index subgroup of the group of relative units in K/F. Consider a fundamental region for the action of G_ι on Z_ι. The image Δ_ι of such a region in the Hilbert modular variety $X_0(N; \mathcal{O}_F)$ is a closed $(n - 1)$-cycle, which is topologically isomorphic to a real $(n - 1)$-dimensional torus.

Conjecture 5.16. *Assume that Δ_ι is homologically trivial. Then the point $\varphi(\Delta_\iota) \in E_1(\mathbf{C})$ is an algebraic point, and is in fact the image of a point in $E(K^{\mathrm{ab}})$ under any embedding $K^{\mathrm{ab}} \longrightarrow \mathbf{C}$ extending $v_1 : F \longrightarrow \mathbf{R}$.*

Remark 5.17. The original formulation of Conjecture 5.16 given in [DL] was phrased in terms of group cohomology. The definition of $\varphi(\Delta_\iota)$ used in Conjecture 5.16, which suggests an analogy between φ and higher Abel-Jacobi maps, was formulated only later, in [CD] (in a context where cusp forms are replaced by Eisenstein series; the elements $\varphi(\Delta_\iota)$ can then be related to Stark units). The equivalence between Conjecture 5.16 and the main conjecture of [DL] is explained in [CD].

Conjecture 5.16 can be formulated more precisely, in a way that makes a prediction about the fields of definition of the points $\varphi(\Delta_\iota)$. It is expected that the system of points
$$\mathrm{HP}(K) := \{\varphi(\Delta_\iota)\}_{\iota : K \longrightarrow M_2(F)},$$
as ι ranges over all possible embeddings, gives rise to an infinite collection of algebraic points in $E(K^{\mathrm{ab}})$ with properties similar to those of the system

of Heegner points defined in Section 5.2. Such a system of points (if its existence, and basic properties, could be established, a tall order at present!) would lead to a proof of Theorem 5.3 for all (modular) elliptic curves defined over totally real fields, not just those that are arithmetically uniformisable.

For more details on Conjecture 5.16, a more precise formulation, and numerical evidence, see Chapter 8 of [Da2], or [DL]. For an explanation of the relation between Conjecture 5.16 and the conjectures of [DL], see [CD].

The Stark-Heegner points attached to Hilbert modular forms that were defined and studied in [DL] and [CD] can be viewed as the *basic prototype* for the general notion of Stark-Heegner points. Here are some further variants that have been explored so far in the literature:

1. If E is an elliptic curve over $\mathbf{Q}$ of conductor $N = pM$ with $p \nmid M$, a p-adic analogue of the map φ of equation (28)—described in terms of group cohomology rather than singular cohomology, following the same approach and in [DL]—is defined in [Da1], by viewing E as uniformised by the "mock Hilbert surface"

$$\Gamma_0(M; \mathbf{Z}[1/p]) \backslash (\mathcal{H}_p \times \mathcal{H}),$$

where $\mathcal{H}_p := \mathbf{C}_p - \mathbf{Q}_p$ is the p-adic upper half plane, and $\Gamma_0(M; \mathbf{Z}[1/p])$ is the group of matrices in $\mathbf{SL}_2(\mathbf{Z}[1/p])$ which are upper-triangular modulo M. The resulting map φ associates a point in $P_\iota \in E(\bar{\mathbf{Q}}_p)$ to any embedding $\iota : K \longrightarrow M_2(\mathbf{Q})$ when K is a *real quadratic* field in which p is inert. The system $\{P_\iota\} \subset E(\bar{\mathbf{Q}}_p)$, as ι ranges over all embeddings of K into $M_2(\mathbf{Q})$, is expected to yield a system of points in $E(K^{\mathrm{ab}})$ with the same properties as the Heegner points attached to an imaginary quadratic base field. This construction is not expected to yield new cases of the Birch and Swinnerton-Dyer over the base field $\mathbf{Q}$—this conjecture is completely known when $\mathrm{ord}_{s=1} L(E, s) \leq 1$, thanks to Theorem 5.3. However, it would give new cases of this conjecture over certain abelian extensions of real quadratic fields, and, more importantly perhaps, it suggests an explicit analytic construction of class fields of real quadratic fields. For more details on Stark-Heegner points attached to real quadratic fields, see [Da1] or Chapter 9 of [Da2]. The article [DP] describes efficient algorithms for calculating the points $\varphi(\Delta_\iota)$, and uses them to gather numerical evidence for the conjectures of [Da1], while [BD] provides some theoretical evidence.

2. The article [Tr] formulates and tests numerically a Stark–Heegner construction that leads to conjectural systems of algebraic points on elliptic curves defined over a quadratic imaginary base field. The details of the construction of [Tr] are explained in the article [Gre2] by Matt Greenberg in this proceedings volume. We remark that there is not a single example of an elliptic curve E genuinely defined over such a field (i.e., which is not isogenous to its Galois conjugate) for which Theorem 5.3 (or even just the Shafarevich–Tate conjecture) has been proved.

References

[BCDT] Breuil, C., Conrad, B., Diamond, F., and Taylor, R., *On the modularity of elliptic curves over Q.* J.A.M.S. 14 (2001), 843-939.

[BD]   Bertolini, M. and Darmon, H., *The rationality of Stark-Heegner points over genus fields of real quadratic fields*, Annals of Mathematics, to appear.

[CD]   Charollois, P., and Darmon, H., *Arguments des unités de Stark et périodes de séries d'Eisenstein*, to appear.

[Chap] Chapdelaine, H., *Non-abelian descent and the generalized Fermat equation*, in this proceedings volume.

[Char] Charollois, P. *Generalized Fermat equations (d'après Halberstadt-Kraus)*, in this proceedings volume.

[CHT]   Clozel, L., Harris, M., and Taylor, R. *Automorphy for some ℓ-adic lifts of automorphic mod ℓ representations.*, to appear.

[CF]   Cassels, J.W.S. and Frohlich, A., Algebraic number theory, Academic Press, London (1967).

[CS]   Arithmetic Geometry, G. Cornell and J.H. Silverman eds. New York: Springer Verlag (1986)

[Cr]   Cremona, J., Private communication.

[Da1]   Darmon, H., *Integration on $\mathcal{H}_p \times \mathcal{H}$ and arithmetic applications.* Ann. of Math. (2) 154 (2001), no. 3, 589–639.

[Da2]   Darmon, H., Rational points on modular elliptic curves, CBMS Regional Conference Series in Mathematics, 101. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2004.

[DDT]   Darmon, H., Diamond,F., and Taylor, R., Fermat's Last Theorem. Current Developments in Mathematics 1, 1995, International Press, pp. 1-157. Reprinted in Elliptic curves, modular forms & Fermat's last theorem (Hong Kong, 1993), 2–140, International Press, Cambridge, MA, 1997.

[DG]   Darmon, H., and Granville, A., *On the equations $z^m = F(x,y)$ and $Ax^p + By^q = Cz^r$*. Bull. London Math. Soc. 27 (1995), no. 6, 513–543.

[DL]   Darmon, H., and Logan, A., *Periods of Hilbert modular forms and rational points on elliptic curves*. Int. Math. Res. Not. 2003, no. 40, 2153–2180.

[DP]   Darmon, H. and Pollack, R., *The efficient calculation of Stark-Heegner points via overconvergent modular symbols*, Israel Journal of Mathematics, 153 (2006), 319–354.

[DV]   Dasgupta, S., and Voight, J. *Some diophantine applications of Heegner points*, in this proceedings volume.

[De]   Deligne, P., *Preuve des conjectures de Tate et Shafarevitch (d'après G. Faltings)*, Séminaire Bourbaki, exposé 616, Volume 1983/84, Asterisque 121-122, 1985.

[El]   Elkies, N.D., *ABC implies Mordell*. Internat. Math. Res. Notices 1991, no. 7, 99–109.

[Fa]   Faltings, G., *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Inventiones Math **73**, 349-366 (1983). English translation in: Arithmetic Geometry, G. Cornell and J.H. Silverman eds. New York: Springer Verlag (1986)

[FW]   Faltings, G., Wüstholz, G., Grunewald, F., Schappacher, N. and Stuhler, U., Rational points. Third edition. Papers from the seminar held at the Max-Planck-Institut für Mathematik, Bonn/Wuppertal,

1983-1984. Aspects of Mathematics, E6. Friedr. Vieweg & Sohn, Braunschweig, 1992.

[Fo]   Fontaine, J-M., *Il n'y a pas de variété abélienne sur Z*. Invent. Math. 81 (1985), no. 3, 515–538.

[Fr1]   Frey, G., *Links between stable elliptic curves and certain diophantine equations*, Ann. Univ. Saraviensis, **1** (1986), 1–40.

[Fr2]   Frey, G., *Links between solutions of $A - B = C$ and elliptic curves*, in Number theory, Ulm 1987, Proceedings, Lecture Notes in Math. **1380**, Springer-Verlag, New York (1989) 31–62.

[Gre1] Greenberg, M., *Computing Heegner points arising from Shimura curve parametrisations*, in this proceedings volume.

[Gre2] Greenberg, M., *Elliptic curves over imaginary quadratic fields and Stark-Heegner points*, in this proceedings volume.

[Gr]   Gross, B.H.,*Kolyvagin's work on modular elliptic curves*, Proc. Durham symposium on L-functions and arithmetic, 1989, to appear.

[GZ]   Gross, B.H. and Zagier, D.B., *Heegner points and derivatives of L-series.* Invent. Math. 84 (1986), 225-320.

[Hel]   Hellegouarch, Y., *Points d'ordre fini sur les courbes elliptiques*, C. R. Acad. Sci. Paris Sér. A-B **273** (1971) A540–A543.

[Ho1]   Howard, B., *Twisted Gross-Zagier theorems.* To appear in Canad. J. Math.

[Ho2]   Howard, B. *Central derivatives of L-functions in Hida families.* Math. Ann. **339**, No.4 (2007)

[HST]   Harris, M., Shepherd-Barron, N., and Taylor, R., *A family of Calabi-Yau varieties and potential automorphy.* preprint.

[Hur]   Hurwitz, A., *Über die diophantische Gleichung $x^3y + y^3z + z^3x = 0$,* Math. Ann. 65 (1908), 428-430.

[Hus]   Husemöller, D., *Elliptic curves*, GTM 111, New York: Springer Verlag (1987)

[Ka]   Kamienny, S., *Torsion points on elliptic curves and q-coefficients of modular forms.* Invent. Math. **109** (1992), no. 2, 221–229.

[Ki]   Kisin, M., *Modularity of 2-adic Barsotti-Tate representations*, preprint. To appear.

[Ko1]   Kolyvagin, V.A., *Finiteness of $E(\mathbf{Q})$ and $Ш(E/\mathbf{Q})$ for a subclass of Weil curves*, Izv. Akad. Nauk. SSSR Ser Mat. 52 (3) (1988), 522-540; Math USSR Izvestiya 32 (1989), 523-541.

[Ko2]   Kolyvagin, V.A., *Euler Systems.* The Grothendieck Festschrift, Vol. II, 435–483, Progr. Math., 87, Birkhäuser Boston, Boston, MA, 1990.

[KW]   Khare, C., and Wintenberger, J.-P., *Serre's Modularity Conjecture (I)*, preprint, to appear.

[Lo]   Longo, M., *On the Birch and Swinnerton-Dyer conjecture for modular elliptic curves over totally real fields*, Annales de l'Institut Fourier, Vol. 56, no. 3 (2006), 689-733.

[Ma1]   Mazur, B., *Modular curves and the Eisenstein ideal*, Publ. Math. IHES 47, 33-186 (1977)

[Ma2]   Mazur, B., *Rational isogenies of prime degree*, Inv. Math. 44, 129-162 (1978)

[Ma3]   Mazur, B., *Arithmetic on curves*, Bulletin of the AMS, Vo. 14, No. 2, April 1986.

[Me]   Merel, L., *Bornes pour la torsion des courbes elliptiques sur les corps de nombres.* Invent. Math. **124** (1996), no. 1-3, 437–449.

[MM]   Murty, M.R. and Murty, V.K., Non-vanishing of L-functions and applications. Progress in Mathematics, 157. Birkhauser Verlag, Basel, 1997.

[Oda]   Oda, T., Periods of Hilbert modular surfaces. Progress in Mathematics, 19. Birkhuser, Boston, Mass., 1982.

[Oe]   Oesterlé, J., *Nouvelles aproches du "théorème" de Fermat"* Séminaire Bourbaki no. 694 (1987-88). Astérisque 161-162, 165-186 (1988)

[Pa1]   Parshin, A.N., *Algebraic curves over function fields I*, Izv. Akad. Nauk. SSSR Ser. Math. 32 (1968), 1191-1219; english transl. in Math. USSR. Izv. 2 (1968)

[Reb]   Rebolledo, M., *Merel's Theorem on the boundedness of the torsion of elliptic curves*, in this proceedings volume.

[Rib1]   Ribet, K., *On modular representations of $Gal(\bar{\mathbf{Q}}/\mathbf{Q})$ arising from modular forms*, Invent. Math. 100, 431-476 (1990).

[Se1]   Serre, J-P., *Propriétés galoisiennes des points d'ordre fini des courbes elliptiques.* Invent. Math. **15** (1972), no. 4, 259–331.

[Se2]   Serre, J.-P., *Sur les représentations modulaires de degré 2 de $Gal(\bar{\mathbf{Q}}/\mathbf{Q})$*, Duke Math. J. Vol. 54, no. 1, 179-230 (1987).

[Sha]   Shafarevich, I.R., *Algebraic number fields*, Proc. ICM (Stockholm, 1962) Inst. Mittag-Leffler, Djursholm, 1963, 163–176; English Transl. in AMS Transl. (2) 31 (1963).

[Shi]   Shimura, G., Introduction to the arithmetic theory of automorphic functions, Princeton Univ. Press, Princeton (1971).

[Si]   Silverman, J., *The arithmetic of elliptic curves*, GTM 106. New York: Springer-Verlag (1986)

[Sz]   Szpiro, L., *La conjecture de Mordell (d'après G. Faltings)*, Séminaire Bourbaki, exposé 619, Volume 1983/84, Asterisque 121-122, 1985.

[Ta]   Tate, J., *Endomorphisms of abelian varieties over finite fields*, Inventiones Math. 2 (1966) 134–144.

[Tay1]   Taylor, R., *Automorphy for some ℓ-adic lifts of automorphic mod ℓ representations. II.* preprint.

[Tay2]   Taylor, R., *On icosahedral Artin representations. II.* Amer. J. Math. **125** (2003), no. 3, 549–566.

[Tr]   Trifkovic, M., *Stark-Heegner points on elliptic curves defined over imaginary quadratic fields.* Duke Math. J. 135 (2006), no. 3, 415–453.

[TW]   Taylor, R. and Wiles, A., *Ring-theoretic properties of certain Hecke algebras.* Ann. of Math. (2) **141** (1995), no. 3, 553–572.

[Vo]   Voight, J., *Shimura curve computations.* In this proceedings volume.

[We]   Weil, A., Variétés abéliennes et courbes algébriques. Hermann, Paris, 1948.

[Wi]   Wiles, A., *Modular elliptic curves and Fermat's last theorem.* Ann. of Math. (2) **141** (1995), no. 3, 443–551.

[Zh1]   Zhang, S., *Heights of Heegner points on Shimura curves.* Ann. of Math. (2) 153 (2001), no. 1, 27–147.

[Zh2]   Zhang, S., *Gross-Zagier formula for* GL_2. Asian J. Math. 5 (2001), no. 2, 183–290.

[ZP]   Zarkhin, Yu. G., and Parshin, A.N., *Finiteness problems in diophantine geometry*, Amer. Math. Soc. Transl. (2), Vol 143, 1989, pp. 35-102.