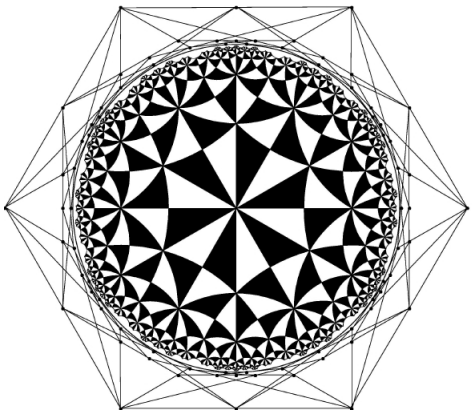
Harvard University, Cambridge, MA

Thesis presented to the Department of Mathematics in partial fulfillment of the requirements for the degree of Bachelor of Arts with Honors

Modular Curves and Mazur's Theorem

by Luca Candelori



March 31, 2008

Contents

Preface

The purpose of this paper is twofold: on the one hand, give an accurate exposition of Mazur's Theorem as it was proved in the 1977 paper *Modular Curves and the Eisenstein Ideal* [Maz]. On the other, I wanted to make the reading accessible for the undergraduate who has not yet received but the barest introduction to the language of schemes, of which Mazur makes heavy use in the paper. Of course, in doing so we had to gloss over some of the more technical aspects of the proof, but where possible we have presented the arguments in a language that can be understood by anyone familiar with the contents of Silverman's wonderful book *The Arithmetic of Elliptic Curves* [Si1]. I hope that this approach in no way will discourage the inexperienced reader to further investigate the topics covered, but it will rather inspire to independently complete the parts of the proof that are referred back to Mazur's paper, and explore the numerous references given.

Acknowledgments

My first and foremost 'thank you' must go to Professor Samit Dasgupta, my thesis adviser. In the months I have devoted to the project, he has spent a countless amount of hours sitting down with me and making sure that I did not leave his office without getting what I came for. It is often rare to find such a dedicated guide, and I could not have hoped for a better adviser. Many thanks to Professor Barry Mazur also, for his helpful insights in the more convoluted parts of the argument and for giving me excellent references. Finally, I want to thank my parents, my sister and Ellen for the moral support they gave me while I was working on the project, and for the joys they give me everyday of my life.

Introduction

1 Why Study Elliptic Curves

One of the most spectacular aspects of number theory is the determination of integer solutions to polynomial equations with coefficients in the rational numbers, a branch of mathematics known as Diophantine analysis. The name comes from Diophantus of Alexandria, a Greek mathematician of the III century AD who is often credited to be the father of the subject. In particular he is credited, together with the Arab mathematician al-Khwārizmī, with the invention of symbolic manipulation, the very foundation of what we may call *Algebra* in a broad sense. Up until Diophantus, in fact, other Greek mathematicians (such as Euclid) would mainly rely on geometric arguments to prove their theorems.

But the study of Diophantine equations is much older than Diophantus himself. The Indian mathematician and priest Baudhayana, for example, independently formalized the Pythagorean Theorem around 800 BC and found approximations of $\sqrt{2}$ by rational numbers. In particular, he noted that any pair of integers (x, y) satisfying:

$$x^2 - 2y^2 = 1 \tag{1}$$

would produce an approximation of $\sqrt{2}$ just by taking x/y. About 1500 years later, the Indian mathematician Brahmagupta described a remarkable algorithm to find integer solutions to equation (1), called the *chakravala method*. The German mathematician Hankel described this method as 'the finest thing achieved in the theory of numbers before Lagrange' [Kay]. Oddly enough, this equation is today known as *Pell's equation*, after Euler attributed its study to a British mathematician of the XVII century.

The work of Diophantus is collected in a book called *Arithmetica*, which became known in Europe only at the end of the XVI Century when it started to be translated into Latin. In 1637, the French lawyer and mathematician Pierre de Fermat, while studying his copy of *Arithmetica*, conjectured that the Diophantine equation

$$x^n + y^n = z^n$$

has no (nontrivial) integer solutions (x, y, z) for any integer $n > 2$. Of course for $n = 2$ this is just the Pythagorean Theorem. Probably nobody would have taken the time to study this equation, except that Fermat had the *brilliant* idea of stating on the margin of his edition of *Arithmetica* that

> "Cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet."

which translates into 'I found a truly marvelous proof of this statement, which this margin is too narrow to contain'. Mathematicians often love challenges, and the challenge thrown by Fermat motivated a lot of the development of number theory in the subsequent centuries. In 1994, thanks to the work of Frey, Ribet, Serre, Taylor and Wiles, a proof of the statement was derived in the broader context of studying the connections between elliptic curves and modular forms. Similarly, many questions in Diophantine analysis can be reduced to questions about elliptic curves, where both the tools of algebraic geometry and algebraic number theory can be applied.

But this is only one of the applications of the study of elliptic curves and, in a sense, a very *selfish* application. A much more concrete one was found independently in 1985 by Neal Koblitz and Victor Miller, when they proposed the use of elliptic curves in public key cryptography. Since then, elliptic curves have become one of the standards of public-key cryptography, and are widely used in everyday applications. The advantage over other systems is the hardness of certain computational problems related to the group structure of an elliptic curve (mainly the discrete logarithm problem), which allows for shorter keys and more efficient implementations. However, as Menezes, Okamoto and Vanstone have shown [MOV] one has to be very careful in implementing cryptographic protocols using elliptic curves. Since these are still fairly unknown objects, there is

always the possibility of finding shortcuts and additional structure that could reduce the hardness of the discrete logarithm problem defined on the group of points of an elliptic curve. A study of the arithmetic properties of elliptic curves, therefore, could be very important even for the applied mathematician who does not get very excited by questions in Diophantine analysis.

Finally, for the pure mathematician the study of elliptic curves is a very rewarding one on its own. The structure of these miracolous mathematical objects is very beautiful, and it allows to bring together many different branches of mathematics that would otherwise struggle in finding a common ground. In this sense, the theory of elliptic curves is a place where ideas are exchaged and problems are formulated that inspire a very broad spectrum of the mathematical community.

2 Statement of the Problem and Outline of Contents

Let $E/\mathbb{Q}$ be an elliptic curve defined over $\mathbb{Q}$. By the Mordell-Weil Theorem [Si1 VIII], the group of points $E(\mathbb{Q})$ is a finitely generated abelian group. Its structure is of the form:

$$E(\mathbb{Q}) = E_{\text{tors}}(\mathbb{Q}) \times \mathbb{Z}^r$$

where r is a positive integer of which very little is known. Whereas computations show that r is usually very small, it is conjectured that there exists curves with arbitrarily high rank (for a discussion, see [Si1 VIII.10]). In particular, a conjecture of Birch and Swinnerton-Dyer relates r to the order of vanishing of a certain function defined over the complex plane.

In this paper, we are interested in the other component of $E(\mathbb{Q})$, i.e. the group of points of $E(\mathbb{Q})$ of finite order. For a given elliptic curve, these are easy to compute thanks to the following theorem of Lutz and Nagell

Theorem (Lutz-Nagell)**.** *Let $E/\mathbb{Q}$ be an elliptic curve with Weierstrass equation*

$$y^2 = x^3 + Ax + B \quad A, B \in \mathbb{Z}$$

Suppose $P \in E(\mathbb{Q})$ is non-zero torsion point. Then

(a) $x(P), y(P) \in \mathbb{Z}$

(b) Either $[2]P = O$*, or else* $y(P)^2$ *divides* $4A^3 + 27B^2$*.*

Proof. This is in [Si1 VIII.7] □

The theorem gives an effective procedure for finding torsion points, especially when the discriminant has only a few divisors.

A much harder question is to completely classify all the possible torsion structures of $E(\mathbb{Q})$. In other words, to give a list of all the possible groups that may occur as $E_{\text{tors}}(\mathbb{Q})$. In 1975 Ogg conjectured the following:

Theorem. *Let* E *be an elliptic curve defined over* $\mathbb{Q}$*. Then* $E_{tors}(\mathbb{Q})$ *is isomorphic to one of the following fifteen groups:*

- $\mathbb{Z}/n\mathbb{Z}$ *for* $1 \leq n \leq 10$ *or* $n = 12$
- $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ *for* $m = 2, 4, 6, 8$

which was proven by Barry Mazur in his 1977 paper *Modular Curves and the Eisenstein Ideal*. In fact, the paper proves a lot more than just Ogg's conjecture, and it is often difficult for the reader only interested in the classification of the rational torsion to isolate the parts needed to prove the statement. One purpose of this paper is precisely to collect the passages that give the proof of Mazur's Theorem in one place and clarify the logical structure of the argument. The other purpose, as stated in the preface, is to give a presentation that can be understood by any undergraduate student who is familiar with the arithmetic theory of elliptic curves at the level of Silverman's book [Si1].

In Chapter I, we introduce the reader to the theory of the modular curves $X_0(N)$, the central object of study of Mazur's paper. We have omitted some of the more basic proofs, which can be found in many texts such as [Si2], and emphasized the calculation of the genus of $X_0(N)$, which is somewhat less of a standard application to be seen in textbooks, but that sheds a lot of light on the structure of these curves.

In Chapter 2, we begin a systematic exposition of Mazur's argument. The first section is devoted to the construction of a certain quotient J of the Jacobian

of $X_0(N)$, called by Mazur the Eisenstein quotient. In the second section we outline some of the scheme-theoretic tools used in the proof of the finitiness of the rational part of the Eisenstein quotient. In third section we apply p^m-descent on $\mathrm{Jac}(X_0(N))$ to prove finitiness of $X_0(N)(\mathbb{Q})$.

In Chapter 3, we use finiteness of $X_0(N)(\mathbb{Q})$ to prove Mazur's Theorem. In this chapter in particular, I have tried to translate the language used by Mazur to a language closer to Silverman's Arithmetic of Elliptic Curves. I hope the experienced reader will forgive some of the imprecisions, and will appreciate the effort in trying to make a clear exposition of a paper which was certainly not meant to be read by undergraduates. Enjoy the reading.

Chapter I

Modular Curves

Mazur's Theorem is in some sense a statement about the set of all elliptic curves, in particular those possessing a rational torsion point. There is a way in which the set of all (isomorphism classes of) elliptic curves can be efficiently studied, and this is done by giving it the structure of an algebraic curve. Similarly, the set of all (isomorphism classes of) elliptic curves with a torsion subgroup of order N can also be parametrized by an algebraic curve. Such curves are called *modular curves*, and are denoted by $X(1)$ and $X_0(N)$ respectively.

This chapter is divided into two sections: in the first one, we construct the modular curves $X(1)$ and $X_0(N)$ as compact Riemann Surfaces and study some of their properties. In particular, we obtain a formula for the genus of $X_0(N)$ in terms of N. In the second section, we outline the connection with the theory of elliptic curves and state some important facts that will be important later in the proof of Mazur's Theorem.

1 Construction of Modular Curves

Definition 1.1. A *topological group* G is a Hausdorff topological space that is also a group, and whose multiplication map $m : G \times G \to G$ and inversion map $i : G \to G$ are continuous. We denote by e the identity element of G.

Suppose now that T is another topological space and that we are given a

continuous map $G \times T \to T$ (denoted by $(g,t) \mapsto gt$) satisfying the following:

(i) $(gh)t = g(ht)$

(ii) $et = t$ for all $t \in T$

Then we say that *G acts continuously on T*. In this situation it makes sense to consider, for each $t \in T$, the set $Gt = \{gt : g \in G\}$, called *the orbit of t under G*, and the set $G\backslash T = \{Gt : t \in T\}$. If this set consists of only one element, we say that *G acts transitively on T*. By considering the surjective map $\pi : T \to G\backslash T$, which sends $t \mapsto Gt$, we can make $G\backslash T$ into a topological space.

1.1 The Modular Curve $X(1)$

Consider first the group $G = \mathbf{GL}_2(\mathbb{C})$ of invertible 2×2 matrices with entries in $\mathbb{C}$. As a subset of $\mathbb{C}^4$, it inherits a subspace topology from the Euclidean topology. Moreover, one can easily check that the multiplication and inversion maps are continuous with respect to this topology, giving the structure of a topological group.

The group $\mathbf{GL}_2(\mathbb{C})$ acts continuously on $\mathbb{C}$ as follows: for any $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ define the function

$$A(z) = \frac{az+b}{cz+d}$$

which is meromorphic on $\mathbb{C}$ with a single pole at $z = -d/c$. Note that multiplication by a constant $\alpha \in \mathbb{C}^\times$ may change the matrix A but it does not change $A(z)$. Therefore, $\mathbf{GL}_2(\mathbb{C})$ actually acts through its quotient $\mathbf{PGL}_2(\mathbb{C}) = \mathbf{GL_2}(\mathbb{C})/\{\alpha I : \alpha \in \mathbb{C}^\times\}$. This action is transitive, for it is a classical result in complex analysis that if z_1, z_2, z_3 and w_1, w_2, w_3 are complex numbers, there is a unique map of the form $A(z)$ mapping z_i to w_i, for $A \in \mathbf{GL}_2(\mathbb{C})$.

If we take any discrete subgroup $\Gamma \subset \mathbf{PGL}_2(\mathbb{C})$ one can show that $\Gamma\backslash\mathbb{C}$ is a Hausdorff space [Shi 1.1]. In particular, consider the discrete subgroup $\mathbf{SL}_2(\mathbb{Z}) \subset \mathbf{GL}_2(\mathbb{C})$, consisting of matrices with integer entries and determinant 1, and its image $\Gamma(1) = \mathbf{SL}_2(\mathbb{Z})/\{\pm I\}$ inside the quotient $\mathbf{PGL}_2(\mathbb{C})$. Note that for any $z \in \mathbb{C}$

and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{GL}_2(\mathbb{C})$, we have

$$\Im[A(z)] = (ad - bc)\frac{\Im[z]}{|cz + d|^2}$$

from which we see that in fact $\Gamma(1)$ acts on the upper-half plane $\mathbb{H} = \{z \in \mathbb{C} : \Im[z] > 0\}$. The space $Y(1) = \Gamma(1)\backslash\mathbb{H}$ is Hausdorff. For geometric applications, however, being Hausdorff might not be enough. We would like to compactify $Y(1)$ much in the same way in which we obatin the Riemann Sphere $\mathbb{P}^1(\mathbb{C})$ from the 1-point compactification of $\mathbb{C}$. One has to be careful, however, about throwing in ∞ without understanding first the action of $\Gamma(1)$ on $\mathbb{H} \cup \{\infty\}$.

We consider first the fixed points of the functions $A(z)$, where $A \in \mathbf{GL}_2(\mathbb{C})$. Since $AB(z) = A(B(z))$, a change of basis does not affect the number of fixed points of $A(z)$. Therefore it suffices to consider only the number of fixed points of transformations that are representative of each conjugacy class. Now, from the theory of Jordan Canonical Form, we know that every matrix A in $\mathbf{GL}_2(\mathbb{C})$ is conjugate to one of:

$$T_\lambda = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \ , \ S_{\lambda,\mu} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

for some choice of $\lambda, \mu \in \mathbb{C}^\times$. We call the first family *parabolic*, and the second *elliptic*. Therefore there are only two possibilities for the number of fixed points, depending on whether A is parabolic or elliptic . Take first the transformation $T_\lambda(z)$ associated to T_λ:

$$T_\lambda(z) = z + \frac{1}{\lambda}$$

Then T_λ has only one fixed point in $\mathbb{C} \cup \{\infty\}$, namely ∞. In general, for any parabolic element A the corresponding transformation $A(z)$ has exactly 1 fixed point. We call these points *cusps*.

Similarly, we can consider the fixed points of the transformation associated to the matrix $S_{\lambda,\mu}$. We have

$$S_{\lambda,\mu}(z) = \frac{\lambda}{\mu}z = z$$

Unless $\lambda = \mu$, in which case $S_{\lambda,\mu}$ is the identity, we must have either $s = 0$ or $s = \infty$, so that $S_{\lambda,\mu}$ has two fixed points. In general any transformation $A(z)$, for

A an elliptic element, has two fixed points in $\mathbb{C} \cup \{\infty\}$. We call these points *elliptic points*.

Definition 1.2. Let G be a subgroup of $\mathbf{GL}_2(\mathbb{C})$. If z is an elliptic point fixed by an element of G, we call z an *elliptic point of G*. If z is a cusp fixed by an element of G, we call z a *cusp of G*.

For each discrete subgroup Γ of $\mathbf{GL}_2(\mathbb{C})$ it is often useful to classify all the cusps and elliptic points of Γ. We begin by analyzing the cusps and elliptic points of $\Gamma(1)$.

Lemma 1.3. *If $z \in \mathbb{C} \cup \{\infty\}$ is an elliptic point of $\Gamma(1)$, then z is conjugate in $\Gamma(1)$ to a primitive n-th root of unity in $\mathbb{C}$, where $n = 3$ or $n = 4$ (i.e. there exists an $A \in \Gamma(1)$ such that $A(z) = \pm i, e^{2\pi i/3}$.*

Proof.

Let $z \in \mathbb{C} \cup \{\infty\}$ and suppose there exists an elliptic element of $\Gamma(1)$ fixing it. Then A is conjugate in $\mathbf{GL}_2(\mathbb{C})$ to a diagonal matrix of the form $S_{\lambda,\mu}$. Since the trace of a matrix is not affected by a change of coordinates, we see that $\operatorname{tr}(A) = \operatorname{tr}(S_{\lambda,\mu}) = \lambda + \mu$. The matrix A has integer entries, so $\operatorname{tr}(A) = \lambda + \mu \in \mathbb{Z}$. Moreover, $\lambda\mu = 1$, $\lambda \neq \mu$, and in particular they cannot be both equal to 1 or -1. It follows that $|\operatorname{tr}(A)| = |\lambda + \mu| < 2 \Rightarrow \operatorname{tr}(A) = 0, \pm 1$. Therefore the characteristic polynomial of A is one of:

$$f_1(x) = x^2 + 1 \quad f_2(x) = x^2 \pm x + 1$$

and therefore A satisfies either $f_1(A) = 0$ or $f_2(A) = 0$.

Case 1:

Suppose $f_1(A) = 0$. Then $A^4 = I_2$, but since $A \equiv -A$ in $\Gamma(1)$ and we are assuming that $A \neq I_2$, we must have that $A^2 = 1$. Now, as a 2x2 matrix with integer coefficients, A has a natural action on the free module $\mathbb{Z}^2$. Including the action of integer scalar matrices, which are compatible with A, we get an action of $\mathbb{Z}[A]$ on $\mathbb{Z}^2$. Since $\mathbb{Z}[A]$ is isomorphic to the ring of Gaussian integers $\mathbb{Z}[i]$, we can define an archimedean absolute value on $\mathbb{Z}[A]$ such that, if $v \in \mathbb{Z}^2$ is such that $(n + mA)v = 0$, we have $|(n + mA)||v| = 0 \Rightarrow |(n^2 + m^2)||v| = 0 \Rightarrow |v| = v = 0$. Therefore $\mathbb{Z}^2$ is a free $\mathbb{Z}[A]$ module of rank

1 and since $\mathbb{Z}[i]$ is a PID we get $\mathbb{Z}^2 = \mathbb{Z}[A]u$ for some $u \in \mathbb{Z}^2$. Now, since u and $v = Au$ are linearly independent over $\mathbb{Z}$, they form a basis for $\mathbb{Z}^2$ over $\mathbb{Z}$. Therefore the matrix $[u\ \ v]$ is invertible over $\mathbb{Z}$ and hence it must have determinant ± 1. Appling A to it we see that:

$$A[u\ \ v] = [u\ \ v] \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

If $\det[u\ \ v] = 1$, then A is conjugate in $\mathbf{SL}_2(\mathbb{Z})$ to $D = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ via $[u\ \ v]$ itself. Otherwise, if $\det[u\ \ v] = -1$, then $\det[v\ \ u] = 1$ and A is conjugate in $\mathbf{SL}_2(\mathbb{Z})$ to $D' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ via $[v\ \ u]$. Since $D \equiv D'$ in $\Gamma(1)$, we see that A is always conjugate to D via an element of $\Gamma(1)$, call it Q. Then, if z is a fixed point of A, we have

$$\begin{aligned} A(z) = QDQ^{-1}(z) &\Rightarrow \\ z = Q \circ D \circ Q^{-1}(z) &\Rightarrow \\ Q^{-1}(z) = D \circ Q^{-1}(z) & \end{aligned}$$

and therefore $Q^{-1}(z)$ is a fixed point of D. In other words, z is conjugate in $\Gamma(1)$ to a fixed point of D, i.e. $\pm i$, a primitive 4-th root of unity.

Case 2:

If $f_2(A) = 0$, then $A^6 = I_2$ and since $A \neq \pm I_2$ we must have $A^3 = I_2$ with A primitive. In this case $\mathbb{Z}[A]$ is isomorphic to $\mathbb{Z}[e^{2\pi i/3}]$ (choosing the root with positive imaginary part), again a PID with an archimedean absolute value, and we can proceed as before. Again $\mathbb{Z}^2 = \mathbb{Z}[A]u$ for some $u \in \mathbb{Z}^2$, and $u, v = Au$ form a basis for $\mathbb{Z}^2$ over $\mathbb{Z}$. We have

$$A[u\ \ v] = [u\ \ v] \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$$

if $\det[u\ \ v] = 1$ or

$$A[v\ \ u] = [v\ \ u] \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$$

if $\det[u\ v] = -1$. Therefore z is conjugate in $\Gamma(1)$ either to a fixed point of $\begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ or to a fixed point of $\begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$ i.e. to either $e^{2\pi i/3}$ or $e^{4\pi i/3}$, which are both primitive 3-rd roots of unity.

□

Lemma 1.3 shows that there are only two types of elliptic points in $\mathbb{C} \cup \{\infty\}$: those fixed by a matrix $A \in \Gamma(1)$ such that $A^2 \equiv I_2$ in $\Gamma(1)$ and those with $A^3 = I_2$. Sometimes we will refer to the first family as *points of order 2* and to the second as *points of order 3*.

The determination of the cusps of $\Gamma(1)$ is more straightforward:

Lemma 1.4. *Let S be the set of cusps of $\Gamma(1)$. Then $S = \mathbf{P_1}(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$.*

Proof.

We see that T_1:

$$T_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

leaves ∞ fixed, therefore $\infty \in S$.

Let now A be any parabolic element of $\Gamma(1)$. Then if s is the fixed point of $A(z)$, it must satisfy

$$A(s) = \frac{as + b}{cs + d} = s$$

which implies

$$cs^2 + s(d - a) - b = 0$$

In other words, s satisfies a polynomial with integer coefficients. If $c = 0$, then $s = \infty$ again (since then $a = d = \pm 1$). Otherwise, we know A has only one fixed point, since A is parabolic, and therefore the discriminant of the polynomial must vanish. In other words, $s \in \mathbb{Q}$.

On the other hand, let $r = p/q \in \mathbb{Q}$. Since $\gcd(p, q) = 1$, we can find integers a, b such that $ap + bq = 1$. Let now

$$A = \begin{pmatrix} p & -b \\ q & a \end{pmatrix}$$

then $A(\infty) = \lim_{z\to\infty} A(z) = p/q = r$. In particular, the parabolic element $B = AT_1A^{-1}$ fixes r. Therefore $S = \mathbf{P_1}(\mathbb{Q})$. □

Remark 1.5. Note that during the proof of Lemma 1.4 we have also shown that every cusp of $\Gamma(1)$ is in the orbit of ∞.

Going back to our compactification of $Y(1)$, we see from Remark 1.5 that adding infinity to $Y(1)$ also requires adding $\mathbb{Q}$. Consider then the set $\widehat{\mathbb{H}} = \mathbb{H} \cup \mathbb{Q} \cup \{\infty\}$ and let $X(1) = \Gamma(1)\backslash\widehat{\mathbb{H}}$. We define a topology on $X(1)$ by specifying a base of neighborhoods as follows:

(i) For any $z \in \mathbb{H}$, an open neighborhood of $z \in X(1)$ is simply an open neighborhood of $z \in Y(1)$ viewed as a subset of $X(1)$.

(ii) For $z \in \mathbb{Q}$, let $U_\epsilon = \{w \in \mathbb{H} : |w - (z + i\epsilon)| < \epsilon\} \cup \{z\}$, for any $\epsilon \in \mathbb{R}_{>0}$. In other words, U_ϵ is the interior of the circle of radius ϵ tangent to the real axis at z union z.

(iii) For $z = \infty$, let $V_\epsilon = \{w \in \mathbb{H} : \Im[w] > \epsilon\} \cup \{\infty\}$ for any $\epsilon \in \mathbb{R}_{>0}$.

Using the fact that $Y(1)$ is already Hausdorff, it is easy to see that $X(1)$ is Hausdorff as well (one needs only to check separability at the cusps). To check that $X(1)$ is compact, we want to identify first a set of representatives for $X(1)$ inside $\widehat{\mathbb{H}}$ (a so called *fundamental region*) which gives a picture of what $X(1)$ looks like.

Proposition 1.6. $\Gamma(1)$ *is generated by the two elements*

$$T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Proof.

For an elementary proof, see [Si2 I.1.6]. □

By Proposition 1.6, the two transformations

$$T(z) = z + 1 \quad S(z) = \frac{-1}{z}$$

generate $\Gamma(1)$. In particular, the orbit of any element $z \in \widehat{\mathbb{H}}$ contains all the translations of its real part by an element of $\mathbb{Z}$. Therefore, to find a set of representatives for $X(1)$, we can restrict our attention to the strip $\{z \in \mathbb{H} : |\Re[z]| \leq \frac{1}{2}\} \cup \{\infty\}$. Also, by using S, we can always invert an element so it suffices to consider all those elements z with $|z| \geq 1$ (for a rigorous argument, see [Si2 I.2.3]).

Proposition 1.7. *$X(1)$ is compact*

Proof.

Call $\mathcal{F}$ the region given by $\{z \in \mathbb{H} : |\Re[z]| \leq \frac{1}{2}\} \cup \{\infty\} \cap \{|z| \geq 1\}$ and consider an open cover $\bigcup_{i\in I} U_i$ of $X(1)$. Denote by $\phi : \widehat{\mathbb{H}} \to X(1)$ the projection map. Then $\bigcup_{i\in I} \phi^{-1}(U_i)$ covers $\widehat{\mathbb{H}}$. One of the $\phi^{-1}(U_i)$, say $\phi^{-1}(U_{i_1})$, contains ∞ so that it is of the form $V_\epsilon = \{w \in \mathbb{H} : \Im[w] > \epsilon\} \cup \{\infty\}$ for some $\epsilon \in \mathbb{R}_{>0}$. Now the set $\mathcal{F} - \phi^{-1}(U_{i_1})$ is closed and bounded, hence there is a finite sub-collection $\phi^{-1}(U_{i_2}), \ldots, \phi^{-1}(U_{i_k})$ of the U_i's covering it. Therefore $\phi^{-1}(U_{i_1}), \ldots, \phi^{-1}(U_{i_k})$ covers all of $\mathcal{F}$. Upon application of ϕ we see that the U_{i_j} cover $X(1)$ and hence $X(1)$ is compact. □

Now that we have a compact Hausdorff (and evidently connected) topological space the next question is whether there is additional structure on $X(1)$ that we should consider. It turns out in fact that we can define on this space the structure of a one-dimensional complex analytic manifold (i.e. a *Riemann surface*).

Remark 1.8. Studying the theory of modular functions, one can define a map:

$$j : X(1) \to \mathbb{P}^1(\mathbb{C})$$

which is an isomorphism of Riemann Surfaces. This gives that $X(1)$ has genus 0. For details see [Si2 I.3].

1.2 The Modular Curves $X_0(N)$

We know look at the action on $\widehat{\mathbb{H}}$ of discrete subgroups of $\mathbf{GL}_2(\mathbb{C})$ other than $\mathbf{SL}_2(\mathbb{Z})$.

Consider first the map $\phi_N : \mathbf{SL}_2(\mathbb{Z}) \to \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$ given by reducing the coefficients of the matrices modulo N, for every positive integer N. It is not immediate at first that this map is surjective, but we can check this fact by elementary means. If we denote by $PC(N)$ the kernel of this map, we have an exact sequence

$$0 \to PC(N) \to \mathbf{SL}_2(\mathbb{Z}) \xrightarrow{\phi_N} \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z}) \to 0$$

The subgroup $PC(N)$, which is normal in $\mathbf{SL}_2(\mathbb{Z})$, is called the *principal congruence subgroup of level N*. It is explicitly given by:

$$PC(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}) : b, c \equiv 0 \mod N \text{ and } a, d \equiv 1 \mod N \right\}$$

This subgroup is a discrete subgroup of $\mathbf{GL}_2(\mathbb{C})$ and the exact sequence shows that it is of finite index inside $\mathbf{SL}_2(\mathbb{Z})$, since this is just the order of $\mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Given the prime factorization of $N = \prod p^e$, we wish to calculate this quantity.

Proposition 1.9. $[\mathbf{SL}_2(\mathbb{Z}) : PC(N)] = N^3 \cdot \prod_{p|N}(1 - p^{-2})$

Proof.

From basic algebra we have

$$\begin{aligned} \mathbb{Z}/N\mathbb{Z} &\cong \prod (\mathbb{Z}/p^e\mathbb{Z}) \\ \mathbf{GL}_2(\mathbb{Z}/N\mathbb{Z}) &\cong \prod (\mathbf{GL}_2(\mathbb{Z}/p^e\mathbb{Z})) \\ \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z}) &\cong \prod (\mathbf{SL}_2(\mathbb{Z}/p^e\mathbb{Z})) \end{aligned}$$

So that in fact it suffices to compute the order of each of the $\mathbf{SL}_2(\mathbb{Z}/p^e\mathbb{Z})$. In order to do so, consider first the exact sequence of groups:

$$1 \to X \to \mathbf{GL}_2(\mathbb{Z}/p^e\mathbb{Z}) \to \mathbf{GL}_2(\mathbb{Z}/p\mathbb{Z}) \to 1$$

where X consists of all the elements of $\mathbf{GL}_2(\mathbb{Z}/p^e\mathbb{Z})$ that are congruent to the identity matrix in $\mathbf{GL}_2(\mathbb{Z}/p\mathbb{Z})$. In other words, each entry of a matrix in X has to be in the congruence class of either 0 or 1 modulo p. The index of each congruency class of p in $\mathbb{Z}/p^e\mathbb{Z}$ is exactly p^{e-1}, therefore there are precisely $p^{4(e-1)}$ elements in X. On the other hand, $\mathbf{GL}_2(\mathbb{Z}/p\mathbb{Z})$ has precisely $(p^2 - 1)(p^2 - p)$ elements (just by

checking conditions on $ad - bc$) and therefore

$$\begin{aligned}\#\mathbf{GL}_2(\mathbb{Z}/p^e\mathbb{Z}) &= \#X \cdot \#\mathbf{GL}_2(\mathbb{Z}/p\mathbb{Z}) \\ &= p^{4(e-1)}(p^2-p)(p^2-1) \\ &= p^{4e}(1-p^{-1})(1-p^{-2})\end{aligned}$$

Now, since mutiplying a row of a matrix by a unit c multiplies its determinant by c, we see that the index of $\mathbf{SL}_2(\mathbb{Z}/p^e\mathbb{Z})$ in $\mathbf{GL}_2(\mathbb{Z}/p^e\mathbb{Z})$ is precisely the number of units in $\mathbb{Z}/p^e\mathbb{Z}$, i.e. $\Phi(p^e) = p^e(1-p^{-1})$ where Φ is Euler's Phi function. We obtain

$$\mathbf{SL}_2(\mathbb{Z}/p^e\mathbb{Z}) = p^{3e}(1-p^{-2})$$

and therefore by multiplying each factor we get

$$[\mathbf{SL}_2(\mathbb{Z}) : PC(N)] = N^3 \cdot \prod_{p|N}(1-p^{-2})$$

as required. □

Any subgroup of $\mathbf{SL}_2(\mathbb{Z})$ containing $PC(N)$ for some N is called a *congruence subgroup*. In particular, for reasons that will become clear later, we want to study:

$$C_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}) : c \equiv 0 \mod N \right\}$$

This subgroup certainly contains $PC(N)$, but it is not normal in $\mathbf{SL}_2(\mathbb{Z})$ (just conjugate it by S), and it therefore does not arise naturally as the kernel of a projection map. We call $\Gamma_0(N)$ the image of $C_0(N)$ in the quotient $\Gamma(1)$. In other words:

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1) : c \equiv 0 \mod N \right\}$$

We wish to use Proposition 1.9 to compute the index $[\Gamma(1) : \Gamma_0(N)]$. We have:

Proposition 1.10. $[\Gamma(1) : \Gamma_0(N)] = N \prod_{p|N}(1+p^{-1})$

Proof.

First of all, note that since $-I_2 \in \Gamma_0(N)$ for all N, we have

$$[\Gamma(1) : \Gamma_0(N)] = [\mathbf{SL}_2(\mathbb{Z}) : C_0(N)]$$

By the multiplicative property of indices, we also have:

$$[\mathbf{SL}_2(\mathbb{Z}) : PC(N)] = [\mathbf{SL}_2(\mathbb{Z}) : C_0(N)][C_0(N) : PC(N)]$$

and since we know the value of the LHS from Proposition 6, it suffices to compute the index $[C_0(N) : PC(N)]$.

Consider the reduction map $\phi_N : \mathbf{SL}_2(\mathbb{Z}) \to \mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$. Under this map, the group $C_0(N)/PC(N)$ is mapped onto the subgroup of $\mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$ consisting of matrices of the form $\begin{pmatrix} a & b \\ 0 & a^{-1} \end{pmatrix}$. Since the number of units in $\mathbb{Z}/N\mathbb{Z}$ is given by $\Phi(N) = N\prod_{p|N}(1-p^{-1})$, the order of this subgroup is given by $N\Phi(N) = N^2\prod_{p|N}(1-p^{-1})$. Applying Proposition 6, we have

$$[\mathbf{SL}_2(\mathbb{Z}) : C_0(N)] = \frac{N^3\prod_{p|N}(1-p^{-2})}{N^2\prod_{p|N}(1-p^{-1})} = N\frac{\prod_{p|N}(1-p^{-1})(1+p^{-1})}{\prod_{p|N}(1-p^{-1})} = N\prod_{p|N}(1+p^{-1})$$

□

In the future, it will also be useful to compute an explicit set of representatives for $\Gamma(1)/\Gamma_0(N)$:

Proposition 1.11. *Let $N > 1$ be an integer and consider the set S of all pairs of integers $\{c, d\}$ satisfying:*

(i) $(c, d) = 1$

(ii) $d \mid N$

(iii) $0 < c \leq N/d$

For each pair $\{c, d\}$ find a, b such that $ad - bc = 1$. Then the set $T = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \{c, d\} \in S \right\}$ is a set of representatives for $\Gamma(1)/\Gamma_0(N)$.

Proof.

It is straightforward to check that every element in T is $\Gamma_0(N)$-inequivalent. But now, since there are precisely N/d c's for each $d \mid N$, we get

$$|T| = N + \frac{N}{d_1} + \ldots + 1 = N(\frac{1}{d_1} + \ldots + \frac{1}{N}) = N \prod_{p|N} \left(1 + p^{-1}\right) = [\Gamma(1) : \Gamma_0(N)]$$

and therefore these are all the representatives. □

As a subgroup of $\Gamma(1)$, the congruence subgroup $\Gamma_0(N)$ inherits an action on the extended upper-half plane $\widehat{\mathbb{H}}$. The quotient $X_0(N) = \Gamma_0(N)\backslash\widehat{\mathbb{H}}$ is again a Hausdorff space since $\Gamma_0(N)$ is discrete. Moreover, we can check as we have done for $X(1)$ that this space is compact and that we can put the complex analytic structure of a Riemann Surface on $X_0(N)$. We will call the $X_0(N)$ the *modular curves*.

In order to compute the genus of $X_0(N)$, note that the inclusion $\Gamma_0(N) \subset \Gamma(1)$ gives rise to a holomorphic map of compact Riemann surfaces

$$\begin{aligned} \pi : X_0(N) &\to X(1) \\ [z] &\longmapsto [z] \end{aligned}$$

given by simple projection of orbits. For any $N > 1$ the map is not constant, hence surjective of finite degree d. By the Riemann-Hurwitz formula:

$$2g(X_0(N)) - 2 = d \cdot (2g(X(1)) - 2) + \sum_{[z] \in X_0(N)} (e_{[z]} - 1)$$

where $e_{[z]}$ is the ramification index of π at $[z]$. We know already that $g(X(1)) = 0$ by Remark 1.8, so in order to compute the genus of $X_0(N)$ it suffices to compute the degree d of π and the various indices of the ramification points.

Proposition 1.12. *Let $\pi : X_0(N) \to X(1)$ be defined as above. Then*

(a) The degree of π is $d = [\Gamma(1) : \Gamma_0(N)]$.

(b) Fix an $s \in \widehat{\mathbb{H}}$. Define:

$$\Gamma(1)_s = \{A \in \Gamma(1) : A(s) = s\}$$

and

$$\Gamma_0(N)_s = \{A \in \Gamma_0(N) : A(s) = s\}$$

Then for every $[z] \in X_0(N)$*, we have* $e_{[z]} = [\Gamma(1)_z : \Gamma_0(N)_z]$.

Proof.

(a) Let $[w] \in X_0(N)$ and suppose $\pi([w]) = [z]$. If $A_1, \ldots, A_k \in \Gamma(1)$ is a sytem of coset representatives for $\Gamma_0(N)$ in $\Gamma(1)$, then $[A_i(w)] \neq [w]$ in $X_0(N)$, but $\pi([A(w)]) = \pi([w]) = [z]$ in $X(1)$. On the other hand, if B is in the coset $\Gamma_0(N)A_i$ for some i, then $A_i^{-1}B \in \Gamma_0(N)$, thus for any $[w] \in X_0(N)$, $[A_i^{-1}B(w)] = [w] \Rightarrow [B(w)] = [A_i(w)]$. Therefore there are at most $[\Gamma(1) : \Gamma_0(N)]$ fibers over $[z] \in X(1)$, given by $[w_1] = [A_1(z)], \ldots, [w_k] = [A_k(z)]$(note that one of the A_i's is the identity). Therefore the map has degree $d = [\Gamma(1) : \Gamma_0(N)]$.

(b) Let $s \in \widehat{\mathbb{H}}$ be a point which is fixed by some element of $\Gamma(1)$ (i.e. s is either a cusp or an elliptic point). Let $P \in \Gamma(1)_s$, but $P \notin \Gamma_0(N)_s$ and suppose that $P \in \Gamma_0(N)A_i$ for some i. Then we saw in part (a) that $[P(s)] = [A_i(s)] = [s]$ in $X_0(N)$, therefore the ramification degree above $[s]$ is at least 2, since this fiber corresponds to the fiber given by the identity. In fact, we see that the ramification index of $[s]$ is exactly the number of coset representatives for $\Gamma_0(N)_s$ in $\Gamma(1)_s$, i.e. $[\Gamma(1)_s : \Gamma_0(N)_s]$ (note that $\Gamma_0(N)_s \subset \Gamma(1)_s$). In particular if the point is unramfied, then it is not fixed by anything in $\Gamma(1)$ and therefore the formula holds trivially.

□

Moreover, we have the following:

Proposition 1.13. *Let* $[x] \in X(1)$ *and let* $\pi^{-1}([x]) = \{[w_1], \ldots, [w_h]\} \subset X_0(N)$*. If* $w_k = A_k(x)$ *with* $A_k \in \Gamma(1)$ *then* $e_{[w_k]} = [\Gamma(1)_x : A_k^{-1}\Gamma_0(N)A_k \cap \Gamma(1)_x]$*. Moreover,* $\Gamma(1) = \bigcup_{k=1}^{h} \Gamma_0(N)A_k\Gamma(1)_x$ *and the union is disjoint (i.e.* $h = \#\Gamma_0(N)\backslash\Gamma(1)/\Gamma(1)_x$*).*

Proof.

Suppose $B \in \Gamma(1)_{w_k}$ and $C \in \Gamma(1)_x$. Then

$$B(w_k) = BA_k(x) = A_k(x) = A_kC(x)$$

and therefore $B(x) = A_kCA_k^{-1}(x)$ for infinitely many $x \in \widehat{\mathbb{H}}$. Since this space is compact and both functions are holomorphic on all of $\widehat{\mathbb{H}}$, they must agree everywhere and therefore $B \in A_k\Gamma(1)_xA_k^{-1}$. Similarly, for any $C \in \Gamma(1)_x$,

$$A_kCA_k^{-1}(w_k) = A_k^{-1}C(x) = A_k(x) = w_k$$

and therefore $A_kCA_k^{-1} \in \Gamma(1)_{w_k}$. This shows that $\Gamma(1)_{w_k} = A_k\Gamma(1)_xA_k^{-1}$ and therefore $\Gamma_0(N)_{w_k} = \Gamma_0(N) \cap A^k\Gamma(1)_xA_k^{-1}$. Applying Proposition 1.12(b):

$$e_{[w_k]} = [\Gamma(1)_{w_k} : \Gamma_0(N)_{w_k}] = [\Gamma(1)_x : A_k^{-1}\Gamma_0(N)A_k \cap \Gamma(1)_x]$$

For the second statement, let $C \in \Gamma(1)$. Then $[C(x)]$ in $X_0(N)$ is equivalent to one of the $[w_k]$'s, since $[C(x)] = [x]$ in $X(1)$. Therefore there exists a $D \in \Gamma_0(N)$ such that $C(x) = D(w_k) = DA_k(x)$. Then $C^{-1}DA_k$ fixes x so $C^{-1}DA_k \in \Gamma(1)_x$. This proves that $\Gamma(1) = \bigcup_{k=1}^{h} \Gamma_0(N)A_k\Gamma(1)_x$. Now if $E \in \Gamma_0(N)A_k\Gamma(1)_x$, then $[E(x)] = [w_k]$ and therefore the union is disjoint. □

By Remark 1.5 we know that there is only 1 $\Gamma(1)$-inequivalent cusp on $X(1)$, namely $[\infty]$, and two elliptic points, namely $[i]$ and $[e^{2\pi i/3}]$, of order 2 and 3 respectively. Suppose $[w_k]$ lies above, say, $[i] \in X(1)$. Then there exists a $A_k \in \Gamma(1)$ such that $A_k(i) = w_k$. By Proposition 1.13, $e_{[w_k]} = [\Gamma(1)_i : A_k^{-1}\Gamma_0(N)A_k \cap \Gamma(1)_i]$. Now, $\Gamma(1)_i$ is isomorphic to the roots of the polynomial $x^2 + 1 = 0$ modulo ± 1, hence it is cyclic of order 2. Therefore $e_{[w_k]} = 1, 2$, depending on the conjugacy class of $A_k^{-1}\Gamma_0(N)A_k$. If $[w_k]$ lies above $[e^{2\pi i/3}]$ instead, then a similar argument shows that $\Gamma(1)_{e^{2\pi i/3}}$ is cyclic of order 3, and therefore $e_{[w_k]} = 1, 3$.

If $[w_k]$ lies above $[\infty]$, then the situation is slightly different. Since $\Gamma(1)_\infty = \left\{ \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} : m \in \mathbb{Z} \right\}$ is infinite, and $e_{[w_k]} = [\Gamma(1)_\infty : A_k^{-1}\Gamma_0(N)A_k \cap \Gamma(1)_\infty]$ must be finite, we must have that $A_k^{-1}\Gamma_0(N)A_k \cap \Gamma(1)_\infty$ is infinite as well. In other words, there is a nontrivial element $A_k^{-1}BA_k$ with $B \in \Gamma_0(N)$ such that:

$$A_k^{-1}BA_k(\infty) = \infty \Rightarrow BA_k(\infty)) = A_k(\infty) \Rightarrow B(w_k) = w_k$$

and therefore w_k is always fixed by an element of $\Gamma_0(N)$.

Remark 1.14. Now we have a complete picture of the ramification of $\pi : X_0(N) \to X(1)$: if $[z] \in X(1)$ and z is neither a cusp or an elliptic point, then $\Gamma(1)_z = \{I_2\}$ and $e_{[w_k]} = 1$ for each point of $X_0(N)$ lying above $[z]$. If z is an elliptic point, and $[w_k]$ lies above $[z]$, then we have two cases: either $A_k^{-1}\Gamma_0(N)A_k \cap \Gamma(1)_z$ is trivial and $e_{[w_k]} = 2, 3$, in which case we deduce that *there is no* $B \in \Gamma_0(N)$ fixing w_k, or $A_k^{-1}\Gamma_0(N)A_k \cap \Gamma(1)_z = \Gamma(1)_z$ and $e_{[w_k]} = 1$, in which case *there is* a $B \in \Gamma_0(N)$ fixing w_k. In the latter case, by analogy we call w_k a *elliptic point of order 2 (or 3)*(of $X_0(N)$) depending on whether B is of order 2 (or 3). Similarly, we say that a point w_k is a *cusp*(of $X_0(N)$) if there exists an element $P \in \Gamma_0(N)$ such that $P(w_k) = w_k$. Note that the discussion in the previous paragraph showed that, as opposed to the case with elliptic points, every point of $X_0(N)$ lying above $[\infty]$ *is* a cusp, and these are all the $\Gamma_0(N)$-inequivalent cusps of $X_0(N)$.

Define now the following constants:

$$\begin{aligned}
\mu &= [\Gamma(1) : \Gamma_0(N)] \\
\nu_2 &= \text{number of } \Gamma_0(N)\text{-inequivalent elliptic points of order 2} \\
\nu_3 &= \text{number of } \Gamma_0(N)\text{-inequivalent elliptic points of order 3} \\
\nu_\infty &= \text{number of } \Gamma_0(N)\text{-inequivalent cusps}
\end{aligned}$$

Then we can derive a formula for the genus of $X_0(N)$ in terms of $\mu, \nu_2, \nu_3, \nu_\infty$:

Proposition 1.15. *Let $N > 1$. Then the genus of $X_0(N)$ is given by:*

$$g(X_0(N)) = 1 + \frac{\mu}{12} - \frac{\nu_2}{4} - \frac{\nu_3}{3} - \frac{\nu_\infty}{2}$$

Proof.

Let $e_1, \ldots, e_h$ be the ramification indices of the points of $X_0(N)$ lying above the elliptic point $[e^{2\pi i/3}]$ of $X(1)$. We have $\mu = e_1 + \ldots e_h$ and $e_i = 1$ if the point is an elliptic point of order 3, and $e_i = 3$ otherwise. The number of i's for which $e_i = 1$ is therefore equal to ν_3. Let ν_3' be such that $h = \nu_3 + \nu_3'$. Then $\mu = \nu_3 + 3\nu_3'$ and therefore

$$\sum_{i=1}^{h} (e_i - 1) = \mu - h = 2\nu_3' = 2(\mu - \nu_3)/3$$

Similarly, for the points lying above $[i]$ we have

$$\sum_{i=1}^{h}(e_i - 1) = (\mu - \nu_2)/2$$

Now, the points lying above $[\infty]$ are *all* the $\Gamma_0(N)$-inequivalent cusps, and therefore $h = \nu_\infty$ which gives:

$$\sum_{i=1}^{h}(e_i - 1) = \mu - \nu_\infty$$

Applying the Riemann-Hurwitz formula:

$$2g(X_0(N)) - 2 = d \cdot (2g(X(1)) - 2) + \sum_{[z]\in X_0(N)} (e_{[z]} - 1)$$

with $g(X(1)) = 0$ we obtain:

$$2g(X_0(N)) - 2 = -2\mu + \frac{(\mu - \nu_2)}{2} + \frac{2(\mu - \nu_3)}{3} + (\mu - \nu_\infty)$$

and therefore

$$\begin{aligned} g(X_0(N)) &= -\mu + \frac{(\mu - \nu_2)}{4} + \frac{2(\mu - \nu_3)}{6} + \frac{\mu - \nu_\infty}{2} + 1 \\ &= 1 + \frac{\mu}{12} - \frac{\nu_2}{4} - \frac{\nu_3}{3} - \frac{\nu_\infty}{2} \end{aligned}$$

□

Since we know the value of μ from Proposition 1.10, in order to compute the genus of $X_0(N)$ in terms of N, it remains to compute the constants v_2, v_3, v_∞. This is done below:

Proposition 1.16. *Let $N > 1$ be an integer and let ν_∞, ν_2, ν_3 denote the set of $\Gamma_0(N)$-inequivalent cusps, elliptic points of order 3, elliptic points of order 3 respectively. If $N = p_1^{e_1} \cdots p_k^{e_k}$ is the prime factorization of N, then:*

(a) $\nu_\infty = \sum_{d|N, d>0} \Phi((d, N/d))$

(b) $\nu_3 = \begin{cases} 0 & \text{if } 9 \mid N \\ \prod_{i=1}^{k} \left(1 + \left(\frac{-3}{p_i}\right)\right) & \text{otherwise} \end{cases}$

(c) $\nu_2 = \begin{cases} 0 & \text{if } 4 \mid N \\ \prod_{i=1}^{k} \left(1 + \left(\frac{-1}{p_i}\right)\right) & \text{otherwise} \end{cases}$

Where Φ is Euler's function and $\left(\frac{\cdot}{p}\right)$ is the Legendre symbol, so that, for any prime p,

$$\left(\frac{-1}{p}\right) = \begin{cases} 0 & \text{if } p = 2 \\ 1 & \text{if } p \equiv 1 \mod 4 \\ -1 & \text{if } p \equiv 3 \mod 4 \end{cases}$$

$$\left(\frac{-3}{p}\right) = \begin{cases} 0 & \text{if } p = 3 \\ 1 & \text{if } p \equiv 1 \mod 3 \\ -1 & \text{if } p \equiv 2 \mod 3 \end{cases}$$

Proof.

(a) By Proposition 1.13, $\nu_\infty = \#\Gamma_0(N)\backslash\Gamma(1)/\Gamma(1)_x$ for any cusp $x \in \widehat{\mathbb{H}}$. We know by Proposition 1.11 that a set of representatives for $\Gamma_0(N)$ inside $\Gamma(1)$ can be given as follows. Take a pair of integers (c, d), $(c, d) = 1$, $d \mid N$, and find integers a, b such that $ad - bc = 1$. As d runs through the divisors of N, and $0 < c \leq N/d$, the matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ are going to give a complete set of representatives for $\Gamma_0(N)$ inside $\Gamma(1)$. Take now any cusp in $\Gamma(1)$, say 0. A generator for $\Gamma(1)_0$ is given by $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and therefore two elements A, A' of $\Gamma_0(N)/\Gamma(1)$ are equivalent under $\Gamma(1)_0$ if and only if they satisfy

$$A' = A \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix}$$

which happens exactly when $d = d'$ and $c' = c + dm$, for some $m > 0$. Fix a d, a divisor of N. By the above, we can always replace c by a residue mod d. Since $(c, d) = 1$, this means that we can replace c by an element of $(\mathbb{Z}/(d))^*$. But c was an element of $\mathbb{Z}/(N/d)$ in the first place, therefore c can be taken as an element of $(\mathbb{Z}/(d, N/d))^*$. There are exactly $\phi((d, N/d))$ elements in this group for each d and therefore

$$\nu_\infty = \sum_{d|N, d>0} \Phi((d, N/d))$$

as required.

(b) Suppose $[w_k]$ lies above $[\zeta = e^{2\pi i/3}]$, so that there exists an $A_k \in \Gamma(1)$ such that $A_k(\zeta) = w_k$. Suppose further that $[w_k]$ is an elliptic point of order 3, i.e. there exists a $B \in \Gamma_0(N)$ such that $B(w_k) = w_k$. Then $A_k^{-1}BA$ fixes ζ and therefore it must be one of:

$$D = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix} \quad D^2 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}$$

In other words, $\Gamma_0(N)_{w_k} = \{I_2, B, B^2\}$. Suppose that B is conjugate to, say, D. Now, let $B' \in \Gamma_0(N)$ be such that $B'(w) = w$ for some w which is $\Gamma_0(N)$-equivalent to w_k. Let C be such that $C(w_k) = w$. Since D and D^2 fix the same elements, suppose without loss of generality that C is conjugate to D. Then $C^{-1}B'C$ fixes w_k, so it must be in $\Gamma_0(N)_{w_k}$, i.e. it must be either B or B^2. But it cannot be B^2, for otherwise we could conjugate D with D^2 in $\Gamma(1)$, which is impossible. On the other hand, suppose that $C = M^{-1}BM$ in $\Gamma_0(N)$. Then C fixes $M^{-1}(w_k)$, which is equivalent to w_k in $X_0(N)$.

If we denote by S_1 the set of all the $B \in \Gamma_0(N)$ which are conjugate in $\Gamma(1)$ to D (resp. S_2 be the set of all the B's conjugate to D^2). Then what we just showed is that ν_3 is precisely the number of conjugacy classes of elements of S_1 (resp. S_2). In order to compute this number, we will put it in a bijection with a certain class of ideals in the ring $\mathbb{Z}[\zeta]$ as follows.

For any $B \in S_1 \cup S_2$, we have an action of $\mathbb{Z}[B]$ on

$$L = \mathbb{Z}^2 = \{(x, y) : x, y \in \mathbb{Z}\} \text{ and } L_N = \{(x, Ny) \in L : x, y \in \mathbb{Z}\}$$

that turns them into $\mathbb{Z}[B]$-modules. Note in particular that

$$\Gamma_0(N) = \{A \in \Gamma(1) : AL_N = L_N\}$$

Now, $\mathbb{Z}[B]$ is isomorphic to $\mathcal{A} = \mathbb{Z}[\zeta]$, and since $\mathcal{A}$ is a PID there exists an isomorphism (of $\mathbb{Z}$-modules) $f : \mathcal{A} \to L$ such that $f(\zeta x) = Bf(x)$ for all $x \in \mathcal{A}$ (just map (ζ) to (B)). On the other hand, the set T of all $\mathbb{Z}$-isomorphisms between $\mathcal{A}$ and L is the disjoint union of

$$T_i = \{f \in T : f(\zeta x) = Bf(x) \;\; B \in S_i\}, \quad i = 1, 2.$$

Note moreover that if $\alpha \in M_2(\mathbb{Z})$ has $\det(\alpha) = -1$, then $f \in T_1 \Leftrightarrow \alpha f \in T_2$. Now pick any $f : \mathcal{A} \to L \in T_1$ and consider the set $\mathcal{I} = f^{-1}(L_N) \subset \mathcal{A}$. Suppose that $B \in \Gamma_0(N)$ and that it satisfies $f(\zeta x) = Bf(x)$ for all $x \in \mathcal{A}$. We want to show that in this case $\mathcal{I}$ is an ideal of $\mathcal{A}$. Closure under addition is obvious. Let $i \in \mathcal{I}$ and $x = a + b\zeta \in \mathcal{A}$. Then

$$f(ix) = f(ia + ib\zeta) = af(i) + bBf(i) \in L_N$$

since B fixes L_N. Conversely, it is clear that if $\mathcal{I}$ is an ideal of $\mathcal{A}$, then B must fix L_N. Therefore $\mathcal{I}$ is an ideal of $\mathcal{A}$ if and only if $B \in \Gamma_0(N)$. Moreover, $\mathcal{A}/\mathcal{I} \cong L/L_N \cong \mathbb{Z}/N\mathbb{Z}$, which implies that

(i) $N_{\mathbb{Q}(\zeta)/\mathbb{Q}}(\mathcal{I}) = N$

(ii) $\mathcal{I}$ is not contained in any rational integer other than 1.

Note that (i) is true by definition and (ii) because otherwise the residue of $\mathcal{A}/\mathcal{I}$ would not be cyclic. On the other hand, any such ideal of $\mathcal{A}$ is such that $\mathcal{A}/\mathcal{I}$ isomorphic to L/L_N and therefore we can always find an element $f \in T$ such that $f(\mathcal{I}) = L_N$. We can also assume $f \in T_1$, for we can always multiply it by a matrix α with determinant -1 to get $\alpha f \in T_2$. We therefore obtain an element $B \in S_1 \cap \Gamma_0(N)$ by letting B be the generator of $(f(\zeta))$ (note that we are using the fact that $\mathcal{A}$ is a PID). We want to show that this correspondence between ideals $\mathcal{I} \subset \mathcal{A}$ satisfying (i) and (ii) and conjugacy classes of elements of $S_1 \cap \Gamma_0(N)$ in $\Gamma_0(N)$ is one-to-one.

Suppose first that for the same ideal $\mathcal{I}$ we find two isomorphisms $f, f' \in T_1$ with $(f(\zeta)) = (B)$ and $(f'(\zeta)) = (B')$. Since f, f' are isomorphisms, inverses are defined and $f \circ f^{-1}$ is a well-defined automorphism of L. Therefore $f \circ f^{-1} \in \mathbf{SL}_2(\mathbb{Z})$ and in particular we can find an element $\gamma \in \Gamma(1)$ such that $f' = \gamma f$. Then $B = \gamma^{-1}B'\gamma$. Moreover, since $f \circ f^{-1}(L_N) = \gamma(L_N) = L_N$, γ is actually in $\Gamma_0(N)$ and B is conjugate to B' in $\Gamma_0(N)$. Conversely, suppose that we have two different ideals $\mathcal{I}$ and $\mathcal{J}$ satisfying properties (i) and (ii). Select isomorphisms f,$f' \in T_1$ such that $f(\mathcal{I}) = L_N$ and $f'(\mathcal{J}) = L_N$ with $(f(\zeta)) = (B)$ and $(f(\zeta)) = (B')$. Suppose moreover that $B = \gamma^{-1}B\gamma$, $\gamma \in \Gamma_0(N)$. Let $h = f^{-1}\gamma f'$. This is a well-defined automorphism of $\mathcal{A}$ with $(h(\zeta)) = (\zeta)$. In particular, h is just multiplication by a unit $\lambda \in \mathcal{A}^\times$. But then $\mathcal{I} = f^{-1}(L_N) = f^{-1}(\gamma L_N) = f^{-1}(\gamma(f'(\mathcal{J}))) = \lambda\mathcal{J} = \mathcal{J}$.

Now that we have established a bijection between ideals $\mathcal{I} \subset \mathcal{A}$ satisfying (i) and (ii) and conjugacy classes in $\Gamma_0(N)$ of elements of $S_1 \cap \Gamma_0(N)$, it remains to compute how many such ideals are there. That number will be ν_3. Suppose $\mathcal{I} = (a + b\zeta) \subset \mathcal{A}$ satisfies properties (i) and (ii). Then $N_{\mathbb{Q}(\zeta)/\mathbb{Q}}(\mathcal{I}) = (a + b\zeta)(a + b\zeta^2) = N$. Factor (N) into prime ideals of $\mathcal{A}$ as $(N) = P_1^{e_1} \cdots P_k^{e_k}$. If P_i is a rational prime, then $P_i \supset (a+b\zeta) = \mathcal{I}$ which violates property (ii). Therefore all the P_i's are primes of $\mathbb{Z}[\zeta]$ lying above a rational prime that splits or that ramifies. Since primes that split come in conjugate pairs in any factorization of a rational integer, and the only prime that ramifies in $\mathcal{A}$ is 3, we can actually write $(N) = \pi^e \cdot P_1^{e_1} \cdots P_r^{e_r} \cdot \sigma(P_1)^{e_1} \cdot \sigma(P_r)^{e_r}$ where σ is the usual conjugation in $\mathbb{C}$ and $(\pi)^2 = 3$. Suppose first $e = 0$. By letting $\mathcal{I} = P_1^{e_1} \cdots P_r^{e_r}$, we obtain an ideal of the desired form. But the ideal will change if we switch an ideal P_i for its conjugate $\sigma(P_i)$. Each spot $i = 1, \ldots r$ can either be taken by a prime P_j or by its conjugate, each time giving a different ideal $\mathcal{I}$. Therefore there are 2^r possiblities, where r is the number of primes dividing N that split. A prime $p \mid N$ splits in $\mathbb{Z}[\zeta]$ if and only if the Legendre symbol $\frac{-3}{p}$ is 1, and therefore we get the desired formula. If $e = 2$, then $3 \mid N$ and no higher powers of 3 do. But conjugation does not change π so its presence in the factorization of N does not affect the total number of ways of writing an $\mathcal{I}$. The formula takes it into account by letting the Legendre symbol of 3 be 0. Note however that if $9 \mid N$ then we have to put at least two powers of π inside $\mathcal{I}$ and that is not possible for then $(3) \supset \mathcal{I}$. Similarly, if $p \mid N$ does not split, then there are no ideals of the required form, since all the $\mathcal{I}$ manufactured as above will be contained in (p). Since for these primes the Legendre symbol is -1, the formula also takes into account their effect.

(c) Proceed as before, by letting $\mathcal{A} = \mathbb{Z}[i]$. Note that in this ring the only prime that ramifies is 2, and a prime splits if and only if $\left(\frac{-1}{p}\right) = 1$.

□

We derive a few simple but meaningful Corollaries:

Corollary 1.17. *Let N be a prime. Then the only cusps of $X_0(N)$ are $[0]$ and $[\infty]$.*

Proof.

First of all, we claim that if $N > 1$, then the two points $[0], [\infty]$ are inequivalent cusps in $X_0(N)$. It is clear that they are cusps, for ∞ is fixed by $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \Gamma_0(N)$ and 0 is fixed by $\begin{pmatrix} 1 & 0 \\ N & 1 \end{pmatrix} \in \Gamma_0(N)$. Suppose that they are equivalent in $\Gamma_0(N)$. Then there exists a matrix $A = \begin{pmatrix} a & b \\ cN & d \end{pmatrix} \in \Gamma_0(N)$ such that $f_A(\infty) = 0 \Rightarrow a/cN = 0 \Rightarrow a = 0$. Then $|\det A| = |bcN| = 1$ but this is impossible since b, c are integers and $N > 1$. Now, from Proposition 1.16 we have that,for N prime,

$$\nu_\infty = \sum_{d|N, d>0} \Phi((d, N/d)) = \Phi((1, N)) + \Phi((N, 1)) = 2\Phi(1) = 2$$

and therefore $[0]$ and $[\infty]$ are the *only* cusps. □

Corollary 1.18. *Let N be a prime number. Then $g(X_0(N)) = 0$ if and only if $N = 2, 3, 5, 7, 13$.*

Proof.

For any prime N, we have that $\mu = N \prod_{p|N}(1 + p^{-1}) = N + 1$ by Proposition 1.10 and that $\nu_\infty = 2$ by Corollary 1.17. If we also plug $g = 0$ in the formula of Proposition 1.15, we get that:

$$N = 3\nu_2 + 4\nu_3 - 1$$

Now, by direct computation, we can check that $g(X_0(N)) = 0$ for $N = 2$ and $N = 3$. For any other prime N, note that N must be congruent to one of $1, 5, 7, 11$ modulo 12. The first case yields $N = 13$, the second $N = 5$, the third $N = 7$ and the fourth $N = -1$, which we discard. □

2 Connection with Elliptic Curves

Now that we have familiarized ourselves with the structure of the modular curves $X_0(N)$, we are going to look at some of their maps, and their connection with the theory of elliptic curves.

We outline first the point of view of modular curves as moduli spaces for elliptic curves. Then, we describe the involution map w_N and the Hecke correspondences T_ℓ defined on $X_0(N)$ and their connection with the theory of elliptic curves.

From now on, we will assume that N is prime, since it will be the only case studied in the proof of Mazur's Theorem.

2.1 Modular Interpretation of $X_0(N)$

Recall from the Uniformization Theorem for elliptic curves that for any lattice $\Lambda \in \mathbb{C}$ there exists a biholomorphic isomorphism:

$$\begin{aligned} \mathbb{C}/\Lambda &\to E_\Lambda \\ z &\longmapsto (\wp(z,\Lambda), \wp'(z,\Lambda)) \end{aligned}$$

where E_Λ is an elliptic curve and $\wp$ is the Weierstrass function. If $E_{\Lambda'}$ is another such curve with associated lattice Λ' and $\phi : E_\Lambda \to E_{\Lambda'}$ is an isomorphism, then ϕ induces an isomorphism $\tilde{\phi} : \Lambda \to \Lambda'$ which is just multiplication by a constant $\alpha \in \mathbb{C}^\times$. Denoting by $\mathcal{L}$ the set of all lattices in $\mathbb{C}$ and by $\mathcal{E}$ the set of all elliptic curves over $\mathbb{C}$, we have a bijection:

$$\mathcal{L}/\mathbb{C}^\times \to \mathcal{E}/\sim$$

Pick a basis ω_1, ω_2 for Λ. We indicate this by $\Lambda = \langle \omega_1, \omega_2 \rangle$. Then a representative for the the coset of Λ inside $\mathcal{L}/\mathbb{C}^\times$ is given by letting $\tau = \omega_1/\omega_2$ in such a way that $\Im[\tau] > 0$. In other words, we have a surjection:

$$\begin{aligned} \mathbb{H} &\to \mathcal{L}/\mathbb{C}^\times \\ \tau &\longmapsto \langle \tau, 1 \rangle \end{aligned}$$

From which we derive the following:

Proposition 2.1. *There is a bijection*

$$Y(1) \to \mathcal{L}/\mathbb{C}^\times$$

Proof.

Suppose τ, τ' map to Λ, Λ' respectively, with $\alpha\Lambda = \Lambda'$. Then $\tau' = a\alpha\tau + b\alpha$ and $1 = c\alpha\tau + d\alpha$. Thus

$$\tau' = \frac{a\tau + b}{c\tau + d}$$

for some $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1)$. On the other hand, suppose $\tau' = A\tau$ for some $A \in \Gamma(1)$. Let $\alpha = c\tau + d$. Then

$$\alpha\Lambda' = \langle c + d\tau, a + b\tau \rangle = \Lambda$$

and therefore $\Lambda \equiv \Lambda'$ in $L/\mathbb{C}^\times$. We therefore have a bijection:

$$Y(1) = \Gamma(1)\backslash\mathbb{H} \to \mathcal{L}/\mathbb{C}^\times$$

□

Composing with the bijection given by the Uniformization Theorem, we see that points in $Y(1)$ correspond bijectively to isomorphism classes of elliptic curves over $\mathbb{C}$. Note in particular that the elliptic points $[i]$ and $[e^{2\pi i/3}]$ correspond to isomorphism of classes of ellipitic curves with complex multiplication by the rings $\mathbb{Z}[i]$ and $\mathbb{Z}[e^{2\pi i/3}]$, respectively.

There is a way to extend this bijection to all of $X(1)$ by making the cusps $[\infty] \cup \mathbb{Q}$ correspond to particular objects called 'generalized elliptic curves'. A discussion of this topic here would take us off track (though see [Del]), but it is enough to say that in this way we can extend the bijection

$$j : \mathcal{E}/\sim \to \mathbb{C}$$

given by the j-invariant to an isomorphism of Riemann surfaces

$$j : X(1) \to \mathbb{P}^1$$

giving again the result that $X(1)$ has genus 0, a fact which was proven earlier giving an informal proof.

Remark 2.2. It turns out that the function $j(\tau)$ generates the field of rational functions of $X(1)$ over $\mathbb{C}$. For details, see [Si2 I.4].

Definition 2.3. Denote by (E, H) the pair given by an elliptic curve and a subgroup H of E cyclic of order N. Two pairs (E', H') are isomorphic if there is an isomorphism $E \to E'$ carrying $H \to H'$.

Proposition 2.4. *Two pairs (E_Λ, H) and $(E_{\Lambda'}, H'$ of elliptic curves over $\mathbb{C}$ are isomorphic if and only if there exists an element $A \in \Gamma_0(N)$ such that $A\Lambda = \Lambda'$.*

Proof.

Suppose that (E_Λ, H) is given by the lattice $\Lambda = \langle \tau, 1 \rangle$ and the cyclic subgroup H of order N is given by $\langle 1/N \rangle$, without loss of generality. Suppose further that $\Lambda' = \langle \tau', 1 \rangle$ and $H' = \langle 1/N \rangle$. If (E_Λ, H) is isomorphic to $(E_{\Lambda'}, H)$ then there exists a $\alpha \in \mathbb{C}^\times$ such that $\alpha\Lambda = \Lambda'$. In other words, there is an element $A \in \Gamma(1)$ such that:

$$\alpha\tau = a\tau' + b$$
$$\alpha \cdot 1 = c\tau' + d$$

In particular, $\alpha \cdot 1/N = c/N\tau' + d/N$ must be contained in $\langle 1/N \rangle$, since H must map to H'. This can only happen if $c \equiv 0(N)$.

On the other hand, suppose that there exists and $A \in \Gamma_0(N)$ such that $A(\tau) = \tau'$. Then the group $\langle 1/N \rangle$ maps to $\langle c/N + d/N \rangle$, which is cyclic of order N, since $c \mid N$.

□

In other words, we have a bijection:

$$Y_0(N) \longrightarrow \{E \in \mathcal{E}/\sim : H \subset E \text{ is cyclic of order N}\}$$

We refer to $Y_0(N)$ as the moduli space of all elliptic curves defined over $\mathbb{C}$ containing a cyclic subgroup of order N. Similarly, $X(1)$ is a moduli space for the space of all elliptic curves defined over $\mathbb{C}$. We will refer to these interpretations of $X_0(N)$ and $X(1)$ as the *modular interpretation*.

Remark 2.5. Similar to the $j(\tau)$ function for $X(1)$, we have a function $j_N(\tau) = j(\tau)$ such that the function field of $X_0(N)$ over $\mathbb{C}$ is precisely $\mathbb{C}(j, j_N)$.

2.2 The involution w_N

Let E be an elliptic curve over $\mathbb{C}$ with a cyclic subgroup H of order N. By [Si1 III.4.13] there exists an elliptic curve E' defined over $\mathbb{C}$ and an isogeny (of degree N) $\phi : E \to E'$ such that $\ker(\phi) = H$. The group structure of E' is given by $E' \cong E/H$. By [Si1 III.6.1] we can also find an isogeny $\hat{\phi} : E' \to E$ (of degree N) such that $\hat{\phi} \circ \phi = [N]$. Since nonconstant isogenies are surjective, it follows that

$$\ker(\hat{\phi}) \cong H' \subset E[N] \cong \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$$

Where H' is of order N, and the inclusion is proper. $H' \subset E'$ is cyclic of order N. Using the modular interpretation, the dual isogeny gives a well defined bijection:

$$\begin{aligned} w_N : Y_0(N) &\to Y_0(N) \\ (E, H) &\longmapsto (E' = E/H, H' = E[N]/H) \end{aligned}$$

which is in fact a morphism of varieties. Note that since $\hat{\hat{\phi}} = \phi$, w_N is of order 2, hence the term 'involution'.

We can in fact calulate an explicit formula for w_N:

$$\begin{aligned} w_N : Y_0(N) &\longrightarrow Y_0(N) \\ [z] &\longmapsto \left[\frac{-1}{Nz}\right] \end{aligned}$$

This is easily seen by writing the 'modular' formula in terms of lattices. Specifically, if $E \cong \mathbb{C}/\langle 1, \tau\rangle$ and H is the subgroup $\langle 1/N\rangle$ then E maps to

$$E/H = \mathbb{C}/\langle \frac{1}{N}, \tau\rangle = \mathbb{C}/\langle 1, \frac{1}{\tau N}\rangle$$

However, $\Im[\frac{1}{\tau N}] < 0$ so we want to take the equivalent lattice $E' = \mathbb{C}/\langle 1, \frac{-1}{\tau N}\rangle$ to stick to our convetion that $\tau \in \widehat{\mathbb{H}}$. Note in particular that this formula can be extended to the cusps to define an involution

$$w_N : X_0(N) \longrightarrow X_0(N)$$

In particular, we have:

Proposition 2.6. *$w_N([0]) = [\infty]$ and $w_N([\infty]) = [0]$, so that w_N interchanges the cusps of $X_0(N)$.*

Proof.

This is clear from the formula $[z] \mapsto \left[\frac{-1}{Nz}\right]$. □

Remark 2.7. If $N = N_1N_2$ is a product of primes, then one can work out a similar definition

$$w_{N_1}(E, H) = (E/H_{N_1}, (H + E[N_1])/H_{N_1})$$

where H_{N_1} denotes the unique subgroup of H of order N_1.

2.3 Hecke Correspondences

Let ℓ be any prime number such that $\ell \neq N$. Consider a pair (E, H) of an elliptic curve defined over $\mathbb{C}$ and a cyclic subgroup of order $N\ell$. By the Sylow theorems, H contains a unique cyclic subgroup of order N, call it H_N (recall our tacit assumption that N is prime). We have a well-defined map

$$\begin{aligned} \pi_{N,\ell} : Y_0(N\ell) &\to Y_0(N) \\ (E, H) &\to (E, H_N) \end{aligned}$$

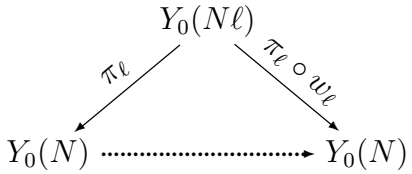
On the other hand, H also contains a unique subgroup of order ℓ, call it H_ℓ. By Remark 2.7 we have

$$w_\ell(E, H) = (E/H_\ell, (H + E[\ell])/H_\ell)$$

Note that the curve E/H_ℓ contains a group $(H_N + H_\ell)/H_\ell$ of order N, since $H_N \cap H_\ell = \emptyset$. Applying $\pi_{N,\ell}$ to E/H_ℓ we get:

$$\pi_{N,\ell} \circ w_\ell(E, H) = (E/H_\ell, (H_N + H_\ell)/H_\ell)$$

We have the following diagram:



which suggests the following correspondence on $Y_0(N)$. Let $[z] = (E, H_N)$ be a point of $Y_0(N)$. For each subgroup H_ℓ^i of E, we can consider the point $[z_i] = (E, H_\ell^i H_N)$ lying above it in $Y_0(N\ell)$. On the other hand, we can map each of the $[z_i]$ down to $Y_0(N)$ using $\pi_{N,\ell} \circ w_\ell$. We define

$$\begin{aligned} T_\ell(z) &= \sum_i \pi_{N,\ell} \circ w_\ell(z_i) \\ &= \sum_i (E/H_\ell^i, (H_N + H_\ell^i)/H_\ell^i) \end{aligned}$$

where the sum runs through all the subgroups of E of order ℓ. If E corresponds to a lattice $\Lambda \subset \mathbb{C}$, then this number is precisely the number of sublattices of Λ of index ℓ, which is $\ell + 1$. Therefore T_ℓ sends a point $[z] \in Y_0(N)$ to a (formal) sum of $\ell + 1$ points in $Y_0(N)$.

It can be shown by looking at what happens to the lattices that the action of T_ℓ is given explicitly on $Y_0(N)$ by the formula:

$$T_\ell([z]) = \ell[z] + \sum_{j=0}^{\ell-1} \left[\frac{z+j}{\ell}\right] \tag{I.1}$$

which again can be extended to the cusps of $X_0(N)$ to obtain a correspondence:

$$T_\ell : X_0(N) \to X_0(N)$$

for each $\ell \neq N$. We call these correspondences *Hecke operators*. The reader is warned that although z represents a complex number, the summation in (1) is by no means related to the summation of points in $\mathbb{C}$. It is just a formal sum of points of $X_0(N)$. In particular, we have

Proposition 2.8.

(a) $T_\ell([\infty]) = (\ell + 1)[\infty]$

(b) $T_\ell[0] = (\ell + 1)[0]$

Proof.

(a) Obvious from (1).

(b) Note that $\gcd(N, \ell) = 1$ so we can find integers a, b such that $a\ell - bN = 1$. Then, the transformations

$$A_j(z) = \frac{az + j}{bNz + \ell}$$

are in $\Gamma_0(N)$ for all $j = 1, \ldots, j - 1$ and $A_j(0) = j/\ell$, which shows that the cusps $[j/\ell]$ are all equivalent to $[0]$ in $X_0(N)$.

□

We have only defined T_ℓ for ℓ a prime, but an anlogous definition can be made for any positive integer n. We will not get into a detailed derivation of the formulas for T_n, n any integer, but we limit ourselves to note the following:

Proposition 2.9. *Let n, m be positive integers such that $(n, m) = 1$. Then $T_n T_m = T_{nm}$*

Proof. See [Si I.9.1] □

Chapter II

The Eisenstein Ideal

In the previous chapter we gave a geometric construction of the modular curves $X_0(N)$ over $\mathbb{C}$ and we studied some of their general properties. We also mentioned (Remark 2.5) that $X_0(N)/\mathbb{C}$ can also be characterized as the curve with function field equal to $\mathbb{C}(j, j_N)$. The advantage of this construction is that it can be used to construct curves $X_0(N)/\mathbb{Q}$ characterized by the fact that their function field is $\mathbb{Q}(j, j_N)$. Since in this chapter we will be mainly interested in questions of $\mathbb{Q}$-rationality, all of our constructions (including the curves $X_0(N)$) will be done over $\mathbb{Q}$. The interested reader may refer to [Roh] for explanations of these constructions.

A central part of Mazur's argument is to prove that the curve $X_0(N)(\mathbb{Q})$ has finitely many points. In this chapter we describe how this is done, following the outline given in [MaSe] though we will refer to [Maz] for the more technical points.

The general plan is to project $X_0(N)(\mathbb{Q})$ onto a certain abelian variety $J(\mathbb{Q})$ via a finite map, and then show that the group $J(\mathbb{Q})$ is finite. We have divided the proof in three secions: first, we build up the definition of J and describe some of its structure; second, we describe the notion of p-admissible filtrations and show how to bound the size of certain cohomology groups using them; third, we apply the method of descent to show finiteness of $J(\mathbb{Q})$, from which finiteness of $X_0(N)(\mathbb{Q})$ immediately follows.

The theorem we are trying to prove requires a deep understanding of scheme

theoretic tools such as group schemes, Néron models and flat cohomology. We could have inserted here a brief introduction to these topics, but nobody would have benefited from it. It would have been too short of an introduction for the beginner to gain an understanding of these objects, and boring for the experienced reader who has already mastered the subject. Wherever it was possible, we have attempted to translate the proofs in a simpler language, and we hope the advanced reader will forgive us if at times this language seems imprecise. Where it was not possible, we have stated the result needed and we have required the reader to either take a leap of faith or to check the reference we give for the proof.

Throughout the chapter, we will be assuming that N is a prime. Moreover we are assuming that $g(X_0(N)) > 0$, so that $N \neq 2, 3, 5, 7, 13$ by Corollary I.1.18.

1 The Eisenstein Quotient

We are going to define a certain abelian variety $J_0(N)$ in which $X_0(N)$ embeds, and then describe one of its quotients J, called by Mazur the *Eisenstein Quotient*. This roughly corresponds to sections II.6-7, and II.10 of [Maz].

1.1 The Jacobian Variety $J_0(N)$

Let C be a nonsingular projective curve of genus g over an algebraically closed field k. Recall from [Si1 II.3] that by a *divisor* on C we mean a formal sum

$$D = \sum_{P \in C} n_P [P]$$

where the n_P's are integers only finitely many of which are nonzero. The set of all such sums forms a free abelian group which we denote by $\mathrm{Div}(C)$. There is a group homomorphism between $\mathrm{Div}(C)$ and $\mathbb{Z}$ given by

$$\deg(D) = \sum_{P \in C} n_P$$

where the sum takes place in $\mathbb{Z}$. We let $\mathrm{Div}^0(C)$ be the kernel of this map (the *degree zero divisors*). For any rational function $f \in k(C)$ (the function field of C)

we can define a divisor by the formula

$$\mathrm{div}(f) = \sum_{P \in C} \mathrm{ord}_f(P)[P]$$

where by $\mathrm{ord}_f(P)$ we just mean the order of vanishing of f at P. We call such divisors *principal divisors*. Note that principal divisors form a subgroup of $\mathrm{Div}^0(C)$, by [Si1 II.3.7], since every principal divisor is of degree 0. However, the converse is not true whenever $g > 0$, and the discrepancy is measured by the group:

$$\mathrm{Pic}^0(C) = \mathrm{Div}^0(C)/\text{principal divisors}$$

which injects into the larger group

$$\mathrm{Pic}(C) = \mathrm{Div}(C)/\text{principal divisors}$$

In other words, we have an exact sequence of abelian groups:

$$0 \to \mathrm{Pic}^0(C) \to \mathrm{Pic}(C) \to \mathbb{Z} \to 0$$

which is the analogous for curves of the ideal class group exact sequence in the study of number fields.

For each point $P \in C$, we can consider its image $[P]$ in the group $\mathrm{Pic}^0(C)$, whenever this is nontrivial (i.e. $g > 0$). If C is an elliptic curve (i.e. g=1), we know this map is actually a bijection ([Si1 III.3.4]) given by $P \mapsto [P] - [O]$, which is an isomorphism between $\mathrm{Pic}^0(C)$ and the group formed by the points of C. But for curves of higher genus the Riemann-Roch theorem tells us that $\mathrm{Pic}^0(C)$ is much larger than C, and we cannot possibly hope to construct such an isomorphism, let alone a bijection. The idea, however, is to define an abelian variety of dimension g (i.e. a variety with morphisms defining an abelian group on the points) in which C embeds, and whose corresponding group is isomorphic to $\mathrm{Pic}^0(C)$.

Proposition 1.1. *Let C be a nonsingular projective curve of genus g defined over an algebraically closed field.*

(a) There exists an abelian variety $\mathrm{Jac}(C)$ of dimension g such that there is an isomorphism of groups:

$$\mathrm{Pic}^0(C) \xrightarrow{\cong} \mathrm{Jac}(C)$$

We call $\mathrm{Jac}(C)$ the Jacobian of C.

(b) There is an embedding $C \hookrightarrow \mathrm{Jac}(C)$.

Proof.

(a) Since the methods of proof of this statement do not find another application in the context of Mazur's Theorem, we will limit ourselves to an overview of how the Jacobian is constructed.

First select a point $O \in C$, whose image will serve as the identity in $\mathrm{Jac}(C)$. Consider C^g, the product of g copies of C, and define a map:

$$\begin{aligned} \phi : C^g &\longrightarrow \mathrm{Pic}^0(C) \\ (P_1, \ldots, P_g) &\longmapsto [P_1] + \ldots + [P_g] - g[O] \end{aligned}$$

Note that for the case $g = 1$ we recover the isomorphism $P \mapsto [P] - [O]$. The problem this time is that this map is evidently not injective. Namely, every permutation of S_g (the symmetric group on g elements) applied to the P_i's will map to the same point of $\mathrm{Pic}^0(C)$. Consider then $C^{(g)} = C^g/S_g$. This seems to be a reasonable guess, and it turns out to be the correct one. Checking injectivity and surjectivity is not very hard. It is the definition of a group law on $C^{(g)}$ isomorphic to the group law of $\mathrm{Pic}^0(C)$ that is somewhat tricky, and we refer to [Weil] for a complete description of how this can be done.

(b) Note that the map

$$\begin{aligned} \psi : C &\longrightarrow \mathrm{Jac}(C) \\ P &\longmapsto [P] - [O] \end{aligned}$$

is an embedding of C into its Jacobian. This is easily seen by viewing ψ as a restriction of ϕ given by

$$\begin{aligned} \phi|_{\{(P,O,\ldots,O) \in C^{(g)} : P \in C\}} C &\longrightarrow \mathrm{Pic}^0(C) \\ (P_1, O, \ldots, O) &\longmapsto [P_1] + [O] + \ldots + [O] - g[O] = [P_1] - [O] \end{aligned}$$

□

Remark 1.2. For arbitrary k, and C defined over k, the only difference is that $\mathrm{Jac}(C)$ will also be defined over k, and that the isomorphism ϕ will be compatible with $\mathrm{Gal}(\overline{k}/k)$.

Going back to our application, we let $J_0(N) = \mathrm{Jac}(X_0(N))$, where we let $O = [\infty]$ (note that $\mathrm{Pic}^0(X_0(N))$ is nontrivial by our assumptions on N). This is an abelian variety defined over $\mathbb{Q}$, and the embedding:

$$\begin{aligned} \psi : X_0(N) &\longrightarrow \mathrm{Jac}(C) \\ [z] &\longmapsto [z] - [\infty] \end{aligned}$$

is defined over $\mathbb{Q}$ and it is compatible with the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$.

Remark 1.3. The maps w_N and T_ℓ defined in $I.2.2$ and $I.2.3$ extend via the embedding ψ to give endomorphisms (defined over $\mathbb{Q}$) of $J_0(N)$. In fact, note that $J_0(N)$ is a much more natural setting for the Hecke operators T_ℓ to act: the formal sums defining them become formal sums of divisors inside $\mathrm{Pic}^0(C)$, which correspond to sums using the group law of $J_0(N)$.

Remark 1.4. Note that, as an abelian variety, $\mathrm{Jac}(C) = C^{(g)}$ is generated by the points of C, an observation which will be useful in the future.

1.2 The Group C

We are going to focus our attention on the point $c = \psi([0]) = [0] - [\infty] \in J_0(N)$. By [Ogg, Prop. 2] the point $[0]$ belongs to $X_0(N)(\mathbb{Q})$ for N prime, and therefore it maps to a point of $J_0(N)(\mathbb{Q})$. We will denote by C the subgroup of $J_0(N)(\mathbb{Q})$ generated by c. We have the following:

Proposition 1.5. *The group C is a cyclic group of order $n = \mathrm{num}\left(\frac{N-1}{12}\right)$*

Proof. This is [Ogg, Theorem]. □

Note in particular that whenever $\mathrm{num}\left(\frac{N-1}{12}\right) \neq 1$ we have $c \neq [\infty] \in J_0(N)$. This is always satisfied in our case, since we are assuming $N \neq 2, 3, 5, 7, 13$. Moreover:

Proposition 1.6.

(a) $w_N(c) = -c$

(b) $T_\ell(c) = (\ell + 1)c$

Proof.

(a) From Chapter 1, Proposition 2.6, we have:

$$w_N(c) = w_N([0]) - w_N([\infty]) = [\infty] - [0] = -c$$

(b) From Chapter 1, Proposition 2.8, we have:

$$T_\ell(c) = T_\ell([0]) - T_\ell([\infty]) = (\ell + 1)([0]) - (\ell + 1)([\infty]) = (\ell + 1)c$$

□

In particular, note that the group C is invariant under the action of T_ℓ and w_N.

1.3 The Hecke Algebra

Denote by $\mathrm{End}(J_0(N))$ the set of all endomorphism of $J_0(N)$ defined over $\mathbb{C}$ ($\mathrm{End}_{\mathbb{Q}}$ will denote the ones defined over $\mathbb{Q}$). Being $J_0(N)$ an abelian variety, we can 'add' two endomorphisms by pointwise addition in the group law and we can 'multiply' them using composition. It is straightforward to check then that $\mathrm{End}(J_0(N))$ forms a ring.

By Remark 1.3, we know already two kinds of endomorphisms: the involution w_N from and the Hecke operators T_ℓ for ℓ a prime $\ell \neq N$. These were defined on $X_0(N)$ but by they both extend to give well-defined $\mathbb{Q}$-endomorphisms of $J_0(N)$.

Definition 1.7. Denote by $\mathbf{T} \subset \mathrm{End}(J_0(N))$ the subring generated by w_N and by T_ℓ for every prime $\ell \neq N$. By extension of scalars, we can form the $\mathbb{Q}$-algebra $\mathbb{Q} \otimes \mathbf{T}$, wich we call the *Hecke algebra*.

Ken Ribet proved the following:

Theorem 1.8. *Let N be a prime. Then $\mathbb{Q} \otimes \mathbf{T} \cong \mathrm{End}(J_0(N)) \otimes \mathbb{Q}$*

Proof.

See [Rib] □

In particular, note that every endomorphism of $J_0(N)$ is defined over $\mathbb{Q}$.

It can be shown by direct computation that the $\mathbb{Q}$-algebra $\mathbb{Q} \otimes \mathbf{T}$ is commutative and free of rank $g = \dim J_0(N) = g(X_0(N))$. Moreover:

Theorem 1.9. *$\mathbb{Q} \otimes \mathbf{T} \cong \prod_\alpha K_\alpha$, where each K_α is a totally real number field.*

Proof. This is a done in [AtLe], Lemmas 13 and 27. □

For each α we have a map $\mathbf{T} \to K_\alpha$ given by projection. Its kernel is an ideal $\mathfrak{p}_\alpha$ which is prime and does not contain any other prime ideal (i.e the $\mathfrak{p}_\alpha$ are *minimal*). Moreover, these are all the minimal prime ideals of $\mathbf{T}$.

Remark 1.10. Following the discussion of Ribet in [Rib] we see that Theorem 1.9 gives us a decomposition up to isogeny of the kind:

$$J_0(N) \cong \prod_\alpha J_\alpha$$

where each J_α is a simple abelian variety, no pair of which is isomorphic, with

$$g_\alpha = \dim(J_\alpha) = [K_\alpha : \mathbb{Q}]$$

and such that

$$\mathbb{Q} \otimes \mathrm{End}_{\mathbb{Q}}(J_\alpha) = \mathbb{Q} \otimes \mathrm{End}(J_\alpha) = K_\alpha$$

the first equality being a consequence of Theorem 1.8.

1.4 The Eisenstein Ideal

Consider now the ideal $\mathfrak{I}$ of $\mathbf{T}$ generated by $1 + w_N$ and all the $1 + \ell - T_\ell$ for each $\ell \neq N$.

Proposition 1.11. $\mathfrak{I}(C) = [0] \in J_0(N)$

Proof.

It suffices to check the equality on the generators of $\mathfrak{I}$ and C. By Proposition 1.6(a), $w_N(c) = -c$ and therefore $(1 + w_N) = c - c = 0$. By Proposition 1.6(b), $T_\ell(c) = (\ell+1)c$ for every prime $\ell \neq N$, and therefore $(1+\ell-T_\ell)(c) = c+\ell c-\ell c-c = 0$. □

Proposition 1.12. *There is an isomorphism* $\mathbf{T}/\mathfrak{I} \leftrightarrow \mathbb{Z}/n\mathbb{Z}$, *where* $n = \operatorname{num}\left(\frac{N-1}{12}\right)$.

Proof. Consider the action of $\mathbf{T}$ on C. For any prime $\ell \neq N$, $T_\ell \in \mathbf{T}$ acts as $\ell + 1$ on C, whereasm w_N acts as -1. Moreover, by Proposition I.2.9, we know that

$$T_\ell T_{\ell'} = T_{\ell\ell'}$$

for any two primes $\ell, \ell' \neq N$ (we are using here the extended definition of Hecke operators T_n for any integer n. See the discussion preceding I.2.9)

Moreover, since C is cyclic of order n (Proposition 1.5) T_ℓ acts in fact as multiplication by $\ell+1 \mod n$. This action factors through $\mathfrak{I}$, by Proposition 1.11, from which one deduces a surjective homomorphism $T/\mathfrak{I} \to \mathbb{Z}/n\mathbb{Z}$. For the proof that this is fact an isomorphism, see [Maz II.9.7]. □

Corollary 1.13. $\mathfrak{I} \neq \mathbf{T}$

Proof.

This is immediate, since our assumption that $N \neq 2, 3, 5, 7, 13$ implies that $n = \operatorname{num}\left(\frac{N-1}{12}\right) \neq 1$. □

Corollary 1.14.

(a) The maximal ideals of $\mathbf{T}$ *containing* $\mathfrak{I}$ *are precisely the ideals* $\mathfrak{B} = \mathfrak{I} + p\mathbf{T}$, *for every* $p \mid n$.

(b) $\mathbf{T}/\mathfrak{B} \cong \mathbb{Z}/p\mathbb{Z}$

Proof.

(a) By the Chinese Remainder theorem, the maximal ideals of $\mathbb{Z}/n\mathbb{Z}$ correspond precisely to the primes p dividing n. Under the isomorphism of Proposition 1.12, which sends $(p) \subset \mathbb{Z}/n\mathbb{Z}$ to $p\mathbf{T}$, the maximal ideals of $\mathbb{Z}/n\mathbb{Z}$ are in bijection with the maximal ideals of $\mathbf{T}/\mathfrak{I}$, i.e. the maximal ideals of $\mathbf{T}$ containing $\mathfrak{I}$.

(b) Clear from part (a).

□

Denote now by $\mathbf{T}_\mathfrak{I}$ the completion of $\mathbf{T}$ by the $\mathfrak{I}$-adic topology. This is a ring with finitely many maximal ideals (i.e. a *semi-local ring*), which correspond to the maximal ideals of $\mathbf{T}$ containing $\mathfrak{I}$. These are precisely the ideals $\mathfrak{B}$ described in Corollary 1.14. We call the ideals $\mathfrak{B}$ the *Eisenstein primes* of $\mathbf{T}$.

Proposition 1.15. *Let $\phi : \mathbf{T} \to \mathbf{T}_\mathfrak{I}$ be the natural homomorphism. For each α, let $\mathfrak{p}_\alpha$ be the kernel of the homomorphism $\mathbf{T} \to K_\alpha$, where K_α is as in Theorem 1.9. Let $\mathfrak{a} = \ker(\phi)$. Then*

$$\mathfrak{a} = \bigcap \{\mathfrak{p}_\alpha : \mathfrak{p}_\alpha + \mathfrak{I} \neq \mathbf{T}\}$$

Proof.

Let $a \in \ker(\phi)$. Then $a \in \mathfrak{I}^r$ for every power $r > 0$. Pick any $\mathfrak{p}_\alpha$ such that $\mathfrak{p}_\alpha + \mathfrak{I} \neq \mathbf{T}$ and consider the ring $\mathbf{T}/\mathfrak{p}_\alpha$. Let $\widetilde{\mathfrak{I}}$ be the image of $\mathfrak{I}$ inside this quotient and consider the completion $(\mathbf{T}/\mathfrak{p}_\alpha)_{\widetilde{\mathfrak{I}}}$. Since $\mathfrak{p}_\alpha$ is prime, the ring $\mathbf{T}/\mathfrak{p}_\alpha$ is an integral domain and since $\mathfrak{p}_\alpha + \mathfrak{I} \neq \mathbf{T}$ this integral domain is nontrivial. It follows that the natural map

$$\psi : \mathbf{T}/\mathfrak{p}_\alpha \longrightarrow (\mathbf{T}/\mathfrak{p}_\alpha)_{\widetilde{\mathfrak{I}}}$$

is an injection. Since $a \in \mathfrak{I}^r$, for every power $r > 0$, its image $\widetilde{a}$ in the quotient $\mathbf{T}/\mathfrak{p}_\alpha$ lies inside $\widetilde{\mathfrak{I}}^r$ for every power $r > 0$. Therefore $\psi(\widetilde{a}) = 0$ in $(\mathbf{T}/\mathfrak{p}_\alpha)_{\widetilde{\mathfrak{I}}}$ and since ψ is injective we have $\widetilde{a} = 0$ in $\mathbf{T}/\mathfrak{p}_\alpha$. This is equivalent to saying that $a \in \mathfrak{p}_\alpha$, which proves the inclusion $\mathfrak{a} \subseteq \bigcap \{\mathfrak{p}_\alpha : \mathfrak{p}_\alpha + \mathfrak{I} \neq \mathbf{T}\}$.

For the reverse inclusion, suppose $\gamma \in \bigcap \{\mathfrak{p}_\alpha : \mathfrak{p}_\alpha + \mathfrak{I} \neq \mathbf{T}\}$. Denote by $\tilde{\gamma}$ the image of γ under the isomorphism of theorem 1.9. Then (after tensoring with $\mathbb{Q}$) we can write $\tilde{\gamma} = (t_1, \ldots, t_k)$ where $t_\alpha \in K_\alpha$. Suppose WLOG that the α such that

$\mathfrak{p}_\alpha + \mathfrak{I} \neq \mathbf{T}$ are precisely the first $k-1$ indices of this decomposition. Since $\gamma \in \mathfrak{p}_\alpha$ for each of these α, $t_1 = \ldots = t_{k-1} = 0$ and we can write

$$\gamma = (0, 0, 0, \ldots, 0, t_n)$$

On the other hand, the element $\tilde{\delta} = (1, \ldots, 1, 0)$ is the image of an element of a $\mathfrak{p}_\alpha$ such that $\mathfrak{p}_\alpha + \mathfrak{I} = \mathbf{T}$. In particular, we can find an element $i \in \mathfrak{I}$ such that $\delta + i = 1$. But then $\delta - 1 = (0, \ldots, -1) \in \mathfrak{I}$ and so is γ. With a similar argument we can show that $\gamma \in \mathfrak{I}^r$ for every r. □

Now, by definition, $\mathfrak{a}$ is an ideal of $\mathbf{T}$. The set

$$\mathfrak{a}J_0(N) = \{[P] \in J_0(N) : \exists [Q] \text{ such that } \psi([Q]) = [P] \text{ for some } \psi \in \mathfrak{a}\}$$

is an abelian subvariety of $J_0(N)$, generated by the images of the endomorphisms of $\mathfrak{a}$.

Definition 1.16. Consider the quotient

$$J_0(N)/\mathfrak{a}J_0(N) = J$$

We call J the *Eisenstein Quotient* of $J_0(N)$.

Remark 1.17. By Remark 1.10, the Eisenstein quotient of $J_0(N)$ has a decomposition (up to isogeny) of the form

$$J = \prod_{\alpha'} J_{\alpha'}$$

where this time α' runs only over the α's of Remark 1.10 such that $\mathfrak{p}_\alpha + \mathfrak{I} \neq \mathbf{T}$. In other words, by taking the Eisenstein quotient we have managed to 'kill' all the simple abelian subvarieties J_α corresponding to prime ideals $\mathfrak{p}_\alpha$ such that $\mathfrak{p}_\alpha$ and $\mathfrak{I}$ are *not contained* by any maximal ideal $\mathfrak{B}$ of corollary 1.14. In other words, each $J_{\alpha'}$ corresponds to an ideal $\mathfrak{p}_{\alpha'}$ contained in some $\mathfrak{B}$.

2 Admissible Filtrations

In this section we are going to develop some tools that will allow us to prove that $J(\mathbb{Q})$, the group of rational points on the Eisenstein quotient, is finite. The strategy is to use a descent method similar, for example, to the one used to prove the weak Mordell-Weil Theorem for elliptic curves [See Si1 VIII, X]. There one wants to bound the size of the Galois cohomology group $H^1(\mathrm{Gal}_{\overline{K}/K}, E[m] : S)$, which is the set of all cocycles ξ such that ξ is trivial in $H^1(I_v, E[m])$ for every v outside S, S finite (here I_v denotes the inertia group at v).

In our specific case, Galois cohomoloy is replaced by a more sophisticated cohomology coming from scheme theory (the $fppf$-cohomology) which has the advantage of picking up the information coming not only from the action of $\mathrm{Gal}_{\overline{K}/K}$, but also from the action of the inertia groups I_v for every place v of K (similar to what $H^1(\cdot, \cdot : S)$ does).

Unfortunately, computations with $fppf$-cohomology can be very hard, and therefore we want to define a certain family of modules (called *admissible*) where these computationsm (or, better, the computations on their corresponding group schemes) become easy. We will see in Section 3 that the piece of torsion of $J(\mathbb{Q})$ we are interested in (corresponding to $E[m]$ in the proof of the weak Mordell-Weil) is in fact admissible.

2.1 Definitions

Let M be a finite abelian module.

Definition 2.1. By a *filtration* of M we mean a finite chain

$$0 = M_1 \subset M_2 \subset \ldots \subset M_r = M$$

of sub-modules of M, where each inclusion is proper. The filtration $\{M_i\}_{i=1}^r$ is *simple* if each sucessive quotient

$$H_i = M_{i+1}/M_i$$

is simple, i.e. if it has no submodules other than 0 and itself. We call each H_i a *constituent* of M. We say that r is the *length* of M.

Remark 2.2. Suppose that there is another simple filtration $\{N_j\}_{j=1}^s$ of M. By the Jordan-Hölder Thoerem [Lang I.3] $s = r$ and there is a bijection between the constituents of M with respect to $\{M_i\}_{i=1}^r$ and the constituents with respect to $\{N_j\}_{j=1}^s$. Therefore the terms *constituent* of M and *length* of M are well defined.

Remark 2.3. Suppose M has a simple filtration $\{M_i\}_{i=1}^r$. Then for each i we have an exact sequence:

$$0 \to M_i \to M_{i+1} \to H_i \to 0$$

Definition 2.4. Suppose that M has exponent p, i.e. every element of M is of order a power of p. By an *admissible filtration of a p-group* we mean a simple filtration of M such that each constituent is isomorphic to either μ_p, the p-th roots of unity, or $\mathbb{Z}/p\mathbb{Z}$.

2.2 $fppf$-Cohomology

Suppose now that $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on M, i.e. the points of M are algebraic numbers (this is also called a *Galois module*). Recall from basic Galois cohomology [see for example Si1 B] that we define:

$$H^0(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), M) = M(\mathbb{Q})$$

i.e the rational points of M. One can also define higher cohomology groups by looking at homomorphisms between G and M. However, for our application the use of Galois cohomology will not suffice. Just as in the case for the weak Mordell Weil Theorem, where one has to impose conditions on the ramification of $E[N]$ and define $H^1(\mathrm{Gal}_{\overline{K}/K}, E[m] : S)$, we need to impose local conditions, (i.e. conditions on the reductions of the points of M).

It turns out that the right object to study is that of a *group scheme*. For an introduction to schemes [Har II-III] is somewhat of a standard reference, and for a discussion of group schemes see [Si2, IV]. On group schemes, one can use the powerful machinery of $fppf$-cohomology (see [EGA IV.2]). Therefore, the strategy is to first build a group scheme for $J_0(N)$ (which will also give subgroup schemes corresponding to subgroups of $J_0(N)$) and then apply $fppf$ to it.

In general, it is not clear how to start with a Galois module M and build a group scheme on it such that the points of this scheme over $\overline{Q}$ will correspond to M. However, we have the following facts:

Fact 2.5. For $M = J_0(N)$, we can define a group scheme $J_0(N)_{/S}$ (called the Néron Model of $J_0(N)$) such that $J_0(N)_{/S}(\overline{\mathbb{Q}}) = J_0(N)(\overline{\mathbb{Q}})$. This construction is essentially unique (see [Si2 IV.5]).

Fact 2.6. For $M = \mathbb{Z}/p$ (resp. μ_p) there is also an essentially unique group scheme $\mathbf{Z/p}$ (resp. $\mu_{\mathbf{p}}$) with the same point-preserving property (see [Maz I.1.6]).

An issue with $fppf$ cohomology is that computations are hard to perform. We therefore want to find suitable subgroup schemes of $J_0(N)$ on which this is easy. The theory of admissible filtrations comes handy in our case.

The notion of an exact sequence carries over for group schemes, and so does the notion of admissible p-group schemes (i.e. a group scheme of order p which has a filtration of subschemes with successive quotients isomorphic to either $\mathbf{Z/p}$ or $\mu_{\mathbf{p}}$). Most importantly

Fact 2.7. Denote by $G_{/S}$ a group scheme of order p defined over S, the set of prime ideals of $\mathbb{Z}$. Then $G_{/S}$ is admissible if and only if its associated Galois module $G(\overline{\mathbb{Q}}) = M(\overline{\mathbb{Q}})$ is admissible. For a proof, see [Maz I.1-2].

For any group scheme $G_{/S}$ with associated Galois module M, we let:

$$H^0(S, G_{/S}) = G(\mathbb{Q}) = M(\mathbb{Q})$$

and

$$H^1(S, G_{/S})$$

be the first cohomology group taken with respect to the $fppf$-cohomology, where S is the set of prime ideals of $\mathbb{Z}$.

Just as in Galois cohomology, we have the following theorem with $fppf$:

Proposition 2.8. *Let*

$$0 \to G_1 \to G_2 \to G_3 \to 0$$

be any short exact sequence of group schemes over S with associated Galois modules M_i. Then there exists a long exact sequence

$$0 \to M_1(\mathbb{Q}) \to M_2(\mathbb{Q}) \to M_3(\mathbb{Q}) \to H^1(S, G_1) \\ \to H^1(S, G_2) \to H^1(S, G_3) \to \dots$$

Let now N be a prime $\neq p$. For any p-group scheme G over $S - \{N\}$ with corresponding Galois p-module M, define:

$$\begin{aligned} h^0(G) &= \log_p(\#H^0(S, G)) = \#M(\mathbb{Q}) \\ h^1(G) &= \log_p(\#H^i(S, G)) \\ \ell(G) &= \log_p(\#M(\overline{\mathbb{Q}})) \\ \delta(G) &= \ell(M) - \log_p(\#\widetilde{M}(\mathbb{F}_N)) \\ \alpha(G) &= \log_p(\#\widetilde{M}(\mathbb{F}_p)) \end{aligned}$$

where we have denoted by $\widetilde{M}$ the reduction of M modulo a prime of S.

Note that all these constants depend only on the Galois module M, except for h^1. However, we have the following bound:

Proposition 2.9. *Let G be an admissible p-group scheme over $S - \{N\}$, and let N be a prime $\neq p$. Then:*

$$h^1(G) - h^0(G) \leq \delta(G) - \alpha(G)$$

Proof.

We discuss an outline of the proof, which is contained in [Maz,p.48-49]. The first step is to prove that if

$$0 \to G_i \to G_{i+1} \to H_i \to 0$$

is the exact sequence of group schemes corresponding to Remark 2.3, then

$$h^1(G_{i+1}) - h^0(G_{i+1}) \leq (h^1(G_i) - h^0(G_i)) + (h^1(H_i) - h^0(H_i))$$

which can be proven by using the long exact sequence of Proposition 2.8. Then, by induction, one only needs to check the inequality for $G_1 = 0$ and for each constituent H_i. Since G is p-admissible, each H_i corresponds to either $\mu_{\mathbf{p}}$ or $\mathbb{Z}/\mathbf{p}\mathbb{Z}$ (not quite, but see [Maz I.1.1b]), so one has to compute the constants h^i, ℓ, δ, α, h^i only for $\mu_{\mathbf{p}}$ and $\mathbb{Z}/\mathbf{p}$ (and their extensions to N). This is done in [Maz p.48] (note however that the first two columns should be switched). □

Remark 2.10. Proposition 2.9 shows that it's easy to bound the first $fppf$ cohomology on admissible p-groups. Moreover, this bound is in terms of constants that can be computed directly from the associated module M. By Fact 2.7, a group scheme is admissible if and only if its associate module is admissible. In other words, if one trusts all the scheme-theoretic machinery to work just as stated, all the computations can be reduced to computations on Galois modules.

3 Finiteness of $J(\mathbb{Q})$

In this section, we prove that the Eisenstein quotient of $J_0(N)$ has only finitely many $\mathbb{Q}$-rational points. Firstly, we show that a certain torsion subgroup $J_0(N)[\mathfrak{B}]$ is admissible, where $\mathfrak{B} = p\mathbf{T} + \mathfrak{I}$ is an Eisenstein prime. By Fact 2.7, the corresponding subscheme inside the Néron model $J_0(N)_{/S}$ (see Fact 2.5) is admissible. Next, we perform p^m-descent on $J_0(N)$ using $fppf$-cohomology on its Néron model. We use admissibility of $J_0(N)[\mathfrak{B}]$ to derive admissibility of certain p^m-torsion subschemes of $J_0(N)_{/S}$, for which we have the bounds given by Proposition 2.9. Those bounds will quickly imply finiteness of $J(\mathbb{Q})$. Finally, we derive the finiteness of $X_0(N)(\mathbb{Q})$ from the finiteness of $J(\mathbb{Q})$.

3.1 Torsion subgroups of $J_0(N)$

For any ideal $\mathfrak{a} \in \operatorname{End} J_0(N)$, define:

$$J_0(N)[\mathfrak{a}] = J_0(N)(\overline{\mathbb{Q}})[\mathfrak{a}] = \{Q \in J_0(N)(\overline{\mathbb{Q}}) : \phi(Q) = [O] \ \ \forall \phi \in \mathfrak{a}\}$$

Then this is a finite abelian module, and the group $\operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ acts on it by acting on the coordinates of the points.

Fact 3.1. The variety $J_0(N)/\mathbb{Q}$ has good reduction everywhere but at N (Igusa 1959).

Proposition 3.2. *Let $\mathfrak{B} = p\mathbf{T} + \mathfrak{I}$. Then $J_0(N)[\mathfrak{B}]$ is admissible.*

Proof.

Note first that the inclusion $\mathfrak{B} \supset p\mathbf{T}$ implies a reverse inclusion $J_0(N)[\mathfrak{B}] \subset J_0(N)[p]$, from which we see that $J_0(N)[\mathfrak{B}]$ has exponent p. Denote by W the direct sum of $J_0(N)[\mathfrak{B}]$ and its Cartier dual (this is just the analogous of the dual $\mathrm{Hom}(E[N], \mu_p)$ induced by the Weil pairing on elliptic curves). Consider the reduction of $J_0(N)[\mathfrak{B}]$ modulo $\mathbb{F}_p$ and let d be its dimension. Then W is a self-dual (with respect to Cartier duality) Galois module of dimension $2d$ over $\mathbb{F}_p$. The action of the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on W factors through a quotient G, which is finite since W is finite. Let now ℓ be a prime $\neq N, p$. The Hecke operator T_ℓ and the Frobenius endomorphism ϕ_ℓ are both elements of the ring $\mathrm{End}(\widetilde{J_0}(N)(\mathbb{F}_\ell))$. By the Eichler-Shimura relation (see [Ste]), these are related in $\mathrm{End}(\widetilde{J_0}(N)(\mathbb{F}_\ell))$ by:

$$\phi_\ell - T_\ell\phi_\ell + \ell = 0$$

Since $J_0(N)$ has good reduction at ℓ, we have an injection:

$$\mathrm{End}(J_0(N)(\overline{\mathbb{Q}})) \hookrightarrow \mathrm{End}(J_0(N)(\overline{\mathbb{F}}_\ell))$$

and therefore the relation holds in $\mathrm{End}(J_0(N)(\overline{\mathbb{Q}}))$ as well. On the other hand, note that $\mathfrak{B} \supset \mathfrak{I}$ and therefore T_ℓ acts as $\ell + 1$ on W, since every element of W is annihilated by $\mathfrak{I}$. We deduce that ϕ_ℓ acts as 1 or ℓ on W, where the values get switched by Cartier duality. Since W is Cartier self-dual of dimension $2d$, ϕ_ℓ must have minimal polynomial equal to $(x-1)^d(x-\ell)^d$.

Let now W' be the $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-module $(\mathbb{Z}/p\mathbb{Z})^d \oplus (\mu_p)^d$. Every element of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ that fixes the elements of W certainly also fixes W', therefore we can assume the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on W' factors through G as well. The characteristic polynomial of $\phi_\ell \in G$ on W' is also given by $(x-1)^d(x-\ell)^d$ and by the Cebotarev density theorem any element of G is the image of ϕ_ℓ for some $\ell \neq p, N$. We can now apply the Brauer-Nesbitt Theorem [CuRe] to deduce that the semisemplification of the representation for W is isomorphic W'. But W' is evidently admissible , and so is W. Since W is the direct sum of $J_0(N)[\mathfrak{B}]$ and its Cartier dual, the module $J_0(N)[\mathfrak{B}]$ is also p-admissible. □

Corollary 3.3. *$J_0(N)[\mathfrak{B}^r]$ is admissible for every $r > 0$.*

Proof.

Let $r > 0$ be an integer and let H be any constituent of an admissible filtration of $J_0(N)[\mathfrak{B}^r]$. Pick t generators $(a_1, \ldots, a_t)$ for the vector space $\mathfrak{B}^r/\mathfrak{B}^{r+1}$. The map

$$\begin{aligned} \frac{J_0(N)[\mathfrak{B}^{r+1}]}{J_0(N)[\mathfrak{B}^r]} &\longrightarrow \bigoplus_{i=1}^{t} J_0(N)(\overline{\mathbb{Q}})[\mathfrak{B}] \\ P &\longmapsto a_1 P \oplus \ldots \oplus a_t P \end{aligned}$$

is an injection, and therefore H is isomorphic to a constituent of $J_0(N)[\mathfrak{B}]$. The Corollary follows from Proposition 3.2. □

Let now p be any prime dividing $n = \operatorname{num}\left(\frac{N-1}{12}\right)$. Consider $\mathbf{T}_p = \mathbb{Z}_p \otimes \mathbf{T}$, the p-adic completion of the Hecke algebra. This is a semi-local ring whose maximal ideals are precisely the maximal ideals containing p. Consequently, $\mathbf{T}_p$ breaks up into a direct product of local rings

$$\mathbf{T}_p \cong \prod_{\mathfrak{a}} T_{\mathfrak{a}}$$

where the product runs through all maximal ideals $\mathfrak{a} \subset \mathbf{T}$ containing p. In particular, one of the $\mathfrak{a}$ is the Eisenstein prime $\mathfrak{B}$ containing p and $\mathfrak{I}$. Let $\varepsilon_{\mathfrak{B}}$ be the idempotent corresponding to this factor, i.e. the map the projects $T_p \to T_{\mathfrak{B}}$.

Consider now the torsion subgroup $J_0(N)[p^m]$, for some integer $m > 0$. The action of $\mathbf{T}$ on it factors through $\mathbf{T}/p^m\mathbf{T}$, since by definition every element of $J_0(N)[p^m]$ is annihilated by p^m. Since $\mathbb{Z}_p$ is simply the inverse limit of the rings $\mathbb{Z}/p^m\mathbb{Z}$, we deduce that there is an action of $\mathbf{T}_p$ on $J_0(N)[p^m]$ for every m. In particular, the idempotent $\varepsilon_{\mathfrak{B}}$ acts on this subgroup and we define:

$$J_0(N)[p^m]_{\mathfrak{B}} = \varepsilon_{\mathfrak{B}} J_0(N)[p^m]$$

Proposition 3.4. *$J_0(N)[p^m]_{\mathfrak{B}}$ is p-admissible*

Proof. This follows by noting that $J_0(N)[p]_{\mathfrak{B}} = J_0(N)[\mathfrak{B}]$ and by Proposition 3.2. □

Now that we have proved that the torsion subgroup $J_0(N)[p^m]_{\mathfrak{B}}$ is admissible, we want to apply $fppf$ cohomology to its corresponding subgroup scheme inside the Néron model $J_0(N)_{/S}$. However, $fppf$ will not work on the Néron model as it is. We need to take the so called 'connected component of the identity' of $J_0(N)_{/S}$. We denote this component by $J_0(N)^0_{/S}$. In terms of points, this corresponds to the component of $J_0(N)_{/S}$ whose points always reduce to smooth points. For a discussion in the case of elliptic curves, see [Si2 IV.5].

Remark 3.5. By Fact 3.1, $J_0(N)$ has good reduction everywhere but at N, so $J_0(N)^0_{/S}$ differs from $J_0(N)_{/S}$ only above N. Moreover, we have that the quotient:

$$\frac{J_0(N)_{/S}}{J_0(N)^0_{/S}}$$

is finite.

Any of the torsion subgroups defined in this section extend uniquely to subgroup schemes of the Néron model $J_0(N)_{/S}$. In particular, let

$$J_0(N)[p^m]_{/S} \subset J_0(N)_{/S}$$

be the subgroup scheme corresponding to $J_0(N)[p^m]$, and denote by $J_0(N)^0[p^m]_{/S}$ its connected component of the identity. Moreover, define

$$J_0(N)^0[p^m]_{\mathfrak{B}/S} = \varepsilon_{\mathfrak{B}} J_0(N)^0[p^m]_{/S}$$

By Proposition 3.4, and Fact 2.7, this is admissible. Moreover, after having taken the connected component of the identity, we are ready to apply $fppf$-cohomology to it.

Recall from Theorem 1.9 that there is a decomposition $\mathbb{Q} \otimes \mathbf{T} \cong \prod_\alpha K_\alpha$, with K_α totally real number fields of degree

$$g_\alpha = [K_\alpha : \mathbb{Q}]$$

Then $g_{\mathfrak{B}}$, be the rank of $\mathbf{T}_{\mathfrak{B}}$ over $\mathbb{Z}_p$, is simply the sum of all the g_α such that the associated $\mathfrak{P}_\alpha$ is contained in $\mathfrak{B}$. Using this, we obtain the following bounds:

Proposition 3.6. *Let $\alpha(\cdot)$ and $\delta(\cdot)$ be the constants defined in Proposition 2.7. Then*

(a) $\delta(J_0(N)^0[p^m]_{\mathfrak{B}/S}) = mg_{\mathfrak{B}} + O(1)$ *as* $m \to \infty$

(b) $\alpha(J_0(N)^0[p^m]_{\mathfrak{B}/S}) = mg_{\mathfrak{B}} + O(1)$ *as* $m \to \infty$

Proof. This is in [Ma1, p.148] □

We conclude with the following very important proposition:

Proposition 3.7. *There exists a constant* $C < \infty$ *such that*

$$\#H^1(S, J_0(N)^0[p^m]_{\mathfrak{B}/S}) \leq C$$

for every integer $m > 0$.

Proof.

By Proposition 3.4 and fact 2.7, $J_0(N)^0[p^m]_{\mathfrak{B}/S}$ is admissible. Therefore, by Proposition 2.9 we can deduce that:

$$h^1(J_0(N)^0[p^m]_{\mathfrak{B}/S}) - h^0(J_0(N)^0[p^m]_{\mathfrak{B}/S}) \leq \delta(J_0(N)^0[p^m]_{\mathfrak{B}/S}) - \alpha(J_0(N)^0[p^m]_{\mathfrak{B}/S})$$

Now, the group $H^0(S, J_0(N)^0[p^m]_{\mathfrak{B}/S}) = J_0(N)^0(\mathbb{Q})[p^m]_{\mathfrak{B}}$ is finite, being a subgroup of the torsion of the Mordell-Weil group of $J_0(N)$. The bounds of Proposition 3.6 then give the desired result. □

Remark 3.8. Proposition 3.7 is the heart of the descent argument. It seems that the very construction of the Eisenstein ideal was motivated in the first place by trying to obtain the finiteness statement of Proposition 3.7. The rest of the descent argument, presented in the next section, is simply an application of cohomological machinery.

3.2 Descent on $J_0(N)$

In this section we prove the following theorem

Theorem 3.9. *The group of rational points on the Eisenstein quotient* $J(\mathbb{Q})$ *is finite.*

In order to prove Theorem 3.9, we first apply p^m-descent on $J_0(N)$ using $fppf$-cohomology:

Proposition 3.10. *The group* $\mathbf{T}_{\mathfrak{B}} \otimes_{\mathbf{T}} J_0(N)(\mathbb{Q})$ *is finite.*

Proof.

For each positive integer $m \geq 1$, consider the exact sequence of group schemes:

$$0 \to J_0(N)^0[p^m]_{/S} \to J_0(N)^0_{/S} \xrightarrow{p^m} J_0(N)^0_{/S} \to 0$$

Applying Proposition 2.8, we obtain a long exact sequence

$$0 \to J_0(N)^0(\mathbb{Q})[p^m] \to J_0(N)^0(\mathbb{Q}) \xrightarrow{p^m} H^1(S, J_0(N)^0[p^m]_{/S}) \\ \to H^1(S, J_0(N)^0_{/S}) \xrightarrow{p^m} H^1(S, J_0(N)^0_{/S}) \to \dots$$

from which we extract an injection:

$$\frac{J_0(N)^0(\mathbb{Q})}{p^m J_0(N)^0(\mathbb{Q})} \hookrightarrow H^1(S, J_0(N)^0[p^m]_{/S})$$

Taking the direct limit with respect to the p^m maps, we obtain an injection:

$$\mathbb{Q}_p/\mathbb{Z}_p \otimes J_0(N)^0(\mathbb{Q}) \longrightarrow \varinjlim H^1(S, J_0(N)^0[p^m]_{/S})$$

(note that duality interchanges inverse and direct limits).

Since the ring $T_{\mathfrak{B}}$ is flat, being the completion of a noetherian ring, we derive an injection of $\mathfrak{B}$-components:

$$\mathbf{T}_{\mathfrak{B}} \otimes_{\mathbf{T}_{\mathfrak{B}}} \left(\mathbb{Q}_p/\mathbb{Z}_p \otimes J_0(N)^0(\mathbb{Q})\right) \longrightarrow \varinjlim H^1(S, J_0(N)^0[p^m]_{\mathfrak{B}/S})$$

By Proposition 3.7, the set on the right is finite, and so the one on the left is finite as well. Now, the group $J_0(N)^0(\mathbb{Q})$ is finitely generated, from which we deduce that $\mathbf{T}_{\mathfrak{B}} \otimes_{\mathbf{T}_{\mathfrak{B}}} J_0(N)^0(\mathbb{Q})$ is finite. By Remark 3.5, the set $J_0(N)/J_0(N)^0$ is finite, from which it follows that $\mathbf{T}_{\mathfrak{B}} \otimes_{\mathbf{T}_{\mathfrak{B}}} J_0(N)(\mathbb{Q})$ is finite as well. □

In order to put Proposition 3.10 to use, we need a few arguments from commutative algebra. Let $\mathfrak{a}$ be any ideal of $\mathbf{T}$ and consider the kernel of the map $\mathbf{T} \to \mathbf{T}_{\mathfrak{a}}$. This is just the intersection $\gamma(\mathfrak{a}) = \cap_r \mathfrak{a}^r$. Let $\mathbf{T}^{(\mathfrak{a})} = \mathbf{T}/\gamma(\mathfrak{a})$ and let

$$J^{(\mathfrak{a})} = J_0(N)/\gamma(\mathfrak{a})J_0(N)$$

(note the parallel with the construction of the Eisenstein quotient, with $\mathfrak{a} = \mathfrak{I}$). Consider also $V = J_0(N)(\mathbb{Q}) \otimes \mathbb{Q}$ as a $\mathbf{T} \otimes \mathbb{Q}$-module, and $V^{(\mathfrak{a})} = J^{(\mathfrak{a})}(\mathbb{Q}) \otimes \mathbb{Q}$ as a $\mathbf{T}^{(\mathfrak{a})} \otimes \mathbb{Q}$ module.

Lemma 3.11. $V^{(\mathfrak{a})} = \mathbf{T}^{(\mathfrak{a})} \otimes_{\mathbf{T}} V$

Proof.

For any exact sequence

$$0 \to A \to B \to C \to$$

of abelian varieties defined over $\mathbb{Q}$, consider the long exact sequence of Galois Cohomology

$$0 \to A(\mathbb{Q}) \to B(\mathbb{Q}) \to C(\mathbb{Q}) \to H^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), A) \to \dots$$

Since $A(\mathbb{Q})$ is finitely-generated and every element $H^1(\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), A)$ is of finite order for any abelian variety over $\mathbb{Q}$, we can tensor with $\mathbb{Q}$ to get an exact sequence:

$$0 \to A(\mathbb{Q}) \to B(\mathbb{Q}) \to C(\mathbb{Q}) \to 0$$

Applying this result to our exact sequence:

$$0 \to \gamma(\mathfrak{a})J_0(N) \to J_0(N) \to J^{(\mathfrak{a})} \to 0$$

we deduce that $V^{(\mathfrak{a})} = V/\gamma(\mathfrak{a})V = \mathbf{T}^{(\mathfrak{a})} \otimes_{\mathbf{T}} V$. □

Lemma 3.12. *If* $\mathbf{T}_{\mathfrak{a}} \otimes_{\mathbf{T}} V = 0$, *then* $J^{(\mathfrak{a})}(\mathbb{Q})$ *is finite.*

Proof.

Let W be the torsion free part of $J_0(N)(\mathbb{Q})$. If $\mathbf{T}_{\mathfrak{a}} \otimes_{\mathbf{T}} W = 0$, as in the hypothesis, then every element of W is 'killed' by a prime ideal in $\mathbf{T}_{\mathfrak{a}}$, i.e. a prime ideal containing $\mathfrak{a}$. On the other hand, every prime ideal containing $\mathfrak{a}$ also contains every prime ideal of $T^{(\mathfrak{a})}$, by construction. Therefore there are only finitely many maximal ideals in $T^{(\mathfrak{a})}$ that do not annihilate M. Therefore $\mathbf{T}^{(\mathfrak{a})} \otimes_{\mathbf{T}} W = 0$, which implies $\mathbf{T}^{(\mathfrak{a})} \otimes_{\mathbf{T}} V = 0$. Applying Lemma 3.11 gives the desired result.

□

We now conlude the proof of Theorem 3.9

Proof of Theorem 3.9.

For any Eisenstein prime $\mathfrak{B}$, Proposition 3.10 gives us that $\mathbf{T}_{\mathfrak{B}} \otimes_{\mathbf{T}} V = 0$. Taking the product over all the Eisenstein primes we deduce that

$$\mathbf{T}_{\mathfrak{J}} \otimes_{\mathbf{T}} V = 0$$

and from Lemma 3.12, we conclude that $J^{(\mathfrak{J})}(\mathbb{Q}) = J(\mathbb{Q})$ is finite. □

3.3 Finiteness of $X_0(N)(\mathbb{Q})$

Theorem 3.13. *Let N be a prime, $N \neq 2, 3, 5, 7, 13$. Then the set $X_0(N)(\mathbb{Q})$ is finite.*

Proof.

Consider the sequence of maps:

$$X_0(N) \hookrightarrow J_0(N) \to J$$

the first being given by $[z] \to [z] - [\infty]$ as in Proposition 1.1(b), and the second by projection onto the Eisenstein quotient. Denote by $\widehat{X}_0(N)$ the projection of $X_0(N)$ inside J. Since $X_0(N)$ is connected of dimension 1, $\widehat{X}_0(N)$ is either a point or a curve. To see that $\widehat{X}_0(N)$ is non-trivial, note that by Remark 1.4, the points of $X_0(N)$ generate $J_0(N)$ as a group, therefore the points of $\widehat{X}_0(N)$ generate J. But $\dim(J) \geq 1$, so it cannot be generated by the trivial element. Therefore $\widehat{X}_0(N)$ is non-empty and it must have dimension 1. We conclude that the map $X_0(N) \to \widehat{X}_0(N)$ is a finite surjective map of curves. By Theorem 3.9, the group $J(\mathbb{Q})$ is finite, and so $\widehat{X}_0(N)(\mathbb{Q}) \subset J(\mathbb{Q})$ is finite as well. But we have just shown that the map $X_0(N) \to \widehat{X}_0(N)$ is finite (and defined over $\mathbb{Q}$), hence $X_0(N)(\mathbb{Q})$ is finite. □

Chapter III

Mazur's Theorem

The results developed in the previous two chapters will allow us to prove here the following theorem, commonly known as Mazur's Theorem.

Theorem (Mazur)**.** *Let E be an elliptic curve defined over $\mathbb{Q}$. Then $E_{tors}(\mathbb{Q})$ is isomorphic to one of the following fifteen groups:*

- $\mathbb{Z}/n\mathbb{Z}$ *for* $1 \leq n \leq 10$ *or* $n = 12$
- $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ *for* $m = 2, 4, 6, 8$

The proof we reproduce here, as originally presented in [Maz,III.5], proceeds by contradiction as follows:

Implication 1: Suppose Mazur's Theorem is false. Then there exists an elliptic curve E defined over $\mathbb{Q}$ with a subgroup $H \subset E_{\text{tors}}(\mathbb{Q})$ of order N, for N a prime number and $N \geq 23$.

Implication 2: Suppose E is an elliptic curve defined over $\mathbb{Q}$ with a subgroup $H \subset E_{\text{tors}}(\mathbb{Q})$ of order N, for N a prime number and $N \geq 23$. Let $K = \mathbb{Q}(\zeta_N)$, for ζ_N a primitive root of unity, and let $L = \mathbb{Q}(E[N])$, obtained by adjoining the coordinates of all N-torsion points of E. Then L/K is unramified at all places of K.

Implication 3: If L/K is unramified at all places of K, then $L = K$.

Implication 4: If $L = K$ then there exists infinitely many points on the curve $X_0(N)(\mathbb{Q})$.

But since N is a prime and $N \geq 23$, Implication 4 contradicts Theorem II.3.13.

1 Kubert's Computations

The first implication is a restatement of the following result, due to D.S. Kubert [Kub, Theorem IV.1.2]

Theorem 1.1 (Kubert). *Let E be an elliptic curve defined over $\mathbb{Q}$. If $E_{tors}(\mathbb{Q})$ is not one of the 15 groups mentioned in Mazur's Theorem, then there exists a prime $N \geq 23$ such that $E_{tors}(\mathbb{Q})$ has a subgroup of order N.*

Proof.

We only include a sketch of the proof, which is given by a case by case analysis. First, it is clear that it suffices to only consider primes, for if $E_{\text{tors}}(\mathbb{Q})$ has a subgroup of order m a product of primes, then by the Sylow theorems it has a subgroup of order p for each of the primes dividing m.

Next, we need to consider all possible structures of the group $E_{\text{tors}}(\mathbb{Q})$. Being a finite abelian group, we must have:

$$E_{\text{tors}}(\mathbb{Q}) = \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \times \ldots \mathbb{Z}/d_r\mathbb{Z}$$

where $d_1 \mid d_2 \mid \ldots \mid d_r$. If $r > 2$, then $E_{\text{tors}}(\mathbb{Q})$ would contain the subgroup $(\mathbb{Z}/d_1\mathbb{Z})^3$, which would also be a subgroup of $E[d_1]$. But $E[d_1] \cong (\mathbb{Z}/d_1\mathbb{Z})^2$, so we must have $r = 1, 2$. Moreover, suppose that $d_1 > 2$. Then $(\mathbb{Z}/d_1\mathbb{Z})^2 \cong E[d_1]$ would be contained in $E[d_1](\mathbb{Q})$, which is impossible since by [Si1 III.8.1.1] this can happen only if $d_1 = 1, 2$. In other words:

$$E_{\text{tors}}(\mathbb{Q}) = \mathbb{Z}/d\mathbb{Z} \text{ or } \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z}$$

where $d \geq 1$ is an integer and $2 \mid d_2$.

Consider first the case when $E_{\text{tors}}(\mathbb{Q})$ is cyiclic of order d. We want to show that d is not divisible by any prime $N < 23$, i.e. that $E_{\text{tors}}(\mathbb{Q})$ has no subgroups

of prime order $N < 23$. If $N \neq 2, 3, 5, 7$, one only needs to show that the curves $Y_0(N)(\mathbb{Q})$ are empty. The cases $N = 11, 17, 19$ are somewhat easier since, from Proposition I.1.15, $g(X_0(N)) = 1$ and one can use the methods of descent developed for elliptic curves. These cases are completely analyzed by Ligozat [Lig]. Mazur and Tate, on the other hand, proved in [MaTa] that there are no rational points of order 13 on elliptic curves defined over $\mathbb{Q}$. If $N = 2, 3, 5, 7$ and $N \mid d$, $N < 23$, we want to prove that $d = 2, 3, 4, 5, 6, 7, 8, 9, 10, 12$. Now, the cases $d = 14, 15, 20, 21, 24, 27, 32, 36, 49$ again correspond to curves $Y_0(d)(\mathbb{Q})$ of genus 1, which Ligozat proved are empty. Lind [Lind] proved that $Y_0(16)(\mathbb{Q})$ is also empty . The cases $d = 18, 25, 35$ are dealt with by Kubert in IV.5, and this covers all the possible cases when $E_{\text{tors}}(\mathbb{Q})$ is cyclic of order d and $N < 23$, $N = 2, 3, 5, 7$ divides d.

It remains to consider the case when $E_{\text{tors}}(\mathbb{Q})$ takes the form $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z}$ where $2 \mid d_2$. The cases $d_2 = 2, 4, 6, 8$ correspond to the structures allowed by Mazur's Theorem. The case $d = 16$ (and hence higher powers of 2) follows from Lind's work. From what we proved in the previous paragraph, we also know that d_2 is not divisibile by any prime N different from $2, 3, 5, 7$, therefore it suffices only to consider the cases $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z}, \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/12\mathbb{Z}, \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/14\mathbb{Z}$. The latter would yield a point on $Y_0(14)(\mathbb{Q})$, which is impossible by Ligozat's work. The other two cases are 2-isogenous to curves E' with a cyclic group of order 20 and 24, respectively, which are contained in $E'_{\text{tors}}(\mathbb{Q})$ since the multiplication-by-2 isogeny is defined over $\mathbb{Q}$. But this would yield points on $Y_0(20)(\mathbb{Q})$ and $Y_0(24)(\mathbb{Q})$, respectively, which again is impossible by Ligozat. □

2 L/K is unramified

In this section we will prove the following proposition:

Proposition 2.1. *Let E be an elliptic curve defined over $\mathbb{Q}$, and such that $E_{tors}(\mathbb{Q})$ contains a cyclic subgroup of order N, for N a prime and $N \geq 23$. Let $K = \mathbb{Q}(\zeta_N)$, where ζ_N is a primitive N-th root of unity, and let $L = \mathbb{Q}(E[n])$ be the field obtained by adjoining all the coordinates of the N-th torsion points of E. Then L/K is unramified at all places of K.*

Note that L is a Galois extension of $\mathbb{Q}$ which contains ζ_N by the nondegenracy of the Weil pairing [Si1 III.8.1.1], and therefore the statement of the proposition makes sense. To be precise, what we prove is actually the following, which is slightly stronger than 2.1:

Proposition 2.2. *Let E, L, K as in Proposition 2.1. Then*

(a) If E has good reduction at a rational prime $q \neq N$, then $L/\mathbb{Q}$ is unramified at q.

(b) If E does not have good reduction at a rational prime $q \neq N$, then $L/\mathbb{Q}$ is unramified at q.

(c) Let $q = N$. Then L/K is unramified above K.

We denote by $\mathbb{Q}_q$ the q-adic completion of $\mathbb{Q}$ and by $\mathbb{Z}_q$ the ring of q-adic integers. As an extension of $\mathbb{Q}_q$, L is given by the same polynomials with coefficients embedded inside $\mathbb{Q}_q$. Moreover, $L/\mathbb{Q}_q$ is unramified if and only if $L/\mathbb{Q}$ is unramified above q, so it suffices to look at each of the extensions $L/\mathbb{Q}_q$, where we can apply the powerful tools of [Si1 VII].

Before proving Proposition 2.2, we need to introduce some notation. Let

$$\pi_q : E/\mathbb{Q}_q \to \widetilde{E}/\mathbb{F}_q$$

be the 'reduction modulo q' map, where $\widetilde{E}$ is the curve defined over $\mathbb{F}_q$ by reducing the coefficients of a minimal Weierstrass equation of E modulo q. This curve might be singular, so let $\widetilde{E}_{ns}(\overline{\mathbb{F}}_q)$ be the group of all its $\overline{\mathbb{F}}_q$-rational smooth points. Moreover, let:

$$E_0(\overline{\mathbb{Q}}_q) = \{P \in E(\overline{\mathbb{Q}}_q) : \pi_q(P) \in \widetilde{E}_{ns}(\overline{\mathbb{F}}_q)\}$$
$$E_1(\overline{\mathbb{Q}}_q) = \{P \in E_0(\overline{\mathbb{Q}}_q) : \pi_q(P) = O \in \widetilde{E}_{ns}(\overline{\mathbb{F}}_q)\}$$

Denote by H the putative subgroup of $E_{\text{tors}}(\mathbb{Q})$ of order N, a prime ≥ 23, and define:

$$H_0(\overline{\mathbb{Q}}_q) = \{P \in H : \pi_q(P) \in \widetilde{E}_{ns}(\overline{\mathbb{F}}_q)\}$$
$$E[N]_0(\overline{\mathbb{Q}}_q) = \{P \in E(\overline{\mathbb{Q}}_q)[N] : \pi_q(P) \in \widetilde{E}_{ns}(\overline{\mathbb{F}}_q)\}.$$

We break down the proof of Proposition 2.2 in three lemmas: Lemma 2.3, Lemma 2.7 and Lemma 2.8. We derive Proposition 2.2 at the end of this section.

Lemma 2.3. *Let E be as in Proposition 2.1. Then E has semi-stable (i.e. good or multiplicative) reduction at all places of $\mathbb{Q}$.*

Proof.

For the sake of contradiction, suppose that there exists a rational prime q such that E has additive reduction at q i.e.

$$\widetilde{E}_{ns}(\overline{\mathbb{F}}_q) = \left(\overline{\mathbb{F}}_q\right)^+$$

We either have $H \cap E_0(\mathbb{Q}_q) = \{O\}$ or $H \subset E_0(\mathbb{Q}_q)$, since H is of prime order. In the former case the image of H in the quotient $E(\mathbb{Q}_q)/E_0(\mathbb{Q}_q)$ would be a subgroup of order $N \geq 23$. But by the classification of Kodaira-Néron [Si1 VII.6.1], we know that $E(\mathbb{Q}_q)/E_0(\mathbb{Q}_q)$ is finite of order 1,2,3 or 4 so this is impossible. Therefore $H \subset E_0(\mathbb{Q}_q)$. In this case, the image of H under π_q is a subgroup of $\tilde{E}_{ns}(\mathbb{F}_q) \cong (\mathbb{F}_q)^+$, therefore it is 0 unless $N = q$, in which case it is the whole group. If $q \neq N$ then, $H \subset E_1(\mathbb{Q}_q)$, which is impossible since $E_1(\mathbb{Q}_q)$ has no m-torsion points for all m relatively prime to q [Si1 VII.3.1] Therefore we must have $N = q$.

Next, we show that even the case $N = q$ is impossible, that is, E does not have additive reduction at N. Recall the following Proposition ([Si1 VII.5.4]):

Proposition 2.4. *Let K be a local field complete with respect to a discrete valuation v, and let π be a generator for the unique maximal ideal of $\mathcal{O}_K$. Let E/K be an elliptic curve. Then there exists a finite extension K'/K such that E has good or split multiplicative reduction over K'. Moreover, the ramification degree of K'/K can be taken to be at most 6.*

Proof. Since it will be the case in our applications, assume $\mathrm{char}(k) \neq 2$, where k is the residue field $\mathcal{O}_K/(\pi)$. If E is given by $y^2 = f(x)$, extend K so that $f(x)$ splits and change coordinates so to obtain an equation for E in Legendre normal form [Si1 III.1.7]:

$$y^2 = x(x-1)(x-\lambda) \;\; \lambda \neq 0, 1$$

Consider then the two quantities c_4, Δ associated to this equation:

$$c_4 = 16(\lambda^2 - \lambda + 1) \;\; , \;\; \Delta = 16\lambda^2(\lambda - 1)^2$$

and apply [Si1 VII.5.1]. If $\lambda \in \mathcal{O}_K$ and $\lambda \not\equiv 0, 1 \ (\pi)$ then $v(\Delta) = 0$ and the curve has good reduction. If $\lambda \in \mathcal{O}_K$ and $\lambda \equiv 0, 1 \ (\pi)$ then $v(\Delta) > 0$ but $v(c_4) = 0$ so the curve has split multiplicative reduction.

The interesting case then is when $\lambda \notin \mathcal{O}_K$. Suppose that $v(\lambda) = -r$ for some positive integer r. Then $v(\pi^r \lambda) = 0$ and the substitution $x \mapsto \pi^{-r} x'$, $y \mapsto \pi^{-3r/2} y'$ gives a Weierstrass equation

$$y' = x'(x - \pi^r)(x' - \pi^r \lambda)$$

that has split multiplicative reduction at (π). Note that in order to perform the change of variables, we need to (possibly) extend our field to $K(\pi^{1/2})$. In the worst case scenario, we will have to extend K twice: first to split $f(x)$, a cubic polynomial, and then to adjoin the square root of π. Since the ramification degree of the first extension is at most 3, and the ramification degree of the second extension is at most 2, we have that the ramification degree of K'/K is at most 6. □

Now let $K'/\mathbb{Q}_N$ be as in Proposition 2.4, so that our curve E has good or split multiplicative reduction at the maximal ideal of $\mathcal{O}_{K'}$. Denote by $\mathbb{F}_{N^k}$ the residue field of $\mathcal{O}_{K'}$. Taken over the ring $\mathcal{O}_{K'}$, the curve E has two different Weierstrass equations defining it, one coming from the original equation with coefficients in $\mathbb{Z}_N$, call it E^1, and the other with coefficients in $\mathcal{O}_{K'}$ obtained by the process of Proposition 2.4, call it E^2. By the Néron Mapping Property [Si2 IV.5], the isomorphism between E^1 and E^2 induces an isogeny ψ (defined over $\mathbb{F}_{N^k}$) between the reductions of the two curves:

$$\begin{array}{ccccc} & & & \overset{\simeq}{\nearrow} & (\overline{\mathbb{F}}_N)^\times \\ (\overline{\mathbb{F}}_N)^+ \cong \widetilde{E}^1_{ns} & \xrightarrow{\psi} & \widetilde{E}^2_{ns} & & \\ & & & \overset{\simeq}{\searrow} & E'/\mathbb{F}_{N^k} \end{array}$$

where by the split arrows we just mean one of the two possibilities for the reduction of $E^2(K')$. Suppose that ψ is nontrivial. If E^2 has multiplicative reduction, then the composition of ψ with the homomorphism given by the top arrow gives a nontrivial homomorphism

$$\psi : (\overline{\mathbb{F}}_N)^+ \longrightarrow (\overline{\mathbb{F}}_N)^\times$$

which is impossible from the following Lemma.

Lemma 2.5. *Let k be an algebraically closed field. Let $\mathbb{G}_a$ be the additive group variety defined by $\mathbb{G}_a = \mathbb{A}^1$ (i.e. $k[X]$ with the usual addition of polynomials) and let $\mathbb{G}_m = k[X,Y]/(XY-1)$ be the multiplicative group variety. Then there are no nontrivial morphisms $\psi : \mathbb{G}_a \to \mathbb{G}_m$.*

Proof.

Note first that by projecting $\mathbb{G}_m$ onto the affine line $\mathbb{G}_a \cong \mathbb{A}^1$, we get that $\mathbb{G}_m \cong \mathbb{A}^1 - \{0\}$. Composing ψ with the injection $\mathbb{A}^1 - \{0\} \hookrightarrow \mathbb{A}^1$ gives a morphism:

$$\mathbb{A}^1 \xrightarrow{\psi} \mathbb{A}^1 - \{0\} \to \mathbb{A}^1$$

which is an endomorphism of $\mathbb{A}^1 \cong \mathbb{G}_a$. Since every nonconstant endomorphism of $\mathbb{G}_a$ must be a nonconstant single variable polynomial, hence surjective, we derive a contradiction, since the map above is clearly not surjective. □

Note that in our case Lemma 2.5 applies with $k = \overline{\mathbb{F}}_N$, $\mathbb{G}_a = \left(\overline{\mathbb{F}}_N\right)^+$, $\mathbb{G}_m = \left(\overline{\mathbb{F}}_N\right)^\times$.

On the other hand, suppose that E^2 has good reduction. Then the composition of ψ with the homomorphism given by the bottom arrow gives a nontrivial homomorphism:

$$\psi : \left(\overline{\mathbb{F}}_N\right)^+ \longrightarrow E'/\mathbb{F}_{N^k}$$

where by $E'/\mathbb{F}_{N^k}$ we just mean an elliptic curve defined over $\mathbb{F}_{N^k}$. This is again impossible from the following:

Lemma 2.6. *Let k be an algebraically closed field and let $\mathbb{G}_a$ as in Lemma 2.5. Then there are no nontrivial morphisms*

$$\phi : \mathbb{G}_a \to E/k$$

where E/k is an elliptic curve defined over k.

Proof.

From the definition, $\mathbb{G}_a \cong \mathbb{A}^1 \cong \mathbb{P}^1 - \{\infty\}$. Therefore, ϕ can be extended to give a morphism $\mathbb{P}^1 \to E$. Since this is nonconstant, it must be surjective [Har II.6.8]. But $\mathbb{P}^1$ has genus 0, whereas E has genus 1, which contradicts the Riemann-Hurwitz formula [Si1 II.5.9]. □

Therefore ψ must be trivial.

Now, the subgroup H contained inside E^1 is isomorphic (over $\mathcal{O}_{K'}$) to a subgroup H' of order N contained inside E^2. Reduction modulo N gives us two corresponding subgroups $\widetilde{H}$ and $\widetilde{H}'$ on $\widetilde{E}^1_{ns}$ and $\widetilde{E}^2_{ns}$ respectively. Since the ramification degree of K'/Q_N is between $0 \leq 6 < N-1$ (recall our assumptions on N) we can apply the results of [Ray] to deduce that the isomorphism between H and H' induces an isomorphism between $\widetilde{H}$ and $\widetilde{H}'$. But we just saw how this is impossible. This conludes the proof of Lemma 2.3 □

Lemma 2.7. *Let E be as in Proposition 2.1. If $q = 2, 3$ then E has multiplicative reduction at q and $H \nsubseteq E_0(\mathbb{Q}_q)$*

Proof.

If E had good reduction at q, then $E_0(\mathbb{Q}_q) = E(\mathbb{Q}_q)$ and in particular $H \subset E_0(\mathbb{Q}_q)$. Since $E_1(\mathbb{Q}_q)$ has no N-torsion, the image of H under π_q would be a subgroup of $\widetilde{E}_{ns}(\mathbb{F}_q)$ of order N. But from the Hasse bound on the size of elliptic curves over finite fields, we know that $\#E(\mathbb{F}_q) \leq 2\sqrt{q} + q + 1$, from which we conclude that $N \leq 2\sqrt{q} + q + 1$. This is impossible, since $N \geq 23$ and $q = 2, 3$. Therefore E cannot have good reduction at q. It cannot have additive reduction either, by Lemma 2.3, so it must have multiplicative reduction.

Suppose then $H \subset E_0(\mathbb{Q}_q)$. By [Si1, Exercise III.3.5] $\widetilde{E}(\mathbb{F}_q) \subset (\mathbb{F}_{q^2})^\times$, so the reduction of H is contained in $(\mathbb{F}_{q^2})^\times$. But $(\mathbb{F}_{q^2})^\times$ has $q^2 - 1$ (i.e. 3 or 8) points, whereas H has at least 23 points. Therefore $H \nsubseteq E_0(\mathbb{Q}_q)$.

□

Lemma 2.8. *Let E be as in Proposition 2.1. If q is a prime where E does not have good reduction then $H \nsubseteq E_0(\mathbb{Q}_q)$.*

Proof.

Suppose this is not the case, so that q is a prime where E does not have good reduction and $H \subset E_0(\mathbb{Q}_q)$. By section 4.2, $q \neq 2, 3$. Moreover, if $q = N$, then by section 4.1 the reduction is multiplicative, and if $H \subset E_0(\mathbb{Q}_N)$ then H would map to a subgroup of order N of $(\mathbb{F}_{N^2})^\times$, which is impossible since $N \nmid N^2 - 1$. Therefore we can assume $q \neq 2, 3, N$.

Now, the curve $X_0(N)(\mathbb{Q})$ has good reduction everywhere but at N, therefore we have a reduction map:

$$\phi_q : X_0(N) \longrightarrow \widetilde{X}_0(N)/\mathbb{F}_q$$

Since E has multiplicative reduction at q, the point (E, H) does not map to any point of $\widetilde{Y}_0(N)$. In fact, under the 'generalized elliptic curve' interpretation, we see that the condition $H \subset E_0(\mathbb{Q}_q)$ implies that

$$\phi_q((E, H)) = [0] \in \widetilde{X}_0(N)/\mathbb{F}_q$$

On the other hand, if we take $q = 3$, then

$$\phi_3((E, H)) = [\infty] \in \widetilde{X}_0(N)/\mathbb{F}_3$$

since E also has multiplicative reduction at 3 but $H \nsubseteq E_0(\mathbb{Q}_3)$, by Lemma 2.7.

Consider now the projection $X_0(N) \to J$, where J is the Eisenstein quotient defined in II.1.16. This is an abelian variety defined over $\mathbb{Q}$, and we can consider its reduction $\widetilde{J}/\mathbb{F}_q$ modulo q. For any prime $\ell \neq 2, N$, by the Oort-Tate classification theorem [OrTa] we have an injection

$$J_{\text{tors}} \hookrightarrow \widetilde{J}(\mathbb{F}_\ell)$$

In particular, by Theorem II.3.9 $J(\mathbb{Q})$ is finite [REF], $J(\mathbb{Q}) \subset J_{\text{tors}}$ and by composing we get an injection

$$J(\mathbb{Q}) \hookrightarrow \widetilde{J}(\mathbb{F}_\ell) \tag{III.1}$$

for any prime $\ell \neq 2, N$. In other words, for any ℓ we have the following commutative diagram:

$$\begin{array}{ccc} X_0(N)(\mathbb{Q}) & \longrightarrow & J(\mathbb{Q}) \\ \downarrow & & \downarrow \\ \widetilde{X}_0(N)(\mathbb{F}_\ell) & \longrightarrow & \widetilde{J}(\mathbb{F}_\ell) \end{array}$$

Let $\ell = q$ and consider the point $(E, H) \in X_0(N)(\mathbb{Q})$. By the above, this point maps to $[\widetilde{0}] \in \widetilde{X}_0(N)(\mathbb{F}_q)$, and therefore it maps to $[\widetilde{0}] \in \widetilde{J}(\mathbb{F}_q)$. By the injection (1),

the preimage of $[\widetilde{0}] \in \widetilde{J}(\mathbb{F}_q)$ must be $[0]$ in $J(\mathbb{Q})$. In other words, (E, H) maps to $[0] \in J(\mathbb{Q})$.

On the other hand, if we take $\ell = 3$, then (E, H) maps to $[\widetilde{\infty}] \in X_0(N)(\mathbb{F}_3)$, and by the same argument we must have that (E, H) maps to $[\infty] \in J(\mathbb{Q})$. We conclude that $[0] = [\infty]$ in $J(\mathbb{Q})$. But $[0] - [\infty]$ has order $n = \text{num}\left(\frac{N-1}{12}\right)$ in $J(\mathbb{Q})$ [Maz III.1.4]. If $n = 1$, then we are forced to conlcude that $N - 1 \leq 12$, which contradicts our initial assumption of $N \geq 23$. This concludes the proof of Lemma 2.8. □

Now we can conclude the proof of Proposition 2.2, which is equivalent to Proposition 2.1.

Proof of Proposition 2.2.

(a) Let $q \neq N$ be any prime where E has good reduction. Then by the criterion of Néron-Ogg-Shafarevich [Si1 VII.7.1], the extension $L/\mathbb{Q}$ is unramified above q. We quickly review the argument here. If E has good reduction, then $E_0 = E$ and we have an exact sequence:

$$0 \to E_1(L/\mathbb{Q}_q) \to E(L/\mathbb{Q}_q) \to \widetilde{E}(k) \to 0$$

where we have indicated by k the residue field of $L/\mathbb{Q}_q$. Now, the group $E_1(L/\mathbb{Q}_q)$ contains no points of order N, whereas $E(L/\mathbb{Q}_q)$ contains all of them, by definition of L. Therefore we have an injection

$$E[N] \hookrightarrow \widetilde{E}(k)$$

But the intertia group $I(L/\mathbb{Q}_q)$ acts trivially on $\widetilde{E}(k)$, and therefore it must act trivially on $E[N]$, by injectivity. Therefore $L/\mathbb{Q}_q$ is unramified, from which it follows immediately that $L/\mathbb{Q}$ (hence L/K) is unramified at q.

(b) Let $q \neq N$ be a prime where E has bad, hence multiplicative (by Lemma 2.3), reduction.

We claim that $E_0[N]$ contains a Galois submodule isomorphic to μ_N, the N-th roots of unity. To see this, consider the exact sequence

$$0 \to H(L/\mathbb{Q}_q) \to E(L/\mathbb{Q}_q)[N] \to \mu_N \to 0$$

of $\mathrm{Gal}(L/\mathbb{Q}_q)$ modules, where the injection is given by inclusion and the surjection by the Weil pairing with one argument fixed. We deduce an exact sequence:

$$0 \to H_0(L/\mathbb{Q}_q) \to E_0(L/\mathbb{Q}_q)[N] \to \mu_N \to 0$$

But by Lemma 2.8, $H \nsubseteq E_0(\mathbb{Q}_q)$, so $H_0(L/\mathbb{Q}_q) = \{O\}$. In particular, $E_0[N] \cong \mu_N$ and therefore $E[N]$ contains a Galois sub-module isomorphic to μ_N. It follows that $E[N] \cong H \oplus \mu_N$. But now, the inertia group $I(L/\mathbb{Q}_q(\zeta_N))_v$ acts trivially on both H and μ_N for any prime v of $\mathbb{Q}_q(\zeta_N)$ lying above q, and therefore it acts trivially on $E[N]$. Therefore $L/\mathbb{Q}_q(\zeta_N)$ is unramified and L/K is unramified at q. It is a well-known fact (see for example [Wash, Proposition 2.1]) that the discriminant of $K/\mathbb{Q}$ is a power of N, and therefore the only rational prime that ramifies in K is N itself. It follows that $K/\mathbb{Q}$ is unramified at q and therefore $L/\mathbb{Q}$ is unramified at q.

(c) Let $q = N$. By Lemma 2.3 E does not have additive reduction at N. By the first paragraph of Lemma 2.8, if E has multiplicative reduction at N, then $H \nsubseteq E_0(\mathbb{Q}_q)$ and we can proceed as in (b). So suppose E has good reduction at N. We have an exact sequence of group schemes:

$$0 \to \mathbf{Z}/\mathbf{p} \to E[N]_{/S} \to \mu_{\mathbf{p}} \to 0$$

Where $\mathbf{Z}/\mathbf{p}$ and $\mu_{\mathbf{p}}$ are defined as in Fact II.2.6 and $E[N]_{/S}$ is the subgroup scheme of the Néron Model $E_{/S}$ for S. If we take the connected component of each group scheme in the sequence, we deduce that $E^0[N]_{/S} = \mu_{\mathbf{P}}$, since $\mathbf{Z}/\mathbf{p}$ consists of p closed points above N, and therefore its connected component is trivial. But again, this shows that we get the desired splitting of $E[N]$.

□

3 Herbrand's Theorem

In this section we prove the following:

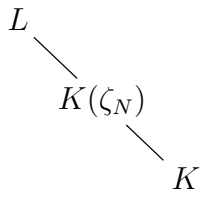
Proposition 3.1. *Suppose E is an elliptic curve defined over $\mathbb{Q}$ with a subgroup $H \subset E_{tors}(\mathbb{Q})$ of order N, for N a prime number and $N \geq 23$. Let $K = \mathbb{Q}(\zeta_N)$, for ζ_N a primitive root of unity, and let $L = \mathbb{Q}(E[N])$, obtained by adjoining the coordinates of all N-torsion points of E. Then $L = K$.*

In light of Proposition 2.1, we know that L/K is unramified. Therefore in order to prove Proposition 3.1 it suffices to prove the following:

Proposition 3.2. *If L/K is everywhere unramified then $L = K$.*

In order to prove Proposition 3.2, we need first a few defitions and a theorem of Herbrand.

Let N be a prime and let L/K be a Galois extension such that L contains ζ_N, a primitive N-th root of unity, and such that $V = \mathrm{Gal}(L/K(\zeta_N))$ is abelian of exponent N (i.e. every element is of order dividing N):



Suppose further that the action of $\mathrm{Gal}(K(\zeta_N)/K)$ on V induced by conjugation in $\mathrm{Gal}(L/K)$ is given by the formula:

$$\tau \cdot v = \chi(\tau)^j \cdot v$$

where $\chi : \mathrm{Gal}(K(\zeta_N)/K) \to (\mathbb{F}_N)^\times$ is the cyclotomic character sending τ to the element $i \in (\mathbb{F}_N)^\times$ such that $\tau(\zeta_N) = \zeta_N^i$. Then we call L a χ^j *- extension of* $K(\zeta_N)$.

The Bernoulli numbers B_i are defined as the coefficients appearing in the power series expansion

$$\frac{z}{e^z - 1} = \sum_{i=0}^{\infty} B_i \frac{z^i}{i!}$$

In particular, $B_0 = 1, B_1 = -1/2, B_2 = 1/6, \ldots$ and $B_i = 0$ for $i \neq 1$ odd. For any prime N, consider all the Bernoulli numbers of the form B_{2k}, $2 \leq 2k < N - 1$, and let $j = 1 - 2k \mod (N - 1)$. The following theorem of Herbrand gives a necessary condition for the existence of an unramified χ^j-extension:

Theorem 3.3 (Herbrand). *If $N \nmid \mathrm{num}\,(B_{2k})$, then there are no nontrivial everywhere unramified χ^j-extensions of $\mathbb{Q}(\zeta_N)$.*

Proof. See [Wash 6.3]. □

In particular, Herbrand's Theorem gives that there are no nontrivial everywhere unramified χ^{-1}-extensions of $K = \mathbb{Q}(\zeta_N)$, since $\mathrm{num}(B_2) = 1$ is not divisible by N. It follows that in order to prove Proposition 3.2 it suffices to prove the following Proposition:

Proposition 3.4. *Suppose E is an elliptic curve defined over $\mathbb{Q}$ with a subgroup $H \subset E_{tors}(\mathbb{Q})$ of order N, for N a prime number and $N \geq 23$. Let $K = \mathbb{Q}(\zeta_N)$, for ζ_N a primitive root of unity, and let $L = \mathbb{Q}(E[N])$, obtained by adjoining the coordinates of all N-torsion points of E. Suppose that $L \neq K$. Then L is a χ^{-1}-extension of K.*

We prove two lemmas first.

Lemma 3.5. *Let E, L, K as in Proposition 3.4. Then there exists a faitfhul representation*

$$\rho : \mathrm{Gal}(L/\mathbb{Q}) \hookrightarrow \mathbf{GL}_2(\mathbb{F}_N)$$

such that

$$\rho(\sigma) = \begin{pmatrix} 1 & b(\sigma) \\ 0 & \chi(\sigma) \end{pmatrix}$$

where $b(\sigma) \in \mathbb{F}_N$ and $\chi : \mathrm{Gal}(L/\mathbb{Q}) \to (\mathbb{F}_N)^\times$ is the cyclotomic character mod N defined by $\sigma(\zeta_N) = \zeta_N^{\chi(\sigma)}$ for σ acting on a primitive N-th root of unity ζ_N.

Proof.

Consider the map

$$e_N(P, \cdot) : E[N] \to \mu_N \subset K$$

given by the Weil pairing e_N. This map is defined over L and surjective with kernel equal to H. Composing with the inclusion $H \subset E[N]$ we get an exact sequence of groups:

$$0 \to H \to E[N] \to \mu_N \to 0$$

which is in fact an exact sequence of $\mathrm{Gal}(L/\mathbb{Q})$-modules, since every map is defined over L.

The exact sequence gives a faithful representation

$$\rho : \mathrm{Gal}(L/\mathbb{Q}) \hookrightarrow \mathbf{GL}_2(\mathbb{F}_N)$$

that works as follows. Pick a $Q \in E[N] - H$, so that P, Q form a basis for $E[N]$. For any $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$, let

$$\rho(\sigma) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

where

$$\begin{aligned} \sigma(P) &= aP + cQ \\ \sigma(Q) &= bP + dQ \quad a, b, c, d \in \mathbb{F}_N \end{aligned}$$

Since σ must permute primitive roots of unity, write $\sigma(\zeta) = \zeta^{\chi(\sigma)}$ for the action of σ on a primitive N-th root of unity ζ_N. Note that χ is simply a homomorphism

$$\mathrm{Gal}(L/\mathbb{Q}) \to (\mathbb{F}_N)^\times$$

which we called in the statement of the lemma the *cyclotomic character mod N*. Note first that $P \in \mathbb{Q}$ and therefore $c = 0$ and $a = 1$. Also, from the non-degeneracy of the Weil pairing we get that $e_N(P, Q)$ is a primitive N-th root of unity, so

$$\sigma(e_N(P, Q)) = \sigma(\zeta_N) = \zeta_N^{\chi(\sigma)}$$

On the other hand, from the Galois-invariance property we have

$$\begin{aligned} \sigma(e_N(P, Q)) &= e_N(\sigma(P), \sigma(Q)) \\ &= e_N(P, bP + dQ) \\ &= e_N(P, bP) e_N(P, dQ) \\ &= e_N(P, Q)^d \\ &= \zeta_N^d \end{aligned}$$

so that $d = \chi(\sigma)$ in $(\mathbb{F}_N)^\times$. In other words, we can write the action of σ on $E[N]$ as the matrix

$$\rho(\sigma) = \begin{pmatrix} 1 & b(\sigma) \\ 0 & \chi(\sigma) \end{pmatrix}$$

□

Lemma 3.6. *Let E, L, K as in Proposition 3.4. Then we either have $L = K$ or L/K cyclic of order N.*

Proof. From the shape of the representation given by Lemma 3.5 we see that $\#\operatorname{Gal}(L/\mathbb{Q}) \mid N(N-1)$. In particular, note that ρ restricted to an element $\tau \in \operatorname{Gal}(L/K) \subset \operatorname{Gal}(L/\mathbb{Q})$ has the form:

$$\rho(\tau) = \begin{pmatrix} 1 & b(\tau) \\ 0 & 1 \end{pmatrix}$$

since the action of χ is trivial on $\operatorname{Gal}(L/K)$. If there exists a $\tau \in \operatorname{Gal}(L/K)$ such that $b(\tau) \neq 0$ (i.e. if $K \neq L$), we have an element of order N in $\operatorname{Gal}(L/K) \subset \operatorname{Gal}(L/\mathbb{Q})$, hence $N \mid \#\operatorname{Gal}(L/\mathbb{Q})$. On the other hand, the tower of fields

$$\mathbb{Q} \subset K = \mathbb{Q}(\zeta_N) \subset L$$

implies that the degree of $L/\mathbb{Q}$ is divisible by $N-1$, since $K/\mathbb{Q}$ is an extension of degree $N-1$. Since $\gcd(N, N_1) = 1$, we either have $\#\operatorname{Gal}(L/\mathbb{Q}) = N(N-1)$, with L/K cyclic of order N, or $\#\operatorname{Gal}(L/\mathbb{Q}) = N-1$, with L/K trivial. □

We now proceed with the proof of Proposition 3.4 by computing the action of $\operatorname{Gal}(K/\mathbb{Q})$ on $\operatorname{Gal}(L/K)$.

Proof of Proposition 3.4.

Let $\sigma \in \operatorname{Gal}(K/\mathbb{Q})$. There are several lifts of this element to an element $\tilde{\sigma} \in \operatorname{Gal}(L/\mathbb{Q})$ (remember we are assuming $L \neq K$) but pick any one for now. For any $\tau \in \operatorname{Gal}(L/K)$, define the action of σ by

$$\sigma(\tau) = \tilde{\sigma}\tau\tilde{\sigma}^{-1}$$

One can check that this formula does not depend on our original choice of $\tilde{\sigma}$, so that the action is well-defined. We want to compute this action explicitly inside the representation given by Lemma 3.5.

Recall from Lemma 3.5 that $\mathrm{Gal}(L/\mathbb{Q})$ has a representation of the form

$$\rho(u) = \begin{pmatrix} 1 & b(u) \\ 0 & \chi(u) \end{pmatrix}$$

whereas from Lemma 3.6 we know that $\mathrm{Gal}(L/K)$ has a representation of the form

$$\rho(\tau) = \begin{pmatrix} 1 & b(\tau) \\ 0 & 1 \end{pmatrix}$$

Let now $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ be given by

$$\rho(\sigma) = \begin{pmatrix} 1 & b(\sigma) \\ 0 & \chi(\sigma) \end{pmatrix}$$

Since we are assuming $L \neq K$, by Lemma 3.6 the extension L/K has degree N. Therefore there is a choice of N lifts for σ to $\mathrm{Gal}(L/\mathbb{Q})$, one for each elemen $b(\sigma) \in \mathbb{F}_N$. For computational convenience, choose $b(\sigma) = 0$ so that $\rho(\tilde{\sigma}) = \begin{pmatrix} 1 & 0 \\ 0 & \chi(\sigma) \end{pmatrix}$.

Then, for any $\tau \in \mathrm{Gal}(L/K)$:

$$\rho(\tilde{\sigma}\tau\tilde{\sigma}^{-1}) = \begin{pmatrix} 1 & 0 \\ 0 & \chi(\sigma) \end{pmatrix} \cdot \begin{pmatrix} 1 & b(\tau) \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & \chi^{-1}(\sigma) \end{pmatrix} = \begin{pmatrix} 1 & b(\tau)\chi^{-1} \\ 0 & 1 \end{pmatrix}$$

In other words, given any $\tau \in \mathrm{Gal}(L/K)$, the action of $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ is given by the simple formula

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

□

This concludes the proof of Proposition 3.4, which together with Herbrand's Theorem proves Proposition 3.2. Proposition 3.2 and Proposition 2.1 together prove Proposition 3.1.

4 End of the Proof

Proposition 3.1 gives us that the extension L/K is trivial. We now proceed in deriving a contradiction from this fact, and conclude the proof of Mazur's Theorem.

Proposition 4.1. *Let E be an elliptic curve defined over $\mathbb{Q}$, and such that $E_{tors}(\mathbb{Q})$ contains a cyclic subgroup H of order N, for N a prime and $N \geq 23$. Let $K = \mathbb{Q}(\zeta_N)$, where ζ_N is a primitive N-th root of unity, and let $L = \mathbb{Q}(E[n])$ be the field obtained by adjoining all the coordinates of the N-th torsion points of E. If $L = K$ there are infinitely many points on the curve $X_0(N)(\mathbb{Q})$.*

Proof.

Consider again the exact sequence:

$$0 \to H \to E[N] \to \mu_N \to 0$$

In view of Lemma 3.5, if $K = L$ then $\mathrm{Gal}(K/\mathbb{Q})$ has a representation:

$$\begin{pmatrix} 1 & 0 \\ 0 & \chi(\sigma) \end{pmatrix}$$

from which we deduce that $E[N]$ splits as $E[N] \cong H \oplus \mu_N$.

In this case, not only does E have a Galois sub-module H, but also a Galois sub-module Φ isomorphic to μ_N. Since Φ is invariant under the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, we get by [Si1 III.4.13.2] that there is a curve E' *defined over* $\mathbb{Q}$ and an isogeny $\psi : E \to E'$, *also defined over* $\mathbb{Q}$, such that the kernel of ψ is Φ. Now, any nonconstant isogeny is a surjective homomorphism, from which we deduce that the group structure of E' is isomorphic to E/Φ. Since $\Phi \cap H = \emptyset$, the image of H under this quotient is again a cyclic group $H' = \psi(H)$ of order N and rational over $\mathbb{Q}$, since H and ψ are. But then the pair (E', H') satisfies the hypotheses of Proposition 3.1, and we can repeat the same argument to find a pair (E'', H'') and an isogeny $\psi' : E' \to E''$ with kernel $\Phi' \cong \mu_N$. Continuing this way we get a chain:

$$E \xrightarrow{\psi} E' \xrightarrow{\psi'} E'' \xrightarrow{\psi''} \ldots \xrightarrow{\psi^{(i-1)}} E^{(i)} \xrightarrow{\psi^{(i)}} \ldots \tag{III.2}$$

of elliptic curves defined over $\mathbb{Q}$ with a cyclic subgroup $H^{(i)}$ of order N, rational over $\mathbb{Q}$. Suppose there exists a pair (i, j) for which $E^{(i)} = E^{(j)}$. The composition of the maps $\psi^{(k)}$ between $E^{(i)}$ and $E^{(j)}$ then gives us an endomorphism of $E^{(i)}$ defined over $\mathbb{Q}$, so it must be one of the multiplication-by-m maps $\phi = [m]$ for some integer m. Since $\ker \phi = E^{(i)}[m]$ contains a group of order N, namely $\Phi^{(i)}$, we must have $m = N^e \cdot r$ for some $e \geq 1$ and $(N, r) = 1$. But then $H^{(i)}$, which is contained in $E^{(i)}[N]$, would also be contained in $\ker(\phi)$, which is impossible by construction.

Therefore no pair (E^i, C^i) appearing in (2) is contained in the same isomorphism class, and we have found infinitely many points on the curve $X_0(N)(\mathbb{Q})$. □

Bibliography

[AtLe] Atkin, A.O.L., Lehner, J. *Hecke operators on* $\Gamma_0(m)$, Math. Ann., 185 (1970). 134-160.

[Del] Deligne, P. *Formes modulaires et représentations l-adiques*. Séminaire Bourbaki 1968/69, no. 355.

[CuRe] Curtis, C.W., Reiner I. *Representation theory of finite groups and associative algebras*, New York, Interscience, 1962.

[EGA IV] Grothendieck, A. and Dieudonné, J. *Étude locale des schémas et des morphismes des schémas* Publ.Math. IHES 20, 24, 28, 32.

[Har] Hartshorne, R. *Algebraic Geometry* Springer GTM 1977.

[Kay] aye, G.R. *Indian Mathematics* Isis, Vol. 2, No. 2. (Sep., 1919), pp. 326-356.

[Kub] Kubert, D.S. *Universal bounds on the torsion of elliptic curves*. Proc. London Math. Soc. (3) 33 (1976),193-237.

[Lang] Lang, S. *Algebra* Springer-Verlag GTM 2002.

[Lig] Ligozat, G. *Fonction L des courbes modulaires*. Séminaire Delange-Pisot-Poitou (1969-70) no. 9

[Lind] Lind, E. *Untersuchungen über die rationalen Punkte der ebenen kubishen Kurven vom Geschlecht Eins*. Appdesbergs Boktryekeriaktiobolag, Uppsala, 1940.

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

[MaSe] Mazur B., Serre J. P. *Points Rationelles des Courbes Modulaires* $X_0(N)$, Seminaire Bourbaki 1974/75 no. 469.

[MaTa] Mazur B., Tate J., *Points of order 13 on elliptic curves*. Invent. Math. 22 (1973) 41-49.

[Ogg] Ogg, A. *Rational oints on certain on certain elliptic modular curves*. Proc. Symp. Pure Math., 24 (1973), AMS, Providence, p. 221-231.

[1] MOV] Menezes, A.J., Okamoto, T., Vanstone S.A. *Reducing elliptic curve logarithms to logarithms in a finite field*. IEEE Trans. Inform. Theory 39(5): 1639-1646, 1993.

[OrTa] Oort F., Tate J., *Group schemes of prime order* Ann. Scient. Ec. Norm. Sup. série 4,3 (1970) 1-21.

[Ray] Raynaud, M. *Schémas en group de type (p,...,p)* Bull.Soc.Math. France 102 (1974) 241-280

[Rib] Ribet K., *Endomorphisms of semi-stable abelian varieties over number fields*, Annals of Math., 101 (1975), p. 555-562.

[Roh] Rohrlich D. *Modular curves, Hecke correspondences, and L-functions*. In Modular Forms and Fermat's Last Theorem, pages 41-100. Springer, 1997.

[Shi] Shimura, G. *Introduction to the Arithmetic Theory of Automorphic Functions*, 1971 Princeton University Press

[Si1] Silverman, J.H. *The Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics, 1986 Springer-Verlag, ISBN 7-5062-0135-6

[Si2] Silverman, J.H. *Advanced Topics in the Arithmetic of Elliptic Curves*, Graduate Texts in Mathematics, 1994 Springer-Verlag

[Ste] Stein, W. *Modular Abelian Varieties* online lecture notes, Harvard University 2003. Available at http://modular.math.washington.edu

[Weil] Weil, A. *Variétés Abéliennes et Courbes Algébriques*, Hermann, Paris, 1948.

[Wash] Washington, L. *Introduction to Cyclotomic Fields* Springer-Verlag 1982