Stark-Heegner points attached to Cartan non-split curves

Juan Ignacio Restrepo

Doctor of Philosophy

Department of Mathematics and Statistics

McGill University

Montreal,Quebec

July 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor in Philosophy

ACKNOWLEDGEMENTS

I want to thank first my advisor, Professor Henri Darmon, whose guidance was vital during the development of this project. All these years, his valuable comments during my talks and our meetings helped me grasp a deeper understanding of mathematics. His patience and calm helped me get courage to ask questions even when I was feeling overwhelmed and pointed me in the right direction.

Next, I must thank Daniel Kohen, who spent a significant amount of time discussing with me his work with Professor Ariel Pacetti, which is a cornerstone of this thesis. He also spent countless hours doing computations with me trying to pin down the correct analogies that led to the main contributions of this project.

I must thank Marc Masdeu and Cameron Franc, who were senior graduate students at the time I started. Their guidance starting to navigate the mathematical world was extremely important. On the same alley, Francesc Castellà explained to me a plethora of things when we shared an office. Luiz Takei, who made me do mathematics in Portuguese while talking with me in a way as down to earth as possible about topics we both found hard many times. And, Andrew Fiori and Nicolas Simard, with whom I had multiple discussions about various topics in Number Theory and computing.

I would be remiss not to express my gratitude towards NSERC, who funded half of my stay at McGill, and the Department of Mathematics and Statistics, McGill University and my advisor for funding the other half.

I would like to thank my friends Ben Smith, Daphna Harel and Katherine Daignault, who provided moral support in moments of distress. The list of people in this category is rather long, so unfortunately I cannot mention every one.

Finally, I want to thank my family for their continued support throughout these years.

ABSTRACT

Let E be an elliptic curve of conductor pq^2, where p and q are prime numbers, and let K be a quadratic extension of $\mathbb{Q}$. If K is imaginary and p and q are split in K, there are Heegner points on the modular curve $X_0(pq^2)$ defined over ring class fields attached to orders in K, which can be mapped to points on E. If q is inert, there are no Heegner points on the modular curve, but points can be obtained from the Cartan non-split curve $X_{ns}^{\varepsilon}(q,p)$. If K is real the panorama is quite different. If p is inert and q is split, Stark-Heegner points have been defined on E, whose field of definition is conjecturally the narrow ring class field attached to an order in K. This work combines these two ideas, defining Stark-Heegner points when p and q are inert in the real quadratic field K, using Cartan non-split curves, which are conjecturally defined over narrow ring class fields attached to orders in K.

ABRÉGÉ

Soit E une courbe elliptique de conducteur pq^2, où p et q sont des nombres premiers, et soit K une extension quadratique de $\mathbb{Q}$. Si K est imaginaire et p et q sont décomposés dans K, on dispose de points de Heegner sur E construits en appliquant la paramétrisation par la courbe modulaire $X_0(pq^2)$ aux points CM attachés à K. Ces points sont définis sur des corps de classes d'anneau (*ring class fields*) attachés à des ordres dans K. Lorsque q est inerte, le système de points algébriques analogue s'obtient en remplaçant $X_0(pq^2)$ par la courbe $X_{ns}^{\varepsilon}(q,p)$ associée au sous-groupe de Cartan non-déployé en q, par une construction rendue explicite par Kohen et Pacetti. Lorsque K et réel, la situation et radicalement différente, du fait que l'on ne dispose plus de la construction des points de Heegner. Néanmoins, si p est inerte et q se décompose dans K, des soi-disant points de *Stark-Heegner* ont été définis sur E, dont le corps de définition est conjecturalement l'anneau de corps de classes au sens restreint attaché à un ordre dans K. Notre travail combine ces deux idées, pour définir des points de Stark-Heegner en niveau pq^2 quand p et q sont inertes dans le corps réel quadratic K, à partir de la cohomologie des groups p-arithmétiques associés à des sous-groupes Cartan non-déployés en q. Il s'agit donc, en fin de compte, d'adapter les constructions de Kohen-Pacetti au cadre des points de Stark-Heegner.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

ABSTRACT iv

ABRÉGÉ v

1 Introduction 1

2 Background 4

2.1 Algebraic Curves 4
2.1.1 Affine Curves 4
2.1.2 Projective Curves 5
2.1.3 M-rational points. 6
2.1.4 Function fields 7
2.1.5 Smoothness 7
2.1.6 Divisors 9
2.1.7 Differentials 10
2.1.8 Genus 11
2.2 Curves of genus zero 13
2.2.1 Conics and Ternary Quadratic Forms 13
2.2.2 Binary Quadratic Forms 19
2.3 Elliptic Curves 25
2.3.1 Group Structure 34
2.3.2 Isogenies 37
2.3.3 L-functions 42

3 Modular Forms 45

3.1 Elliptic Curves arising from Modular Forms 50
3.2 The Modular Curve $X_0(N)$. 58
3.3 Hecke operators 67
3.4 L-functions associated to Modular Forms 71

3.5 Modularity 73

4 Heegner and Stark-Heegner points; the classical case 76

4.1 Complex Multiplication and Class Field Theory 76
4.2 Heegner points 81
4.3 Stark-Heegner points 87
4.3.1 p-adic measures and p-adic line integrals 88
4.3.2 Modular Forms on $\Gamma_{p,M}$ 91
4.3.3 Measures associated to an Elliptic Curve and the double integrals 93
4.3.4 Tate's uniformization 98
4.3.5 The Stark-Heegner point 102
4.3.6 Computational remarks 111
4.4 Heegner points attached to Cartan Non-Split curves 111
4.4.1 Cartan Non-split curves 112
4.4.2 Modular Forms over $\Gamma_{ns}^{\varepsilon}(p)$ 115
4.4.3 Modular Parametrization 118
4.4.4 Higher levels 121
4.4.5 Heegner points 122

5 Stark-Heegner points attached to Cartan Non-split curves 125

5.1 The group 126
5.1.1 Cusps 132
5.2 Modular Forms 138
5.3 Measures, double Integrals and semi-indefinite Integrals 147
5.4 The Stark-Heegner point 154
5.5 Setup for computations 156
5.5.1 The case $q = 3$. 156
5.5.2 More general values of q. 159

6 Further directions 176

References 182

Chapter 1
Introduction

Finding rational points on elliptic curves is not as easy as finding rational points on conic sections. Actually, determining their existence, or lack thereof, is also a much harder problem. For conic sections we have the Hasse-Minkowski Principle at our disposal (see [Ser73]) which helps us determine readily when a conic section has rational points. Moreover, once we have a rational point on the conic, finding all of the points can easily be done using the chord and tangent method. The counterpart for elliptic curves is much harder; the techniques are much more sophisticated and are not even guaranteed to succeed (see [ST92]). However, often they do work and we manage to characterize the set (a posteriori, group) of rational points.

One of the most used techniques in order to find rational points on elliptic curves stems from the work of Heegner on the class-number problem for imaginary quadratic fields. He defined some points, under the hypothesis that all the primes dividing N are split in K (or ramify, with a further restriction), on the modular curves $X_0(N)$, which can be seen as the quotient of the complex upper half-plane by the congruence group $\Gamma_0(N)$. These points are called *Heegner points.* Birch recognized that the modular parametrization mapped these points into points on the elliptic curve defined over an algebraic extension of the base field. Upon taking the trace of one of these points, we can obtain a point with rational coordinates (or, with

coordinates in the field with which we started) and under some further conditions, they can be shown to be non-torsion (see [Gro84] and [GZ86]).

Darmon proceeded to extend Birch's ideas to cases where the base field is not an imaginary quadratic field (see [Dar04]). The so-obtained points do not lie on the complex upper half-plane, so the usual modular parametrization will not yield any points. Hence, Darmon used a p-adic version of the modular parametrization and found that these points share a lot of similarities with the points Birch found before. Again, we need suitable conditions for these points to exist. The setup is a quadratic real field K and an elliptic curve of conductor $N = pM$, where p is inert in K and all the primes dividing M are split in K. Since Darmon drew a parallel between this generalization and Stark units relative to elliptic or circular units, he called these *Stark-Heegner points,* although sometimes, in the literature, these points are also referred to as *Darmon points.*

Recently, Kohen and Pacetti (see [KP14]) explored cases where the suitable conditions required for Birch's and Darmon's constructions do not hold. This approach allowed for the finding of new points which could not be obtained with the previous methods, as the classical modular curves do not have Heegner points when a prime dividing the conductor is inert. Their setup involves a quadratic imaginary field K and an Elliptic Curve of conductor $N = n^2m$, where n is square-free, all the primes dividing n are inert in K and all the primes dividing m are split in K. Their construction is based on the consideration of different modular curves called *Cartan non-split curves,* which can also be seen as the quotient of the complex upper half-plane by a congruence subgroup. However, the construction of these curves can

also be presented (as is the case with the classical modular curves) as the solution to a moduli problem, allowing for a possible analogous generalization to the one of Darmon above.

This thesis emulates Darmon's generalization of the classical case in conditions similar to those where Kohen and Pacetti obtained new points. Specifically, for an elliptic curve of conductor pq^2 and a quadratic real field K where the primes p and q are inert, we describe a construction that yields points on E, which are conjecturally defined over narrow ring class fields attached to orders in K.

Chapter 2
Background

In this chapter we provide a brief introduction to the objects that will be used throughout the work. We will be looking for special points on curves. The most basic invariant about a curve is its genus, which we define in Section 2.1.8 below. This chapter mainly follows [Sil86] and [Har77].

2.1 Algebraic Curves

2.1.1 Affine Curves

Let K be a field and L an algebraic closure. Let $S \subseteq K[x_1, \ldots, x_n]$ be a set and let $Z(S) = \{P \in L^n | \forall f \in S, f(P) = 0\}$. Let $R = L[x_1, \ldots, x_n]$. The set $Z(S)$ is called an *algebraic set*. Clearly $Z(S) = Z((S))$, where (S) is the ideal generated by S in R. Also, if $S_1 \subseteq S_2 \subseteq R$, $Z(S_1) \supseteq Z(S_2)$. Since R is a noetherian ring, every ideal has a finite generating set, so we can safely assume S is finite (or an ideal.) If S_1 and S_2 are two ideals of R, $Z(S_1) \cup Z(S_2) = Z(S_1 S_2)$. If $\{S_i\}_{i \in I}$ is a collection of subsets of R, $\bigcap_{i \in I} Z(S_i) = Z\left(\bigcup_{i \in I} S_i\right)$. Also, $L^n = Z(0)$ and $\varnothing = Z(1)$. This means that we can endow L^n with a topology using $\{Z(S) | S \subseteq R\}$, the algebraic sets, as the collection of closed sets. This topology is referred to as the *Zariski* topology.

Now, let $T \subseteq L^n$ and let $I(T) = \{f \in R | \forall P \in T, f(P) = 0\}$. This is seen to be an ideal of R. If $T_1 \subseteq T_2 \subseteq L^n$, $I(T_1) \supseteq I(T_2)$. To every ideal, we are associating an algebraic set and to every algebraic set we are associating an ideal. We have the relations $Z(I(T)) = \overline{T}$ and $I(Z(S)) = \sqrt{S}$ (Hilbert's Nullstellensatz), where $\overline{T}$ is the

closure of T in L^n and $\sqrt{S}$ is the radical of the ideal generated by S. This establishes a correspondence between algebraic sets and radical ideals.

An algebraic set Z is said to be *irreducible* if whenever $Z = Z_1 \cup Z_2$ with Z_1 and Z_2 algebraic sets, we have $Z = Z_1$ or $Z = Z_2$. An irreducible algebraic set is called an *algebraic variety.* It should be noted that the defining ideal of an algebraic variety can always be taken to be prime. Every point $(a_1, \ldots, a_n) = Z(x_1 - a_1, \ldots, x_n - a_n)$ is an irreducible set. An *affine algebraic curve* is an irreducible algebraic set whose only irreducible subsets are points.

2.1.2 Projective Curves

We define the projective space $\mathbb{P}^n(L) = (L^{n+1} - \{0\})/\sim$, where $(x_0, \ldots, x_n) \sim (y_0, \ldots, y_n)$ if and only if there exists $\lambda \in L^\times$ such that $x_i = \lambda y_i$ for $i = 0, \ldots, n$. It does not make sense to evaluate a polynomial in $R = L[x_0, \ldots, x_n]$ at a point in $\mathbb{P}^n(L)$, as the polynomial does not evaluate to the same element at every representative of it. However, if we only restrict to homogeneous polynomials, we have that $f(\lambda(x_0, \ldots, x_n)) = \lambda^d f(x_0, \ldots, x_n)$, where d is the degree of the polynomial. This does not ensure the value is well defined, but it does ensure the vanishing of the polynomial at a point in the projective space is well defined. Let $S \subseteq L[x_0, \ldots, x_n]$ be a subset of homogeneous polynomials. A *projective algebraic set* is a set of the form $Z(S) = \{P \in \mathbb{P}^n(L) | \forall f \in S, f(P) = 0\}$. The ideal (S) generated by S is called a *homogeneous ideal.* The discussion above for affine algebraic curves follows almost verbatim, except for the fact that the homogeneous ideals we deal with must be contained in the ideal $R_+ = (x_0, \ldots, x_n)$ to have the correct correspondence. In this case, we call it a *projective algebraic curve.* Because most curves we consider in this

thesis are algebraic, the use of the word *curve* will refer to algebraic curves unless explicitly stated otherwise.

Every projective curve C can be covered by affine patches. Let $C_i \subseteq C$ be the subset of points where $x_i \neq 0$. The dehomogenization map $\varphi_i : C_i \longrightarrow L^n$ given by

$$(x_0, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \mapsto (x_0/x_i, \ldots, x_{i-1}/x_i, x_{i+1}/x_i, \ldots, x_n/x_i)$$

maps C_i isomorphically into an affine curve, which is one of the said affine patches. Every affine curve C' is an affine patch of a projective curve via the homogenization map $\phi_i : C' \longrightarrow \mathbb{P}^n(L)$ given by

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

Note further that these maps are inverses of each other, i.e., $\phi_i \circ \varphi_i = id_{C_i}$ and $\varphi_i \circ \phi_i = id_{C'}$. Often, when dealing with projective curves, we work with the affine patch C_0 and understand tacitly that the actual curve we are referring to is $\phi_0(C_0)$ together with the points that satisfy the same equations and $x_0 = 0$.

2.1.3 M-rational points.

The polynomials we chose have coefficients in K, so we say that the curve C is *defined over* K. For any algebraic extension M of K together with an embedding $M \hookrightarrow L$, we have natural embeddings $M^n \hookrightarrow L^n$ and $\mathbb{P}^n(M) \hookrightarrow \mathbb{P}^n(L)$. The Galois group $\mathrm{Gal}(L/M)$ acts on the left (in contrast to Silverman's notation, where it acts on the right) on L^n and $\mathbb{P}^n(M)$ and this action restricts to C. We denote by $C(M)$ the set of M-rational points of C, i.e., the set of points in L^n (or $\mathbb{P}^n(L)$ if the curve is projective) that are invariant under the action of $\mathrm{Gal}(L/M)$. This is also the set

of points in L^n (or $\mathbb{P}^n(L)$) in the image of the respective embedding which satisfy the polynomial equations defining the curve C.

2.1.4 Function fields

Let C be a curve defined over K. We define the *coordinate ring* $K[C]$ to be the quotient of the ring $K[x_1, \ldots, x_n]$ (or $K[x_0, \ldots, x_n]$ if C is projective) by $I(C)$. We note that if C is affine, $K[C]$ is isomorphic to the ring of *regular* functions on C, this is, the functions that can be locally represented as the quotient of two polynomials in $K[x_1, \ldots, x_n]$. On the other hand, if C is projective $K[C]$ is isomorphic to K.

The *function field* of C, denoted $K(C)$, is the field defined by equivalence classes of regular functions defined on open subsets of C, where we say that two classes are equivalent if the functions agree on a nonempty open set contained in the overlap of the domains. We remark that if C is affine $K(C)$ is isomorphic to the field of fractions of $K[C]$. (Since C is irreducible, $I(C)$ is prime, making $K[C]$ an integral domain.) When C is projective, $K(C)$ is isomorphic to the function field of an affine patch of C, which is also isomorphic to the degree zero elements of the field of fractions of the coordinate ring $K[C]$. Note that when the denominator does not vanish, these quotients evaluate to some value in K, despite the homogeneous coordinates having many representatives. This is because the degrees being equal will cancel the λ value attesting for equivalence, and the value the function takes on at each point (where it is defined) is independent of the choice of homogeneous coordinates.

2.1.5 Smoothness

Let $C \subseteq L^n$ be an affine curve with defining ideal $\langle f_1, \ldots, f_r \rangle$ and let $P \in C$ be a point. The *Jacobian matrix at* P is defined as the matrix of partial derivatives of

the f_j with respect to each of the variables evaluated at P. The Jacobian matrix has dimensions $n \times r$ and its rank can be at most $n-1$. We say that C is smooth (or nonsingular) at P if this rank is actually equal to $n-1$. Note that in the case of a plane curve defined by only one polynomial, we have a 2×1 matrix and the rank is at most 1. Smoothness is given by this matrix having a nonzero entry. The curve C is said to be smooth it is smooth at every point $P \in C$.

For a projective curve $C \subseteq \mathbb{P}^n(L)$ things are not very different. We say the curve is smooth at P if an affine patch (and therefore any) containing P is smooth. Also, we can check the smoothness at P via a similar test to the one employed for affine curves. Let the defining ideal be $\langle f_1, \ldots, f_r \rangle$ and choose a fixed representative of P. This Jacobian matrix has dimensions $(n+1) \times r$ but its rank can be at most $n-1$ as well. The projective curve is smooth at P when this rank attains its maximum possible value, $n-1$ again.

In contrast with the affine case, elements of the function field of smooth projective curves have a very special property relating its zeroes and poles. Let $P \in C$ and denote by m_P the ideal in $L[C]$ consisting of functions vanishing at P. The localization of $L[C]$ away from m_P is a local ring with maximal ideal m_P. Furthermore, it is a discrete valuation ring and $\bigcap_{n=1}^{\infty} m_P^n = 0$. Let π_P be an element of m_P which does not lie in m_P^2. (This element, called a *uniformizer*, exists because an equivalent condition for smoothness of a curve is that the vector space m_P/m_P^2 be a one-dimensional L-vector space.) For a nonzero element $f \in L[C]$, we define the *order of vanishing of f at P*, and denote it $\mathrm{ord}_P(f)$, to be the largest integer n such that $f \in m_P^n$. Since the intersection of all the powers of m_P is just 0, this n must

exist. Note that this integer can be 0. Now, let $t \in L(C)^\times$ and choose $f, g \in L[C]$ of the same degree such that $t = f/g$. We extend the definition of $\mathrm{ord}_P(\cdot)$ to $L(C)^\times$ by defining $\mathrm{ord}_P(t) = \mathrm{ord}(f) - \mathrm{ord}_P(g)$. It should be noted that this definition is independent of the choice of uniformizer and the representation chosen for t. If $\mathrm{ord}_P(t) > 0$ we say that t *vanishes* or has a *zero* at P, and if $\mathrm{ord}_P(t) < 0$ we say that t has a *pole* at P. If $\mathrm{ord}_P(t) \geq 0$ we say that t is *defined* at P. For any $t \in L(C)^\times$, $\mathrm{ord}_P(t) = 0$ for all but finitely many P. Moreover, the sum $\sum_{P \in C} \mathrm{ord}_P(t) = 0$ and if t has no poles (and therefore no zeroes) $t \in L$.

2.1.6 Divisors

Let C be a curve. Let $\mathrm{Div}(C)$ be the abelian free group generated by the points of C. The elements of $\mathrm{Div}(C)$ can be seen as formal sums indexed by the points of C, this is

$$\mathrm{Div}(C) = \left\{ \sum_{P \in C} n_P[P] : n_P \in \mathbb{Z} \text{ and } n_P = 0 \text{ for all but finitely many } P \right\}.$$

There is a natural group homomorphism $\deg\colon \mathrm{Div}(C) \longrightarrow \mathbb{Z}$ defined by

$$\deg\left(\sum_{P \in C} n_P[P] \right) = \sum_{P \in C} n_P.$$

This sum is well defined because the n_P are almost all 0 and we refer to this homomorphism as the *degree map*. Its kernel, $\ker(\deg)$, is denoted $\mathrm{Div}^0(C)$.

$\mathrm{Div}(C)$ has a natural action of the Galois group $\mathrm{Gal}(L/M)$ defined by

$$\sigma\left(\sum_{P \in C} n_P[P] \right) = \sum_{P \in C} n_P[\sigma(P)].$$

A divisor is said to be defined over M if it is invariant under the action of $\mathrm{Gal}(L/M)$. In particular, divisors where $n_P = 0$ for all $P \in C(L) - C(M)$ are defined over M, but these are not the only ones. It suffices (and it is enough) that for every point $P \in C$, the values of the n_Q be equal to n_P for all $Q \in C$ which are conjugates of P under some automorphism in $\mathrm{Gal}(L/M)$. The set of divisors of C defined over M is denoted $\mathrm{Div}_M(C)$, and $\mathrm{Div}^0_M(C)$ is its degree zero counterpart.

For an element $f \in L(C)$ we let $\mathrm{div}(f) = \sum_{P \in C} \mathrm{ord}_P(f)[P]$ be the divisor whose coefficient at P is the order of vanishing of f at P. Note that this divisor is well defined, as the order of vanishing has finite support for a fixed f. The divisors associated to an element of the function field of C will be referred to as *principal divisors*. We say that two divisors P and Q are *linearly equivalent* if there exists a principal divisor $\mathrm{div}(f)$ such that $P - Q = \mathrm{div}(f)$. The *Picard group* of C, $\mathrm{Pic}(C)$, is the quotient $\mathrm{Div}(C)$ modulo linear equivalence. The remark at the end of the last subsection asserts that $\mathrm{div}(f) \in \mathrm{Div}^0(C)$ for all $f \in L(C)$. This implies that this quotient descends to $\mathrm{Pic}^0(C)$, $\mathrm{Div}^0(C)$ modulo linear equivalence, which is also known as the *Jacobian* of C. We define $\mathrm{Pic}_M(C)$ and $\mathrm{Pic}^0_M(C)$ as the subgroups of $\mathrm{Pic}(C)$ and $\mathrm{Pic}^0(C)$ fixed by $\mathrm{Gal}(L/M)$, respectively.

2.1.7 Differentials

Let $\Omega(C)$ be the $L(C)$-vector space generated by the formal elements dx, $x \in L(C)$, quotiented by the subspace generated by the elements of the form da for all $a \in L$, $d(x+y) - dx - dy$ and $d(xy) - xdy - ydx$ for all $x, y \in L(C)$. $\Omega(C)$ happens to be a 1-dimensional $L(C)$-vector space (Theorem II.4.2(a) in [Sil86]), so it is generated by any nonzero element.

Let $P \in C$ and let $\pi_P \in L(C)$ be a uniformizer at P. For a given $\omega \in \Omega(C)$, we have a unique element $g \in L(C)$ such that $\omega = gd\pi_P$ (since $\Omega(C)$ is 1-dimensional). Denote this g by $\omega/d\pi_P$. We define the *order of* ω *at* P as $\mathrm{ord}_P(\omega) = \mathrm{ord}_P(\omega/d\pi_P)$. This order is independent of the uniformizer π_P.

Like in the case of elements in the function field, $\mathrm{ord}_P(\omega) = 0$ for all but finitely many $P \in C$. This allows us to define $\mathrm{div}(\omega) = \sum_{P \in C} \mathrm{ord}_P(\omega)[P]$ and this will be an element of $\mathrm{Div}(C)$. For $\omega \in \Omega(C)$ we say that the differential is *holomorphic* (or *regular*) if $\mathrm{ord}_P(\omega) \geq 0$ for all $P \in C$. We say that it is *nonvanishing* if $\mathrm{ord}_P(\omega) \leq 0$ for all $P \in C$.

Notice that if ω_1 and ω_2 are two nonzero differentials, there must exist some $f \in L(C)^\times$ such that $\omega_1 = f\omega_2$, because of the 1-dimensionality of the vector space of differential forms in C. Hence , whenever $\omega_2 = gd\pi_P$, we have that $\omega_1 = f\omega_2 = fgd\pi_P$, so $\mathrm{ord}_P(\omega_1) = \mathrm{ord}_P(f) + \mathrm{ord}_P(\omega_2)$. This equality holds for all $P \in C$, which implies that $\mathrm{div}(\omega_1) = \mathrm{div}(f) + \mathrm{div}(\omega_2)$, making $\mathrm{div}(\omega_1)$ and $\mathrm{div}(\omega_2)$ linearly equivalent, which in turns implies that the image of $\mathrm{div}(\omega_1)$ and $\mathrm{div}(\omega_2)$ in the Picard group is the same. This allows for a definition: the *canonical divisor class on* C is the image in $\mathrm{Pic}(C)$ of any nonzero differential $\omega \in \Omega(C)$.

2.1.8 Genus

Now we have enough material to define the invariant we mentioned at the beginning of the introduction.

If $D^1, D^2 \in \mathrm{Div}(C)$, we say that $D^1 \geq D^2$ whenever $n_P^1 \geq n_P^2$ for all $P \in C$. For a divisor $D \in \mathrm{Div}(C)$ we associate the following two objects: the vector space

$$\mathcal{L}(D) = \{f \in L(C)^\times : \mathrm{div}(f) \geq -D\} \cup \{0\}$$

and its dimension $\ell(D) = \dim_L(\mathcal{L}(D)) < \infty$.

Note that if $\deg D < 0$, this vector space is trivial (and hence, its dimension is 0) because whenever $f \in \mathcal{L}(D)$ and $f \neq 0$ we have that

$$0 = \deg(\operatorname{div}(f)) \geq \deg(-D) = -\deg(D).$$

Also, if D^1, D^2 are linearly equivalent, the corresponding vector spaces are isomorphic via the isomorphism

$$\mathcal{L}(D^1) \longrightarrow \mathcal{L}(D^2), \qquad f \longmapsto fg,$$

where g is such that $D^1 = D^2 + \operatorname{div}(g)$. In particular, if $K_C \in \operatorname{Div}(C)$ is a canonical divisor on C, say $\mathcal{L}(C) \cong \{\omega \in \Omega(C) : \omega \text{ is holomorphic}\}$. The *genus* of the curve C is the number $\ell(K_C)$.

Theorem 2.1 (Riemann-Roch)**.** *Let C be a curve and K_C a canonical divisor on C. There is an integer $g' \geq 0$ such that for every divisor $D \in \operatorname{Div}(C)$,*

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

Applying Riemann-Roch to $D = 0$ we obtain $g = \ell(K_C) = g'$, so g' is actually the genus of C and we will drop the prime henceforth. Also, applying Riemann-Roch to $D = K_C$ (together with what was just mentioned) we obtain $\deg(K_C) = 2g - 2$. Finally, if $\deg(D) > 2g - 2$, then $\deg(K_C - D) < 0$, making $\ell(K_C - D) = 0$, yielding $\ell(D) = \deg(D) - g + 1$. We record these in the following corollary.

Corollary 2.2. *Let C be a curve, K_C be a canonical divisor on C and g' the integer given by the Riemann-Roch Theorem.*

a) The genus of C, $\ell(K_C)$, is equal to g'.

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

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

2.2 Curves of genus zero

If we classify curves according to their genus, the first natural example to look at is curves of genus 0. This theory has been thoroughly studied and there are techniques to easily decide whether or not there are K-rational points on these curves, and how to find them. (See [Ser73].)

2.2.1 Conics and Ternary Quadratic Forms

We keep the notations from the previous section. For any algebraically closed field, there is essentially just one curve of genus 0. Furthermore, any smooth ternary quadratic form over K (a conic) is isomorphic over L to this curve.

Theorem 2.3. *The only curve of genus 0 over* L *is* $\mathbb{P}^1(L)$.

Proof. First, let us show that the genus of $\mathbb{P}^1(L)$ is actually 0. $\mathbb{P}^1(L) = Z(0)$, making $\mathbb{P}^1(L)$ a projective variety (as 0 is a homogeneous prime ideal in $L[x, y]$). Let $\infty = (1, 0)$, $0 = (0, 1)$ and $P = (a, b)$. Let $t = x/y$ be a uniformizer at 0, $\pi_\infty = y/x = 1/t$ be a uniformizer at ∞ and $\pi_P = (xb - ay)/by = t - a/b$. We can see that these lie in $L(\mathbb{P}^1(L))$, being the quotient of homogeneous polynomials of the same degree in x and y. Let $\omega = dt$, which is a nonzero differential. $d\pi_P = d(t - a/b) = dt - d(a/b) = \omega$. Also,

$$0 = d(1) = d\left(t \cdot \frac{1}{t}\right) = td\left(\frac{1}{t}\right) + \frac{1}{t}dt = td\pi_\infty + \frac{1}{t} \cdot \omega,$$

whence $\omega = -t^2 d\pi_\infty$. This shows that $\mathrm{ord}_0(\omega) = \mathrm{ord}_P(\omega) = 0$, as $\omega/dt = \omega/d\pi_P = 1$ and $\mathrm{ord}_\infty(\omega) = -2$, as $\omega/d\pi_\infty = -t^2 = -\pi_\infty^{-2}$. This implies that a canonical divisor

for $\mathbb{P}^1(L)$ is $-2[\infty]$, whose degree is -2. We know that the degree of a canonical divisor is $2g-2$, so we may conclude that $g=0$.

For the converse, let C be a curve of genus 0 with canonical divisor K_C and let P and Q be two distinct points of C. Consider the divisor $D=[P]-[Q]$ and apply Riemann-Roch to it. Note that $\deg(K_C-D)=\deg(K_C)-\deg(D)=-2<0$, so $\ell(K_C-D)=0$ and we obtain $\ell(D)=\ell(D)-\ell(K_C-D)=\deg(D)-0+1=1$. Let $f\in\mathcal{L}(D)$. We have that $\sum_{R\in C} n_R[R]=\operatorname{div}(f)\geq[Q]-[P]$, so $n_Q\geq 1$, $n_P\geq -1$ and $n_R\geq 0$ for all other $R\in C$. Adding all these inequalities we find that

$$0=n_Q+n_P+{\sum_{R\in C}}' n_R\geq 1+(-1)+0=0,$$

where the primed sum ignores the points P and Q. In order to satisfy this, we must have equality at every point, implying that $\operatorname{div}(f)=D$.

This means that $D=0$ in $\operatorname{Pic}(C)$. If we let D' be any divisor of degree 0, we can write it as a finite sum of divisors of the form $[P]-[Q]$ (for different points P and Q), which makes $D'=0$ in $\operatorname{Pic}(C)$ as well. This implies that every degree 0 divisor is principal, making $\operatorname{Pic}(C)\cong\mathbb{Z}$, which in turn implies that $C\cong\mathbb{P}^1(L)$. □

Theorem 2.4. *Assume the characteristic of K is not 2. Every smooth ternary quadratic form over K is isomorphic over L to $\mathbb{P}^1(\mathbb{Q})$.*

Proof. Consider a homogeneous smooth ternary quadratic form Q in $\mathbb{P}^2(K)$,

$$Q(x,y,z)=ax^2+by^2+cz^2+dxy+eyz+fzx=\begin{pmatrix}x & y & z\end{pmatrix}\begin{pmatrix}a & d/2 & f/2\\ d/2 & b & e/2\\ f/2 & e/2 & c\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}.$$

If $d = e = f = 0$ we say that Q is *diagonal.* Denote by

$$A = \begin{pmatrix} a & d/2 & f/2 \\ d/2 & b & e/2 \\ f/2 & e/2 & c \end{pmatrix}.$$

Since the curve is smooth, we find that the system

$$\begin{cases} 2ax + dy + fz = 0 \\ dx + 2by + ez = 0 \\ fx + ey + 2cz = 0 \end{cases} \qquad \text{or,} \qquad 2A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

does not have nontrivial solutions. This clearly amounts to the matrix A being non-singular.

A linear change of variables can be represented by an invertible 3×3 matrix P, with coefficients in K, multiplying the vector on the left:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = P \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \qquad \text{or, equivalently,} \qquad P^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

This way, the form Q_P defined by

$$Q_P(X, Y, Z) = \begin{pmatrix} X & Y & Z \end{pmatrix} P^T A P \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} x & y & z \end{pmatrix} A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = Q(x, y, z)$$

is isomorphic to Q.

If $a \neq 0$, we can complete squares to obtain a quadratic form with no cross terms involving x. This can be obtained via the matrix

$$P = \begin{pmatrix} 1 & -d/2a & -f/2a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \text{ since } P^T AP = \begin{pmatrix} a & 0 & 0 \\ 0 & b - d^2/4a & e/2 - df/4a \\ 0 & e/2 - df/4a & c - f^2/4a \end{pmatrix},$$

so

$$Q_P(X, Y, Z) = aX^2 + (b - d^2/4a)Y^2 + (c - f^2/4a)Z^2 + (e - df/2a)YZ.$$

Note that Q_P is the sum of a unary quadratic form and a binary quadratic form, which is closer to diagonal. If $b \neq 0$ or $c \neq 0$, we can carry out a similar process to get rid of cross terms involving y or z, so, if at least one amongst a, b and c is different from 0, we can eliminate cross terms of one of the variables and express the quadratic form as the sum of a unary quadratic form and a binary quadratic form.

If $a = b = c = 0$, $\det A = def/4$, so $d \neq 0$ for $\det A$ to be nonzero. We set

$$P = \begin{pmatrix} 1 & 1 & -(e+f)/2d \\ 1 & -1 & -(e+f)/2d \\ 0 & 0 & 1 \end{pmatrix}, \text{ whence } P^T AP = \begin{pmatrix} d & 0 & 0 \\ 0 & -d & (f-e)/2 \\ 0 & (f-e)/2 & -(f+e)^2/4d \end{pmatrix}$$

and we obtain a form Q_P isomorphic to Q which is the sum of a unary quadratic form and a binary quadratic form.

So far, we have shown that Q is isomorphic to a quadratic form with matrix

$$\begin{pmatrix} h & 0 & 0 \\ 0 & k & l/2 \\ 0 & l/2 & m \end{pmatrix},$$

so assume henceforth that $Q(x,y,z) = hx^2 + ky^2 + mz^2 + lyz$. If $k \neq 0$, we can diagonalize in the same way as before via the matrix

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -l/2k \\ 0 & 0 & 1 \end{pmatrix}, \text{ since } P^T AP = \begin{pmatrix} h & 0 & 0 \\ 0 & k & 0 \\ 0 & 0 & m - l^2/4k \end{pmatrix},$$

obtaining the diagonal quadratic form $Q_P(X,Y,Z) = hX^2 + kY^2 + (m - l^2/4k)Z^2$. If $m \neq 0$ we can still do this prioritizing the third variable. Hence, the only problem may arise when $k = m = 0$, in which case $\det A = -hl^2/4$, so $l \neq 0$. In this situation, we set

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \text{whence} \quad P^T AP = \begin{pmatrix} h & 0 & 0 \\ 0 & l & 0 \\ 0 & 0 & -l \end{pmatrix},$$

obtaining again a diagonal quadratic form. All the previous isomorphisms are over K. Without loss of generality, we may assume that $Q(x,y,z) = ax^2 + by^2 + cz^2$, where $abc \neq 0$. If we want to consider also isomorphisms over L, $Q(x,y,z) = (\sqrt{a}x)^2 + (\sqrt{b}y)^2 - (\sqrt{-c}z)^2$, so all quadratic forms are isomorphic to $x^2 + y^2 - z^2$. Consider the curve C in $\mathbb{P}^2(L)$ defined by $x^2 + y^2 - z^2$ (which is smooth, as the rank of the

matrix $\begin{pmatrix} 2x & 2y & -2z \end{pmatrix}$ is 1). Let

$$\begin{aligned} \phi\colon \mathbb{P}^1(L) &\longrightarrow C \\ (m,n) &\longmapsto (m^2-n^2, 2mn, m^2+n^2), \end{aligned}$$

and

$$\begin{aligned} \psi\colon C &\longrightarrow \mathbb{P}^1(L) \\ (x,y,z) &\longmapsto (x+z,y) = (y, z-x). \end{aligned}$$

Notice that $(m^2-n^2)^2+(2mn)^2-(m^2+n^2)^2 = 0$, and if $m^2-n^2 = m^2+n^2 = 0$ then $m = n = 0$ so ϕ is well defined. If $x+z = y = 0$ then we have the point $(1,0,-1)$, in which case $z-x \neq 0$. Also, $(x+z,y) = (y(x+z), y^2) = (y(x+z), z^2-x^2) = (y, z-x)$, so ψ is also well defined. Finally it is easy to check that $\psi\circ\phi = \mathrm{id}_{\mathbb{P}^1(L)}$ and $\phi\circ\psi = \mathrm{id}_C$, as

$$\begin{aligned} (m,n) &\mapsto (m^2-n^2, 2mn, m^2+n^2) \mapsto (2m^2, 2mn) = (m,n) \\ (m,n) &\mapsto (m^2-n^2, 2mn, m^2+n^2) \mapsto (2mn, 2n^2) = (m,n) \end{aligned}$$

and, using $y^2 = z^2 - x^2$ on C,

$$\begin{aligned} (x,y,z) \mapsto (x+z,y) &\mapsto ((x+z)^2-(z^2-x^2), 2(x+z)y, (x+z)^2+(z^2-x^2)) \\ &\quad (2x(x+z), 2y(x+z), 2z(x+z)) = (x,y,z) \\ (x,y,z) \mapsto (y,z-x) &\mapsto ((z^2-x^2)-(z-x)^2, 2y(z-x), (z^2-x^2)+(z-x)^2) \\ &\quad (2x(z-x), 2y(z-x), 2z(z-x)) = (x,y,z), \end{aligned}$$

so the maps are inverses of each other, making both curves isomorphic. □

This implies that all smooth ternary quadratic forms have genus 0, and these are the only curves over K with this property.

2.2.2 Binary Quadratic Forms

This section follows [Cox13]. A binary quadratic form is a homogeneous polynomial of degree two in two variables. We will be interested in binary quadratic forms of the form

$$F(x, y) = ax^2 + bxy + cy^2, \qquad a, b, c \in \mathbb{Z}, \qquad \gcd(a, b, c) = 1,$$

which are referred to as *integral primitive binary quadratic forms*. To it, we associate the quantity $D = b^2 - 4ac$, which we call the *discriminant* of F. Note that the discriminant is the additive inverse of the determinant of the matrix $\left(\begin{smallmatrix} 2a & b \\ b & 2c \end{smallmatrix}\right)$ and it is always congruent to 1 or 0 modulo 4. It may happen that $D = dk^2$ where d is another discriminant. If the largest such k is 1, we say that D is *fundamental*.

An integer n is said to be *represented* by F if there exist integers x, y such that $F(x, y) = n$. If $\gcd(x, y) = 1$, we say that n is *properly represented* by F. If F represents n, we have that

$$(2ax + by)^2 - Dy^2 = 4an.$$

If $D < 0$ we can see that $4an \geq 0$, so F only represents integers of the same sign. Because of this, if $D < 0$ we say that F is *definite*. In addition, if $a > 0$ we say that F is *positive definite* and if $a < 0$ we say that F is *negative definite*. Studying negative definite quadratic forms is equivalent to studying positive definite quadratic forms

so in the case of definite binary quadratic forms we will restrict ourselves to positive definite binary quadratic forms. If $D > 0$, F represents integers of both signs, so we say that F is *indefinite.*

If the matrix $\alpha = \left(\begin{smallmatrix} t & u \\ v & w \end{smallmatrix}\right) \in \mathbf{GL}_2(\mathbb{Z})$, then left multiplication by α provides a bijection from $\mathbb{Z}^2$ to itself (viewed as column vectors), which descends to the subset of pairs of integers which are relatively prime, for if $\gcd(x, y) = 1$ we have integers A and B such that $Ax + By = 1$, so

$$\begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 1 \Rightarrow \begin{pmatrix} A & B \end{pmatrix} \alpha^{-1} \alpha \begin{pmatrix} x \\ y \end{pmatrix} = 1,$$

whence, denoting by $(A'\ B')$ the matrix $(A\ B)\alpha^{-1}$ we find that $A'x' + B'y' = 1$, where the pair x', y' is the image of the pair x, y by α.

Define an action of $\mathbf{GL}_2(\mathbb{Z})$ on the set of binary quadratic forms by

$$\begin{aligned} (F \cdot \alpha)(x, y) &= F((x, y)\alpha^T) = F(tx + uy, vx + wy) \\ &= a(tx + uy)^2 + b(tx + uy)(vx + wy) + c(vx + wy)^2 \\ &= F(t, v)x^2 + (2atu + btw + buv + 2cvw)xy + F(u, w)y^2. \end{aligned}$$

It is a right action because we can write

$$F(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

so if α, $\beta \in \mathbf{GL}_2(\mathbb{Z})$ we have

$$\begin{aligned}(F \cdot (\alpha\beta))(x,y) &= \begin{pmatrix} x & y \end{pmatrix} (\alpha\beta)^T \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \alpha\beta \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} x & y \end{pmatrix} \beta^T\alpha^T \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} \alpha\beta \begin{pmatrix} x \\ y \end{pmatrix} \\ &= ((F \cdot \alpha) \cdot \beta)(x,y).\end{aligned}$$

By the preceding comments, the binary quadratic forms F and $F \cdot \alpha$ (properly) represent the same set of integers. We say that F and $F \cdot \alpha$ are *equivalent*. If, furthermore, $\alpha \in \mathbf{SL}_2(\mathbb{Z})$, we say that F and $F \cdot \alpha$ are *properly equivalent*. The orbits of the action of $\mathbf{SL}_2(\mathbb{Z})$ are referred to as *classes*, as the action induces an equivalence relation in the set of binary quadratic forms. Note that the discriminant of two equivalent forms is the same (as the determinant will be multiplied by the determinants of α and α^T, both equal to 1), so each class is comprised of forms with the same discriminant. For a fixed D we may partition the set of binary quadratic forms of discriminant D into the classes mentioned above. This set is denoted $\mathrm{Cl}(D)$ and it can be endowed with a group structure called the *composition law*, due to Gauss. See [Cox13] for details.

A positive binary quadratic form $ax^2 + bxy + cy^2$ is said to be *reduced* if

$$|b| \leq a \leq c, \quad \text{and} \quad b \geq 0 \quad \text{if either} \quad |b| = a \quad \text{or} \quad a = c.$$

Then, every positive definite binary quadratic form is properly equivalent to a unique reduced form, yielding canonical representatives for the classes of $\mathrm{Cl}(D)$ when $D < 0$.

Furthermore, if a positive binary quadratic form is reduced, we have that

$$-D = 4ac - b^2 \geq 4a^2 - b^2 \geq 3a^2,$$

whence $a \leq \sqrt{\frac{-D}{3}}$, implying that there are only finitely many positive definite binary quadratic forms with a fixed discriminant, as this bounds the value of b as well, and c is the quotient of $b^2 - D$ and $4a$. (In particular, we just need $4a$ to divide $b^2 - D$ at this point.)

For example, for $D = -164$, $a \leq \sqrt{164/3} < 8$. Since $b^2 - 4ac = -164$, b is an even integer whose absolute value is bounded by a. If $a = 1$, b can only be 0, making $c = 41$ and yielding the form $x^2 + 41y^2$. If $a = 2$, b can only be 0 or 2. In the former case, $4a$ does not divide $b^2 - D$ and and in the latter we obtain $c = 21$, giving only the form $2x^2 + 2xy + 21y^2$. For $a = 3$, b can be -2, 0 and 2. If $b = \pm 2$, $c = 14$ and if $b = 0$, c would not be an integer. We obtain the forms $3x^2 \pm 2xy + 14y^2$. For $a = 4$, the possible values of b are -2, 0, 2 and 4, but none of these make $b^2 + 164$ divisible by 16, so there are no forms with $a = 4$. $a = 5$ has ± 4, ± 2 and 0 as possible values for b. The only ones that work are $b = \pm 4$, where $c = 9$. This gives us the forms $5a^2 \pm 4xy + 9y^2$. If $a = 6$, the possible values of b are ± 4, ± 2, 0 and 6. For $b = \pm 2$ we obtain $c = 7$ and the other values do not give any integral values of c, yielding the forms $6x^2 \pm 2xy + 7y^2$. Finally, if $a = 7$, b can be ± 6, ± 4, ± 2 and 0. For $b = \pm 2$, we would get $c = 6$, which is smaller than 7. For the other values, the divisibility relation is not satisfied, so there are no reduced positive definite binary quadratic

forms in this case. Summarizing, we have that

$$\begin{aligned}\mathrm{Cl}(-164) = \{&[x^2+41y^2], [2x^2+2xy+21y^2], [3x^2+2xy+14y^2], [3x^2-2xy+14y^2] \\ &[5x^2+4xy+9y^2], [5x^2-4xy+9y^2], [6x^2+2xy+7y^2], [6x^2-2xy+7y^2]\}\end{aligned}$$

The case of indefinite forms is a little bit more subtle. The remaining of this subsection follows [Fla89]. A form $F(x,y) = ax^2+bxy+cy^2$ of discriminant $D > 0$ is said to be reduced if

$$0 < b < \sqrt{D} \qquad \text{and} \qquad \sqrt{D} - b < |2a| < \sqrt{D} + b.$$

Every indefinite binary quadratic form is properly equivalent to a reduced form, but it is not unique. However, all the reduced forms equivalent to a given form are of a very particular kind. Define the *right neighbor* RF of F to be the form $cx^2+Bxy+Cy^2$ of the same discriminant as F such that B is the integer in the interval $(\sqrt{D}-|2c|, \sqrt{D})$ satisfying $b+B$ being divisible by $2c$. From $2c \mid b+B$ we can see that b and B have the same parity, so $2 \mid b-B$, further implying that $4c \mid b^2 - B^2 = D + 4ac - B^2$. Whence $4c \mid B^2 - D$ and C is just the corresponding quotient, an integer, which shows that the right neighbor is also an integral binary quadratic form. Moreover, if $b+B = 2c\delta$, $RF = F \cdot \left(\begin{smallmatrix} 0 & -1 \\ 1 & \delta \end{smallmatrix}\right)$, so RF and F are properly equivalent. We have the following theorem:

Theorem 2.5. *Let F and G be two indefinite integral binary quadratic forms of discriminant D. The sequences defined by*

$$\{R^n F\}_{n\geq 0} \qquad \textit{and} \qquad \{R^n G\}_{n\geq 0}$$

are periodic. Furthermore, all the forms appearing in the periodic part of the sequences are reduced and F and G are properly equivalent if and only if the sequences overlap.

This gives us an algorithm to compute representatives for $\mathrm{Cl}(D)$, albeit not canonical. The bounds imposed on the definition of being reduced imply that b and a have a finite range of values, and the discriminant gives c also a finite range of values, implying that $\mathrm{Cl}(D)$ is finite.

For example, for $D = 145$, $0 < b < \sqrt{145} < 13$, and since b and D have the same parity, b can only take on the values 1, 3, 5, 7, 9 and 11. We summarize the possible values a can take on in each case in the following table:

b	$\lfloor\sqrt{D} - b\rfloor$	$\lfloor\sqrt{D} + b\rfloor$	possible a's
1	11	13	± 6
3	9	15	± 5, ± 6, ± 7
5	7	17	± 4, ± 5, ± 6, ± 7, ± 8
7	5	19	± 3, ± 4, ± 5, ± 6, ± 7, ± 8, ± 9
9	3	21	± 2, ± 3, ± 4, ± 5, ± 6, ± 7, ± 8, ± 9, ± 10
11	1	23	± 1, ± 2, ± 3, ± 4, ± 5, ± 6, ± 7, ± 8, ± 9, ± 10, ± 11

The only possible combinations such that $2a$ divides $b^2 - 145$ are recorded in the following table:

b	a's	$b^2 - D$	respective c's
1	± 6	-144	∓ 6
3	$-$	-136	$-$
5	± 5, ± 6	-120	∓ 6, ∓ 5
7	± 3, ± 4, ± 6, ± 8	-96	∓ 8, ∓ 6, ∓ 4, ∓ 3
9	± 2, ± 4, ± 8	-64	∓ 8, ∓ 4, ∓ 2
11	± 1, ± 2, ± 3, ± 6	-24	∓ 6, ∓ 3, ∓ 2, ∓ 1

This gives us 28 possible forms. After finding the right neighbor of all of them, we can find the different orbits and we just need to take one element per orbit. We end up with

$$\mathrm{Cl}(145) = \{[6x^2 + xy - 6y^2], [5x^2 + 5xy - 6y^2], [3x^2 + 7xy - 8y^2], [8x^2 + 7xy - 3y^2]\}.$$

2.3 Elliptic Curves

After dealing with curves of genus 0, the next natural example is curves of genus 1. This theory already poses more problems than the previous case, as there are no known algorithms that can guarantee the existence of K-rational points on the curve or find all of the points provided that one already has one point. However, there are algorithms that often work (see [ST92]). In this section we follow mainly [Sil86].

A genus 1 curve together with a point is referred to as an *Elliptic Curve*. Elliptic curves, as in the case of genus 0 curves, always have planar models. Let us start our brief discussion about these by establishing the nature of their equations. In order to do this we will use again the Riemann-Roch theorem. Denote the special point (the

one that is part of the definition) by $\mathcal{O}$ and keep the notations from the previous sections.

Consider the spaces $\mathcal{L}(n[\mathcal{O}])$ for $n \in \{1, 2, 3, 4, 5, 6\}$. Since $\deg(n[\mathcal{O}]) = n \geq 0 = 2 \cdot 1 - 2$, Corollary 2.2(c) applies and $\ell(n[\mathcal{O}]) = n - 1 + 1 = n$. The function $1 \in \mathcal{L}([\mathcal{O}])$, as $\operatorname{div}(1) = 0 \geq -[\mathcal{O}]$, and since $\ell([\mathcal{O}]) = 1$, we find that 1 is a basis for $\mathcal{L}([\mathcal{O}])$. Since $\ell(2[\mathcal{O}]) = 2$ and $\ell(3[\mathcal{O}]) = 3$, let $x \in \mathcal{L}(2[\mathcal{O}])$ such that $x \notin \mathcal{L}([\mathcal{O}])$ and $y \in \mathcal{L}(3[\mathcal{O}])$ such that $x \notin \mathcal{L}(2[\mathcal{O}])$. Note that x and y have poles of order exactly 2 and 3 at $\mathcal{O}$, respectively, and $\{1, x\}$ and $\{1, x, y\}$ are bases for $\mathcal{L}(2[\mathcal{O}])$ and $\mathcal{L}(3[\mathcal{O}])$, respectively. The functions x^2 and xy have poles of order exactly 4 and 5, respectively, and the dimensions given by Corollary 2.2(c) give that $\{1, x, y, x^2\}$ and $\{1, x, y, x^2, xy\}$ are bases for $\mathcal{L}(4[\mathcal{O}])$ and $\mathcal{L}(5[\mathcal{O}])$, respectively. Finally, the elements x^3 and y^2 both lie in $\mathcal{L}(6[\mathcal{O}])$, which has dimension 6, making the set $\{1, x, y, x^2, xy, x^3, y^2\}$ linearly dependent, whence, there are elements $l_i \in K$, $i \in \{1, 2, 3, 4, 5, 6, 7\}$, such that $l_6 l_7 \neq 0$ and

$$l_1 + l_2 x + l_3 y + l_4 x^2 + l_5 xy + l_6 x^3 + l_7 y^2 = 0.$$

After multiplying by $l_6^2 l_7^3$ we obtain

$$l_1 l_6^2 l_7^3 + l_2 l_6 l_7^2 (l_6 l_7 x) + l_3 l_6 l_7 (l_6 l_7^2 y) + l_4 l_7 (l_6 l_7 x)^2 + l_5 (l_6 l_7 x)(l_6 l_7^2 y) + (l_6 l_7 x)^3 + (l_6 l_7^2 y)^2 = 0,$$

so, after rescaling we can obtain an equation of the form

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6, \tag{2.1}$$

where $a_1, a_2, a_3, a_4, a_6 \in K$. Thus, we can construct an isomorphism from C into a subset of $\mathbb{P}^2(K)$ via

$$\begin{aligned} C &\longrightarrow \qquad \mathbb{P}^2(K) \\ P &\longmapsto (x(P), y(P), 1), \end{aligned}$$

where the point $\mathcal{O}$ is mapped to $\left(\frac{x}{y}(\mathcal{O}), 1(\mathcal{O}), \frac{1}{y}(\mathcal{O})\right) = (0, 1, 0)$. Equation (2.1) is called the *Weierstrass equation* of the curve C. Often, we refer to the pair $(E, \mathcal{O})$ just by E; in this case we have a Weierstrass equation in mind in which $\mathcal{O}$ is the point at infinity.

When dealing with fields of characteristic different from 2, we can multiply by 4 and complete the square on the left side to obtain

$$(2y + a_1x + a_3)^2 = 4x^3 + (4a_2 + a_1^2)x^2 + 2(2a_4 + a_1a_3)x + (4a_6 + a_3^2),$$

so the change of variables

$$\begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = P_1 \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}, \qquad \text{where } P_1 = \begin{pmatrix} 1 & 0 & 0 \\ a_1 & 2 & a_3 \\ 0 & 0 & 1 \end{pmatrix},$$

yields the equation

$$Y^2 = 4X^3 + b_2X^2 + 2b_4X + b_6,$$

where

$$b_2 = 4a_2 + a_1^2, \qquad b_4 = 2a_4 + a_1a_3 \qquad \text{and} \qquad b_6 = 4a_6 + a_3^2.$$

Furthermore, if the characteristic is different from 3, we can multiply by $2^4 3^6$ to obtain

$$\begin{aligned}(2^2 3^3 Y)^2 &= (2^2 3^2 X)^3 + 3(2^2 3^2 X)^2 (3b_2) + 2^5 3^6 b_4 X + 2^4 3^6 b_6 \\ &= \big((2^2 3^2 X)^3 + 3(2^2 3^2 X)^2 (3b_2) + 3(2^2 3^2 X)(3b_2)^2 + (3b_2)^3\big) \\ &\quad + 2^5 3^6 b_4 X + 2^4 3^6 b_6 - \big(3(2^2 3^2 X)(3b_2)^2 + (3b_2)^3\big) \\ &= (2^2 3^3 X + 3b_2)^3 + (2^2 3^2 X + 3b_2)(2^3 3^4 b_4 - 3^3 b_2^2) \\ &\quad + \big(2^4 3^6 b_6 - 3^3 b_2^3 - 3b_2(2^3 3^4 b_4 - 3^3 b_2^2)\big) \\ &= (36X + 3b_2)^3 + 27(24b_4 - b_2^2)(36X + 3b_2) + 54(216b_6 - 36b_2 b_4 + b_2^3),\end{aligned}$$

so the change of variables

$$\begin{pmatrix} \mathcal{X} \\ \mathcal{Y} \\ 1 \end{pmatrix} = P_2 \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}, \qquad \text{where } P_2 = \begin{pmatrix} 36 & 0 & 3b_2 \\ 0 & 108 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

yields the equation

$$\mathcal{Y}^2 = \mathcal{X}^3 - 27c_4\mathcal{X} - 54c_6,$$

where

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

For simplicity of notation, write $\mathcal{X} = x$ and $\mathcal{Y} = y$ and let us work with the equation

$$y^2 = x^3 - 27c_4 x - 54c_6, \tag{2.2}$$

which is referred to as the *short Weierstrass equation* of the curve C.

Note that all these transformations preserve the field where the coefficients lie and the point at infinity $(0, 1, 0)$. After homogeneizing we obtain that C can be modeled by the zero set of the polynomial

$$y^2z + 27c_4xz^2 + 54c_6z^3 - x^3,$$

whose matrix of partial derivatives is

$$\begin{pmatrix} 27c_4z^2 - 3x^2 & 2yz & y^2 + 54c_4xz + 162c_6z^2 \end{pmatrix}.$$

For this matrix to not have rank 1, we require $2yz = 0$ in particular. If $z = 0$, we have the point at infinity $\mathcal{O}$, so $x = 0$ and $y = 1$, yielding the matrix $\begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$, with rank 1. (In particular, the point at infinity is smooth.) If $y = 0$, we have that $z = 1$ and the simultaneous equations

$$x^2 = 9c_4, \qquad c_4x = -3c_6.$$

Squaring the latter, multiplying by c_4^2 the former and equating yields $9c_4^3 = 9c_6^2$. Notice that the discriminant of the cubic polynomial (in x) defining the elliptic curve is $\Delta' = 2^2 3^9 (c_4^3 - c_6^2)$, so the discriminant of this polynomial must be nonzero.

If we start with a short Weirstrass equation already, namely, a curve E where $a_1 = a_2 = a_3 = 0$, there is no need to apply the transformations we applied in order to obtain a short Weierstrass model. However, after applying all these operations, we obtain the equation

$$y^2 = x^3 + 6^4a_4x + 6^6a_6.$$

The discriminant of this cubic polynomial is $2^{12}3^{12}$ times the discriminant of the cubic polynomial associated to the original equation. It makes sense that in order to define the discriminant of the elliptic curve we take this into account, so we define the discriminant of E as

$$\Delta = 16 \cdot \frac{-4(-27c_4)^3 - 27(-54c_6)^2}{2^{12}3^{12}} = \frac{c_4^3 - c_6^2}{1728},$$

where the extra 16 is introduced in order to preserve integrality of the discriminant when all the a_i are in a specific ring. If the a_i are all integers, it can be shown that $c_4^3 - c_6^2$ is always divisible by 1728 and that this is the largest possible integer with this property. (It just suffices to expand the quotient in terms of the a_i to see that 1728 divides, and plugging in the tuples $(0, -1, -1, 0, 0)$ and $(1, 0, -1, 0, 0)$ gives the values -11 and -28 respectively, which do not share any common factors.) In the case of starting with the short Weierstrass equation $y^2 = x^3 + ax + b$, $\Delta = -16(4a^2 + 27b^3)$.

Note that even if we start with coefficients in a number ring, we can reduce modulo its primes, resulting in a curve in the quotient ring. The value of the discriminant will be reduced modulo the same prime, and if it is nonzero, we will obtain an elliptic curve as well and the curve is said to have *good reduction* modulo this prime. If the discriminant reduces to 0, the curve is said to have *bad reduction* modulo this prime.

We define the *j-invariant* of the curve as

$$j = \frac{1728c_4^3}{c_4^3 - c_6^2} = \frac{c_4^3}{\Delta}.$$

The j-invariant does not depend on the model chosen for the curve and since $\Delta \neq 0$, this value is always in K. It classifies elliptic curves up to L-isomorphism, as we will

explain next. Also, its importance will become even more prominent when we talk about modular parametrizations. For simplitiy of the argument, let us consider only the case where $\mathrm{char}(K) \neq 2, 3$. Let E and E' be two elliptic curves with the same j-invariant and let c_4, c_6 and c_4', c_6' be the corresponding values obtained after finding a short Weierstrass model for E and E', respectively. This is,

$$E\colon y^2 = x^3 - 27c_4x - 54c_6 \quad \text{and} \quad E'\colon y^2 = x^3 - 27c_4x - 54c_6.$$

If $j = 0$, $c_4 = c_4' = 0$. Note that if $c_4c_4' = 0$, $j = 0$ as well. Since $\Delta\Delta' \neq 0$, $c_6c_6' \neq 0$. Let u be a sixth root of the quotient c_6/c_6'. After multiplying the equation for E by u^6 we obtain

$$(u^3y)^2 = (u^2x)^3 - 54u^6c_6 = (u^2x)^3 - 54c_6',$$

so the map $E \longrightarrow E'$, $(x, y) \longmapsto (u^2x, u^3y)$ is an isomorphism. If $j = 1728$, $c_4 = c_4^3 - c_6^2$, implying $c_6 = 0$. Likewise for E', so $c_6' = 0$ as well. Note that if $c_6c_6' = 0$, $j = 1728$. Let u be a fourth root of the quotient c_6/c_4'. After multiplying the equation for E by u^6 we obtain

$$(u^3y)^2 = (u^2x)^3 - 27u^6c_4x = (u^2x)^3 - 27c_4'(u^2x),$$

so the map $E \longrightarrow E'$, $(x, y) \longmapsto (u^2x, u^3y)$ is an isomorphism again. If $j \neq 0, 1728$, after equating the equations for the j-invariants, cross-multiplying and canceling, we obtain the equation

$$c_4^3c_6'^2 = c_4'^3c_6^2, \text{ or, since } c_4c_4'c_6c_6' \neq 0, \left(\frac{c_4'}{c_4}\right)^3 = \left(\frac{c_6'}{c_6}\right)^2.$$

Let $u' \in L$ be a fourth root of c_4'/c_4. Its twelfth power is $u'^{12} = \left(\frac{c_4'}{c_4}\right)^3 = \left(\frac{c_6'}{c_6}\right)^2$, so $u'^6 = \pm c_6'/c_6$. Let $u = u'$ if this sign is positive and $-u'$ if it is negative. We will still have $u^4 = c_4'/c_4$ but now we also have that $u^6 = c_6'/c_6$. The same isomorphism as before will do the job.

We define the *invariant differential* associated to the curve as

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

The functions $y^2 + a_1xy + a_3y$ and $x^3 + a_2x^2 + a_4x + a_6$ are equal in $K(C)$, so after applying the formal symbol d we obtain the same differential. It follows that

$$\begin{aligned} d(y^2 + a_1xy + a_3y) &= d(x^3 + a_2x^2 + a_4x + a_6) \\ 2ydy + a_1xdy + a_1ydx + a_3dy &= 3x^2dx + 2a_2xdx + a_4dx \\ \implies (2y + a_1x + a_3)dy &= (3x^2 + 2a_2x + a_4 - a_1y)dx \\ \implies \frac{dy}{3x^2 + 2a_2x + a_4 - a_1y} &= \frac{dx}{2y + a_1x + a_3}, \end{aligned}$$

showing that both differentials given in the definition of ω are the same. The curve is given by the polynomial $f = y^2z + a_1xyz + a_3yz^2 - x^3 - a_2x^2z - a_4xz^2 - a_6z^3$, so the matrix $\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix}$, where

$$\begin{aligned} \frac{\partial f}{\partial x} &= -(3x^2 + 2a_2xz + a_4z^2 - a_1yz) \\ \frac{\partial f}{\partial y} &= 2yz + a_1xz + a_3z^2 \\ \frac{\partial f}{\partial z} &= y^2 + a_1xy + 2a_3yz - a_2x^2 - 2a_4xz - 3a_6z^2, \end{aligned}$$

has rank 1. Notice that

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = 3f,$$

so, if $3x^2 + 2a_2xz + a_4z^2 - a_1yz = 2yz + a_1xz + a_3z^2 = 0$, since $f = 0$, it follows that $z\dfrac{\partial f}{\partial z} = 0$, so for points such that $z \neq 0$ it would follow that the matrix has rank 0, yielding a contradiction. This means that the two polynomials in the definition of ω cannot vanish simultaneously at finite points.

Take a point $P = (x_0, y_0) \in C$ where $\dfrac{\partial f}{\partial y} \neq 0$, namely, $2y_0 + a_1x_0 + a_3 \neq 0$. We can rewrite the polynomial $x^3 + a_2x^2 + a_4x + a_6$ as $(x - x_0)^3 + A_2(x - x_0)^2 + A_4(x - x_0) + A_6$ for some A_2, A_4, $A_6 \in L$. Also, $y^2 + a_1xy + a_3y = y^2 + a_1(x - x_0)y + (a_3 + a_1x_0)y$, so putting both together and evaluating at P we find that $y_0^2 + (a_3 + a_1x_0)y_0 = A_6$. This means that

$$y^2 + (a_3 + a_1x_0)y - A_6 = (y - y_0)(y + y_0 + a_1x_0 + a_3) \in \langle x - x_0 \rangle \text{ in } L[C].$$

The polynomial $y + y_0 + a_1x_0 + a_3$ does not vanish at P, which means that $y - y_0 \in \langle x - x_0 \rangle$ in the localization of $L[C]$ away from m_P, making $x - x_0$ a uniformizer at P, implying that $\mathrm{ord}_P(\omega) = 0$.

If $\dfrac{\partial f}{\partial y} = 0$ at P, we have that $\dfrac{\partial f}{\partial x} \neq 0$.

$$\begin{aligned} y^2 + a_1xy + a_3y &= (y^2 - 2yy_0 + y_0^2) + a_1x(y - y_0) + a_3y_0 + 2yy_0 - y_0^2 + a_1xy_0 \\ &= (y - y_0)^2 + a_1x(y - y_0) + a_3y_0 + 2y_0(y - y_0) + y_0^2 + a_1xy_0 \\ &= x^3 + a_2x^2 + a_4x + a_6, \end{aligned}$$

so

$$x^3 + a_2x^2 + a_4x + a_6 - a_1xy_0 - (y_0^2 + a_3y_0) =$$
$$x^3 + a_2x^2 + a_4x + a_6 - a_1xy_0 - (x_0^3 + a_2x_0^2 + a_4x_0 + a_6 - a_1x_0y_0) =$$
$$(x^3 - x_0^3) + a_2(x^2 - x_0^2) + a_4(x - x_0) - a_1y_0(x - x_0) =$$
$$(x - x_0)(x^2 + xx_0 + x_0^2 + a_2(x + x_0) + a_4 - a_1y_0).$$

Putting both equations together we find that

$$(y - y_0)^2 + (a_1x + 2y_0)(y - y_0) = (x - x_0)(x^2 + xx_0 + x_0^2 + a_2(x + x_0) + a_4 - a_1y_0),$$

which lies in the ideal $\langle y - y_0 \rangle$ in $L[C]$. The polynomial $x^2 + xx_0 + x_0^2 + a_2(x + x_0) + a_4 - a_1y_0$ does not vanish at P so $x - x_0 \in \langle y - y_0 \rangle$ in the localization of $L[C]$ away from m_P, making $y - y_0$ a uniformizer at P, implying again that $\mathrm{ord}_P(\omega) = 0$.

Since the curve has genus 1, Corollary 2.2(b) states that the degree of a canonical divisor is 0, implying that $\mathrm{ord}_{\mathcal{O}}(\omega) = 0$ as well, so the differential ω is holomorphic and nonvanishing. (It is also easy to take $\pi_{\mathcal{O}} = x/y$ as a uniformizer, compute ω as $gd\pi_{\mathcal{O}}$ for some $g \in K(C)$ and compute $\mathrm{ord}_{\mathcal{O}}(g)$.)

2.3.1 Group Structure

When we were dealing with a curve of genus 0, C, we showed that for two different points, say P and Q, there exists an element in $L(C)^\times$ such that $\mathrm{div}(f) = [P] - [Q]$. In big contrast with this idea, when we have a curve of genus 1, $[P]$ cannot even be linearly equivalent to $[Q]$. We record this in the following lemma.

Lemma 2.6. *Let $(E, \mathcal{O})$ be an elliptic curve and let $P, Q \in E$. Then*

$$[P] \sim [Q] \text{ in } \mathrm{Pic}(E) \qquad \Longleftrightarrow \qquad P = Q.$$

Proof. One direction is trivial so we will focus on the other one. Let $f \in L(E)^\times$ such that $\mathrm{div}(f) = [P] - [Q]$. Since $[P] - [Q] \geq -[Q]$, this implies that $f \in \mathcal{L}([Q])$. By 2.2(c), $\ell([Q]) = \deg([Q]) = 1$, so f actually generates $\mathcal{L}([Q])$. Since constant functions are in $\mathcal{L}([Q])$, this shows that f is constant, $\mathrm{div}(f) = 0$ and $[P] = [Q]$ in $\mathrm{Div}(E)$, so $P = Q$. □

Proposition 2.7. *Let $(E, \mathcal{O})$ be an elliptic curve. For every $D \in \mathrm{Div}^0(E)$ there exists a unique point $P \in E$ such that $D \sim [P] - [\mathcal{O}]$ in $\mathrm{Pic}^0(E)$. The map from $\mathrm{Div}^0(E) \longrightarrow E$ that maps D to $[P]$ as described above induces a bijection between the sets $\mathrm{Pic}^0(E)$ and E.*

Proof. Consider the divisor $D + [\mathcal{O}]$, of degree 1. By 2.2(c), $\ell(D + [\mathcal{O}]) = 1$, so, let $0 \neq f \in \mathcal{L}(D + [\mathcal{O}])$. This implies that $\mathrm{div}(f) \geq -D - [\mathcal{O}]$, but $\deg(\mathrm{div}(f)) = 0$ and $\deg(D - [\mathcal{O}]) = -1$, so there must exist a point $P \in E$ such that $\mathrm{div}(f) = -D - [\mathcal{O}] + [P]$, whence $D \sim [P] - [\mathcal{O}]$ as claimed. Uniqueness follows from Lemma 2.6, as $[P] - [\mathcal{O}] \sim [P'] - [\mathcal{O}]$ implies $[P] \sim [P']$. This maps descends to $\mathrm{Pic}^0(E)$ because if $D_1 \sim D_2$, $D_1 \mapsto P_1$ and $D_2 \mapsto P_2$, we have that

$$[P_1] - [\mathcal{O}] \sim D_1 \sim D_2 \sim [P_2] - [\mathcal{O}],$$

so $[P_1] \sim [P_2]$ and this implies $P_1 = P_2$, where it is also clear that the map is injective in this quotient. Surjectivity follows from mapping $[P] - [\mathcal{O}] \mapsto P$ for every $P \in E$. □

Proposition 2.7 allows us to assign a group structure to E, as it is in bijection with an abelian group. The identity of $\mathrm{Pic}^0(E)$ is $0 = [\mathcal{O}] - [\mathcal{O}]$, so $\mathcal{O}$ is the identity of E. If P and Q are two points in E, the bijection assigns to the point $P + Q$ the point corresponding to the degree-zero divisor $[P] + [Q] - 2[\mathcal{O}]$. Recall that $\mathcal{O}$ is a K-rational point, so, if we start with a divisor D defined over M, where $K \subseteq M \subseteq L$ is an intermediate field, the divisor $D + [\mathcal{O}]$ is also defined over M. Let P be the point that corresponds to D and $f \in L(E)^\times$ such that

$$\mathrm{div}(f) = -D - [\mathcal{O}] + [P].$$

Let $\sigma \in \mathrm{Gal}(L/M)$ and apply σ to this equality to obtain

$$\mathrm{div}(\sigma(f)) = \sigma(\mathrm{div}(f)) = -\sigma(D) - \sigma([\mathcal{O}]) + \sigma([P]) = -D - [\mathcal{O}] + [\sigma(P)].$$

Since $\sigma(f) \in L(E)^\times$, this implies that $D \sim [\sigma(P)] - [\mathcal{O}]$, and by Lemma 2.6 we find that $P = \sigma(P)$. This implies that $P \in E(M)$, so starting with an element of $\mathrm{Div}^0_M(E)$ produces a point defined over M. From this, it is clear that if $P, Q \in E(M)$, $P + Q \in E(M)$ as well, as the divisor $[P] + [Q] - 2[\mathcal{O}] \in \mathrm{Div}^0_M(E)$. Likewise, the divisor $-[P] + [\mathcal{O}]$ is defined over M, so its corresponding point, call it $-P$, is defined over M as well and $P + (-P)$ is the point corresponding to

$$[P] + [-P] - 2[\mathcal{O}] = ([P] - [\mathcal{O}]) + ([-P] - [\mathcal{O}]) \sim ([P] - [\mathcal{O}]) + (-[P] + [\mathcal{O}]) = 0,$$

whence $P + (-P) = \mathcal{O}$. We conclude that $-P \in E(M)$ and that $-P$ is the inverse of P, making $E(M)$ a subgroup of E (and thus, giving it a group structure as well.) If K is a number field, the group $E(K)$ always has at least one element (namely, the

point $\mathcal{O}$). It may consist of a finite or an infinite number of elements. Nevertheless, we have the following structure theorem:

Theorem 2.8 (Mordell)**.** *Let K be a number field and let $(E, \mathcal{O})$ be an elliptic curve defined over K. The abelian group $E(K)$ is finitely generated. This is, there exist a finite group T and and a nonnegative integer r such that*

$$E(K) \cong T \times \mathbb{Z}^r.$$

The integer r of Mordell's theorem is called the *algebraic rank* of E over K, and T is the *torsion* subgroup of E.

2.3.2 Isogenies

Let E_1 and E_2 be two elliptic curves defined over K. An element $f \in K(E_2)$ can be seen as a map $f\colon E_2 \longrightarrow \mathbb{P}^1(L)$. A map $\phi\colon E_1 \longrightarrow E_2$ induces a pull-back map $\phi^*\colon K(E_2) \longrightarrow K(E_1)$ defined by precomposing by ϕ, i.e., $\phi^* f = f \circ \phi$. The following diagram illustrates the definition.

$$\begin{array}{ccc} E_1 & \xrightarrow{\phi} & E_2 \\ & \searrow^{\phi^* f} & \downarrow f \\ & & \mathbb{P}^1(L) \end{array}$$

If the map is a morphism of algebraic varieties and a homomorphism of groups, we say that ϕ is an *isogeny* from E_1 to E_2 and that E_1 and E_2 are *isogenous*. Clearly $\phi^* 1 = 1$, $\phi^*(f+g) = \phi^* f + \phi^* g$ and $\phi^*(f \cdot g) = \phi^* f \cdot \phi^* g$, so ϕ^* is a field homomorphism, and for $\phi \neq 0$, $K(E_1)$ is a finite extension of $\phi^*(K(E_2))$ (Theorem II.2.4, [Sil86]). We define the degree of ϕ to be 0 if $\phi = 0$ or if $\phi \neq 0$, the index $[L(E_1) : \phi^* L(E_2)]$ (where L is again an algebraic closure of K) and we denote it by $\deg \phi$. If we have

$\phi\colon E_1 \longrightarrow E_2$ and $\psi\colon E_2 \longrightarrow E_3$, we get the compositions

$$E_1 \xrightarrow{\phi} E_2 \xrightarrow{\psi} E_3$$

$$L(E_1) \xleftarrow[\phi^*]{} L(E_2) \xleftarrow[\psi^*]{} L(E_3),$$

so $L(E_1)$ is an algebraic field extension of $\phi^* \circ \psi^* L(E_3)$ and

$$\begin{aligned}[L(E_1) : (\phi^* \circ \psi^*)(L(E_3))] &= [L(E_1) : \phi^*(L(E_2))] \cdot [\phi^*(L(E_2)) : (\phi^* \circ \psi^*)(L(E_3))] \\ &= [L(E_1) : \phi^*(L(E_2))] \cdot [L(E_2) : \psi^*(L(E_3))],\end{aligned}$$

where the last equality follows from the fact that a field is isomorphic to its image under any field homomorphism. We can see that $\deg(\psi \circ \phi) = \deg \phi \cdot \deg \psi$.

An isogeny is either constant or surjective (Theorem II.2.3, [Sil86]), so for any $Q \in E_2$ there is at least one $P \in E_1$ with $\phi(P) = Q$. Let π_P and π_Q be uniformizers at P and Q, respectively. Theorems II.2.6 and III.4.10 in [Sil86] can be adapted to our needs as follows: If ϕ is a separable map ($K(E_1)$ is a separable extension of $\phi^*(K(E_2))$, which is always the case when $\mathrm{char}(K) = 0$), then the ramification index $e_\phi(P) = \mathrm{ord}_P(\phi^* \pi_Q) = 1$ for all $P \in \phi^{-1}(Q)$. For every $Q \in E_2$ we have

$$\sum_{P \in \phi^{-1}(Q)} e_\phi(P) = \#\phi^{-1}(Q) = \deg \phi.$$

This means that every point has the same number of preimages, and that this number is also equal to the degree of the isogeny. In particular, $\deg \phi = \# \ker \phi$.

If the source and the target of the isogeny are the same elliptic curve E, we can make two isogenies interact with each other to obtain new isogenies. Denote

by $\mathrm{End}(E)$ the set of isogenies from E to itself, together with the zero map. For $\phi, \psi \in \mathrm{End}(E)$, we define $(\phi + \psi)(P) = \phi(P) + \psi(P)$. If $\varphi \in \mathrm{End}(E)$, we have that

$$\begin{aligned}(\varphi \circ (\phi + \psi))(P) &= \varphi((\phi + \psi)(P)) = \varphi(\phi(P) + \psi(P)) \\ &= \varphi(\phi(P)) + \varphi(\psi(P)) = (\varphi \circ \phi)(P) + (\varphi \circ \psi)(P),\end{aligned}$$

so $\varphi \circ (\phi + \psi) = \varphi \circ \phi + \varphi \circ \psi$. Likewise, we have $(\phi + \psi) \circ \varphi = \phi \circ \varphi + \psi \circ \varphi$. Denote by $[0]\colon E \longrightarrow E$ and $[1]\colon E \longrightarrow E$ the maps defined by $[0](P) = \mathcal{O}$ and $[1](P) = P$ for all $P \in E$. They are clearly isogenies and the sum and composition endow $\mathrm{End}(P)$ with a ring structure.

For a positive integer m we have a map

$$[m]\colon E \longrightarrow E \text{ defined as } P \longmapsto \underbrace{P + \cdots + P}_{m \text{ times}} = mP$$

called the *multiplication by* m map. Its kernel is denoted by $E[m]$ and we refer to it as the *m-torsion* of E. We also have the map $[-1]\colon E \longrightarrow E$ which simply assigns the inverse of P to P. $[-1]$ is a homomorphism, as E is abelian. If m is a negative integer, we define $[m] := [-m] \circ [-1] = [-1] \circ [-m]$. With these definitions, we can easily see that $[m] + [n] = [m+n]$ and $[m \cdot n] = [m] \circ [n]$ for all $m, n \in \mathbb{Z}$. This defines a ring homomorphism from $\mathbb{Z} \longrightarrow \mathrm{End}(E)$. Proposition III.4.2 in [Sil86] shows that this is an injective homomorphism and that $\mathrm{End}(E)$ has no zero divisors.

Let us get back to the case where we have an isogeny $\phi\colon E_1 \longrightarrow E_2$. As mentioned above, this induces a map $\phi^*\colon K(E_2) \longrightarrow K(E_1)$, which can be extended to a map $\phi^*\colon L(E_2) \longrightarrow L(E_1)$. This isogeny also defines a pull-back at the level of divisor

groups, which by abuse of notation we also denote

$$\phi^*\colon \operatorname{Div}(E_2) \longrightarrow \operatorname{Div}(E_1) \text{ where } \phi^*[Q] = \sum_{P \in \phi^{-1}(Q)} e_\phi(P)[P].$$

The definition of the ramification index implies that $\phi^*\pi_Q$ lies in the $e_\phi(P)$-th power of the maximal ideal m_P, but not in its $(e_\phi(P)+1)$-st power. Since the powers of m_P are principal, there exists $t \in L(E_1)^\times_{m_P}$ such that $\phi^*\pi_Q = \pi_P^{e_\phi(P)} \cdot t$, as π_P is any generator of m_P. Let $f \in K(E_2)$ and $m = \operatorname{ord}_Q(f)$, so $f = \pi_Q^m \cdot g$, where $g \in K(E_2)^\times_{m_Q}$. Then, we have

$$\phi^* f = \phi^*(\pi_Q^m \cdot g) = (\phi^*\pi_Q)^m \cdot \phi^* g = \pi_P^{e_\phi(P)\cdot m} \cdot t^m \cdot \phi^* g.$$

Since g is defined and does not vanish at Q, $\phi^* g$ is defined and does not vanish at P. This implies that $\operatorname{ord}_P(\phi^* f) = e_\phi(P)\operatorname{ord}_Q(f)$. It follows that

$$\begin{aligned}
\operatorname{div}(\phi^* f) &= \sum_{P \in E_1} \operatorname{ord}_P(\phi^* f)[P] = \sum_{P \in E_1} e_\phi(P)\operatorname{ord}_Q(f)[P] \\
&= \sum_{Q \in E_2} \sum_{P \in \phi^{-1}(Q)} e_\phi(P)\operatorname{ord}_Q(f)[P] = \sum_{Q \in E_2} \operatorname{ord}_Q(f) \sum_{P \in \phi^{-1}(Q)} e_\phi(P)[P] \\
&= \sum_{Q \in E_2} \operatorname{ord}_Q(f)\phi^*[Q] = \phi^*\left(\sum_{Q \in E_2} \operatorname{ord}_Q(f)[Q]\right) = \phi^*(\operatorname{div} f),
\end{aligned}$$

so ϕ^* maps principal divisors to principal divisors. Moreover,

$$\deg \phi^*[Q] = \deg\left(\sum_{P \in \phi^{-1}(Q)} e_\phi(P)[P]\right) = \sum_{P \in \phi^{-1}(Q)} e_\phi(P) = \deg\phi = \deg\phi \deg[Q],$$

so $\deg(\phi^* D) = \deg\phi \deg D$ and ϕ^* maps degree zero divisors to degree zero divisors. Putting both things together, we find that the isogeny ϕ induces a homomorphism

$\phi^*\colon \mathrm{Pic}^0(E_2) \longrightarrow \mathrm{Pic}^0(E_1)$. Using this homomorphism, we construct $\hat{\phi}$ via the following commutative diagram:

$$\begin{array}{ccc} E_2 & \xrightarrow{\sim} & \mathrm{Pic}^0(E_2) \\ \vdots & & \downarrow{\scriptstyle \phi^*} \\ E_1 & \xleftarrow{\sim} & \mathrm{Pic}^0(E_1) \end{array}$$

This map can be realized as follows. Let $Q \in E_2$. The top isomorphism maps Q to the class of $[Q] - [\mathcal{O}]$, which is mapped via ϕ^*, in the unramified case we are dealing with, to

$$\sum_{P \in \phi^{-1}(Q)} [P] - \sum_{T \in \phi^{-1}(\mathcal{O})} [T] = \sum_{P \in \phi^{-1}(Q)} ([P] - [\mathcal{O}]) - \sum_{T \in \phi^{-1}(\mathcal{O})} ([T] - [\mathcal{O}]),$$

whose class is mapped to

$$\sum_{P \in \phi^{-1}(Q)} P - \sum_{T \in \phi^{-1}(\mathcal{O})} T.$$

Fix a point $P \in \phi^{-1}(Q)$. For any other point in $\phi^{-1}(Q)$, say P_i, we have that $\phi(P_i - P) = \phi(P_i) - \phi(P) = Q - Q = \mathcal{O}$, so there exists $T_i \in \phi^{-1}(\mathcal{O})$ such that $P_i - P = T_i$. This shows that

$$\sum_{P \in \phi^{-1}(Q)} P - \sum_{T \in \phi^{-1}(\mathcal{O})} T = \sum_{i=1}^{\deg \phi} (P_i - T_i) = [\deg \phi] P.$$

In particular, we find that $(\hat{\phi} \circ \phi)(P) = \hat{\phi}(Q) = [\deg \phi]P$. It can be shown that $\hat{\phi}$ is an isogeny as well, that $\deg \hat{\phi} = \deg \phi$ and that $\hat{\hat{\phi}} = \phi$. The isogeny $\hat{\phi}$ is referred to as the *dual isogeny* to ϕ.

From above, it follows that $\phi \circ \hat{\phi} = [\deg \phi]$ on E_1 and $\hat{\phi} \circ \phi = [\deg \phi]$ on E_2. Theorem III.6.2 in [Sil86] shows that if $\phi, \psi\colon E_1 \longrightarrow E_2$ are two isogenies, $\widehat{\phi + \psi} =$

$\hat{\phi}+\hat{\psi}$. Since $[\hat{1}] = [1]$, it follows that $\widehat{[m+1]} = \widehat{[m]}+[1]$, so if we have that $\widehat{[m]} = [m]$, we find that $\widehat{[m+1]} = [m+1]$. Now, $[\deg[m]] = \widehat{[m]} \circ [m] = [m] \circ [m] = [m^2]$, so $\deg[m] = m^2$, as $\mathbb{Z}$ embeds in $\mathrm{End}(E_1)$. This last fact allows us to show a structure theorem for the torsion points, which is Theorem III.6.4 in [Sil86].

Theorem 2.9. *Let $(E, \mathcal{O})$ be an elliptic curve defined over a field K, p be a prime number and r a positive integer. We have that*

$$E[p^r] = E[p^r](\bar{K}) \cong \begin{cases} \mathbb{Z}/p^r\mathbb{Z} \times \mathbb{Z}/p^r\mathbb{Z} & \text{if } \mathrm{char}(K) \neq p \\ \mathbb{Z}/p^r\mathbb{Z} \text{ or } 0 & \text{if } \mathrm{char}(K) = p \end{cases}$$

When $\mathrm{char}(K) = p$, we can see that there are two options. We say that E is *ordinary* in the former case and *supersingular* in the latter. (The term supersingular does not imply that E is singular, despite the name.)

In an abelian group, when two torison elements have relatively prime orders the order of their product is the product of their orders, i.e., if $\mathrm{ord}\, x = h$ and $\mathrm{ord}\, y = k$ with $\gcd(h, k) = 1$, we have that $\mathrm{ord}(xy) = hk$. This follows from trivial computations. In order to understand the structure of $E[m]$, we just need to understand the structure of $E[p^r]$ for each p^r dividing m. In particular, if $\mathrm{char}(K) \nmid m$ or $\mathrm{char}(K) \neq 0$, we have that $E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$.

2.3.3 *L*-functions

The Riemann ζ-function is defined on the right half-plane $\Re(z) > 1$ by the Dirichlet series

$$\zeta(s) = \sum_{n=1}^{\infty} n^{-s} = \prod_{p \in \mathcal{P}} \left(1 - p^{-s}\right)^{-1},$$

where $\mathcal{P}$ is the set of rational primes. The last equality follows from the Fundamental Theorem of Arithmetic and the absolute convergence of the infinite series, and it is called the Euler product of $\zeta(s)$. Each factor is called the *local factor* at p. More generally, for a number field K, the Dedeking ζ-function is defined on the same right half-plane as the Riemann ζ-function by the Dirichlet series

$$\zeta(K,s) = \sum_{n=1}^{\infty} a_n n^{-s} = \sum_{I \subseteq \mathcal{O}_K} N_{K/\mathbb{Q}}(I)^{-s} = \prod_{\mathfrak{p} \subseteq \mathcal{O}_K} \left(1 - N_{K/\mathbb{Q}}(\mathfrak{p})^{-s}\right)^{-1},$$

where a_n is the number of ideals of norm n in $\mathcal{O}_K$, the ring of algebraic integers of K, and the last product runs over prime ideals in $\mathcal{O}_K$. The last equality follows from the fact that every ideal can be uniquely decomposed as a product of prime ideals, plus the absolute convergence of the series in the convergence half-plane. These functions have a simple pole at $s = 1$, can be extended holomorphically to the rest of the complex plane and have a functional equation which gives the function some sort of symmetry with respect to the vertical line $s = 1/2$.

For a Dirichlet character χ we can also define a Dirichlet L-function

$$L(\chi, s) = \sum_{n=1}^{\infty} \chi(n) n^{-s} = \prod_{p \in \mathcal{P}} (1 - \chi(p) p^{-s})^{-1},$$

and these also have functional equations and can be analytically continued to the whole plane, with a possible pole at 1 and/or 0.

It turns out that it is possible to attach to the elliptic curve E defined over $\mathbb{Q}$ a function with very similar properties. In order to define these, for a prime number p let us define the curve E_p as the curve obtained by reducing a *minimal Weierstrass model* of E modulo p. When $p \nmid \Delta$, this results in an elliptic curve over $\mathbb{F}_p$, and

since $\mathbb{P}^2(\mathbb{F}_p)$ is a finite set, we can count the number of elements in $E_p(\mathbb{F}_p)$, N_p. Let $a_p = p + 1 - N_p$, referred to as the *trace of Frobenius at p*. Define the local factor at p to be $L_p(s) = 1 - a_p p^{-s} + p^{1-2s}$. When $p \mid \Delta$, we define the local factor at p to be $1 + p^{-s}$, $1 - p^{-s}$ or 1, depending on the type of bad reduction at the prime p. We define the Hasse-Weil function as

$$L(E, s) = \prod_{p \in \mathcal{P}} L_p(s)^{-1} = \sum_{n=1}^{\infty} a_n n^{-s},$$

where the a_p are the traces of the Frobenii for prime numers p, follow specific recurrence relations for prime powers, and are multiplicative. We have the *Hasse* bound, which states that $|a_p| \leq 2\sqrt{p}$. Thanks to it, these expressions are known to converge for $\Re(s) > 3/2$, and they can be extended holomorphically to the whole complex plane (see Section 3.4 and Section 3.5). They have similar functional equations to the ζ-functions discussed above and share some symmetry with respect to the vertical line $s = 1$. The value of $L(E, s)$ at $s = 1$ is of utmost importance. The order of the zero at $s = 1$ is called the *analytic rank* of the elliptic curve E.

Conjecture 2.10 (Birch-Swinnerton-Dyer)**.** *Let E be an elliptic curve over $\mathbb{Q}$. Its algebraic rank and its analytic rank coincide.*

Chapter 3
Modular Forms

In this chapter we will introduce modular forms from a classical point of view, following mainly [DS05]. Denote by $\mathcal{H}$ the complex upper half-plane throughout. For a positive integer N we define the *principal congruence subgroup* of *level* N to be

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

Let $\Gamma(N) \subseteq \Gamma \subseteq \mathbf{SL}_2(\mathbb{Z})$ be another group. We say that Γ is a *congruence subgroup* of *level* N. The two most important examples of level N congruence subgroups are

$$\Gamma_0(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} * & * \\ 0 & * \end{pmatrix} \pmod{N} \right\}$$

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}) : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{N} \right\}.$$

Let $f\colon \mathcal{H} \longrightarrow \mathbb{C}$ be a function and k be an integer. There is a left action of $\mathbf{GL}_2(\mathbb{Q})^+$ on $\mathcal{H}$ (actually on $\mathbb{C} \cup \{\infty\}$, the Riemann Sphere) via fractional linear

transformations, i.e.,

$$\alpha(\tau) = \frac{a\tau + b}{c\tau + d}, \text{ where } \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Define the *weight k slash operator*, $|_k$, as

$$f|_k[\alpha](\tau) = \det(\alpha)^{k-1}(c\tau + d)^{-k} f\left(\frac{a\tau + b}{c\tau + d}\right), \text{ where } \alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

which gives a right action of $\mathbf{GL}_2(\mathbb{Q})^+$ on the set of functions with domain $\mathcal{H}$ and codomain $\mathbb{C}$. We say that f is *weakly modular of weight k with respect to* Γ if f is meromorphic and

$$f|_k[\alpha] = f \text{ for all } \alpha \in \Gamma.$$

Since $\Gamma \supseteq \Gamma(N)$, we know that the matrix $\left(\begin{smallmatrix} 1 & N \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$. Let h be the smallest positive integer such that $\left(\begin{smallmatrix} 1 & h \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$. For a weakly modular form of any weight with respect to Γ, we have that $f(\tau) = f(\tau + h)$, so f is periodic with period h. The map $\tau \mapsto e^{2\pi i\tau/h} = q_h$ wraps $\mathcal{H}$ into the punctured unit disk of radius 1 and is also periodic of period h. This allows us to see f as a map with the punctured disk as its domain and the meromorphicity gives us a Laurent expansion, so

$$f(\tau) = \sum_{n \in \mathbb{Z}} a_n q_h^n,$$

which is referred to as the *q-expansion* of f at the *cusp* ∞. We say that f is *holomorphic at* ∞ if $a_n = 0$ for all $n < 0$. Furthermore, we say that f *vanishes at* ∞

if $a_0 = 0$. For a rational number r/s, the matrix $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ acts by

$$\alpha(r/s) = \frac{a \cdot r/s + b}{c \cdot r/s + d} = \frac{ar + bs}{cr + ds} \quad \text{and} \quad \alpha(\infty) = \frac{a}{c},$$

where every time a fraction has denominator 0 we consider it to be ∞. Restricted to $\mathbf{SL}_2(\mathbb{Z})$ this action is transitive on the set $\mathbb{Q} \cup \{\infty\}$, as for every reduced fraction a/c, we can find $b, c \in \mathbb{Z}$ such that $ad - bc = 1$ and

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} (\infty) = a/c \qquad \text{and} \qquad \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} (a/c) = \infty,$$

so for any two rational numbers r_1/s_1 and r_2/s_2 we can find matrices such that $r_1/s_1 \mapsto \infty \mapsto r_2/s_2$. From this, we can see that the action of Γ partitions the set $\mathbb{Q} \cup \{\infty\}$ into equivalence classes (the number of such bounded by the index $[\mathbf{SL}_2(\mathbb{Z}) : \Gamma]$). For a rational number r/s, let $\alpha \in \mathbf{SL}_2(\mathbb{Z})$ such that $\infty \mapsto r/s$. If f is weakly modular with respect to Γ, the function $f|_k[\alpha]$ is weakly modular with respect to $\alpha^{-1}\Gamma\alpha$, and if the latter is holomorphic (resp. vanishes) at ∞ we say that f is holomorphic (resp. vanishes) at r/s. This is independent of the matrix and representative chosen, as f is invariant under the action of Γ.

A *Modular Form of weight k with respect to* Γ is a holomorphic function that is weakly modular of weight k with respect to Γ and that is holomorphic at the cusps. A modular form is said to be *cuspidal* if it vanishes at every cusp. We also call these just *cusp forms*. The set of modular forms of weight k with respect to Γ forms a (finite dimensional) complex vector space, which we denote by $M_k(\Gamma)$. The cusp forms form a subspace, denoted by $S_k(\Gamma)$. If we drop the weight index, we obtain a

graded algebra, as the product of two modular forms is again a modular form, the weight being the sum of the two weights.

The first examples of modular forms come from the so-called Eisenstein series. For $k \geq 4$, we let

$$G_k(\tau) = \sum_{(m,n)\in\mathbb{Z}^2}{}'(m\tau + n)^{-k},$$

where the primed summation means we exclude the term corresponding to $(m, n) = (0, 0)$. Holomorphicity follows from convergence theorems. Weak modularity comes from the fact that every element of $\mathbf{SL}_2(\mathbb{Z})$ induces a bijection on $\mathbb{Z}^2 - (0, 0)$ (and convergence theorems as well), and the exponent we have determines the weight k. There is only one cusp, and holomorphicity at ∞ follows from the fact that $G_k(\tau)$ is uniformly bounded near ∞. Some manipulations show that

$$G_k(t) = 2\zeta(k) + 2\frac{(2\pi i)^k}{(k-1)!}\sum_{n=1}^{\infty}\sigma_{k-1}(n)q^n,$$

where $\sigma_{k-1}(n)$ denotes the sum of the $(k-1)$-st powers of the positive divisors of n. If we let $g_2(\tau) = 60G_4(\tau)$ and $g_3(\tau) = 140G_6(\tau)$, we obtain that $g_2(\tau)^3 - 27g_3(\tau)^2$ is a cusp form, as each term is a weight 12 modular form, and the leading terms are $60^3 \cdot 2^3 \cdot \zeta(4)^3$ and $27 \cdot 140^2 \cdot 2^2 \cdot \zeta(6)^2$, both equal to $(2\pi)^{12}/1728$. This cusp form is called the *discriminant function*, it is denoted $\Delta(\tau)$ and its q-expansion admits the representation

$$\Delta(\tau) = (2\pi)^{12}q\prod_{n=1}^{\infty}(1-q^n)^{24} = (2\pi)^{12}(q - 24q^2 + 252q^3 - 1472q^4 + \mathcal{O}(q^5)),$$

nowhere vanishing on $\mathcal{H}$.

Finally, the function

$$j\colon \mathcal{H} \longrightarrow \mathbb{C}, \quad j(\tau) = 1728\frac{g_2(\tau)^3}{\Delta(\tau)} = \frac{1}{q} + 744 + 196884q + 21493760q^2 + \mathcal{O}(q^3)$$

is weakly modular of weight zero, and it is holomorphic on $\mathcal{H}$. However, there is a simple pole at the cusp, with residue 1. j is referred to as the *modular function*, often called the *j-invariant* as well.

Consider $\chi\colon \mathbb{Z} \longrightarrow \mathbb{C}$ a Dirichlet character modulo N. Since $\Gamma_1(N) \subseteq \Gamma_0(N)$ we have that $M_k(\Gamma_0(N)) \subseteq M_k(\Gamma_1(N))$. Let $f \in M_k(\Gamma_1(N))$. Slashing f by an element $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ of $\Gamma_0(N)$ will most likely not return f. Consider the subspace $M_k(N, \chi) \subset M_k(\Gamma_1(N))$ of modular forms such that $f|_k[\alpha] = \chi(d)f$ (notice that since $ad - bc = 1$ and $N \mid c$, $\gcd(d, N) = 1$) for every $\alpha \in \Gamma_0(N)$. We can verify that

$$M_k(\Gamma_1(N)) = \bigoplus_{\chi \in \widehat{(\mathbb{Z}/N\mathbb{Z})^\times}} M_k(N, \chi),$$

which gives us a decomposition of $M_k(\Gamma_1(N))$ into eigenspaces. In particular, the eigenspace of the trivial character is $M_k(\Gamma_0(N))$.

We also have interaction between different levels. Since $\Gamma_0(Nd) \subseteq \Gamma_0(N)$ for all positive integers d, we have that $M_k(\Gamma_0(Nd)) \supseteq M_k(\Gamma_0(N))$, the same containment holding for cusp forms. Furthermore, if $f \in M_k(\Gamma_0(N))$, $f|_k[\left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)] \in M_k(\Gamma_0(Nd))$, as

$$\begin{aligned}(f|_k[\left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)])\,|_k[\left(\begin{smallmatrix} a & b \\ cNd & \delta \end{smallmatrix}\right)] &= f|_k[\left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} a & b \\ cNd & \delta \end{smallmatrix}\right)] = f|_k[\left(\begin{smallmatrix} a & bd \\ cN & \delta \end{smallmatrix}\right)\left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)] \\ &= (f|_k[\left(\begin{smallmatrix} a & bd \\ cN & \delta \end{smallmatrix}\right)])\,|_k[\left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)] = f|_k[\left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)]\end{aligned}$$

when $\left(\begin{smallmatrix} a & b \\ cNd & \delta \end{smallmatrix}\right) \in \Gamma_0(Nd)$ (which implies $\left(\begin{smallmatrix} a & bd \\ cN & \delta \end{smallmatrix}\right) \in \Gamma_0(N)$). The computation extends to cusp forms, and even to spaces of weakly modular forms in which the holomorphicity conditions are relaxed to just being meromorphic (called Automorphic forms instead).

These remarks prompt the following definition:

Definition 3.1. Let $N > 1$ be an integer and let $S_{k,M,d} = \{f|_k[\left(\begin{smallmatrix} d & 0 \\ 0 & 1 \end{smallmatrix}\right)] | f \in S_k(\Gamma_0(M))\}$. Define

$$S_k(\Gamma_0(N))^{\text{old}} = \langle S_{k,M,d} : M \text{ proper divisor of } N, d \text{ a divisor of } N/M \rangle \subseteq S_k(\Gamma_0(N))$$

as the subspace of $S_k(\Gamma_0(N))$ that comes from cusp forms of lower levels. We refer to its elements as *oldforms*.

Analogously, we can define $M_k(\Gamma_0(N))^{\text{old}}$ and even put a character modulo N there (which also implies we can define $S_k(\Gamma_1(N))^{\text{old}}$ and $M_k(\Gamma_1(N))^{\text{old}}$).

3.1 Elliptic Curves arising from Modular Forms

Let Λ be a lattice in $\mathbb{C}$, with generators ω_1 and ω_2. We may assume that $\omega_1/\omega_2 \in \mathcal{H}$ without loss of generality.

Consider the Weierstrass $\wp_\Lambda$ function attached to the lattice Λ, defined as

$$\begin{aligned} \wp_\Lambda \colon \mathbb{C} - \Lambda &\longrightarrow \mathbb{C} \\ z &\longmapsto \frac{1}{z^2} + \sum_{\omega \in \Lambda}{}' \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right), \end{aligned}$$

where the primed summation denotes addition over all nonzero elements of the lattice. This even function is not defined at any point $\omega \in \Lambda$ because of the vanishing denominators. However, it converges absolutely and uniformly on compact subsets

away from the lattice, yielding a meromorphic function with poles only at the elements of the lattice. Its derivative,

$$\begin{aligned} \wp'_\Lambda \colon \mathbb{C} - \Lambda &\longrightarrow \mathbb{C} \\ z &\longmapsto -\sum_{\omega \in \Lambda} \frac{2}{(z-\omega)^3} \end{aligned}$$

is clearly Λ-periodic (thanks to the absolute convergence), which means that $\wp_\Lambda(z+\omega_j) - \wp_\Lambda(z)$ is constant for $j = 1, 2$. Plugging in $z = -\omega_j/2$ and using the fact $\wp_\Lambda$ is even, we can see this constant is 0, and hence $\wp_\Lambda$ is also Λ-periodic. All the terms of the primed summation are holomorphic on a small neighborhood around 0 (as the pole at 0 comes from the first term), so $\wp_\Lambda$ has a double pole at the origin, and hence, at every point of the lattice.

The Λ-periodicity makes $\wp_\Lambda$ and $\wp'_\Lambda$ descend to functions from $\mathbb{C}/\Lambda$ to $\hat{\mathbb{C}}$, with a double pole and a triple pole at the class of 0, respectively. This is reminiscent of the functions x and y introduced in section 2.3, which had a pole of order two and three, respectively, at $\mathcal{O}$. We can actually find their Laurent expansion, but we need a little bit more of notation for this.

For $k \geq 4$ an integer, define the Eisenstein function of weight k as

$$G_k(\Lambda) = \sum_{\omega \in \Lambda}{}' \omega^{-k},$$

where the primed summation means the same as above. This function is absolutely convergent so we can rearrange terms. When we specialize to the lattice $\Lambda_\tau = \tau\mathbb{Z} + \mathbb{Z}$, we obtain the $G_k(\tau)$ we defined in the previous section, so they generalize the modular forms we had before. Since those had special transformation properties, we can

expect these to have special transformation properties as well, although here, they will be more transparent. Consider a nonzero complex number λ and the lattice $c\Lambda$. Then

$$G_k(\lambda\Lambda) = \sum_{\omega\in\lambda\Lambda}{}'\omega^{-k} = \sum_{\omega\in\Lambda}{}'(\lambda\omega)^{-k} = \lambda^{-k}\sum_{\omega\in\Lambda}{}'\omega^{-k} = \lambda^{-k}G_k(\Lambda),$$

Every matrix in $\mathbf{SL}_2(\mathbb{Z})$ acts on Λ in a simple transitive way, so each matrix in $\mathbf{SL}_2(\mathbb{Z})$ is only reorganizing the terms of the sum. If we let $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbb{Z})$, $\tau \in \mathcal{H}$ and $\lambda = (c\tau + d)^{-1}$, keeping in mind that $\alpha\Lambda = \Lambda$, we obtain

$$G_k(\alpha(\tau)) = G_k(\Lambda_{\alpha(\tau)}) = G_k(\lambda(\alpha\Lambda_\tau)) = \lambda^{-k}G_k(\Lambda_\tau) = (c\tau + d)^k G_k(\tau),$$

which is the crucial weak modularity property.

Let us go back to $\wp_\Lambda$ and focus on $1/(z-\omega)^2$. We have

$$\frac{1}{(z-\omega)^2} = \frac{-\omega^2}{(1-(z/\omega))^2} = \omega^{-2}\sum_{n=0}^{\infty}(n+1)\left(\frac{z}{\omega}\right)^n = \omega^{-2} + \sum_{n=1}^{\infty}(n+1)\omega^{-(n+2)}z^n,$$

where we used the identity $1/(1-x)^2 = \sum_{n=0}^{\infty}(n+1)x^n$ from calculus. This implies that

$$\begin{aligned}\wp_\Lambda(z) &= \frac{1}{z^2} + \sum_{\omega\in\Lambda}{}'\sum_{n=1}^{\infty}(n+1)\omega^{-(n+2)}z^n = \frac{1}{z^2} + \sum_{n=1}^{\infty}\sum_{\omega\in\Lambda}{}'(n+1)\omega^{-(n+2)}z^n \\ &= \frac{1}{z^2} + \sum_{n=1}^{\infty}(n+1)\sum_{\omega\in\Lambda}{}'\omega^{-(n+2)}z^n = \frac{1}{z^2} + \sum_{n=1}^{\infty}(n+1)G_{n+2}(\Lambda)z^n.\end{aligned}$$

Notice that $G_k(\Lambda) = 0$ for odd values of k due to cancellation, so we actually have the Laurent expansions

$$\wp_\Lambda(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty}(2n+1)G_{2n+2}(\Lambda)z^{2n} = \frac{1}{z^2} + 3G_4(\Lambda)z^2 + 5G_6(\Lambda)z^4 + \mathcal{O}(z^6)$$
$$\wp'_\Lambda(z) = \frac{-2}{z^3} + \sum_{n=1}^{\infty}2n(2n+1)G_{2n+2}(\Lambda)z^{2n-1} = \frac{-2}{z^3} + 6G_4(\Lambda)z + 20G_6(\Lambda)z^3 + \mathcal{O}(z^5).$$

Also, squaring $\wp'_\Lambda(z)$ and cubing $\wp_\Lambda(z)$ gives a pole of order 6 at $z = 0$, yielding

$$\wp'_\Lambda(z)^2 = \frac{4}{z^6} - 24G_4(\Lambda)z^{-2} - 80G_6(\Lambda) + \mathcal{O}(z)$$
$$4\wp_\Lambda(z)^3 = \frac{4}{z^6} + 36G_4(\Lambda)z^{-2} + 60G_6(\Lambda) + \mathcal{O}(z),$$

so

$$\wp'_\Lambda(z)^2 - 4\wp_\Lambda(z)^3 = -60G_4(\Lambda)z^{-2} - 140G_6(\Lambda) + \mathcal{O}(z),$$

and hence, we find that

$$\wp'_\Lambda(z)^2 - 4\wp_\Lambda(z)^3 + 60G_4(\Lambda)\wp_\Lambda(z) + 140G_6(\Lambda) = \mathcal{O}(z).$$

If we restrict it to a fundamental parallelogram, the left hand side is bounded on its closure (it being compact), so the Λ-periodicity implies the left hand side is bounded on $\mathbb{C}$. The right hand side indicates the function is holomorphic, hence, Liouville's Theorem implies that it is a constant. As $z \to 0$, the right hand side tends to 0 as well, implying said constant is 0. We conclude that

$$\wp'_\Lambda(z) = 4\wp_\Lambda(z)^3 - 60G_4(\Lambda)\wp_\Lambda(z) - 140G_6(\Lambda),$$

or, if we adopt the terminology from the previous section,

$$\wp'_\Lambda(z) = 4\wp_\Lambda(z)^3 - g_2(\Lambda)\wp_\Lambda(z) - g_3(\Lambda).$$

The map that goes from $\mathbb{C} \longrightarrow \mathbb{P}^2(\mathbb{C})$ defined as

$$z \longmapsto (\wp_\Lambda(z), \wp'_\Lambda(z), 1) = \left(\frac{\wp_\Lambda(z)}{\wp'_\Lambda(z)}, 1, \frac{1}{\wp'_\Lambda(z)}\right)$$

can be factored through the quotient $\mathbb{C}/\Lambda$, and its target can be restricted to $E_\Lambda(\mathbb{C})$, where E has a Weierstrass model given by

$$y^2 = 4x^3 - g_2(\Lambda)x - g_3(\Lambda).$$

This map turns out to be a group isomorphism, yielding the so-called Weierstrass uniformization

$$\begin{aligned} \mathbb{C}/\Lambda &\xrightarrow{\sim} E_\Lambda(\mathbb{C}) \\ z \pmod{\Lambda} &\longmapsto (\wp_\Lambda(z), \wp'_\Lambda(z), 1). \end{aligned} \tag{3.1}$$

When the lattice is Λ_τ, we denote this elliptic curve by E_τ. The group operation on $\mathbb{C}/\Lambda$ is quite easy to understand. The kernel of the multiplication by m map, denoted by $E_\tau[m]$, is isomorphic to $\mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$ and can be seen as $\langle 1/N, \tau/N\rangle$. The subgroup $E_\tau[m]$ will be very important in the next section.

Let $\omega_3 = \omega_1 + \omega_2$ and let $z_j = \omega_j/2$. We can see that the z_j constitute the elements of order exactly 2 in $\mathbb{C}/\Lambda$. Notice that $\wp'_\Lambda(z_j) = 0$, as $z_j \equiv -z_j \pmod{\Lambda}$ and $\wp'_\Lambda$ is an odd function. This implies that $\wp_\Lambda(z_j)$ is a root of the polynomial $4x^3 - g_2(\Lambda)x - g_3(\Lambda)$. Let $f\colon \mathbb{C} \longrightarrow \hat{\mathbb{C}}$ be a non-constant meromorphic function that

is Λ-periodic and let P be the fundamental parallelogram of Λ. (P can be seen as the convex hull of the set $\{0, \omega_1, \omega_2, \omega_3\}$, or the set $\{x_1\omega_1 + x_2\omega_2 \colon x_1, x_2 \in [0,1]\}$.) Let ∂P be the boundary of P and t be a complex number such that $t + \partial P$ does not contain any zero or pole. Such t exists because the meromorphicity of f gives it finitely many poles and zeroes on $\mathbb{C}/\Lambda$. It is straightforward to compute the integrals

$$\frac{1}{2\pi i}\int_{t+\partial P} f(z)dz = 0 \qquad \text{and} \qquad \frac{1}{2\pi i}\int_{t+\partial P} \frac{f'(z)}{f(z)}dz = 0,$$

which allow us to conclude two things. Firstly, thanks to the Residue Theorem, there are no functions f with a simple pole (as the sum of the residues is 0). Secondly, thanks to the Argument Principle, we can see that f has the same number of poles and zeroes, counting multiplicity, which in turn, allows us to conclude that every value is taken on the same number of times. Hence, $\wp_\Lambda$ takes on every value twice, as it only has one double pole. Since $\wp'_\Lambda(z_j) = 0$, the value $\wp_\Lambda(z_j)$ is taken on only once, with multiplicity two, implying that the three values $\wp_\Lambda(z_j)$ are distinct. This implies that the cubic has nonzero discriminant, hence $g_2(\Lambda)^3 - 27g_3(\Lambda)^2 \neq 0$. Also, this justifies the name of the weight 12 cusp form Δ, as it is the discriminant of the elliptic curve obtained this way. Furthermore, we can see that the j-invariant also coincides with the modular function j from the previous section, justifying again the terminology.

So, every lattice $\Lambda \in \mathbb{C}$ produces an elliptic curve over $\mathbb{C}$. The converse is also true and its based on the fact that the modular function j is surjective, which follows from a similar argument to the one exposed above to show $\wp_\Lambda$ takes on every value twice. The end of the proof is virtually the same as when we showed in the previous

chapter that two elliptic curves with the same j-invariant are isomorphic over the algebraic closure.

Let $\phi\colon \mathbb{C}/\Lambda \longrightarrow \mathbb{C}/\Lambda'$ be a (holomorphic) non-constant map of elliptic curves (which we refer to as an *isogeny*). Using topological arguments, this map lifts to a (holomorphic) map $\tilde{\phi}\colon \mathbb{C} \longrightarrow \mathbb{C}$, which maps Λ to Λ' (as the class $0+\Lambda$ needs to be mapped to the class $0+\Lambda'$). Moreover, if z_1 and z_2 differ by $\lambda \in \Lambda$, $\tilde{\phi}(z_1)$ and $\tilde{\phi}(z_2)$ differ by $\lambda' \in \Lambda'$, as $\tilde{\phi}$ lifts a map between the quotients. For a fixed $\lambda \in \Lambda$, the difference $\tilde{\phi}(z+\lambda)-\tilde{\phi}(z) \in \Lambda'$, which is a discrete set. Continuity implies it has to be constant. Upon differentiation, we find that $\tilde{\phi}'(z+\lambda)-\tilde{\phi}'(z)=0$, so $\tilde{\phi}'$ is Λ-periodic, making it bounded, and then constant by Liouville's Theorem. Upon integration, we find that $\tilde{\phi}(z)=mz+b$, with $m \neq 0$ and $b \in \Lambda'$ (by what was mentioned at the beginning of the paragraph and the fact ϕ is not constant). $\tilde{\phi}-b$ also lifts ϕ, so we can assume without loss of generality that $\tilde{\phi}(z)=mz$. Since $m\omega_1, m\omega_2 \in \Lambda'$, we find that $m\Lambda \subseteq \Lambda'$. Conversely, if we have $m \in \mathbb{C}^\times$ such that $m\Lambda \subseteq \Lambda'$, we clearly have a map $\phi\colon \mathbb{C}/\Lambda \longrightarrow \mathbb{C}/\Lambda'$ defined by $\phi(z)=mz$.

If $m\Lambda \neq \Lambda'$, there exists $\lambda' \in \Lambda' - m\Lambda$, so if we let $z=\lambda'/m$ we find that $\phi(z)=\lambda'=0$ in $\mathbb{C}/\Lambda'$, but $z \notin \Lambda$ by definition. This implies that ϕ is not injective, so, ϕ is not an isomorphism. If $m\Lambda=\Lambda'$, we have that $\frac{1}{m}\Lambda'=\Lambda \subseteq \Lambda$, so there is a map $\psi\colon \mathbb{C}/\Lambda' \longrightarrow \mathbb{C}/\Lambda$, and we can see that $\psi \circ \phi$ and $\phi \circ \psi$ are both the identity map on the corresponding torus, implying that ϕ is an isomorphism. If we denote by $\tau=\omega_1/\omega_2$ and we let $m=1/\omega_2$ we find that $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda_\tau$ are isomorphic. Furthermore, if $\alpha=\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbb{Z})$, the two lattices $\tau\mathbb{Z}+\mathbb{Z}$ and $(a\tau+b)\mathbb{Z}+(c\tau+d)\mathbb{Z}$ are the same, yielding the same torus.

If we let $m = (c\tau + d)^{-1}$, we find that $\mathbb{C}/\Lambda_\tau \cong \mathbb{C}/\Lambda_{\alpha(\tau)}$, so the orbit of $\tau \in \mathcal{H}$ under the action of $\mathbf{SL}_2(\mathbb{Z})$ gives us a set $\{E_{\tau'} : \tau' \in \mathrm{Orb}_{\mathbf{SL}_2(\mathbb{Z})}(\tau)\}$, which is comprised of isomorphic elliptic curves. If $E_\tau \cong E_{\tau'}$, there exists an $m \in \mathbb{C}^\times$ such that $\Lambda_\tau = m\Lambda_{\tau'}$, so there exist integers a, b, c, d such that

$$m\tau' = a\tau + b \qquad \text{and} \qquad m = c\tau + d.$$

Multiplying the first equation by d, the second one by b and subtracting, and multiplying the first equation by c, the second one by a and subtracting, we obtain

$$m(d\tau' - b) = (ad - bc)\tau \qquad \text{and} \qquad m(-c\tau' + a) = (ad - bc),$$

respectively. If $ad - bc = 0$ we would have $b = d = 0$ and $a = c = 0$, as τ' and 1 form a basis for Λ'. Dividing both equations by $m(ad - bc)$, we find that

$$\frac{\tau}{m} = \frac{d}{ad - bc}\tau' - \frac{b}{ad - bc} \qquad \text{and} \qquad \frac{1}{m} = \frac{-c}{ad - bc}\tau' + \frac{a}{ad - bc},$$

which implies that $ad - bc$ divides a, b, c and d. We can see that the determinant of $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ divides each of its entries, so we may divide each entry by it and still obtain an integer as the new determinant. This means that $1/(ad - bc)$ is an integer, so $ad - bc = \pm 1$. Dividing the first two equations we get

$$\tau' = \frac{a\tau + b}{c\tau + d}, \text{ so } \Im(\tau') = \Im\left(\frac{(a\tau + b)(\overline{c\tau + d})}{|c\tau + d|}\right) = \Im\left(\frac{ad\tau + bc\overline{\tau}}{|c\tau + d|}\right) = \frac{(ad - bc)\Im(\tau)}{|c\tau + d|}$$

and $\Im(\tau'), \Im(\tau) > 0$ show that $ad - bc > 0$. We conclude that $E_\tau \cong E_{\tau'}$ implies that τ and τ' lie in the same orbit induced by the action of $\mathbf{SL}_2(\mathbb{Z})$. Let us collect this into a proposition:

Proposition 3.2. *Let E be an elliptic curve over $\mathbb{C}$. There exists a number $\tau \in \mathcal{H}$ such that $E \cong E_\tau$. Moreover, for $\tau, \tau' \in \mathcal{H}$, $E_\tau \cong E_{\tau'}$ if an only if there exists a matrix*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}) \quad \textit{such that} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}(\tau) = \frac{a\tau + b}{c\tau + d} = \tau'.$$

Proof. See above. □

3.2 The Modular Curve $X_0(N)$.

Consider the quotient $Y_0(N) := \Gamma_0(N)\backslash\mathcal{H} = \{\Gamma_0(N)\tau \colon \tau \in \mathcal{H}\}$ of the complex upper half-plane by the action of the level N congruence subgroup of integral matrices of determinant 1 which are upper-triangular modulo N. We can define a topology on $Y_0(N)$ by declaring a subset of $Y_0(N)$ open if and only if its inverse image in $\mathcal{H}$ is open. More concretely, if $\pi \colon \mathcal{H} \longrightarrow Y_0(N)$ is the natural projection, $\tilde{U}$ is open if and only if $U = \pi^{-1}(\tilde{U})$ is open.

There is a slight complication at points $\pi(\tau)$ such that the stabilizer of τ in $\Gamma_0(N)$ is not trivial (and, by trivial, we mean $\{\pm I\}$). For any other point, just take a neighborhood small enough such that it is mapped bijectively to a neighborhood of $\pi(\tau)$. When τ is fixed by an element $\gamma \in \Gamma_0(N)$, no neighborhood will have a property as above. These are called *elliptic* points. If τ is an elliptic point, the matrix $\delta_\tau = \left(\begin{smallmatrix} 1 & \tau \\ 1 & -\bar{\tau} \end{smallmatrix}\right)$, with determinant $\tau - \bar{\tau} \neq 0$, maps $\tau \mapsto 0$ and $\bar{\tau} \mapsto \infty$. The stabilizer of τ also stabilizes $\bar{\tau}$, and conjugated by δ_τ stabilizes 0 and ∞. Quotienting by $\pm I$ we obtain the a subgroup of distinct transformations which stabilize τ (and 0 and ∞ after conjugating). This means that all of its elements are of the form az, and it being finite, we can see that these are all rotations about the origin. The

chart can be fixed by the local coordinate z^{h_τ}, where h_τ is the number of these rotations (also called the period of τ). The charts loosely described above endow $Y_0(N)$ with a Riemann surface structure. (Note that all we mentioned applies also to any congruence group.)

It would be desirable that $Y_0(N)$ be compact, but it turns out that we need to add some points for this to happen. Intuitively, we need to add ∞ to $\mathcal{H}$. As we do this for $N = 1$, we see that we also would need to add every element in the orbit of ∞ under the action of $\mathbf{SL}_2(\mathbb{Z})$. This orbit is precisely all rational numbers, as we described in the introduction of this chapter. We compactify $\mathcal{H}$ by adding all these elements (called *cusps*, just as before), so $\mathcal{H}^* := \mathcal{H} \cup \mathbb{Q} \cup \{\infty\}$. We define $X_0(N) := \Gamma_0(N)\backslash\mathcal{H}^* = \{\Gamma_0(N)\tau \colon \tau \in \mathcal{H}^*\}$. The map π extends to $\mathcal{H}^*$, and, in order to endow $X_0(N)$ with a topology, we need to endow $\mathcal{H}^*$ with one as well. The only points that are new are the cusps. Neighborhoods about ∞ will be, as expected, $\{\infty\} \cup \{\tau \in \mathcal{H} \colon \Im(\tau) > c\}$, for any $c \in \mathbb{R}^+$. Neighborhoods about any other cusp will be images under $\mathbf{SL}_2(\mathbb{Z})$ of the neighborhoods we just described. For the charts, if $s \in \mathbb{Q}$ we can bring it up to ∞ with a marix in $\mathbf{SL}_2(\mathbb{Z})$ and look at the stabilizer over there. The stabilizer of ∞ in $\mathbf{SL}_2(\mathbb{Z})$ is generated by the two matrices $\left(\begin{smallmatrix} \pm 1 & 1 \\ 0 & \pm 1 \end{smallmatrix}\right)$, so, in this conjugate of $\Gamma_0(N)$ it will be a subgroup of the group generated by them. Let h be the smallest positive element that can occur as a top right entry in one of these matrices. (For some groups, it could happen that $\left(\begin{smallmatrix} -1 & h \\ 0 & -1 \end{smallmatrix}\right)$ lies in Γ but $\left(\begin{smallmatrix} 1 & h \\ 0 & 1 \end{smallmatrix}\right) \in \Gamma$ does not.) The map $e^{2\pi i\tau/h}$ maps this neighborhood of ∞ biholomorphically into a disk centered at 0, which works as the sought chart.

The compactification just mentioned turns $X_0(N)$ into a compact Riemann surface. The appendix of [GH94] contains the following theorem:

Theorem 3.3. *Suppose C is a compact Riemann surface. Then there exists an immersion*

$$f\colon C \longrightarrow \mathbb{P}^2(\mathbb{C})$$

such that $f(C)$ has at most ordinary double points.

In particular, this means that compact Riemann surfaces are algebraic curves. In the previous section, we saw how the Riemann surfaces $\mathbb{C}/\Lambda$, for a lattice Λ, could be described as a curve. We will now follow the same approach as before, albeit it has a few extra complications. We need to find two meromorphic functions with the same poles and try to find an algebraic relation between the two of them. Unfortunately, modular forms of weight 0 do not exist, as the lack of poles in a compact Riemann surface implies the function is constant. However, we have the j-invariant, which is a level 1 weakly holomorphic modular form (meromorphic at the cusps) of weight 0, i.e., a modular function. Since it is invariant under the action of $\mathbf{SL}_2(\mathbb{Z})$, it is certainly invariant under the action of $\Gamma_0(N)$. This time the derivative j' will not be invariant under the action of $\Gamma_0(N)$ so we need to find a different function.

Define $j_N(\tau) := j(N\tau)$. By the comments preceding Definition 3.1, we know j_N is a meromorphic function on $X_0(N)$, as it is weakly modular of weight 0 with a pole at the cusps. We have the following theorem from [DS05].

Theorem 3.4. *The field of functions of $X_0(N)$ is $\mathbb{C}(j, j_N)$. It is a degree 1 transcendental extension of $\mathbb{C}$. The polynomial $\varphi_N(x, y)$ such that $\varphi_N(j, j_N) = 0$ has integral coefficients.*

The polynomial equations for $\varphi_N(x,y)$ have huge coefficients and are highly singular. The following method can be applied to find such polynomials. Assume N is prime, to reduce the necessary computations.

Notice that $\Gamma_0(N)$ only has two cusps, ∞ and 0. The cusp ∞ has width 1, as $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right) \in \Gamma_0(N)$. The matrix $\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)$ brings 0 to ∞ and conjugation by it transforms $\Gamma_0(N)$ into $\Gamma^0(N)$ (lower-triangular modulo N). The stabilizer of ∞ here is generated by $\left(\begin{smallmatrix}\pm1&N\\0&\pm1\end{smallmatrix}\right)$, with index N, hence, the cusp 0 has width N.

Since j is invariant under $\mathbf{SL}_2(\mathbb{Z})$, we have that

$$j\left(\frac{-1}{\tau}\right) = j(\tau), \quad \text{so} \quad j_N|_0[\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)](\tau) = j\left(N\cdot\frac{-1}{\tau}\right) = j\left(\frac{-1}{\tau/N}\right) = j(\tau/N).$$

We conclude that

$$j(\tau) = \frac{1}{q} + 744 + \mathcal{O}(q), \quad j|_0[\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)](\tau) = \frac{1}{q} + 744 + \mathcal{O}(q) = \frac{1}{q_N^N} + 744 + \mathcal{O}(q_N^N)$$
$$j_N(\tau) = \frac{1}{q^N} + 744 + \mathcal{O}(q^N), \quad j_N|_0[\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)](\tau) = j(\tau/N) = \frac{1}{q_N} + 744 + \mathcal{O}(q_N),$$

which show that j has a pole of order N at 0 and a simple pole at ∞ and that j_N has a simple pole at 0 and a pole of order N at ∞. (When $\Gamma_0(N)$ has more than two cusps, we need more equations like the ones mentioned above.)

Now, in order to find a relation between j and j_N, the goal is to find an algebraic combination of them such the poles at 0 and ∞ disappear (the same way we did with $\wp$ and $\wp'$). For a function $\varphi \in \mathbb{C}(X_0(N))$ with poles only at 0 and ∞, we denote by $\varphi = \mathcal{O}(m,n)$ the fact that φ has a pole of order at most m at ∞ and a pole of order at most n at 0. The function $j^a j_N^b$ has a pole of order $a+Nb$ at ∞ and a pole of order $Na+b$ at 0, so in the previous notation, $j^a j_N^b = \mathcal{O}(a+Nb, Na+b)$.

If $\varphi = \mathcal{O}(m_1, n_1)$ and $\phi = \mathcal{O}(m_2, n_2)$, then $\varphi + \phi = \mathcal{O}(\max\{m_1, m_2\}, \max\{n_1, n_2\})$, as the pole of larger order prevails. Potentially, if $m_1 = m_2$ or $n_1 = n_2$, we will be able to reduce the order of the pole (if the corresponding leading coefficients add up to 0). In order to start reducing the order of the pole, we need to find three pairs (a_1, b_1), (a_2, b_2), (a_3, b_3) such that:

$$\begin{aligned} a_1 + Nb_1 &= a_2 + Nb_2 \qquad & Na_1 + b_1 &= Na_3 + b_3 \\ a_1 + Nb_1 &> a_3 + Nb_3 \qquad & Na_1 + b_1 &> Na_2 + b_2, \end{aligned}$$

which will ensure that the order of the pole at ∞ of $j^{a_1} j_N^{b_1}$ coincides with the one of $j^{a_2} j_N^{b_2}$ and is greater than the one of $j^{a_3} j_N^{b_3}$, and the pole at 0 of $j^{a_1} j_N^{b_1}$ coincides with the one of $j^{a_3} j_N^{b_3}$ and is greater than the one of $j^{a_2} j_N^{b_2}$. With a linear programming approach, we add the variables x and y to obtain the system

$$\begin{aligned} a_1 + Nb_1 &= a_2 + Nb_2 \qquad & Na_1 + b_1 &= Na_3 + b_3 \\ a_1 + Nb_1 &= a_3 + Nb_3 + x \qquad & Na_1 + b_1 &= Na_2 + b_2 + y, \end{aligned}$$

whose solution is given by

$$\begin{aligned} a_1 &= a_3 - x/(N^2 - 1) \\ b_1 &= b_3 + Nx/(N^2 - 1) \\ a_2 &= a_3 - x/(N^2 - 1) - Ny/(N^2 - 1) \\ b_2 &= b_3 + Nx/(N^2 - 1) + y/(N^2 - 1). \end{aligned}$$

The integrality conditions imply that both x and y are divisible by $N^2 - 1$ and positive. The smallest such solution would have them both be equal to $N^2 - 1$

precisely, so

$$(a_1, b_1) = (N, N), \quad (a_2, b_2) = (0, N+1), \quad (a_3, b_3) = (N+1, 0).$$

This means we should consider the functions $j^N j_N^N, j^{N+1}$ and j_N^{N+1}, and we have

$$\begin{aligned} j^N j_N^N &= \mathcal{O}(N(N+1), N(N+1)) \\ j^{N+1} &= \mathcal{O}(N+1, N(N+1)) \\ j_N^{N+1} &= \mathcal{O}(N(N+1), N+1). \end{aligned}$$

Notice that the leading term of the functions of the form $j^a j_N^b$ is 1, so we will never need to divide in the process of finding this polynomial. We illustrate how to proceed with $N = 2$, which will yield a relation of degree 4 with coefficients that already display a very large size.

We start with the function $\varphi = j^2 j_2^2 = q^{-6} + 1488q^{-5} + \mathcal{O}(q^{-4})$, which is $\mathcal{O}(6, 6)$. The functions j_2^3 and j^3 are $\mathcal{O}(6, 3)$ and $\mathcal{O}(3, 6)$ respectively, so we update φ by subtracting $j_2^3 + j^3$. We are left with $\varphi = 1488q^{-5} + 946569q^{-4} + \mathcal{O}(q^{-3})$, which is $\mathcal{O}(5, 5)$.

Now, we take the functions jj_2^2 and $j^2 j_2$, which are $\mathcal{O}(5, 4)$ and $\mathcal{O}(4, 5)$, respectively, and we update φ by subtracting $1488(jj_2^2 + j^2 j_2)$. This time we are left with $\varphi = -162000q^{-4} + 40773375q^{-3} + \mathcal{O}(q^{-2})$, which is $\mathcal{O}(4, 4)$.

We proceed by taking the functions j_2^2 and j^2, which are $\mathcal{O}(4, 2)$ and $\mathcal{O}(2, 4)$, respectively, and we further update φ by adding $162000(j_2^2 + j^2)$. We obtain $\varphi = 40773375q^{-3} + 39083391000q^{-2} + \mathcal{O}(q^{-1})$, which is $\mathcal{O}(3, 3)$.

At this point, we can consider the function jj_2, which is $\mathcal{O}(3,3)$ as well, and update φ by subtracting $40773375jj_2$, thus obtaining $\varphi = 8748000000q^{-2}+8748000000q^{-1}+\mathcal{O}(1)$, which is $\mathcal{O}(2,2)$.

Finally, we use the functions j_2 and j, which are $\mathcal{O}(2,1)$ and $\mathcal{O}(1,2)$, respectively, and after updating φ subtracting $8748000000(j_2 + j)$ we obtain $\varphi = -157464000000000 + \mathcal{O}(q^{43})$, which is $\mathcal{O}(0,0)$. A priori, we expected φ to be $\mathcal{O}(1,1)$ but the order of the pole dropped by 2. Equations (11.14) and (11.15) in [Cox13] assert that the degree of the polynomial in this case is $N+1$ in each of the variables and monic, so the equation has degree at most $2N$. This means that this procedure must produce the sought polynomial. As we cannot find $j^a j_2^b$ that is $\mathcal{O}(1,1), \mathcal{O}(1,0)$ or $\mathcal{O}(0,1)$, we actually expect to reduce the order of the pole by a larger amount in this case. This will happen multiple times when the value of N is larger. The fact that the current φ is $\mathcal{O}(0,0)$ indicates that it has no poles at 0 or ∞, and being an algebraic combination of j and j_2, it has no poles anywhere on the compact Riemann surface $X_0(2)$, making it actually constant.

By collecting all the performed updates to φ, letting $X = j$ and $Y = j_N$, we obtain the relation

$$X^2Y^2 - X^3 - Y^3 - 1488XY(X+Y) + 162000(X^2+Y^2) - \\ 40773375XY - 8748000000(X+Y) = -157464000000000.$$

It is customary to write it as

$$\begin{aligned}\Phi_2(X,Y) = \; & X^3+Y^3-X^2Y^2+1488XY(X+Y)-162000(X^2+Y^2) \\ & +40773375XY+8748000000(X+Y)-157464000000000.\end{aligned}$$

The algorithm, for N prime, goes as follows. The object $\varphi = \mathcal{O}(m,n)$ has attributes $\varphi.m$ and $\varphi.n$, which denote the order of the pole at ∞ and 0, respectively, and the attribute $\varphi.lead$, which denotes the leading coefficient of the q-expansion.

Algorithm 1 Find Modular Polynomial

```
1: function FindExponents(φ, N)
2:     a = φ.m mod N
3:     b = φ.n/N                \\integer division
4:     return (a, b)
5: procedure ModularPolynomial(N)
6:     φ = j^N − j_N^N          \\initialize φ at a function O(N(N+1), N(N+1))
7:     coeffs = [(N, N, −1)]    \\initialize coeffs at an array with triple (N, N, −1)
8:     while φ.m ≠ 0 and φ.n ≠ 0 do
9:         a, b = FindExponents(φ, N)
10:        append to coeffs triple (b, a, −φ.lead)
11:        if a == b then
12:            φ = φ − φ.lead · j^a j_N^b
13:        else
14:            φ = φ − φ.lead · (j^a j_N^b + j^b j_N^a)
15:    append to coeffs triple (0, 0, φ.lead)
16:    return coeffs
```

As we mentioned before, this algorithm could potentially get stuck not being able to find values of a and b that will continue reducing the order of the poles. For example, when $N = 3$ and $\varphi = \mathcal{O}(6,6)$ we obtain $(a,b) = (0,2)$, which yields j_3^2 and j^2, which are $\mathcal{O}(6,2)$ and $\mathcal{O}(2,6)$. After subtracting, we could expect φ to be $\mathcal{O}(5,5)$, which would give us $(a,b) = (2,1)$, requiring us to subtract a scalar multiple of jj_3^2

(and of $j^2 j_3$), but this function is $\mathcal{O}(5,7)$ (and $\mathcal{O}(7,5)$) and this increases the order of the poles. However, φ turns out to be $\mathcal{O}(4,4)$ at this point, with $(a,b) = (1,1)$ and $jj_3 = \mathcal{O}(4,4)$ as well. The restriction on the degree of the polynomial guarantees that the procedure will finish.

The equations obtained above describe the curves $Y_0(N)$. Upon homogenizing, we obtain equations for the curves $X_0(N)$. These are highly singular, and the method is not very practical itself. For a more detailed account of how to find these polynomials (with more sophisticated methods), see [Elk98] or [CL05]. Nonetheless, there are ways of finding smooth models. We will not delve into how to accomplish this. There is another way of thinking about these curves, which is the so-called modular interpretation. We follow [DS05] for this.

Let E be an elliptic curve over $\mathbb{C}$ and let S be a cyclic subgroup of order N. We know that E is isomorphic to some $E_{\tau'}$ by Proposition 3.2, and the N-torsion of $E_{\tau'}$ is given by a quotient of the super-lattice $\frac{1}{N}\Lambda_{\tau'} = \left\langle \frac{\tau'}{N}, \frac{1}{N} \right\rangle = \frac{\tau'}{N}\mathbb{Z} + \frac{1}{N}\mathbb{Z} \supseteq \Lambda_{\tau'}$, which is isomorphic to $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$, as predicted by Theorem 2.9 and the comments following it. In order to describe S, we need to find an element of order N in $S \hookrightarrow \left\langle \frac{\tau'}{N}, \frac{1}{N} \right\rangle \cong \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$, which will be the generator.

Denote the image of such generator by (c,d), and let $g = \gcd(c,d,N)$. Adding (c,d) to itself N/g times yields $(c/g, d/g) = (0,0)$ in $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$, so $g = 1$. Hence, there exist $a, b, k \in \mathbb{Z}$ such that $ad - bc - kN = 1$. Take $\alpha = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, which reduces modulo N to a matrix in $\mathbf{SL}_2(\mathbb{Z}/N\mathbb{Z})$, and hence lifts to a matrix in $\mathbf{SL}_2(\mathbb{Z})$, so we can assume without loss of generality that $\alpha \in \mathbf{SL}_2(\mathbb{Z})$. If $\tau = \alpha(\tau')$, we have that

$E_{\tau'} \cong E_\tau$ and the image of (c, d) under this isomorphism is $(0, 1)$, so the pair (E, S) is isomorphic to the pair $(E_\tau, \langle 1/N \rangle)$ for some $\tau \in \mathcal{H}$.

Now, if $\tau, \tau' \in \mathcal{H}$ lie on the same coset of $\Gamma_0(N)$, $\tau = \alpha(\tau')$ keeping the notation of the previous paragraph, with $\alpha \in \Gamma_0(N)$. E_τ and $E_{\tau'}$ are isomorphic and the isomorphism maps $1/N$ in E_τ to $(c\tau' + d)/N$ in $E_{\tau'}$, where $N \mid c$, so it is the same as d/N. Since $ad \equiv 1 \pmod{N}$, $\gcd(N, d) = 1$ and $1/N \in \langle d/N \rangle$ inside the N-torsion of $E_{\tau'}$, so $(E_\tau, \langle 1/N \rangle) \cong (E_{\tau'}, \langle 1/N \rangle)$.

On the other hand, if $\tau, \tau' \in \mathcal{H}$ are such that $(E_\tau, \langle 1/N \rangle) \cong (E_{\tau'}, \langle 1/N \rangle)$. The isomorphism of elliptic curves implies the existence of $\alpha \in \mathbf{SL}_2(\mathbb{Z})$ as in the previous paragraphs. It maps $1/N \in E_\tau$ to $(c\tau' + d)/N \in E_{\tau'}$. Since the isomorphism maps $\langle 1/N \rangle \subseteq E_\tau$ to $\langle 1/N \rangle \subseteq E_{\tau'}$, $\langle 1/N \rangle = \langle (c\tau' + d)/N \rangle$, implying $N \mid c$, so $\alpha \in \Gamma_0(N)$.

These paragraphs show that $Y_0(N)$ can be seen as the moduli space corresponding to the data *pairs of elliptic curves over* $\mathbb{C}$ *together with a cyclic subgroup of order* N. This construction is actually functorial, so in general, we can drop the requirement that the elliptic curve be defined over $\mathbb{C}$ (as long as N is invertible, which for us will always be the case, given that we are working on number fields). For a number field K, the K-points of $Y_0(N)$ can be identified with isomorphism classes of pairs (E, S) where E is an elliptic curve defined over K and S is a cyclic subgroup of $E[N]$. For more details, see [?].

3.3 Hecke operators

For the vector space $M_k(\Gamma_1(N))$, we can define a commuting family of operators, called *Hecke operators*, which encode arithmetic properties. There are several equivalent ways of defining them, but we will limit ourselves to the more *down to*

earth definition, i.e., the one acting on q-expansions. For the equivalent definitions, see [DS05].

Let p be a prime, and let $\beta_j = \left(\begin{smallmatrix} 1 & j \\ 0 & p \end{smallmatrix}\right)$, for $j \in \{0, 1, \ldots, p-1\}$, $\beta_\infty = \left(\begin{smallmatrix} p & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $\gamma = \left(\begin{smallmatrix} m & n \\ N & p \end{smallmatrix}\right)$, where m, n are integers such that $mp - nN = 1$. Let $f \in M_k(\Gamma_1(N))$. Define

$$T_p f = \begin{cases} \displaystyle\sum_{j=0}^{p-1} f|_k[\beta_j] + f|_k[\gamma\beta_\infty] & p \nmid N \\ \displaystyle\sum_{j=0}^{p-1} f|_k[\beta_j] & p \mid N. \end{cases}$$

It can be seen that $T_p f \in M_k(\Gamma_1(N))$, and the operator descends to an operator on $S_k(\Gamma_1(N))$. Moreover, if $f = \sum_{n\geq 0} a_n q^n \in M_k(N, \chi)$, then $T_p f \in M_k(N, \chi)$ and

$$T_p f = \sum_{n\geq 0} a_{np} q^n + \chi(p) p^{k-1} \sum_{n \geq 0} a_n q^{np}$$

and, naturally, the operator still descends to cusp forms.

Define T_1 as the identity map, and for $r \geq 2$, define

$$T_{p^r} = T_p T_{p^{r-1}} - p^{k-1}\chi(p) T_{p^{r-2}}.$$

Simple computations show that $T_p T_q = T_q T_p$ where p and q are primes, so, it makes sense to define

$$T_n = \prod T_{p_i^{e_i}}, \quad \text{where } n = \prod p_i^{e_i}$$

The commutativity of the operators indexed by primes implies the commutativity of the operators for every index.

For an integer d, we also define the *diamond operator* $\langle d \rangle$, which acts by slashing a modular form by any matrix in $\Gamma_0(N)$ whose lower-right entry is congruent to d

modulo N. When $\gcd(N,d) > 1$, there is no such matrix, so $\langle d \rangle := 0$. Notice that for all integers d_1 and d_2 we have that $\langle d_1 d_2 \rangle = \langle d_1 \rangle \langle d_2 \rangle$.

For a congruence subgroup $\Gamma \in \mathbf{SL}_2(\mathbb{Z})$ we can define the Petersson inner product on the space $S_k(\Gamma)$,

$$\begin{aligned} \langle \cdot, \cdot \rangle_\Gamma \colon S_k(\Gamma) \times S_k(\Gamma) &\longrightarrow \mathbb{C} \\ (f,g) \qquad &\longmapsto \langle f,g \rangle_\Gamma = \frac{1}{V_\Gamma} \int_{X(\Gamma)} f(\tau)\overline{g(\tau)}(\Im(\tau))^{k-2} dxdy, \end{aligned}$$

where $X(\Gamma) := \Gamma \backslash \mathcal{H}^*$ (so the integral runs over any lift of $X(\Gamma)$ on $\mathcal{H}^*$), $\tau = x + iy$ and $V_\Gamma = \displaystyle\int_{X(\Gamma)} \Im(\tau)^{-2} dxdy$. (See [DS05] for details.)

Specializing to $\Gamma = \Gamma_1(N)$ we obtain that for $p \nmid N$ the diamond operators and the Hecke operators satisfy the relations

$$\langle p \rangle^* = \langle p \rangle^{-1} \qquad \text{and} \qquad T_p^* = \langle p \rangle^{-1} T_p,$$

where T^* is the adjoint operator of T (i.e., the unique operator such that $\langle Tf, g \rangle_\Gamma = \langle f, T^* g \rangle_\Gamma$). These relations make such diamond operators and Hecke operators normal on $S_k(\Gamma_1(N))$ (and on $S_k(\Gamma_0(N))$). A family of commuting normal operators on a finite-dimensional inner product space is simultaneously diagonalizable. If we have two vectors, say f and g, that have at least one different eigenvalue, say $Tf = \lambda f$ and $Tg = \mu g$, where T is one member of the family, we have that

$$\lambda \langle f,g \rangle = \langle \lambda f, g \rangle = \langle Tf, g \rangle = \langle f, T^* g \rangle = \langle f, \overline{\mu} g \rangle = \mu \langle f, g \rangle,$$

so f and g are orthogonal. For subspaces generated by eigenvectors that share all eigenvalues at each T, we can orthogonalize them and thus we obtain an orthogonal basis of simultaneous eigenvectors. We refer to these as *eigenforms*.

Definition 3.1 showed us how to obtain modular forms of higher level coming from lower levels. It turns out that the spaces $S_k(\Gamma_0(N))^{\text{old}}$ and $S_k(\Gamma_1(N))^{\text{old}}$ are stable under the action of the Hecke operators and the diamond operators. The spaces $S_k(\Gamma_0(N))^{\text{new}}$ and $S_k(\Gamma_0(N))^{\text{new}}$ are defined as the orthogonal complement with respect to the Petersson inner product of $S_k(\Gamma_0(N))^{\text{old}}$ and $S_k(\Gamma_1(N))^{\text{old}}$ in $S_k(\Gamma_0(N))$ and $S_k(\Gamma_1(N))$, respectively. Furthermore, both spaces are also stable under the action of the Hecke operators and the diamond operators, so all of the spaces mentioned above have an orthogonal basis of eigenforms for the Hecke operators (and diamond operators for $\Gamma_1(N)$) away from the level, i.e., for the set $\{T_n, \langle n\rangle \colon \gcd(n, N) = 1\}$.

If $f \in S_k(\Gamma_1(N))$ has a q-expansion $\sum_{n\geq 1} a_n q^n$ such that $a_n = 0$ for all n with $\gcd(n, N) = 1$, it turns out that $f \in S_k(\Gamma_1(N))^{\text{old}}$. This is not trivial and it is a result due to Atkin and Lehner. Its proof can be found in several sources, for example [AL70] or [DS05]. If $f \in S_k(\Gamma_1(N))^{\text{new}}$, the panorama is quite amicable. The following is Theorem 5.8.2 in [DS05].

Theorem 3.5. *Let $f \in S_k(\Gamma_1(N))^{\text{new}}$ be an eigenform for all the Hecke operators and diamond operators away from the level. Then*

a) f is an eigenform for all Hecke operators and diamond operators.

b) If $g \in S_k(\Gamma_1(N))^{\text{new}}$ is another eigenform for all Hecke operators and all diamond operators away from the level whose corresponding eigenvalues (away from the level) match those of f, then f and g are scalar multiples of each other.

Such an eigenform can be normalized in such a way that the coefficient of q in its q-expansion is equal to 1, and is referred to as a *newform*. Part b) of Theorem 3.5 is referred to as the *Multiplicity One* theorem, which is inherent of newforms. Also, being normalized this way, it is easy to compute that $T_n f = a_n f$, so the coefficients of the q-expansion are precisely the eigenvalues of the Hecke operators. To conclude this section, we mention the existence of some important involutions. For each prime q dividing N, let α be the positive integer such that $p^\alpha \mid N$ but $p^{\alpha+1} \nmid N$. Let x, y, z, w be any integers satisfying $q^\alpha xw - (N/q^\alpha)yx = 1$ and let $W_q = \left(\begin{smallmatrix} q^\alpha x & y \\ Nz & q^\alpha w \end{smallmatrix}\right)$. Slashing by the W_q induces an involution on $S_k(\Gamma_0(N))$ and it does not depend on the values x, y, z, w. These are known as the *Atkin-Lehner* involutions. Their product becomes equivalent to slashing by the matrix $\left(\begin{smallmatrix} 0 & -1 \\ N & 0 \end{smallmatrix}\right)$, which induces an involution on $S_k(\Gamma_1(N))$ (which is stable on the eigenspaces). Each newform has an associated number w_f corresponding to whether or not the involution flips the sign. This involution is referred to as the *Fricke* involution. See [AL70] or [Cre92].

3.4 L-functions associated to Modular Forms

Let $\tau = it$, where t is a positive real number (so $\tau \in \mathcal{H}$). The Mellin transform of $q^n = e^{2\pi i \tau n}$ can be computed as

$$\int_0^\infty e^{2\pi i \tau n} t^{s-1} dt = \int_0^\infty e^{-2\pi t n} t^{s-1} dt = (2\pi)^{-s}\Gamma(s) n^{-s}.$$

This means that if $f(\tau) = \sum_{n\geq 1} a_n q^n$ is a cusp form, its Mellin transform will be

$$\int_0^\infty f(it) t^{s-1} dt = (2\pi)^{-s}\Gamma(s) \sum_{n=1}^{\infty} a_n n^{-s}.$$

For $f \in S_k(\Gamma_1(N))$ as above, define the functions

$$L(f,s) = \sum_{n=1}^{\infty} a_n n^{-s} \qquad \text{and} \qquad \Lambda(f,s) = (2\pi)^{-s} N^{s/2} \Gamma(s) L(f,s),$$

which are referred to as the *L-function* and the *completed L-function*. The growth condition at the cusps puts a bound of the growth of the coefficients, which makes the integral converge for $\Re(s) > 1 + k/2$ (and the series defining the L-function). By carefully splitting the integral in the Mellin transform and exploiting the modularity of f, the integral can be shown to have an analytic continuation to the whole complex plane, and hence, so does $\Lambda(f,s)$. Since $(2\pi)^s$, $N^{-s/2}$ and $1/\Gamma(s)$ are entire, we can affirm the same about $L(f,s)$. Furthermore, for a newform we find the functional equation

$$\Lambda(f, k-s) = (-1)^{k/2} w_f \Lambda(f,s).$$

The number $w = (-1)^{k/2} w_f$ is referred to as the *sign* of the functional equation associated to f. This equation naturally translates into an equation with the L-function without completing it, but it's much less clean.

When $w = -1$ and $s = k/2$ we can see that $\Lambda(f,s) = -\Lambda(f,s)$, so we conclude that $\Lambda(f,s) = 0$, which also implies that $L(f,s) = 0$, (since the extra factors don't vanish when $\Re(s) > 0$).

For the L-function of a newform in $S_k(N,\chi)$ we have an Euler product given by

$$\begin{aligned} L(f,s) &= \sum_{n=1}^{\infty} a_n n^{-s} = \prod_{p \in \mathcal{P}} (1 - a_p p^{-s} + \chi(p) p^{k-1-2s})^{-1} \\ &= \prod_{p \nmid N} (1 - a_p p^{-s} + \chi(p) p^{k-1-2s})^{-1} \prod_{p \mid N} (1 - a_p p^{-s})^{-1}, \end{aligned}$$

which follows from the relations between the T_n, the fact that a_n is the eigenvalue of T_n and the absolute convergence for $\Re(s) > 1 + k/2$.

3.5 Modularity

Let $f \in S_2(\Gamma_0(N))$ be a newform. The curve $X_0(N)$ is a Riemann surface of genus g, which is a g-holed torus with $H_1(X_0(N), \mathbb{Z}) = \mathbb{Z}^{2g}$. If we choose a cusp as the base point of the Riemann surface, every loop γ will lift to a path on $\mathcal{H}^*$, which allows us to define the map

$$\begin{aligned} \langle \cdot, f \rangle \colon H_1(X_0(N), \mathbb{Z}) &\longrightarrow \mathbb{C} \\ \gamma \qquad &\longmapsto \langle \gamma, f \rangle = \int_\gamma 2\pi i f(\tau) d\tau, \end{aligned}$$

where the convergence of the integral is ensured by the fact that f is a cusp form. If f has rational coefficients, the image of this map is a rank 2 lattice in $\mathbb{C}$, denoted by Λ_f. Then, the elliptic curve isomorphic to $\mathbb{C}/\Lambda_f$, E_f, is known to be defined over $\mathbb{Q}$, has conductor N and its Hasse-Weil L-function coincides with the L-function attached to f. This is due to Eichler and Shimura; see [Kna92] for theoretical details and [Cre92] for computational details.

Slightly more generally, choose a basis $\{\gamma_1, \ldots, \gamma_{2g}\}$ of the $\mathbb{Z}$-homology of $X_0(N)$ and basis $\{f_1, \ldots, f_g\}$ of eigenforms (extending a basis of newforms). Let $\omega_j = 2\pi i f_j(\tau) d\tau$ and let $\Omega(X_0(N)) = \langle \omega_1, \ldots, \omega_g \rangle$ be the space of holomorphic differentials in $X_0(N)$. Let

$$\Omega_j = \left(\int_{\gamma_j} \omega_1, \ldots, \int_{\gamma_j} \omega_g \right) \in \mathbb{C}^g \qquad \text{and} \qquad \Lambda = \bigoplus_{j=1}^{2g} \Omega_j \mathbb{Z} \subseteq \mathbb{C}^g.$$

This lattice is discrete and $\Lambda \otimes \mathbb{R} = \mathbb{C}^g$. The *Jacobian* of $X_0(N)$, denoted $J(X_0(N))$, is isomorphic to $\mathbb{C}^g/\Lambda$. If $\tau \in \mathcal{H}^*$, the map

$$\begin{array}{ccccc} \mathcal{H}^* & \longrightarrow & \mathbb{C}^g & \longrightarrow & J(X_0(N)) \\ \tau & \longmapsto & \left(\int_{i\infty}^{\tau} \omega_1, \ldots, \int_{i\infty}^{\tau} \omega_g\right) & \longmapsto & \left(\int_{i\infty}^{\tau} \omega_1, \ldots, \int_{i\infty}^{\tau} \omega_g\right) \bmod \Lambda \end{array}$$

is independent of the path and descends to a map from the quotient by $\Gamma_0(N)$ of $\mathcal{H}$, so we obtain a map $X_0(N) \longrightarrow J(X_0(N))$ from the modular curve into its Jacobian. This is the so-called *Abel-Jacobi* map.

The elliptic curve E_f is a quotient of the Jacobian of $X_0(N)$ (after projecting on the appropriate component, and rescaling by the *Manin constant*, denoted by c) so we further obtain a map

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

This map can be made very explicit using the q-expansion of f. For $\tau \in \mathcal{H}^*$, let

$$z_\tau = c\int_{i\infty}^{\tau} 2\pi i f(z)dz = c\int_{i\infty}^{\tau} 2\pi i \sum_{n=1}^{\infty} a_n \mathfrak{q}^n dz = c\int_0^q \sum_{n=1}^{\infty} a_n \mathfrak{q}^{n-1} d\mathfrak{q} = c\sum_{n=1}^{\infty} \frac{a_n}{n} q^n,$$

after performing the substitution $\mathfrak{q} = e^{2\pi i z}$, $d\mathfrak{q} = 2\pi i \mathfrak{q} dz$. z_τ depends on τ modulo $\Gamma_0(N)$ as mentioned above. Using the Weierstrass $\wp$-function we obtain

$$\begin{array}{ccc} \varphi\colon X_0(N) & \longrightarrow & E_f \\ \tau & \longmapsto & (\wp(z_\tau), \wp'(z_\tau), 1). \end{array}$$

This map is referred to as the *modular parametrization* of E_f. Despite its apparent analytic description, this is a map of algebraic curves defined over $\mathbb{Q}$.

Conversely, for an elliptic curve E defined over $\mathbb{Q}$ of conductor N, there exists a weight 2 newform f of level N, whose L-function coincides with the Hasse-Weil L-function for which E_f is isogenous to E. The map described above composed with this isogeny equips us with a map

$$X_0(N) \longrightarrow E_f \longrightarrow E. \tag{3.2}$$

This was first known as the Shimura-Taniyama conjecture. Later on, Weil's name was added. It was first partially proved by Wiles in the early 1990s for semi-stable elliptic curves in [Wil95] and his worked was extended to all elliptic curves over $\mathbb{Q}$ by Breuil, Conrad, Diamond and Taylor in [BCDT01]. It is referred to as the *modularity theorem*.

Chapter 4
Heegner and Stark-Heegner points; the classical case

As we mentioned in the introduction, finding algebraic points on elliptic curves is not as easy as finding them on conics. Currently, the best methods that produce algebraic points rely on the existence of algebraic points on the modular curve and, under suitable conditions, the modular parametrization often yields non trivial points on the elliptic curve. These points on the modular curve are called *Heegner points*, and, by extension, the resulting points on the elliptic curve inherit the same name.

Class Field Theory together with the theory of Complex Multiplication for elliptic curves gives an important algebraicity result, which is key for the Heegner construction.

4.1 Complex Multiplication and Class Field Theory

This section gives a quick survey of the results we need. For a more detailed explanation, see [Cox13] and [Sil94] as well as the references provided there.

In section 2.3.2 we defined the endomorphism ring $\mathrm{End}(E)$ for an elliptic curve E together with an embedding of $\mathbb{Z}$ into it. In most cases, this embedding is in fact an isomorphism. When this embedding is not surjective we say that E has *complex multiplication*, and often we just say that E has *CM*. The origin of this term is that when an elliptic curve has CM, its endomorphism ring can be embedded into the ring of integers of an imaginary quadratic extension of $\mathbb{Q}$.

Theorem 4.1. *Let E be an elliptic curve over a field of characteristic 0. Then either* $\mathrm{End}(E) = \mathbb{Z}$ *or there exist an imaginary quadratic extension K and a positive integer c such that* $\mathrm{End}(E) = \mathbb{Z} + c\mathcal{O}_K$*, where $\mathcal{O}_K$ denotes the ring of integers of K.*

The latter kind of subrings of K are called *orders*. The integer c is called the *conductor* of the order. The set $\mathcal{O} = \mathbb{Z} + c\mathcal{O}_K$ forms a subring of K and we refer to it as the order of conductor c in K.

For a given fractional $\mathcal{O}$-ideal $\mathfrak{a}$, we can consider the subset of K of elements α with the property that $\alpha\mathfrak{a} \subseteq \mathfrak{a}$. Since $\mathfrak{a}$ is a fractional $\mathcal{O}$-ideal, $\mathcal{O}$ is clearly contained in this subset, and it is closed under multiplication and addition, inheriting a ring structure. This makes it into an order in K as well. We say that $\mathfrak{a}$ is a *proper* $\mathcal{O}$-ideal if this order is still $\mathcal{O}$ and not a larger order. Denote by $I(\mathcal{O})$ the set of proper $\mathcal{O}$-ideals, which can be endowed with an abelian group structure. Every principal $\mathcal{O}$-ideal is proper, so the set of principal $\mathcal{O}$-ideals, denoted by $P(\mathcal{O})$, forms a subgroup of $I(\mathcal{O})$. Furthermore, if we only take ideals which can be generated by an element $\alpha \in K$ such that under every real embedding $\sigma : K \longrightarrow \mathbb{R}$ we have $\sigma(\alpha) > 0$, we obtain a subgroup of $P(\mathcal{O})$, denoted $P^+(\mathcal{O})$. For K imaginary quadratic, this requirement is automatic, so $P^+(\mathcal{O}) = P(\mathcal{O})$. If K is real quadratic and its fundamental unit is of norm -1, every principal ideal can be generated by a totally positive element and in this case $P^+(\mathcal{O}) = P(\mathcal{O})$ as well. If the fundamental unit has norm 1, then $[P(\mathcal{O})\colon P^+(\mathcal{O})] = 2$.

The *class group* and the *narrow class group* of $\mathcal{O}$ are defined as

$$\mathrm{Cl}(\mathcal{O}) := I(\mathcal{O})/P(\mathcal{O}) \qquad \text{and} \qquad \mathrm{Cl}^+(\mathcal{O}) := I(\mathcal{O})/P^+(\mathcal{O}),$$

respectively. If $\mathcal{O} = \mathcal{O}_K$, the class group measures the failure of the ring $\mathcal{O}_K$ being a principal ideal domain. The number of elements of $\mathrm{Cl}(\mathcal{O}_K)$ is called the *class number* of K and the number of elements of $\mathrm{Cl}^+(\mathcal{O}_K)$ is called the *narrow class number* of K. As mentioned before, in some instances the narrow class group coincides with the class group, but sometimes it will result in a finer invariant.

We have the following theorem:

Theorem 4.2. *Let D be a non-square discriminant and let*

$$\mathcal{O} = \begin{cases} \mathbb{Z}\left[\frac{\sqrt{D}}{2}\right] & D \equiv 0 \pmod 4 \\ \mathbb{Z}\left[\frac{1+\sqrt{D}}{2}\right] & D \equiv 1 \pmod 4. \end{cases}$$

The groups $\mathrm{Cl}^+(\mathcal{O})$ and $\mathrm{Cl}(D)$ are isomorphic.

To put these class groups in a more suitable language for class field theory, we need a few definitions. For an order $\mathcal{O} \subseteq K$ of conductor c, let $I_K(c)$ be the subgroup of fractional ideals in K generated by integral ideals which are prime to c, i.e., ideals $\mathfrak{a}$ such that $\mathfrak{a} + c\mathcal{O}_K = \mathcal{O}_K$. (Equivalently, the norm of the ideal is relatively prime to c.) Let $P_{K,\mathbb{Z}}(c)$ be the subgroup of $I_K(c)$ generated by principal ideals $\alpha\mathcal{O}_K$ where $\alpha \in \mathcal{O}_K$ is such that $\alpha - a \in c\mathcal{O}_K$ for some integer a relatively prime to c. Finally, let $P_{K,1}(c)$ be the subgroup of $P_{K,\mathbb{Z}}(c)$ generated by principal ideals as above, in which $a = 1$.

Fix a number field K. Let $\mathfrak{c}$ be a finite formal product of places in K (possibly empty) such that the real infinite places appear at most once and the complex infinite places do not appear. Such $\mathfrak{c}$ is referred to as a *modulus*. The *ray class field* of K modulo $\mathfrak{c}$ is the unique field extension $K_\mathfrak{c}/K$ with the following properties:

- It is abelian
- It is unramified outside of $\mathfrak{c}$
- Its conductor divides $\mathfrak{c}$
- It is maximal with respect to the three properties above.

A modulus $\mathfrak{c}$ can be written as $\mathfrak{c}_0\mathfrak{c}_\infty$, where $\mathfrak{c}_0$ and $\mathfrak{c}_\infty$ are comprised of the finite and infinite places, respectively. As before, we define $I_K(\mathfrak{c})$ as the subgroup of fractional ideals in K generated by ideals which are prime to $\mathfrak{c}_0$. In contrast with $P_{K,1}(c)$, we define $P_{K,1}(\mathfrak{c})$ as the subgroup of $I_K(\mathfrak{c})$ generated by principal ideals $\alpha\mathcal{O}_K$ where $\alpha \in \mathcal{O}_K$ is such that $\alpha - 1 \in \mathfrak{c}_\mathfrak{o}$ and $\sigma(\alpha) > 0$ for every σ dividing $\mathfrak{c}_\infty$.

Let H be a subgroup of $I_K(\mathfrak{c})$ that contains $P_{K,1}(\mathfrak{c})$. Let L be the fixed field of $K_\mathfrak{c}$ by $H/P_{K,1}(\mathfrak{c})$. By Galois theory, we have the following diagram

$$\begin{array}{ccc} & K_\mathfrak{c} & \\ {\scriptstyle I_K/P_{K,1}(\mathfrak{c})}\big/ & | & \\ & L = K_\mathfrak{c}^{H/P_{K,1}(\mathfrak{c})} & \\ & | {\scriptstyle I_K(\mathfrak{c})/H} & \\ & K & \end{array}$$

which yields a field extension whose Galois group is $I_K(\mathfrak{c})/H$.

Let K be a quadratic field and $\mathcal{O} = \mathbb{Z} + c\mathcal{O}_K$ be the order of conductor c. Let $\mathfrak{c}$ be the modulus corresponding to c (i.e., $\mathfrak{c}$ looks the same as the factorization of (c) into primes ideals in K) and $\mathfrak{c}^+$ the modulus corresponding to c and the infinite real places. Let H_c and H_{c^+} be the fixed fields of $K_\mathfrak{c}$ and $K_{\mathfrak{c}^+}$ by $P_{K,\mathbb{Z}}(c)$ and $P_{K,\mathbb{Z}}(c^+)$, respectively. They are referred to as the *ring class field* and *narrow ring class field*

of K of conductor c. The Galois groups of these extensions are isomorphic to $\mathrm{Cl}(\mathcal{O})$ and $\mathrm{Cl}^+(\mathcal{O})$, respectively.

When $\mathcal{O} = \mathcal{O}_K$, $\mathfrak{c} = 1$ is an empty product and $\mathfrak{c}^+$ is just the infinite places. In this case, K_1 is referred to as the *Hilbert class field* of K and K_{1^+} as the *Hilbert narrow class field* of K.

Now, the most important result of this section:

Theorem 4.3. *Let E be an elliptic curve with complex multiplication by the ring $\mathcal{O} \subseteq K$ of conductor c, where K is an imaginary quadratic field. Then $j(E)$ is an algebraic integer. Furthermore, $j(E) \in H_{\mathfrak{c}}$, where $H_{\mathfrak{c}}$ is the ring class field of conductor $\mathfrak{c}$, and $[\mathbb{Q}(j(E))\colon \mathbb{Q}] = [K(j(E))\colon K] = [H_{\mathfrak{c}}\colon K]$.*

When E is defined over $\mathbb{C}$, we know that it is isomorphic to E_τ for some $\tau \in \mathcal{H}$, and all endomorphisms of E_τ are given by a complex number α such that $\alpha\Lambda_\tau \subseteq \Lambda_\tau$, by doing $z \mapsto \alpha z$ in $\mathbb{C}/\Lambda_\tau$. For this to happen, we need the equations

$$\alpha \cdot \tau = a\tau + b$$

$$\alpha \cdot 1 = c\tau + d$$

to hold for some a, b, c, $d \in \mathbb{Z}$. If the endomorphism is not the trivial endomorphism, $\alpha \neq 0$ and we can divide these two equations to find that τ is fixed by the Möbius transformation induced by the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, which yields an equation in τ. Upon solving it, we find that τ is the solution of a quadratic equation with integral coefficients (and, a fortiori, so is α).

Conversely, if $\tau \in \mathcal{H} \cap K$ is a quadratic number, we can see that E_τ has complex multiplication by the order comprised of numbers of the form $c\tau + d$ as above. The

minimal equation of τ is of the form $A\tau^2 + B\tau + C = 0$ where A, B, C are relatively prime. Every quadratic equation with integer coefficients where τ vanishes is a scalar multiple of this one so we have the polynomial relation

$$cX^2 + (d-a)X - b = AkX^2 + BkX + Ck,$$

whence $c = Ak$, $d - a = Bk$ and $b = -Ck$. This shows that E_τ has complex multiplication by the ring $\mathbb{Z}[A\tau] \subseteq K$. We can readily see that we can determine exactly in what field $j(\tau)$ lies for quadratic numbers.

Corollary 4.4. *Let $\tau \in \mathcal{H} \cap K$ be a quadratic number and let A be the leading coefficient of the minimal polynomial of τ. Let c be the conductor of the order $\mathcal{O} = \mathbb{Z}[A\tau] \subseteq K$. Then E_τ has complex multiplication by $\mathcal{O}$ and the special value $j(\tau)$ is an algebraic integer of degree $[H_\mathfrak{c} \colon K]$ which lies in the ring class field $H_\mathfrak{c}$.*

4.2 Heegner points

If τ and $N\tau$ are quadratic numbers in $\mathcal{H}$ such that the elliptic curves E_τ and $E_{N\tau}$ have complex multiplication by the same order $\mathcal{O} \subseteq K$ of conductor c, the previous section implies that their j-invariants will lie in the same algebraic extension of $K_\mathfrak{c}/\mathbb{Q}$, where $K_\mathfrak{c}$ is the ray class field of conductor $\mathfrak{c}$. Then, Equation (3.2) ensures that the point $(j(\tau), j(N\tau))$ of $X_0(N)$ is mapped to a point defined over the same field, thus obtaining a point on an elliptic curve of conductor N defined over this ray class field.

In order to find a τ as above, let us analyze the minimal equations of τ and $N\tau$. Let D be a negative integer congruent to 1 or 0 modulo 4 and let a, b, c be relatively prime integers such that $b^2 - 4ac = D$. Let $N\tau \in \mathcal{H}$ be the solution of the quadratic equation $ax^2 + bx + c = 0$ in the upper half-plane. Let $g = \gcd(aN^2, bN, c)$,

so $\gcd(aN^2/g, bN/g, c/g) = 1$ and

$$(aN^2/g)\tau^2 + (bN/g)\tau + (c/g) = (a(N\tau)^2 + b(N\tau) + c)/g = 0,$$

making $(aN^2/g)x^2 + (bN/g)x + (c/g) = 0$ the minimal equation of τ. By the first part of Corollary 4.4, we find that

$$\mathrm{End}(E_\tau) = \mathbb{Z}[aN^2\tau/g] \qquad \text{and} \qquad \mathrm{End}(E_{N\tau}) = \mathbb{Z}[aN\tau].$$

For these two orders to be equal, we must have $g = N$, and hence, the minimal polynomial of τ becomes $aNx^2 + bx + C$, where $C = c/N$ and $N \mid c$. Its discriminant is $b^2 - 4aNC = b^2 - 4ac = D$ so the two binary quadratic forms

$$ax^2 + bxy + cy^2 \qquad \text{and} \qquad aNx^2 + bxy + Cy^2$$

are in the same class group. Hence, we need to find a binary quadratic form $F(x, y)$ with discriminant D whose leading term is divisible by N and take $\tau \in \mathcal{H}$ such that $F(\tau, 1) = 0$.

Let us flip things around and start with a discriminant D together with a binary quadratic form $Ax^2 + Bxy + Cy^2$ of discriminant D and let us try to find en element in the same equivalence class in $\mathrm{Cl}(D)$ whose x^2 coefficient is divisible by N. This is, we need a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbb{Z})$ such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \left(Ax^2 + Bxy + Cy^2\right) = (Aa^2 + Bac + Cc^2)x^2 +$$

$$(2Aab + B(ad + bc) + 2Ccd)xy + (Ab^2 + Bbd + Cd^2)y^2$$

yields an x^2 coefficient divisible by N. This boils down to finding a non-trivial solution to the original quadratic form congruent to 0 modulo N.

By the Chinese remainder theorem, we need to find non-trivial solutions modulo each of the prime powers dividing N. Let $p \mid N$ be odd. If $p \nmid A$, we have

$$p \mid Aa^2 + Bac + Cc^2 \iff p \mid 4A^2a^2 + 4ABac + 4ACc^2 = (2Aa + Bc)^2 - c^2D.$$

Then $p \nmid c$ because otherwise we'd conclude that $p \mid a$, so we can assume WLOG that $c = 1$ (modulo p). This implies that we have a solution modulo p if and only if D is a square modulo p, so p splits (or ramifies) in $K = \mathbb{Q}(\sqrt{D})$. If $p \mid A$ then $D \equiv B^2 \pmod{p}$ so p still splits (or ramifies) in K.

Let us look at a couple of examples:

- Let $D = -11$ and $E = 15a1$ (following Cremona labels). We have that the conductor of E is $N = 15$ and the primes 3 and 5 split in $K = \mathbb{Q}(\sqrt{-11})$. The cardinality of $\mathrm{Cl}(-11)$ is 1 so we only have the identity quadratic form, namely $x^2 + xy + 3y^2$. The pair $(a, c) = (3, 1)$ vanishes at this quadratic form modulo 15. We need to find a pair (b, d) such that $ad - bc = 1$, which can easily be achieved by $(2, 1)$. Then

$$\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \left(x^2 + xy + 3y^2\right) = 15x^2 + 23xy + 9y^2,$$

  which vanishes at $(\tau, 1)$ with $\tau = \dfrac{-23 + \sqrt{-11}}{30}$. The minimal polynomials of τ and 15τ are $15x^2+23x+9$ and $x^2+23x+135$, respectively, so both E_τ and $E_{15\tau}$

have CM by $\mathbb{Z}[15\tau] = \mathcal{O}_K$, the maximal order in K. The modular parametrization yields the point $\left(\frac{3-\sqrt{-11}}{2}, \frac{-3-3\sqrt{-11}}{2}\right)$. It can be readily verified that it satisfies the equation $y^2 + xy + y = x^3 + x^2 - 10x - 10$ corresponding to E. Since the class number is 1, we expected this point to belong to $E(K)$, as it happened.

- Let $D = -15$ and $E = 17a1$. The conductor of E is $N = 17$, which splits in $K = \mathbb{Q}(\sqrt{-15})$. The cardinality of $\mathrm{Cl}(-15)$ is 2 so we have two quadratic forms this time, namely, $x^2 + xy + 4y^2$ and $2x^2 + xy + 2y^2$. The pairs $(a_1, c_1) = (11, 0)$ and $(a_2, c_2) = (14, 1)$ vanish at these quadratics modulo 17, respectively. We can complete them with the pairs $(b_1, d_1) = (10, 1)$ and $(b_2, d_2) = (13, 1)$, thus obtaining

$$\begin{pmatrix} 11 & 10 \\ 1 & 1 \end{pmatrix} \left(x^2 + xy + 4y^2\right) = 136x^2 + 249xy + 114y^2$$
$$\begin{pmatrix} 14 & 13 \\ 1 & 1 \end{pmatrix} \left(2x^2 + xy + 2y^2\right) = 408x^2 + 759xy + 353y^2,$$

which vanish at $(\tau_1, 1)$ and $(\tau_2, 1)$, respectively, with $\tau_1 = \frac{-249+\sqrt{-15}}{272}$ and $\tau_2 = \frac{-759+\sqrt{-15}}{816}$. Since these denominators are rather large, many more coefficients of the L-series are needed in order to compute the modular parametrization, as the convergence is much slower. τ_1 yields, approximately, the point

$$(-0.381966 + 1.73205i, 0.427051 - 3.7948939i)$$

and τ_2

$$(-2.6180339 - 1.73205i, -2.927051 + 5.7313855i)$$

which add up to

$$(1.0000 + 3.8729833i, -4.500000 + 6.809475i)$$

The point associated to τ_1 can be recognized as

$$\left(-\frac{3}{2} + \sqrt{-3} + \frac{\sqrt{5}}{2}, -\frac{5}{4} - \frac{11\sqrt{-3}}{4} + \frac{3\sqrt{5}}{4} + \frac{\sqrt{-15}}{4}\right)$$

and the one associated to τ_2 as

$$\left(-\frac{3}{2} - \sqrt{-3} - \frac{\sqrt{5}}{2}, -\frac{5}{4} + \frac{11\sqrt{-3}}{4} - \frac{3\sqrt{5}}{4} + \frac{\sqrt{-15}}{4}\right)$$

which can be verified to lie on $E(\mathbb{Q}(\sqrt{-3}, \sqrt{5}))$ as expected ($\mathbb{Q}(\sqrt{-3}, \sqrt{5})$ being the Hilbert class field of $\mathbb{Q}(\sqrt{-15})$) and they are conjugates (under $\sigma \in \mathrm{Gal}\left(\mathbb{Q}(\sqrt{-3}, \sqrt{5})/K\right)$ given by $\sigma(\sqrt{-3}) = -\sqrt{-3}$ and $\sigma(\sqrt{5}) = -\sqrt{5}$). Their sum is

$$\left(1 + \sqrt{-15}, \frac{-9 + 3\sqrt{-15}}{2}\right)$$

which lies on $E(\mathbb{Q}(\sqrt{-15}))$, and coincides with the approximated sum obtained before.

- Let $D = -164$ and $E = 11a1$. The conductor $N = 11$ splits in $K = \mathbb{Q}(\sqrt{-41})$ and the cardinality of $\mathrm{Cl}(D)$ is 8. For each of the eight reduced binary quadratic forms of discriminant -41 found in section 2.2.2 we can find a representative in their class whose x^2 coefficient is divisible by 11, compute its fixed point and push it to the elliptic curve via the modular parametrization. Upon adding

these eight points we obtain the point

$$\left(\frac{-32+6\sqrt{-41}}{25}, \frac{192-36\sqrt{-41}}{125}\right)$$

- If in the last example we consider the curve $E = 37a1$, which has rank 1, more things can be said. The conductor $N = 37$ splits over K again so we can find representatives for each of the classes whose x^2 term is divisible by 37, yielding eight points on $E(L)$, where L is the splitting field of the polynomial

$$x^8 + 10x^7 + 36x^6 + 8x^5 - 47x^4 - 54x^3 - 10x^2 - 12x + 4,$$

where all the x coordinates of the Heegner points obtained vanish. The y coordinates all vanish at

$$x^8 + 24x^7 + 243x^6 + 524x^5 + 95x^4 - 324x^3 + 149x^2 - 24x - 32,$$

which splits in L. The Hilbert class field of K is defined by the polynomial

$$x^8 - 3x^7 + x^6 + 4x^5 - 4x^4 + 4x^3 + x^2 - 3x + 1,$$

which also splits in L, so the Hilbert class field is a quadratic extension of L. (Actually, the compositum of K and L.) After adding all of the points, which are conjugates in L, we obtain the point $(0, 0)$, which, needless to say, belongs to $E(\mathbb{Q})$.

In the light of the last example, we have the following theorem, due to Gross-Zagier. See [GZ86].

Theorem 4.5. *Let $E/\mathbb{Q}$ be an elliptic curve of conductor N and K be an imaginary quadratic field such that every prime dividing N splits in K. Let h be the class number of K. Assume that the L-series of E has a zero of order 1 at $s = 1$. Then the sum of the h Heegner points is a point of infinite order in $E(\mathbb{Q})$.*

4.3 Stark-Heegner points

The construction of Heegner points relies heavily on the modular parametrization and the algebraicity of the j values at imaginary quadratic numbers. A natural question to ask is if there is an analogous process that yields algebraic points on elliptic curves by starting with real quadratic numbers. At this point, the classical modular parametrization seems to be of no use, since j cannot be evaluated at real numbers, so the analogy will have to rely on slightly different techniques.

Let K be a quadratic extension of $\mathbb{Q}$. Demanding that K be imaginary is equivalent to saying that the place ∞ does not split in K. If we are dealing with a real quadratic extension of $\mathbb{Q}$ (which will be our case), we need to find a different place to play the role of ∞. We choose a prime p dividing the conductor N of the elliptic curve E such that p is inert in K. This prime, naturally, does not always exist, but if we restrict ourselves to the case where N and the discriminant of K are relatively prime and the sign of the functional equation of $L(E/K, s)$ is -1, we can guarantee its existence. (See Theorem 3.17 in [Dar04].) From now on, write $N = pM$, with $\gcd(p, M) = 1$ and p inert in K.

Define the p-adic upper half-plane $\mathcal{H}_p$ to be $\mathbb{P}^1(\mathbb{C}_p) - \mathbb{P}^1(\mathbb{Q}_p)$. In opposition with its archimedean counterpart, the p-adic upper half-plane does not split in a natural way into two disjoint components and there is no canonical choice, so we just keep

all of them. The topology comes from *affinoids* and *annuli*, which are inverse images of special subsets of $\mathbb{P}^1(\bar{\mathbb{F}}_p)$ under the reduction map red : $\mathbb{P}^1(\mathbb{C}_p) \longrightarrow \mathbb{P}^1(\bar{\mathbb{F}}_p)$. If Γ is a discrete subgroup of $\mathbf{SL}_2(\mathbb{Q}_p)$ the quotient $\Gamma\backslash\mathcal{H}_p$ is equipped with the structure of a rigid analytic curve over $\mathbb{Q}_p$. For a quick survey, see section 5.1 on [Dar04].

Previously, we had the group $\Gamma_0(N)$ acting discretely on $\mathcal{H}$. Let $\Gamma_{p,M}$ be the subgroup of $\mathbf{SL}_2(\mathbb{Z}[1/p])$ of matrices which are upper triangular modulo M. The action of $\Gamma_{p,M}$ is not discrete on either $\mathcal{H}_p$ or $\mathcal{H}$, but it turns out to be discrete on the product $\mathcal{H}_p \times \mathcal{H}$. We will be dealing with $\Gamma_{p,M}$ instead of the usual $\Gamma_0(N)$.

4.3.1 p-adic measures and p-adic line integrals

This section largely follows sections 5.2 and 5.3 of [Dar04]. The boundary of $\mathcal{H}_p$ is $\mathbb{P}^1(\mathbb{Q}_p)$, which is endowed with its natural p-adic topology generated by the open balls

$$B(a,r) = \{t \text{ such that } |t-a| < p^{-r}\}$$
$$B(\infty,r) = \{t \text{ such that } |t| > p^{r}\},$$

for every $a \in \mathbb{Q}_p$. Any compact open subset of $\mathbb{P}^1(\mathbb{Q}_p)$ is a finite disjoint union of open balls and any finite disjoint union of open balls is compact and open. By assigning values in $\mathbb{C}_p$ in a coherent way to open balls we obtain a function μ which assigns a value in $\mathbb{C}_p$ to every compact open subset of $\mathbb{P}^1(\mathbb{Q}_p)$ and is finitely additive. If we further require that $\mu(\mathbb{P}^1(\mathbb{Q}_p))$ be 0, we refer to μ as a *p-adic distribution* on $\mathbb{P}^1(\mathbb{Q}_p)$. Additionally, if the p-adic distribution is bounded (i.e., the set of values of $|\mu(U)|_p$ is bounded in the usual sense) we say that μ is a *measure* on $\mathbb{P}^1(\mathbb{Q}_p)$.

For a p-adic measure μ, we can define a $\mathbb{C}_p$-linear operator that assigns values in $\mathbb{C}_p$ to continuous $\mathbb{C}_p$-valued functions on $\mathbb{P}^1(\mathbb{Q}_p)$ as follows:

$$\int_{\mathbb{P}^1(\mathbb{Q}_p)} \lambda(t) d\mu(t) = \lim \sum_{\alpha} \lambda(t_\alpha)\mu(U_\alpha), \tag{4.1}$$

where $\{U_\alpha\}_\alpha$ is a disjoint cover of $\mathbb{P}^1(\mathbb{Q}_p)$ by compact open subsets, the limit is taken over increasingly finer such covers and $t_\alpha \in U_\alpha$ is any sample point. The finite-additivity and boundedness of μ are the necessary ingredients to show this is a well defined operator. If λ is the characteristic function of a compact open set U, we write

$$\mu(U) = \mu(\lambda) = \int_{\mathbb{P}^1(\mathbb{Q}_p)} \lambda(t) d\mu(t) = \int_U d\mu(t).$$

For a discrete subgroup Γ of $\mathbf{SL}_2(\mathbb{Q}_p)$ we say that μ is Γ-invariant if $\mu(\gamma U) = \mu(U)$ for all $\gamma \in \Gamma$ and U open compact subset of $\mathbb{P}^1(\mathbb{Q}_p)$.

For a Γ-invariant measure μ we define

$$\begin{aligned} f_\mu \colon \mathcal{H}_p &\longrightarrow \mathbb{C}_p \\ z &\longmapsto \int_{\mathbb{P}^1(\mathbb{Q}_p)} \left(\frac{1}{z-t}\right) d\mu(t). \end{aligned}$$

Section 5.2 of [Dar04] shows that f_μ is rigid analytic and it is Γ-invariant of weight 2 as per Definition 5.5 in [Dar04]. It is a theorem of Schneider and Teitelbaum (Theorem 5.9 in [Dar04]) that the map $\mu \mapsto f_\mu$ is an isomorphism from the set of Γ-invariant p-adic measures on $\mathbb{P}^1(\mathbb{Q}_p)$ to the set of weight 2 Γ-invariant rigid analytic functions on $\mathcal{H}_p$.

The function

$$\log(1-z) = \sum_{n=1}^{\infty} \frac{z^n}{n}$$

has as domain the open disc in $\mathbb{C}_p$ of radius 1 centered at 1. For small real numbers x and y, we know that $\log((1-x)(1-y)) = \log(1-x) + \log(1-y)$ from the properties of the natural logarithm, so we have an equality of formal power series which lets us conclude the same is true for z in said open disc. Choosing $\pi \in \mathbb{C}_p^\times$ such that $|\pi|_p < 1$ and setting $\log(\pi) = 0$, we can extend log to $\mathbb{C}_p^\times$ by demanding that log be a homomorphism from $\mathbb{C}_p^\times$ to $\mathbb{C}_p$. This is referred to as choosing a branch of the logarithm.

For a rational differential $f(z)dz$ on $\mathbb{P}^1(\mathbb{C}_p)$, we can assign a formal antiderivative

$$F(z) = R(z) + \sum_{j=1}^{t} \lambda_j \log(z - z_j),$$

where $R(z)$ is a rational function, the λ_j's are some constants in $\mathbb{C}_p$, and the sum is taken over all the poles of $f(z)dz$. This antiderivative is unique up to an additive constant. In the case of $f(z) = 1/(z-t)$ one simply has $F(z) = \log(z-t)$.

For a Γ-invariant rigid analytic function f on $\mathcal{H}_p$ we have an associated measure μ_f (given by Schneider and Teitelbaum's isomorphism; see sketch of proof of Theorem 5.9 in [Dar04] for details). Define the *line integral* on $\mathcal{H}_p$ from τ_1 to τ_2 to be

$$\int_{\tau_1}^{\tau_2} f(z)dz = \int_{\mathbb{P}^1(\mathbb{Q}_p)} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f(t),$$

which can be motivated by the formal computation

$$\begin{aligned}\int_{\mathbb{P}^1(\mathbb{Q}_p)} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f(t) &= \int_{\mathbb{P}^1(\mathbb{Q}_p)} \log\left(\frac{\tau_2-t}{\tau_1-t}\right) d\mu_f(t) \\ &= \int_{\mathbb{P}^1(\mathbb{Q}_p)} \int \frac{dz}{z-t}\Big|_{\tau_1}^{\tau_2} d\mu_f(t) \\ &= \int_{\tau_1}^{\tau_2} \int_{\mathbb{P}^1(\mathbb{Q}_p)} \frac{1}{z-t} d\mu_f(t) dz \\ &= \int_{\tau_1}^{\tau_2} f(z)dz.\end{aligned}$$

Note that this line integral has the expected property

$$\int_{\tau_1}^{\tau_2} f(z)dz + \int_{\tau_2}^{\tau_3} f(z)dz = \int_{\tau_1}^{\tau_3} f(z)dz \text{ for all } \tau_1,\, \tau_2,\, \tau_3 \in \mathcal{H}_p,$$

which follows from linearity of the integral operator and the fact that the logarithm is a homomorphism.

4.3.2 Modular Forms on $\Gamma_{p,M}$

In order to extend the notion of modular forms, we need to mention the *Bruhat-Tits* tree of $\mathbf{GL}_2(\mathbb{Q}_p)$, which we denote by $\mathcal{T}$. The set of vertices, denoted $\mathcal{T}_0$, is comprised of $\mathbb{Z}_p$ lattices $\Lambda \subseteq \mathbb{Q}_p^2$ such that $\Lambda \otimes \mathbb{Q}_p = \mathbb{Q}_p^2$, where we identify two such lattices Λ_1 and Λ_2 if there exists $\alpha \in \mathbb{Q}_p$ such that $\alpha\Lambda_1 = \Lambda_2$. Two vertices, v_1 and v_2, are connected if there exist representatives Λ_1 and Λ_2 of v_1 and v_2 such that $p\Lambda_1 \subset \Lambda_2 \subset \Lambda_1$, where the both inclusions are proper. Note that this implies that $p\Lambda_2 \subset p\Lambda_1 \subset \Lambda_2$, so the graph is indeed undirected. The set of unordered edges is denoted $\mathcal{T}_1$. An ordered edge e is an ordered pair of adjacent vertices (v_1, v_2). The *source* of e, denoted $s(e)$ is v_1 and the *target*, denoted $t(e)$ is v_2. Let $\mathcal{E}(\mathcal{T})$ denote the set of ordered edges of $\mathcal{T}$. For each oriented edge e, let $\bar{e}$ denote the edge obtained by

interchanging the source and the target of e. The tree $\mathcal{T}$ is regular of degree $p+1$. The class of $\mathbb{Z}_p^2$ is denoted v_0 and gives us a distinguished vertex of $\mathcal{T}$ which lets us see it as a rooted tree. The $p+1$ adjacent vertices of v_0 are index p sublattices of $\mathbb{Z}_p^2$, which can be put in bijection with $\mathbb{P}^1(\mathbb{F}_p)$ via $k \mapsto \{(a,b) \in \mathbb{Z}_p^2 \text{ such that } ak + b \in p\mathbb{Z}_p\}$ (for $k = \infty$ this amounts to $p\mathbb{Z}_p \times \mathbb{Z}_p$). The edges connecting v_0 to each of its neighbors are labeled $e_0, e_1, \ldots, e_{p-1}, e_\infty$. The action of $\mathbf{GL}_2(\mathbb{Q}_p)$ on $\mathcal{T}$ induces a graph automorphism.

Let $\Gamma = \Gamma_{p,M}$ as above. A cusp form of weight 2 for Γ is a function

$$f \colon \mathcal{E}(\mathcal{T}) \times \mathcal{H} \longrightarrow \mathbb{C}$$

satisfying the following three properties:

(1) $f(\gamma e, \gamma\tau) = (c\tau + d)^2 f(e,\tau)$ for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$.

(2) For each vertex $v \in T_0$ we have

$$\sum_{s(e)=v} f(e,\tau) = 0,$$

and for each ordered edge $e \in \mathcal{E}(\mathcal{T})$ we have $f(\bar{e},\tau) = -f(e,\tau)$.

(3) For each fixed oriented edge $e \in \mathcal{E}(\mathcal{T})$ the function $f(e,\tau)$, denoted $f_e(\tau)$, is a weight 2 cusp form in the usual sense for the stabilizer of e in Γ. This group is denoted Γ_e.

The space of weight 2 cusp forms for Γ is denoted by $S_2(\mathcal{T},\Gamma)$. The action of Γ on $\mathcal{T}$ has only two orbits in $\mathcal{T}_0$ and one orbit in $\mathcal{T}_1$. The stabilizer of e_0 is $\Gamma_0(N)$ and the stabilizer of v_0 is $\Gamma_0(M)$. The properties spelled out above imply that f is determined

by $f_0 := f_{e_0}$, which is a weight 2 cusp form for $\Gamma_0(N)$. We have the following very important lemma, which is Lemma 9.2 in [Dar04].

Lemma 4.6. *The map*

$$\begin{aligned} S_2(\mathcal{T},\Gamma) &\longrightarrow S_2(\Gamma_0(N)) \\ f &\longmapsto f_0 \end{aligned}$$

is injective. Furthermore, the image is the p-new part of $S_2(\Gamma_0(N))$.

4.3.3 Measures associated to an Elliptic Curve and the double integrals

In particular, Lemma 4.6 implies that to an elliptic curve E of conductor N there is an associated weight 2 cusp form $f \in S_2(\mathcal{T},\Gamma)$ such that f_0 is the usual normalized weight 2 eigenform associated to E.

Every edge $e \in \mathcal{T}_1$ has an associated annulus in $\mathcal{H}_p$, whose closure in $\mathbb{P}^1(\mathbb{C}_p)$ can be intersected with $\mathbb{P}^1(\mathbb{Q}_p)$ to give a compact open subset of $\mathbb{P}^1(\mathbb{Q}_p)$ denoted by U_e (see Theorem 5.9 in [Dar04]). For $x, y \in \mathbb{P}^1(\mathbb{Q})$ we define the complex distribution $\tilde{\mu}_f\{x \to y\}$ by

$$\tilde{\mu}_f\{x \to y\}(U_e) = c \cdot 2\pi i \int_x^y f_e(z)dz,$$

where c is the Manin constant associated to E. These values can all be expressed in terms of the periods of f_0, which lie on a lattice $\Lambda_E \subseteq \mathbb{C}$. This lattice can be generated by a real period Ω_+ and a purely imaginary period Ω_- (or contained with index 2 in such a lattice, so Ω_+ and Ω_- are half-periods). We can write

$$c \cdot 2\pi i \int_x^y f_e(z)dz = \kappa_f^+\{x \to y\}(e) \cdot \Omega_+ + \kappa_f^-\{x \to y\}(e) \cdot \Omega_-,$$

where $\kappa_f^{\pm}\{x \to y\}$ takes on integral values when fed an oriented edge $e \in \mathcal{E}(\mathcal{T})$. Upon a choice of sign $w_\infty = \pm 1$, we obtained an integral measure

$$\mu_f\{x \to y\}(U_e) = \kappa_f^{w_\infty}\{x \to y\}(e).$$

Since this measure takes on integral values, it can be seen as a p-adic measure on $\mathbb{P}^1(\mathbb{Q}_p)$. Using the line integral defined in the previous section and this p-adic distribution we define the *double integral* attached to f from τ_1 to τ_2, in $\mathcal{H}_p$, and from x to y, in $\mathbb{P}^1(\mathbb{Q})$, as

$$\int_{\tau_1}^{\tau_2}\int_x^y \omega_f = \int_{\mathbb{P}^1(\mathbb{Q}_p)} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{x \to y\}(t),$$

which is a well defined number in $\mathbb{C}_p$. In virtue of the integrality of the values of the measure, we define its multiplicative counterpart by formally exponentiating, ultimately resulting in a finer invariant (which can be recovered by taking logarithms, and the logarithm is not an injective function).

Definition 4.7. For τ_1 and τ_2 in $\mathcal{H}_p$ and x and y in $\mathbb{P}^1(\mathbb{Q})$, the multiplicative double integral attached to f is defined as

$$\prod_{\tau_1}^{\tau_2}\int_x^y \omega_f = \prod_{\mathbb{P}^1(\mathbb{Q}_p)} \left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{x \to y\}(t) = \lim \prod_\alpha \left(\frac{t-\tau_2}{t-\tau_1}\right)^{\mu_f\{x\to y\}(U_\alpha)},$$

where $\{U_\alpha\}_\alpha$ is a disjoint cover of $\mathbb{P}^1(\mathbb{Q}_p)$ by compact open subsets, the limit is taken over increasingly finer such covers and $t_\alpha \in U_\alpha$ is any sample point.

From the additivity of the p-adic line integral, we can expect a multiplicativity property in the outer integral. From the additivity of the system of measures, which comes from complex line integrals, we can expect a multiplicativity

property in the inner integral. From the Γ-invariance property in $S_2(\mathcal{T},\Gamma)$, as $f(\gamma e,\gamma z)=(cz+d)^2 f(e,z)$ can be rephrased as $f(\gamma e,\gamma z)d(\gamma z)=f(e,z)dz$, we can expect a Γ-invariance property. Likewise, the additive double integral has analogous properties. More explicitly, we have the following lemma:

Lemma 4.8. *For all $\tau_1,\tau_2,\tau_3\in\mathcal{H}_p$, $x,y,z\in\mathbb{P}^1(\mathbb{Q})$ and $\gamma\in\Gamma$, the double integrals satisfy*

(1)

$$\oint_{\tau_1}^{\tau_3}\int_x^y \omega_f=\oint_{\tau_1}^{\tau_2}\int_x^y \omega_f\cdot\oint_{\tau_2}^{\tau_3}\int_x^y \omega_f \quad \textit{and} \quad \int_{\tau_1}^{\tau_3}\int_x^y \omega_f=\int_{\tau_1}^{\tau_2}\int_x^y \omega_f+\int_{\tau_2}^{\tau_3}\int_x^y \omega_f$$

(2)

$$\oint_{\tau_1}^{\tau_2}\int_x^z \omega_f=\oint_{\tau_1}^{\tau_2}\int_x^y \omega_f\cdot\oint_{\tau_1}^{\tau_2}\int_y^z \omega_f \quad \textit{and} \quad \int_{\tau_1}^{\tau_2}\int_x^z \omega_f=\int_{\tau_1}^{\tau_2}\int_x^y \omega_f+\int_{\tau_1}^{\tau_2}\int_y^z \omega_f$$

(3)

$$\oint_{\gamma\tau_1}^{\gamma\tau_2}\int_{\gamma x}^{\gamma y} \omega_f=\oint_{\tau_1}^{\tau_2}\int_x^y \omega_f \quad \textit{and} \quad \int_{\gamma\tau_1}^{\gamma\tau_2}\int_{\gamma x}^{\gamma y} \omega_f=\int_{\tau_1}^{\tau_2}\int_x^y \omega_f$$

Let $\mathbb{Q}_{p^2}$ be the quadratic unramified extension of $\mathbb{Q}_p$ and $\mathcal{O}$ its ring of integers. The reduction map of the beginning of the section can be restricted to $\mathbb{P}^1(\mathbb{Q}_{p^2})$ and its image falls in $\mathbb{P}^1(\mathbb{F}_{p^2})$. Let $\mathcal{H}_p^0=\{\tau\in\mathbb{P}^1(\mathbb{Q}_{p^2})$ such that $\mathrm{red}(\tau)\notin\mathbb{P}^1(\mathbb{F}_p)\}$. When $\tau_1,\tau_2\in\mathcal{H}_p^0$, they must have valuation 0. If $t\in\mathbb{Q}_p$ also has valuation 0, $(t-\tau_1)$ and $(t-\tau_2)$ do as well (otherwise, $\mathrm{red}(\tau)\in\mathbb{P}^1(\mathbb{F}_p)$), so the quotient $(t-\tau_2)/(t-\tau_1)$ does too. If t has positive valuation, clearly numerator and denominator have valuation 0, so the quotient has valuation 0 again. If t has negative valuation, each term has the same valuation and hence the quotient has valuation 0. This implies that all the terms in the Riemann product associated to the multiplicative double integral when

τ_1 and τ_1 are in $\mathcal{H}_p^0$ have valuation 0, hence, lie in $\mathcal{O}^\times$. Hence, we can conclude that

$$\not\int_{\tau_1}^{\tau_2} \int_x^y \omega_f \in \mathcal{O}^\times.$$

Moreover, to compute it to an accuracy of $p^{-\alpha}$ it suffices to compute

$$\not\int_{\tau_1}^{\tau_2} \int_x^y \omega_f \pmod{p}$$

and

$$\int_{\tau_1}^{\tau_2} \int_x^y \omega_f \pmod{p^\alpha}$$

This is Lemma 1.5 in [DP06]. [DP06] also shows how to compute these integrals in the case of $M = 1$. The procedure described there consists of splitting the integral at ∞, using the additivity, and computing each integral separately. This is,

$$\int_{\tau_1}^{\tau_2} \int_x^y \omega_f = \int_{\tau_1}^{\tau_2} \int_\infty^y \omega_f - \int_{\tau_1}^{\tau_2} \int_\infty^x \omega_f,$$

reducing this to the computation of integrals of the form

$$\int_{\tau_1}^{\tau_2} \int_\infty^x \omega_f,$$

where $x \in \mathbb{Q}$. The sequence of convergents of x obtained from its continued fraction expansion yields rational numbers p_j/q_j such that the matrix

$$M_j = \begin{pmatrix} (-1)^{j-1} p_j & p_{j-1} \\ (-1)^{j-1} q_j & q_{j-1} \end{pmatrix}$$

is in $\mathbf{SL}_2(\mathbb{Z})$, where $p_{-1} = 1$ and $q_{-1} = 0$. Hence,

$$\int_{\tau_1}^{\tau_2}\int_{\infty}^{x}\omega_f = \sum_{j=0}^{k}\int_{\tau_1}^{\tau_2}\int_{p_{j-1}/q_{j-1}}^{p_j/q_j}\omega_f = \sum_{j=0}^{k}\int_{\tau_1}^{\tau_2}\int_{M_j(0)}^{M_j(\infty)}\omega_f = \sum_{j=0}^{k}\int_{M_j^{-1}\tau_1}^{M_j^{-1}\tau_2}\int_{0}^{\infty}\omega_f,$$

reducing the problem to the computation of integrals of the form

$$\int_{\tau_1}^{\tau_2}\int_{0}^{\infty}\omega_f$$

for any $\tau_1, \tau_2 \in \mathcal{H}_p^0$, as $\mathcal{H}_p^0$ is stable under the action of $\mathbf{SL}_2(\mathbb{Z})$. It might be necessary to compute more integrals if the level is not p, as not every M_j will lie in Γ. In this case, right coset representatives of Γ in $\mathbf{SL}_2(\mathbb{Z}[1/p])$ can be chosen and the computation is reduced to computing finitely many integrals instead of just one.

Define

$$J_\infty(\tau_1, \tau_2) = \int_{\mathbb{P}^1(\mathbb{Q}_p)-\mathbb{Z}_p} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{0 \to \infty\}(t)$$

and

$$J_a(\tau_1, \tau_2) = \int_{a+p\mathbb{Z}_p} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{0 \to \infty\}(t),$$

so

$$\int_{\tau_1}^{\tau_2}\int_{0}^{\infty}\omega_f \ = J_\infty(\tau_1, \tau_2) + \sum_{a=0}^{p-1} J_a(\tau_1, \tau_2).$$

Using the power series of the logarithm, each of these integrals can be easily expressed as the difference of a power series in τ_1 and a power series in τ_2 (centered at a, or at 0 in the case of ∞) whose coefficients are the moments of the measure $\mu_f\{0 \to \infty\}$ (divided by n, where n is the exponent of τ_1 or τ_2). These can be effectively computed via the so-called overconvergent modular symbols. See the end

of Section 1.3 in [DP06] for details on how to obtain these power series and Chapter 2 for details about the overconvergent modular symbols.

4.3.4 Tate's uniformization

This subsection mainly follows [Sil94], chapters 1 and 5.

Weierstass uniformization (Equation (3.1)) assigns to a complex number z a point in the elliptic curve $E(\mathbb{C})$ via the quotient $\mathbb{C}/\Lambda$, where Λ is a rank two $\mathbb{Z}$-module. In the p-adic realm, rank two $\mathbb{Z}$-modules are dense, so we do not have the same lattices. However, if we exponentiate first, we can salvage the situation. In the complex case, the map $\mathbb{C} \longrightarrow \mathbb{C}^\times$ given by exponentiation (after multiplying by $2\pi i$) has $\mathbb{Z}$ as kernel. Let $q = e^{2\pi i\tau}$ and consider the lattice $\Lambda_\tau \subseteq \mathbb{C}$. Denote by $q^{\mathbb{Z}}$ the multiplicative group generated by q. Now, the composition of the exponential map and projecting into $q^{\mathbb{Z}}$ has as kernel all complex numbers z such that $e^{2\pi iz} = q^n = e^{2\pi in\tau}$ for some $n \in \mathbb{Z}$. This is, $z - n\tau \in \mathbb{Z}$, implying $z \in \Lambda_\tau$, so the composition factors through $\mathbb{C}/\Lambda_\tau$. Since the composition is clearly surjective, the resulting map is actually a group isomorphism between $\mathbb{C}/\Lambda_\tau$ and $\mathbb{C}^\times/q^{\mathbb{Z}}$. We can exploit the isomorphism given by the Weierstrass uniformization to obtain an isomorphism between $\mathbb{C}^\times/q^{\mathbb{Z}}$ and E_τ. To find out what the map from $\mathbb{C}^\times/q^{\mathbb{Z}}$ to $E_\tau(\mathbb{C})$ looks like, we need to understand what $\wp$ and $\wp'$ look like in terms of $u = e^{2\pi iz}$ and q instead of z and τ. It turns out that

$$\frac{1}{(2\pi i)^2}\wp_\tau(z) = \sum_{n\in\mathbb{Z}} \frac{q^n u}{(1-q^n u)^2} + \frac{1}{12} - 2\sum_{n\geq 1}\frac{q^n}{(1-q^n)^2}$$
$$\frac{1}{(2\pi i)^3}\wp'_\tau(z) = \sum_{n\in\mathbb{Z}} \frac{q^n u(1+q^n u)}{(1-q^n u)^3},$$

which is Theorem 6.2, Chapter 1 on [Sil94]. This yields an explicit isomorphism from $\mathbb{C}^\times/q^{\mathbb{Z}}$ to $E_\tau(\mathbb{C})$ given by

$$u \mapsto \left((2\pi i)^2 \left(\sum_{n\in\mathbb{Z}} \frac{q^n u}{(1-q^n u)^2} + \frac{1}{12} - 2\sum_{n\geq 1} \frac{q^n}{(1-q^n)^2} \right), (2\pi i)^3 \sum_{n\in\mathbb{Z}} \frac{q^n u(1+q^n u)}{(1-q^n u)^3}, 1 \right),$$

which is transcendental. Recall that E_τ is given by

$$y^2 = 4x^3 - g_2(\tau)x - g_3(\tau), \tag{4.2}$$

where

$$g_2(\tau) = (2\pi i)^4 \left(\frac{1}{12} + 20\sum_{n\geq 1} \sigma_3(n)q^n \right) = \frac{(2\pi i)^4}{12}\big(1 + 240 s_3(q)\big)$$
$$g_3(\tau) = (2\pi i)^6 \left(\frac{-1}{216} + \frac{7}{3}\sum_{n\geq 1} \sigma_5(n)q^n \right) = \frac{(2\pi i)^6}{216}\big(-1 + 504 s_5(q)\big),$$

where $s_k(q) = \sum_{n\geq 1} \sigma_k(n)q^n$. Let $X = x/(2\pi i)^2$ and $Y = y/(2\pi i)^3$. Dividing Equation (4.2) by $(2\pi i)^6$ we obtain

$$Y^2 = 4X^3 - \frac{1 + 240 s_3(q)}{12} X - \frac{-1 + 504 s_5(q)}{216}.$$

Let x' and y' be such that $Y = 2y' + x'$ and $X = x' + \frac{1}{12}$. Substituting and simplifying, we find that

$$y'^2 + y'x' = x'^3 - 5s_3(q)x' - \frac{5s_3(q) + 7s_5(q)}{12}.$$

Solving for x' and y' we obtain

$$x' = X - \frac{1}{12} = \sum_{n\in\mathbb{Z}} \frac{q^n u}{(1-q^n u)^2} - 2\sum_{n\geq 1} \frac{q^n}{(1-q^n)^2}$$
$$y' = \frac{Y-X}{2} + \frac{1}{24} = \sum_{n\in\mathbb{Z}} \frac{(q^n u)^2}{(1-q^n u)^3} + \sum_{n\geq 1} \frac{q^n}{(1-q^n)^2}.$$

Denote x' and y' by $X_q(u)$ and $Y_q(u)$, respectively. Let

$$a_4(q) = -5s_3(q) \qquad \text{and} \qquad a_6(q) = -\frac{5s_3(q) + 7s_5(q)}{12},$$

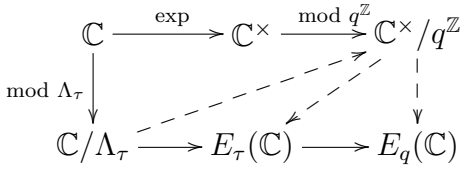
and let

$$E_q : y^2 + xy = x^3 + a_4(q)x + a_6(q).$$

Notice that $5m^3+7m^5 \equiv 0 \pmod{12}$ for all integers m, as from $m^3 \equiv m \pmod 3$ we obtain $5m^3+7m^5 \equiv -m+m \pmod 3$, for even m clearly $5m^3+7m^5 \equiv 0 \pmod{12}$ and for odd m, from $m^2 \equiv 1 \pmod 4$ we obtain $5m^3+7m^5 \equiv 5m+7m \equiv 0 \pmod 4$. This implies that the coefficients of $a_6(q)$ are all integers, whence the coefficients of E_q are in $\mathbb{Z}[[q]]$. E_q is referred to as the *Tate curve*; its discriminant and j-invariant are given by

$$\Delta(q) = q\prod_{n\geq 1}(1-q^n)^{24} \quad \text{and} \quad j(E_q) = j(t) = \frac{1}{q} + 744 + 196884q + \mathcal{O}(q^2),$$

as expected. To summarize, by following the diagram



we have obtained the so-called Tate parametrization, given by

$$\begin{aligned} \mathbb{C}^\times/q^{\mathbb{Z}} &\xrightarrow{\sim} E_q(\mathbb{C}) \\ u &\longmapsto (X_q(u), Y_q(u), 1). \end{aligned}$$

This construction generalizes to other complete fields. This theorem is due to Tate.

Theorem 4.9 (Tate)**.** *Let K be a p-adic field and let $q \in K^\times$ with $|q| < 1$.*

(a) The series defining $a_4(q)$ and $a_6(q)$ converge in K.

(b) The Tate curve E_q is an elliptic curve defined over K with discriminant and j-invariant as above.

(c) $E_q(\bar{K})$ is parametrized via

$$\begin{aligned} \bar{K}^\times/q^n &\xrightarrow{\sim} E_q(\bar{K}) \\ u &\longmapsto (X_q(u), Y_q(u), 1). \end{aligned}$$

(d) If L is an algebraic extension of K then the parametrization above restricts to $L^\times$ in the natural way, i.e.,

$$\begin{aligned} L^\times/q^n &\xrightarrow{\sim} E_q(L) \\ u &\longmapsto (X_q(u), Y_q(u), 1). \end{aligned}$$

Notice that $|j(E_q)| > 1$, as the leading term is $1/q$ and $q < 1$ is necessary to ensure convergence of the series defining E_q. We have this other theorem of Tate

Theorem 4.10 (Tate)**.** *Let K be a p-adic field and let E/K be an elliptic curve with $|j(E)| > 1$. Then there is a unique $q \in K^\times$ with $|q| < 1$ such that E is isomorphic*

to E_q over $\bar{K}$. Moreover, this isomorphism is over K if and only if E has split multiplicative reduction.

4.3.5 The Stark-Heegner point

This section follows Sections 9.4-9.6 of [Dar04].

For E and K as above this section, let q be Tate's p-adic period attached to E. Consider the integral

$$\kappa_{\tau,x}(\gamma_1,\gamma_2) = \oint_\tau^{\gamma_1\tau} \int_{\gamma_1 x}^{\gamma_1\gamma_2 x} \omega_f \pmod{q^{\mathbb{Z}}},$$

seen as a 2-cochain in $C^2(\Gamma, \mathbb{C}_p^\times/q^{\mathbb{Z}})$, where the action of Γ on $\mathbb{C}_p^\times/q^{\mathbb{Z}}$ is trivial. Applying the d operator (mapping 2-cochains to 3-cochains) we obtain

$$\begin{aligned}
d\kappa_{\tau,x}(\gamma_1,\gamma_2,\gamma_3) &= \gamma_1\kappa_\tau(\gamma_2,\gamma_3) \div \kappa_\tau(\gamma_1\gamma_2,\gamma_3)\cdot\kappa_\tau(\gamma_1,\gamma_2\gamma_3) \div \kappa_\tau(\gamma_1,\gamma_2)\\
&= \oint_\tau^{\gamma_2\tau}\int_{\gamma_2 x}^{\gamma_2\gamma_3 x} \omega_f \div \oint_\tau^{\gamma_1\gamma_2\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2\gamma_3 x} \omega_f \cdot \oint_\tau^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1\gamma_2\gamma_3 x} \omega_f \div \oint_\tau^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1\gamma_2 x} \omega_f\\
&= \oint_\tau^{\gamma_2\tau}\int_{\gamma_2 x}^{\gamma_2\gamma_3 x} \omega_f \cdot \oint_{\gamma_1\gamma_2\tau}^{\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2\gamma_3 x} \omega_f \cdot \oint_\tau^{\gamma_1\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2 x} \omega_f \cdot \oint_\tau^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1\gamma_2\gamma_3 x} \omega_f\\
&= \oint_\tau^{\gamma_2\tau}\int_{\gamma_2 x}^{\gamma_2\gamma_3 x} \omega_f \cdot \oint_{\gamma_1\gamma_2\tau}^{\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2\gamma_3 x} \omega_f \cdot \oint_\tau^{\gamma_1\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2\gamma_3 x} \omega_f\\
&= \oint_\tau^{\gamma_2\tau}\int_{\gamma_2 x}^{\gamma_2\gamma_3 x} \omega_f \cdot \oint_{\gamma_1\gamma_2\tau}^{\gamma_1\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2\gamma_3 x} \omega_f\\
&= \oint_\tau^{\gamma_2\tau}\int_{\gamma_2 x}^{\gamma_2\gamma_3 x} \omega_f \cdot \oint_{\gamma_2\tau}^{\tau}\int_{\gamma_2 x}^{\gamma_2\gamma_3 x} \omega_f = 1,
\end{aligned}$$

showing that in fact, κ_τ is a 2-cocycle. A priori, it would seem like this cocyle depends on the base point x, and of τ. Define the 1-cochain

$$\rho_{x,y}(\gamma) = \oint_\tau^{\gamma\tau}\int_{\gamma x}^{\gamma y} \omega_f \pmod{q^{\mathbb{Z}}} \in C^1(\Gamma, \mathbb{C}_p^\times/q^{\mathbb{Z}}).$$

Applying the d operator (mapping 1-cochains to 2-cochains) we obtain

$$\begin{aligned} d\rho_{x,y}(\gamma_1,\gamma_2) &= \gamma_1\rho_{x,y}(\gamma_2) \div \rho_{x,y}(\gamma_1\gamma_2)\cdot\rho_{x,y}(\gamma_1) \\ &= \oint_{\tau}^{\gamma_2\tau}\int_{\gamma_2 x}^{\gamma_2 y}\omega_f \div \oint_{\tau}^{\gamma_1\gamma_2\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2 y}\omega_f \cdot \oint_{\tau}^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1 y}\omega_f \\ &= \oint_{\gamma_1\tau}^{\gamma_1\gamma_2\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2 y}\omega_f \cdot \oint_{\gamma_1\gamma_2\tau}^{\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2 y}\omega_f \cdot \oint_{\tau}^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1 y}\omega_f \\ &= \oint_{\gamma_1\tau}^{\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2 y}\omega_f \cdot \oint_{\tau}^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1 y}\omega_f, \end{aligned}$$

so

$$\begin{aligned} \kappa_{\tau,x}(\gamma_1,\gamma_2) \div \kappa_{\tau,y}(\gamma_1,\gamma_2) &= \oint_{\tau}^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1\gamma_2 x}\omega_f \div \oint_{\tau}^{\gamma_1\tau}\int_{\gamma_1 y}^{\gamma_1\gamma_2 y}\omega_f \\ &= \oint_{\tau}^{\gamma_1\tau}\int_{\gamma_1 x}^{\gamma_1 y}\omega_f \div \oint_{\tau}^{\gamma_1\tau}\int_{\gamma_1\gamma_2 x}^{\gamma_1\gamma_2 y}\omega_f \\ &= d\rho_{x,y}(\gamma_1,\gamma_2), \end{aligned}$$

implying that the class of $\kappa_{\tau,x}$ in $H^2(\Gamma,\mathbb{C}^\times/q^{\mathbb{Z}})$ does not depend on the choice of base point, as they differ by a coboundary. Furthermore, if we define the 1-cochain

$$\rho_{\tau_1,\tau_2}(\gamma) = \oint_{\tau_1}^{\tau_2}\int_{\gamma x}^{x} \pmod{q^{\mathbb{Z}}} \in C^1(\Gamma,\mathbb{C}^\times/q^{\mathbb{Z}}),$$

we have

$$\begin{aligned} d\rho_{\tau_1,\tau_2}(\gamma_1,\gamma_2) &= \gamma_1\rho_{\tau_1,\tau_2}(\gamma_2) \div \rho_{\tau_1,\tau_2}(\gamma_1\gamma_2)\cdot\rho_{\tau_1,\tau_2}(\gamma_1) \\ &= \oint_{\tau_1}^{\tau_2}\int_{\gamma_2 x}^{x}\omega_f \div \oint_{\tau_1}^{\tau_2}\int_{\gamma_1\gamma_2 x}^{x}\omega_f \cdot \oint_{\tau_1}^{\tau_2}\int_{\gamma_1 x}^{x}\omega_f \\ &= \oint_{\tau_1}^{\tau_2}\int_{\gamma_1 x}^{x}\omega_f \cdot \oint_{\tau_1}^{\tau_2}\int_{x}^{\gamma_1\gamma_2 x}\omega_f \cdot \oint_{\tau_1}^{\tau_2}\int_{\gamma_2 x}^{x}\omega_f \\ &= \oint_{\tau_1}^{\tau_2}\int_{\gamma_1 x}^{\gamma_1\gamma_2 x}\omega_f \cdot \oint_{\tau_1}^{\tau_2}\int_{\gamma_2 x}^{x}\omega_f, \end{aligned}$$

so

$$\begin{aligned}\kappa_{\tau_1}(\gamma_1,\gamma_2) \div \kappa_{\tau_2}(\gamma_1,\gamma_2) &= \oint_{\tau_1}^{\gamma_1\tau_1} \int_{\gamma_1 x}^{\gamma_1\gamma_2 x} \omega_f \div \oint_{\tau_2}^{\gamma_1\tau_2} \int_{\gamma_1 x}^{\gamma_1\gamma_2 x} \omega_f \\ &= \oint_{\tau_1}^{\tau_2} \int_{\gamma_1 x}^{\gamma_1\gamma_2 x} \omega_f \div \oint_{\gamma_1\tau_1}^{\gamma_1\tau_2} \int_{\gamma_1 x}^{\gamma_1\gamma_2 x} \omega_f \\ &= \oint_{\tau_1}^{\tau_2} \int_{\gamma_1 x}^{\gamma_1\gamma_2 x} \omega_f \div \oint_{\tau_1}^{\tau_2} \int_{x}^{\gamma_2 x} \omega_f,\end{aligned}$$

whence, the class of κ_τ in $H^2(\Gamma, \mathbb{C}^\times/q^{\mathbb{Z}})$ does not depend on τ either.

Conjecture 4.11. *The class of κ_τ is trivial in $H^2(\Gamma, \mathbb{C}_p^\times/q^{\mathbb{Z}})$.*

In [Dar01] it is shown that Conjecture 4.11 is a refinement of the Exceptional Zero Conjecture (in [MTT86]), which was proven by Greenberg and Stevens in [GS93]. Nevertheless, the vanishing of this cocycle is still conjectural, and corresponds to Conjecture 9.10 in [Dar04]. Henceforth, assume that Conjecture 4.11 holds.

Define $\kappa_\tau^\#(\gamma_1,\gamma_2) = \kappa_\tau(\gamma_2^{-1},\gamma_1^{-1})$, a related 2-cocycle. It is clear that the class of $\kappa_\tau^\#$ in $H^2(\Gamma, \mathbb{C}_p^\times/q^{\mathbb{Z}})$ is trivial if and only if that of κ_τ is, so Conjecture 4.11 implies that the class of $\kappa_\tau^\#$ is trivial.

Denote by $\mathcal{M}$ the group of functions

$$\begin{aligned} m\colon \mathbb{P}^1(\mathbb{Q}) \times \mathbb{P}^1(\mathbb{Q}) &\longrightarrow \quad \mathbb{C}_p^\times/q^{\mathbb{Z}} \\ (x,y) \quad &\longmapsto \quad m\{x \to y\}\end{aligned}$$

such that $m\{x \to y\} \cdot m\{y \to z\} = m\{x \to z\}$ and $m\{y \to x\} = m\{x \to y\}^{-1}$. These are the so-called *modular symbols*. Fix a cusp x and denote by $\mathcal{M}_0$ the group

arising from restriction of modular symbols to $\Gamma x \times \Gamma x$. Let $\mathcal{F} = \mathcal{F}(\mathbb{C}_p^\times/q^{\mathbb{Z}})$ be the group of $\mathbb{C}_p^\times/q^{\mathbb{Z}}$-valued functions on Γx.

Both $\mathcal{M}_0$ and $\mathcal{F}$ have a natural left-action of Γ, given by

$$(\gamma m)\{y \to z\} = m\{\gamma^{-1}y \to \gamma^{-1}z\} \qquad \text{and} \qquad (\gamma g)(y) = g(\gamma^{-1}y).$$

Define maps

$$\mathbb{C}_p^\times/q^{\mathbb{Z}} \xrightarrow{\ \iota\ } \mathcal{F} \xrightarrow{\ \Delta\ } \mathcal{M}_0$$

by

$$\iota(z_p) = (y \mapsto z_p) \qquad \text{and} \qquad \Delta(g)(y,z) = g(y) \div g(z).$$

We can readily check that both, Δ and ι, commute with the action of Γ and that $\Delta(g)$ is indeed a modular symbol. ι is clearly an injection. For any $m \in \mathcal{M}_0$, define $g_m(y) = m\{y \to x\}$. We have that

$$\Delta(g_m)(y,z) = m\{y \to x\} \div m\{z \to x\} = m\{y \to x\} \cdot m\{x \to z\} = m\{y \to z\},$$

which shows that Δ is surjective. Clearly

$$\forall y, z, \Delta(g)(y,z) = 1 \iff \forall x, y, g(y) = g(z) \iff \exists z_p \colon g = \iota(z_p),$$

so $\ker(\Delta) = \mathrm{im}(\iota)$ and we have a short exact sequence

$$0 \longrightarrow \mathbb{C}_p^\times/q^{\mathbb{Z}} \longrightarrow \mathcal{F} \longrightarrow \mathcal{M}_0 \longrightarrow 0.$$

This sequence induces a long exact sequence of cohomology groups, and, in particular, we have the connecting homomorphism

$$\delta\colon H^1(\Gamma, \mathcal{M}_0) \longrightarrow H^2(\Gamma, \mathbb{C}_p^\times/q^{\mathbb{Z}}).$$

Lemma 4.8 shows that the map c_τ from Γ to $\mathcal{M}_0$ defined by

$$c_\tau(\gamma)\{y \to z\} = \oint_\tau^{\gamma\tau} \int_y^z \omega_f$$

is indeed a modular symbol. The d operator applied to c_τ yields the 2-cochain

$$\begin{aligned} dc_\tau(\gamma_1, \gamma_2) &= \gamma_1 c_\tau(\gamma_2) \div c_\tau(\gamma_1\gamma_2) \cdot c_\tau(\gamma_1) \\ &= \oint_\tau^{\gamma_1\tau} \int_y^z \omega_f \cdot \oint_\tau^{\gamma_2\tau} \int_{\gamma_1^{-1}y}^{\gamma_1^{-1}z} \omega_f \div \oint_\tau^{\gamma_1\gamma_2\tau} \int_y^z \omega_f \\ &= \oint_\tau^{\gamma_1\tau} \int_y^z \omega_f \cdot \oint_{\gamma_1\tau}^{\gamma_1\gamma_2\tau} \int_y^z \omega_f \div \oint_\tau^{\gamma_1\gamma_2\tau} \int_y^z \omega_f \\ &= \oint_\tau^{\gamma_1\gamma_2\tau} \int_y^z \omega_f \div \oint_\tau^{\gamma_1\gamma_2\tau} \int_y^z \omega_f = 1 \end{aligned}$$

implying that c_τ is a 1-cocycle, and as such, it sits naturally in $H^1(\Gamma, \mathcal{M}_0)$. The slightly cumbersome computation of δ applied to the class of c_τ turns out to be $\kappa_\tau^\#$ (see the comments preceding Conjecture 9.14 in [Dar04]). Since the class of $\kappa_\tau^\#$ is trivial in $H^2(\Gamma, \mathbb{C}_p^\times/q^{\mathbb{Z}})$, it is not entirely unreasonable to expect the class of c_τ in $H^1(\Gamma, \mathcal{M}_0)$ to be trivial as well. We have the following strengthening of Conjecture 4.11.

Conjecture 4.12. *The class of c_τ is trivial in $H^1(\Gamma, \mathcal{M}_0)$.*

Conjecture 4.12 implies that c_τ is a coboundary, so, there exists a 0-cochain (i.e., an element of $\mathcal{M}_0$) $\tilde{\eta}_\tau$ such that $c_\tau = d\tilde{\eta}_\tau$, i.e.,

$$\int_\tau^{\gamma\tau} \int_y^z \omega_f = c_\tau(\gamma)\{y \to z\} = \tilde{\eta}_\tau\{\gamma^{-1}y \to \gamma^{-1}z\} \div \tilde{\eta}_\tau\{y \to z\}. \tag{4.3}$$

The 0-cochain $\tilde{\eta}_\tau$ can be multiplied by any 0-cocycle without affecting the property mentioned in the equation above. The set $Z^0(\Gamma, \mathcal{M}_0)$ of 0-cocycles is precisely the set of modular symbols in $\mathcal{M}_0$ which are invariant under the action of Γ, which we denote $\mathcal{M}_0^\Gamma$. Consider the group homomorphism

$$\begin{aligned} h\colon \mathcal{M}_0^\Gamma &\longrightarrow \quad \mathrm{Hom}(\Gamma, \mathbb{C}_p^\times/q^{\mathbb{Z}}) \\ m &\longmapsto (\gamma \mapsto m\{x \to \gamma x\}). \end{aligned}$$

Verifying that $h(m)$ is indeed a homomorphism amounts to show that

$$\begin{aligned} h(m)(\gamma_1\gamma_2) &= m\{x \to \gamma_1\gamma_2 x\} = m\{x \to \gamma_1 x\} \cdot m\{\gamma_1 x \to \gamma_1\gamma_2 x\} \\ &= m\{x \to \gamma_1 x\} \cdot m\{x \to \gamma_2 x\} = h(m)(\gamma_1) \cdot h(m)(\gamma_2). \end{aligned}$$

The restriction of a modular symbol m to $\Gamma x \times \Gamma x$ is completely determined by $h(m)$, as for all $y, z \in \Gamma x$ there exist $\gamma_1, \gamma_2 \in \Gamma$ such that $y = \gamma_1 x$ and $z = \gamma_2 x$, and

$$m\{y \to z\} = m\{\gamma_1 x \to \gamma_2 x\} = m\{x \to \gamma_2 x\} \div m\{x \to \gamma_1 x\},$$

whence h is an injection.

Notice that $h(m)$ does not depend on the cusp x. If we choose any other cusp $y \in \mathbb{P}^1(\mathbb{Q})$, we have that

$$\begin{aligned} m\{y \to \gamma y\} &= m\{y \to x\} \cdot m\{x \to \gamma x\} \cdot m\{\gamma x \to \gamma y\} \\ &= m\{y \to x\} \cdot m\{x \to \gamma x\} \cdot m\{x \to y\} \\ &= m\{y \to x\} \cdot m\{x \to \gamma x\} \div m\{y \to x\} = m\{x \to \gamma x\}. \end{aligned}$$

Since the target of $h(m)$ is abelian, $h(m)$ vanishes at every commutator. Furthermore, if γ is any element of Γ such that $\gamma y = y$ for some cusp $y \in \mathbb{P}^1(\mathbb{Q})$, we have that $h(m)(\gamma) = m\{y \to \gamma y\} = m\{y \to y\} = 1$, so $h(m)$ also vanishes at every such γ as well.

Let $\Gamma' \subseteq \Gamma$ be the normal closure of the subset comprised of all the commutators and the elements that fix cusps, as above. Then, $\Gamma' \subseteq \ker(h(m))$, so we have the following commutative diagram

$$\begin{array}{ccc} \Gamma & \xrightarrow{h(m)} & \mathbb{C}_p^\times/q^{\mathbb{Z}} \\ {\scriptstyle \pi_\Gamma}\downarrow & \nearrow_{h'(m)} & \\ \Gamma/\Gamma' & & \end{array}$$

where $h(m) = h'(m) \circ \pi_\Gamma$.

Lemma 4.13. *The map*

$$\begin{aligned} h' \colon \mathcal{M}_0^\Gamma &\longrightarrow \quad \mathrm{Hom}(\Gamma/\Gamma', \mathbb{C}_p^\times/q^{\mathbb{Z}}) \\ m &\longmapsto (\gamma\Gamma' \mapsto m\{x \to \gamma x\}) \end{aligned}$$

is a monomorphism.

Proof. See the comments leading to this statement. □

Lemma 4.14. *The group* Γ' *is of finite index in* Γ.

Proof. See Lemma 3.5 in [Dar01], Theorem 2 in [Men67] or Theorem 3 in [Ser70]. □

Let e_Γ be the exponent of Γ/Γ', which is a finite number by Lemma 4.14. Then, Lemma 4.13 implies that every modular symbol in $\mathcal{M}_0^\Gamma$ is annihilated when raised to the e_Γ-th power, whence the modular symbol $\tilde{\eta}_\tau^{e_\Gamma}$ does not depend on the choice of $\tilde{\eta}_\tau$ satisfying Equation (4.3). Putting all of the above together, we obtain the following map, whose existence is conjectural due to the assumption of Conjecture 4.12.

Conjecture 4.15. *There exists a unique function*

$$\begin{array}{ccc} \mathcal{H}_p(\mathbb{C}_p) \times \Gamma x \times \Gamma x \longrightarrow & & \mathbb{C}_p/q^{\mathbb{Z}} \\ (\tau, r, s) & \longmapsto & \not\int_r^\tau \int_r^s e_\Gamma \omega_f, \end{array}$$

such that, for all τ_1, $\tau_2 \in \mathcal{H}_p(\mathbb{C}_p)$, r, s, $t \in \Gamma x$ *and* $\gamma \in \Gamma$, *we have*

(1)

$$\not\int_r^\tau \int_r^s e_\Gamma \omega_f \times \not\int_s^\tau \int_s^t e_\Gamma \omega_f = \not\int_r^\tau \int_r^t e_\Gamma \omega_f$$

(2)

$$\not\int_r^{\tau_2} \int_r^s e_\Gamma \omega_f \div \not\int_r^{\tau_1} \int_r^s e_\Gamma \omega_f = \left(\not\int_{\tau_1}^{\tau_2} \int_r^s \omega_f \right)^{e_\Gamma}$$

(3)

$$\not\int_{\gamma r}^{\gamma \tau} \int_{\gamma r}^{\gamma s} e_\Gamma \omega_f = \not\int_r^\tau \int_r^s e_\Gamma \omega_f$$

This map is referred to as the semi-indefinite integral.

Let $K = \mathbb{Q}(\sqrt{D})$ (where $D > 0$ and D is 0 or 1 modulo 4) and let $\tau \in \mathcal{H}_p \cap K$. Let $M_0(M)[1/p]$ be the ring of 2×2 matrices with entries in $\mathbb{Z}[1/p]$ which are upper triangular modulo M. Note that $\Gamma = \{\gamma \in M_0(M)[1/p] \colon \det(\gamma) = 1\}$. Let $\mathcal{O}_\tau \subseteq M_0(M)[1/p]$ be the order of matrices which leave τ invariant under the usual action. For this order to be non-trivial, we need every prime dividing M to split or ramify in K. Since N and the discriminant of K are relatively prime, we need every prime factor of M to split in K. Every matrix in $\mathcal{O}_\tau$ has as eigenvector the column vector $(\tau, 1)$ and it can be identified with its eigenvalue, giving an isomorphism to a $\mathbb{Z}[1/p]$-order in K, playing the role of CM. The units of norm one of $\mathcal{O}_\tau$ form the stabilizer of τ in Γ. The set of units of a $\mathbb{Z}[1/p]$-order in K has rank one, so we can find a generator, γ_τ of $\mathcal{O}_{\tau,1}^\times / \langle \pm 1 \rangle$.

For each τ we can define $P_\tau \in E(\mathbb{Q}_{p^2})$ as follows: The semi-indefinite integral

$$\not\int_\tau \int_r^{\gamma_\tau r} e_\Gamma \omega_f$$

depends only on τ, as

$$\begin{aligned} \not\int_\tau \int_r^{\gamma_\tau r} e_\Gamma \omega_f \div \not\int_\tau \int_s^{\gamma_\tau s} e_\Gamma \omega_f &= \not\int_\tau \int_r^s e_\Gamma \omega_f \div \not\int_\tau \int_{\gamma_\tau r}^{\gamma_\tau s} e_\Gamma \omega_f \\ &= \not\int_\tau \int_r^s e_\Gamma \omega_f \div \not\int_{\gamma_\tau^{-1}\tau}^{s} \int_r^s e_\Gamma \omega_f = 1, \end{aligned}$$

since $\gamma_\tau^{-1}\tau = \tau$. The image of the semi-indefinite integral is $\mathbb{Q}_{p^2}^\times / q^{\mathbb{Z}}$, so it can be mapped to a point $P_\tau \in E(\mathbb{Q}_{p^2})$ via Tate's uniformization.

Conjecture 4.16 (Darmon)**.** *Let τ and $\mathcal{O}_\tau$ as above. Let H^+ denote the narrow ring class field of K attached to $\mathcal{O}_\tau$. Then $P_\tau \in E(H^+)$.*

This conjecture has been tested numerically in many cases, but current tools are still unable to provide theoretical reasons besides the analogies followed by Darmon to formulate it. Many examples and experimental verifications can be found in [DP06] and [GM15].

4.3.6 Computational remarks

As with the double integral, this computation can be reduced to computing integrals of the form

$$\fint^{\tau} \int_0^{\infty} n\omega_f$$

when $M = 1$ (and finitely many ones, coming from the right cosets of Γ in $\mathbf{SL}_2(\mathbb{Z}[1/p])$). Again, for $M = 1$, the computation of the semi-indefinite integral can be expressed in terms of double integrals by

$$\begin{aligned}\fint^{\tau} \int_0^{\infty} n\omega_f &= \fint^{\tau} \int_0^{1} n\omega_f \times \fint^{\tau} \int_1^{\infty} n\omega_f = \fint^{-1/\tau} \int_{\infty}^{-1} n\omega_f \times \fint^{\tau-1} \int_0^{\infty} n\omega_f \\ &= \fint^{\tau-1} \int_0^{\infty} n\omega_f \times \fint^{1-1/\tau} \int_{\infty}^{0} n\omega_f = \fint_{1-1/\tau}^{\tau-1} \int_0^{\infty} n\omega_f.\end{aligned}$$

In the case of $M > 1$, Guitart and Masdeu in [GM15] developed a different technique for the computation of some of these integrals. It will be explained with some detail in the next chapter.

4.4 Heegner points attached to Cartan Non-Split curves

This section is entirely based on [KP14]. It deals with an extension of the Heegner hypothesis to cases where there are no Heegner points on $X_0(N)$. The original setting, as described in section 4.2, is to take E an elliptic curve of conductor N and K a quadratic imaginary field such that all primes dividing N split in K. When N

is square-free, this implies that the sign of the functional equation of $L(E/K, s)$ is -1, but when N is not square-free, this implication is not quite correct.

Consider an elliptic curve E of conductor p^2 (p odd), where p is inert in K. The sign of $L(E/K, s)$ is still -1 but the curve $X_0(p^2)$ will not have Heegner points, so we cannot hope to push them through the classical modular parametrization to obtain points in E. However, there is a modular curve, the so-called *Cartan Non-split* curve, where we might find Heegner points.

4.4.1 Cartan Non-split curves

Let ε be a quadratic nonresidue modulo p. The Cartan non-split open modular curve of level p associated to ε, denoted $Y_{ns}^{\varepsilon}(p)$, has the following modular interpretation: it classifies elliptic curves together with an $\mathbb{F}_p$-linear endomorphism of the p torsion whose square is multiplication by ε. Two such pairs (E, ϕ) and (E', ϕ') are said to be equivalent if there exists an isogeny $\psi\colon E \longrightarrow E'$ such that the diagram

$$\begin{array}{ccc} E[p] & \xrightarrow{\phi} & E[p] \\ {\scriptstyle\psi}\downarrow & & \downarrow{\scriptstyle\psi} \\ E'[p] & \xrightarrow{\phi'} & E'[p] \end{array}$$

commutes. The *normalizer* of the Cartan Non-split of level p, denoted $Y_{ns}^{+}(p)$ classifies the same pairs, but the diagram can either commute or anticommute (i.e., $\psi\phi = \pm\phi'\psi$ as opposed to $\psi\phi = \phi'\psi$). The compactifications obtained by adding the cusps are denoted by $X_{ns}^{\varepsilon}(p)$ and $X_{ns}^{+}(p)$, respectively. An alternate moduli interpretation can be found in [RW14].

The map $(E, \phi) \mapsto (E, -\phi)$ is an involution on $X_{ns}^{\varepsilon}(p)$ whose fixed points can be identified with $X_{ns}^{+}(p)$.

For computations, it is useful to have the complex model of the curve. Let $M_{ns}^{\varepsilon}(p)$ be the ring of 2×2 matrices with integer entries such that

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix} \equiv \begin{pmatrix} a & b \\ b\varepsilon & a \end{pmatrix} \pmod{p}.$$

Its normalizer, denoted by $M_{ns}^{+}(p)$, also contains the 2×2 matrices such that

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix} \equiv \begin{pmatrix} a & b \\ -b\varepsilon & -a \end{pmatrix} \pmod{p}.$$

The groups $\Gamma_{ns}^{\varepsilon}(p)$ and $\Gamma_{ns}^{+}(p)$ are the elements in $M_{ns}^{\varepsilon}(p)$ and $M_{ns}^{+}(p)$ of determinant 1, respectively. Both groups contain $\Gamma(p)$, so they are level p congruence subgroups. The normalizer of $\Gamma_{ns}^{\varepsilon}(p)$ in $\mathbf{SL}_2(\mathbb{Z})$ is precisely $\Gamma_{ns}^{+}(p)$, justifying the terminology.

The complex points of the Cartan Non-split and its normalizer can be identified with quotients of the complex upper half-plane,

$$Y_{ns}^{\varepsilon}(p)(\mathbb{C}) = \Gamma_{ns}^{\varepsilon}(p)\backslash\mathcal{H} \qquad \text{and} \qquad Y_{ns}^{+}(p)(\mathbb{C}) = \Gamma_{ns}^{+}(p)\backslash\mathcal{H},$$

as follows:

For $\tau \in \mathcal{H}$ associate the pair (E_τ, ϕ_τ), where E_τ is the usual elliptic curve associated to $\tau \in \mathcal{H}$ and ϕ_τ is the $\mathbb{F}_p$-endomorphism of $E_\tau[p]$ whose matrix in the basis $\{1/p, \tau/p\}$ is $\left(\begin{smallmatrix} 0 & 1 \\ \varepsilon & 0 \end{smallmatrix}\right)$.

Notice that $(E_\tau, \phi_\tau) \sim (E_{\tau'}, \phi_{\tau'})$ if and only if there exists a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbf{SL}_2(\mathbb{Z})$ such that the induced isomorphism $\psi\colon E_\tau \longrightarrow E_{\tau'}$ when restricted to the p-torsion satisfies $\phi_{\tau'}\psi = \psi\phi_\tau$ (resp. $\pm\psi\phi_\tau$ if we are considering the normalizer).

Since

$$\psi\left(\phi_\tau\left(\frac{1}{p}\right)\right) = \psi\left(\frac{\tau\varepsilon}{p}\right) = \frac{(a\tau'+b)\varepsilon}{p}$$
$$\phi_{\tau'}\left(\psi\left(\frac{1}{p}\right)\right) = \phi_{\tau'}\left(\frac{c\tau'+d\varepsilon}{p}\right) = \frac{c+d\tau'\varepsilon}{p},$$

the two points will be equivalent in $Y^{\epsilon}_{ns}(p)(\mathbb{C})$ (resp. $Y^{+}_{ns}(p)(\mathbb{C})$) if and only if

$$d \equiv \quad a \pmod{p} \quad \text{and} \quad c \equiv \quad b\varepsilon \pmod{p}$$
$$(\text{resp. or } d \equiv -a \pmod{p} \quad \text{and} \quad c \equiv -b\varepsilon \pmod{p}),$$

which shows $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma^{\varepsilon}_{ns}(p)$ (resp. $\Gamma^{+}_{ns}(p)$), as expected.

The existence of a pair of the form (E_τ, ϕ_τ) for any $(E/\mathbb{C}, \phi)$ is a bit more subtle. The details can be found in [KP14].

For $a = (r, s) \in \mathbb{Q}^2$ and $\tau \in \mathcal{H}$, let

$$f_a(\tau) = \frac{g_2(\tau)g_3(\tau)}{\Delta(\tau)}\wp_\tau(r\tau + s),$$

which satisfies the transformation property $f_a(\gamma z) = f_{a\gamma}(z)$ for every $\gamma \in \mathbf{SL}_2(\mathbb{Z})$. The Λ_τ-periodicity of $\wp_\tau$ ensures that if $(r', s') \equiv (r, s) \pmod 1$ then $f_{(r,s)} = f_{(r',s')}$. If $pa \in \mathbb{Z}^2$ and $\gamma \in \Gamma(p)$, we can see that $a \equiv a\gamma \pmod 1$, so f_a becomes invariant under the action of $\Gamma(p)$. Proposition 7.5.1 in [DS05] states that the function field of $X(p)$ over $\mathbb{C}$ is comprised of the j-invariant together with all the f_a such that $pa \in \mathbb{Z}^2$ (and we can choose a finite subset by the comments above). Choose representatives β_i for $\pm\Gamma(p)\backslash\Gamma^{\varepsilon}_{ns}(p)$. The function field of $X^{\varepsilon}_{ns}(p)$ is then

$$\mathbb{C}(X^{\varepsilon}_{ns}(p)) = \mathbb{C}\left(j, \sum_i f_{a\beta_i}\right),$$

where a runs over the finite set above.

An important difference with the modular curve for $\Gamma_0(N)$ is that the cusps are rational there, while here they are defined over the cyclotomic extension $\mathbb{Q}(\xi_p)$, where $\xi_p^p = 1$. There are $p-1$ cusps and they are all Galois conjugates of each other. For the normalizer, there are just $(p-1)/2$, they are defined over $\mathbb{Q}(\xi + \xi^{-1})$, and they are also Galois conjugates of each other. (See [Ser97], Appendix 5, or [BB12] Section 6.1.)

4.4.2 Modular Forms over $\Gamma_{ns}^{\varepsilon}(p)$

Let $\alpha_p = \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$. Conjugating $\Gamma(p)$ by α_p gives an isomorphism

$$\Gamma(p) \xrightarrow{\sim} \alpha_p^{-1}\Gamma(p)\alpha_p = \Gamma_0(p^2) \cap \Gamma_1(p),$$

which yields an isomorphism between the weight two cusp form spaces given by

$$S_2(\Gamma(p)) \xrightarrow{\sim} S_2(\Gamma_0(p^2) \cap \Gamma_1(p))$$
$$f \longmapsto f|_2[\alpha_p].$$

Since $\Gamma_{ns}^{\varepsilon}(p) \supseteq \Gamma(p)$, we have $S_2(\Gamma_{ns}^{\varepsilon}(p)) \subseteq S_2(\Gamma(p))$. Using the decomposition $S_2(\Gamma_0(p^2) \cap \Gamma_1(p)) = \bigoplus_{\chi} S_2(\Gamma_0(p^2), \chi)$, where the sum is taken over all the (even) characters of conductor dividing p, for any weight two cusp form $f \in S_2(\Gamma_{ns}^{\varepsilon}(p))$, $f|_2[\alpha_p]$ can be written (uniquely) as the sum of weight two cusp forms of level p^2 with nebentypus χ, for χ as above.

For ℓ a prime number different from p, we can define Hecke operators $\mathcal{T}_\ell^{\varepsilon}$ using the double coset $\Gamma_{ns}^{\varepsilon}(p)\alpha_\ell\Gamma_{ns}^{\varepsilon}(p)$ (resp. $\Gamma_{ns}^{+}(p)\alpha_\ell\Gamma_{ns}^{+}(p)$), where α_ℓ is any matrix in

$M_{ns}^{\varepsilon}(p)$ with determinant ℓ. Notice that this means that when $\ell \equiv 1 \pmod{p}$, we can choose $\alpha_\ell = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \ell \end{smallmatrix}\right)$, and it can be shown that this makes it coincide with the classical Hecke operator on $S_2(\Gamma(p))$. Likewise, if $\ell \equiv -1 \pmod{p}$ and we are considering the operator induced by the double coset $\Gamma_{ns}^{+}(p)\alpha_\ell\Gamma_{ns}^{+}(p)$, we can choose $\alpha_\ell = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \ell \end{smallmatrix}\right)$ and the same conclusion can be drawn.

The following is Theorem 1.12 in [KP14], which compiles results from Theorem 1 in [Che98], Theorem 1.1 in [Edi96] and Theorem 2 [dSE00].

Theorem 4.17 (Chen-Edixhoven)**.** *The new part of* $\mathrm{Jac}(X_0^{+}(p^2))$ *is isogenous to* $\mathrm{Jac}(X_{ns}^{+}(p))$. *Also the new part of* $\mathrm{Jac}(X_0(p^2))$ *is isogenous to* $\mathrm{Jac}(X_{ns}^{\varepsilon}(p))$. *Furthermore, these isogenies are Hecke equivariant.*

Theorem 4.17 associates to every normalized newform $g \in S_2(\Gamma_0(p^2))^{\mathrm{new}}$ a form $f \in S_2(\Gamma_{ns}^{\varepsilon}(p))$ such that f and g have the same eigenvalues for all $\ell \neq p$, i.e., if $T_\ell g = \lambda_\ell g$, then $\mathcal{T}_\ell^{\varepsilon} f = \lambda_\ell f$ for all $\ell \neq p$.

Using the Fourier expansion of g we can compute the Fourier expansion of f. For χ a character of conductor p, denote by $g \otimes \chi$ the twist of g by χ, which lives in $S_2(\Gamma_0(p^2), \chi^2)$. The following is Theorem 1.14 in [KP14], which relates the two Fourier expansions.

Theorem 4.18. *Let* f *and* g *as above. Let* π_p *be the local automorphic representation of* g *at* p. *Then*

- *If* π_p *is supercuspidal* $g \otimes \chi$ *is a newform in* $S_2(\Gamma_0(p^2), \chi^2)$ *when* χ *has conductor* p *and there exist* $\alpha_\chi \in \mathbb{C}$ *for every* χ *of conductor dividing* p *such that*

$$f|_2[\alpha_p] = \sum_{\chi} a_\chi \cdot g \otimes \chi$$

- *If π_p is Steinberg, there exists a newform $h \in S_2(\Gamma_0(p))$ such that g is the twist of h by the quadratic character of conductor p and there exist $a, a_\chi \in \mathbb{C}$ such that*

$$f|_2[\alpha_p] = a \cdot h + \sum_\chi a_\chi \cdot g \otimes \chi$$

- *If π_p is a ramified Principal Series, there exist a non-quadratic character θ_p of conductor p, newforms $h \in S_2(\Gamma_0(p), \bar{\theta}^2)$ and $\bar{h} \in S_2(\Gamma_0(p), \theta^2)$ with $h \otimes \theta_p = \bar{h} \otimes \bar{\theta_p} = g$, and $a_1, a_2, a_\chi \in \mathbb{C}$ such that*

$$f|_2[\alpha_p] = a_1 \cdot h + a_2 \cdot \bar{h} + \sum_\chi a_\chi \cdot g \otimes \chi$$

Rational Modular Forms

We know that a newform in $\Gamma_0(N)$ is normalized in the sense that its first coefficient is equal to 1. Multiplicity-one for the space of newforms states that given $g \in S_2(\Gamma_0(p^2))^{\text{new}}$ with a prescribed packet of eigenvalues outside of p, all other forms with these eigenvalues will be scalar multiples of g. Theorem 4.17 implies that the same holds for forms in $S_2(\Gamma^\varepsilon_{ns}(p))$, but, the natural question that arises is how to normalize f. We start by pre-normalizing f following Theorem 1.22 in [KP14].

Theorem 4.19. *Let $f \in S_2(\Gamma^\varepsilon_{ns}(p))$ be an eigenform which has the same eigenvalues as a rational newform $g \in S_2(\Gamma(p^2))$. Normalize $f|_2[\alpha_p]$ in such a way that the first Fourier coefficient is rational. Then f (and $f|_2[\alpha_p]$) has a q-expansion whose coefficients lie in $\mathbb{Q}(\xi_p)$.*

At the end of section 4.4.1 we mentioned that there are $p-1$ cusps in $X^\varepsilon_{ns}(p)$ defined over $\mathbb{Q}(\xi_p)$ which are conjugate of each other under the Galois action of the

cyclotomic extension. For $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(\xi_p)/\mathbb{Q})$ there exists a matrix $A_\ell \in \mathbf{SL}_2(\mathbb{Z})$ such that $\sigma_\ell(\infty) = A \cdot \infty$. The q-expansion of f at $A_\ell \cdot \infty$ is given by the q-expansion of $f|_2[A_\ell^{-1}]$. We say that f is *rational* if for every $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(\xi_p)/\mathbb{Q})$ the q-expansion of f at $\sigma_\ell(\infty)$ equals $\sigma_\ell^{-1}(f) = \sum_{n\geq 1} \sigma_\ell^{-1}(a_n(f))q_h^n$ (which is equivalent to saying $f|_2[A_\ell] = \sigma_\ell(f)$ for every $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(\xi_p)/\mathbb{Q})$).

Thanks to Theorem 4.19, we can see that any scalar $c \in \mathbb{Q}(\xi_p)^\times$ will yield q-expansions for cf with coefficients in $\mathbb{Q}(\xi_p)$, but we further want f to be rational. If there exists a c such that cf is rational, every rational multiple of c will turn cf into a rational modular form. Furthermore, if $c_1 f$ and $c_2 f$ are rational, then c_1/c_2 must be a rational number (as $\sigma_\ell(c_1/c_2) = c_1/c_2$ for all $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(\xi_p)/\mathbb{Q})$). It is the existence of such c that is more subtle. This is the content of Theorem 1.27 in [KP14].

Theorem 4.20. *Let $f \in S_2(\Gamma_{ns}^\varepsilon(p))$ be an eigenform with rational eigenvalues. There exists a unique $c \in \mathbb{Q}(\xi_p)^\times$ (modulo $\mathbb{Q}^\times$) such that cf is rational.*

Recall that modular forms have bounded denominators. This means that we can choose the constant in Theorem 4.20 in such a way that all the coefficients of the q-expansion are algebraic integers and that for every integer $n \geq 2$ at least one coefficient of cf/n is not an algebraic integer. Unfortunately, this still requires a choice of "sign," as there will be two such constants (up to sign). Both choices are equally good.

4.4.3 Modular Parametrization

At this point, we can emulate the construction in Section 3.5 with $f \in S_2(\Gamma_{ns}^\varepsilon(p))$ a normalized eigenform. Let g be the genus of $X_{ns}^\varepsilon(p)$. Choose a basis $\{\gamma_1, \ldots, \gamma_{2g}\}$

of the $\mathbb{Z}$-homology of $X_{ns}^{\varepsilon}(p)$. The map that assigns to every loop its corresponding period

$$\int_{\gamma} f(q)\frac{dq}{q} = \int_{\gamma} \frac{2\pi i}{p} f(\tau)d\tau,$$

where $q = e^{2\pi i\tau/p}$, forms a rank 2 lattice $\Lambda_f \subseteq \mathbb{C}$ when f has rational eigenvalues (which is the case for us, as we started with a newform corresponding to an elliptic curve of conductor p^2). The curve $E_f = \mathbb{C}/\Lambda_f$ is isogenous to the elliptic curve we started with. Lemma 1.33 in [KP14] shows that E_f does not depend on the choice of ε.

As before, the dimension of $S_2(\Gamma_{ns}^{\varepsilon}(p))$ is equal to the genus. Take $\{f_1, \ldots, f_g\}$ a basis of $S_2(\Gamma_{ns}^{\varepsilon}(p))$ comprised of eigenforms, with $f_1 = f$. Let

$$\omega_j = f_j(q)\frac{dq}{q} = \frac{2\pi i}{p} f_j(\tau)d\tau = \frac{f_j(\tau)}{pj'(\tau)/2\pi i}dj$$

(which is a rational holomorphic differential form when f_j is rational) and let $\Omega(X_{ns}^{\varepsilon}(p)) = \langle \omega_1, \ldots, \omega_g \rangle$ be the space of holomorphic differentials in $X_{ns}^{\varepsilon}(p)$. For $j = 1, \ldots, 2g$, let

$$\Omega_j = \left(\int_{\gamma_j} \omega_1, \ldots, \int_{\gamma_j} \omega_g \right) \in \mathbb{C}^g \qquad \text{and} \qquad \Lambda = \bigoplus_{j=1}^{2g} \Omega_j \mathbb{Z} \subseteq \mathbb{C}^g,$$

which again, forms a discrete full rank lattice in $\mathbb{C}^g$.

The Abel-Jacobi map is not rational anymore because the cusps of $X_{ns}^{\varepsilon}(p)$ are not rational, but the projection map from the Jacobian to E_f is, so this modular parametrization is rational over $\mathbb{Q}(\xi_p)$. To remedy this, let us analyze a bit further the Abel-Jacobi map.

The map

$$\begin{array}{ccccc} \mathcal{H}^* & \longrightarrow & \mathbb{C}^g & \longrightarrow & J(X_{ns}^{\varepsilon}(p)) \\ \tau & \longmapsto & \left(\displaystyle\int_{i\infty}^{\tau}\omega_1,\ldots,\int_{i\infty}^{\tau}\omega_g\right) & \longmapsto & \left(\displaystyle\int_{i\infty}^{\tau}\omega_1,\ldots,\int_{i\infty}^{\tau}\omega_g\right) \mod \Lambda \end{array}$$

descends to the quotient by $\Gamma_{ns}^{\varepsilon}(p)$, but we are choosing, in a rather arbitrary fashion, the cusp ∞. For any $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(\xi_p)/\mathbb{Q})$ we can define an Abel-Jacobi map based at the cusp $\sigma_\ell(\infty)$

$$\begin{array}{ccccc} \mathcal{H}^* & \longrightarrow & \mathbb{C}^g & \longrightarrow & J(X_{ns}^{\varepsilon}(p)) \\ \tau & \longmapsto & \left(\displaystyle\int_{\sigma_\ell(\infty)}^{\tau}\omega_1,\ldots,\int_{\sigma_\ell(\infty)}^{\tau}\omega_g\right) & \longmapsto & \left(\displaystyle\int_{\sigma_\ell(\infty)}^{\tau}\omega_1,\ldots,\int_{\sigma_\ell(\infty)}^{\tau}\omega_g\right) \mod \Lambda, \end{array}$$

so, each σ_ℓ yields, after projecting into E_f, it being a quotient of the Jacobian, a modular parametrization

$$\begin{array}{ccc} X_{ns}^{\varepsilon}(p) & \xrightarrow{\text{A-J}_{\sigma_\ell}} & J(X_{ns}^{\varepsilon}(p)) \\ & \searrow^{\Phi_{\sigma_\ell}} & \downarrow \pi \\ & & E_f \end{array} .$$

The map A-J $= \sum_{\sigma_\ell}$ A-J$_{\sigma_\ell}$ will be invariant under the action of $\mathrm{Gal}(\mathbb{Q}(\xi_p)/\mathbb{Q})$, which makes it a map defined over $\mathbb{Q}$. Hence, we obtain a map

$$\begin{array}{ccc} X_{ns}^{\varepsilon}(p) & \xrightarrow{\text{A-J}} & J(X_{ns}^{\varepsilon}(p)) \\ & \searrow^{\Phi} & \downarrow \pi \\ & & E_f \end{array} , \tag{4.4}$$

where $\Phi = \sum_{\sigma_\ell} \Phi_{\sigma_\ell}$, and this modular parametrization is rational.

This map can also be made explicit, as in Section 3.5. The Manin constant is a rational number (independent from ε), so we define

$$z_\tau = c\left(\frac{2\pi i}{p}\sum_{\sigma_\ell}\int_{i\infty}^{A_\ell^{-1}\tau}\sigma_\ell(f)(z)dz\right),$$

where the sum is taken over all $\sigma_\ell \in \mathrm{Gal}(\mathbb{Q}(\xi_p)/\mathbb{Q})$. Thus, we obtain an analogue for Equation (3.2)

$$\begin{aligned}\varphi\colon X_{ns}^{\varepsilon}(p) &\longrightarrow \qquad E_f\\ \tau &\longmapsto (\wp(z_\tau), \wp'(z_\tau), 1),\end{aligned}$$

where $\wp$ is the Weierstrass $\wp$-function of the lattice Λ_f. As before, this map can be composed with the isogeny relating E_f and E, yielding a map

$$X_{ns}^{\varepsilon}(p) \longrightarrow E_f \dashrightarrow E. \tag{4.5}$$

4.4.4 Higher levels

We will be very brief here, as we will only use the definition of the Cartan non-split group in what follows. It is included for the sake of completeness, and the details can be found in Section 2 of [KP14].

Let $N = n^2m$, with n square free and $\gcd(n, 2m) = 1$. Write $n = p_1 \cdots p_r$. For every p_j, let ε_j be a quadratic nonresidue modulo p_j and denote by $\vec{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_r)$. The Cartan non-split congruence subgroup associated to these data is

$$\Gamma_{ns}^{\vec{\varepsilon}}(n, m) = \bigcap_{j=1}^{r}\Gamma_{ns}^{\varepsilon_j}(p_j) \cap \Gamma_0(m),$$

which is the determinant 1 elements of the ring

$$M_{ns}^{\vec{\varepsilon}}(n,m) = \bigcap_{j=1}^{r} M_{ns}^{\varepsilon_j}(p_j) \cap M_0(m).$$

The Chen-Edixhoven Theorem takes the following form, which is Theorem 2.1 in [KP14].

Theorem 4.21. *The n^2-new part of $J(X_0(n^2m))$ is isogenous to $J(X_{ns}^{\vec{\varepsilon}}(n,m))$. The isogeny is Hecke equivariant.*

For an elliptic curve E of level N, there is a weight 2 newform g of level N, for which there exists an eigenform $f \in S_2(\Gamma_{ns}^{\vec{\varepsilon}}(n,m))$ with the same eigenvalues as g away from n.

The Abel-Jacobi map this time will be defined over $\mathbb{Q}(\xi_{p_1}, \ldots, \xi_{p_r})$, but the same averaging as before will bring it down to a rational map.

4.4.5 Heegner points

As before, we will try to embed orders of quadratic fields into rings of matrices and consider fixed points. Let E be an elliptic curve of conductor $N = n^2m$ as in the previous section. Let K be a number field and $\mathcal{O} \subseteq K$ an order. Assume that

(1) The discriminant of $\mathcal{O}$ is relatively prime to nm.

(2) Every prime dividing m is split in K.

(3) Every prime dividing n is inert in K.

Let D_0 be the fundamental discriminant of K and let $D = D_0 r^2$. $\mathcal{O}$ can be seen as the $\mathbb{Z}$-module generated by 1 and $\omega = \omega_D = \dfrac{D + \sqrt{D}}{2}$. In order to find an embedding of $\mathcal{O} \hookrightarrow M_{ns}^{\vec{\varepsilon}}(n,m)$ it suffices to find a matrix where we can map ω, which is, a matrix with the same trace as the trace of ω and determinant as the norm of ω. More

explicitly, we need a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ such that $a + d = D$ and $ad - bc = \dfrac{D(D-1)}{4}$. Substituting the former in the latter we obtain

$$a(D-a) - bc = \frac{D(D-1)}{4}.$$

We need to guarantee local conditions and glue them using the Chinese Remainder Theorem. The main problem will be solving the quadratic equation.

Let p be a prime dividing m. The local condition at p states that

$$a(D-a) \equiv \frac{D(D-1)}{4} \pmod{p},$$

which can be rewritten as

$$(2a - D)^2 \equiv D \pmod{p}.$$

Condition (2) ensures that p splits in $K = \mathbb{Q}(\sqrt{D})$, so this equation has solution modulo p. Condition (3) implies that we cannot find Heegner points directly via the ideas in Section 4.2, because this equation would not have a local solution.

Now, let p be a prime dividing n. The local condition at p states that

$$\frac{D}{2}\left(D - \frac{D}{2}\right) - b^2\varepsilon \equiv \frac{D(D-1)}{4} \pmod{p},$$

which can be rewritten as

$$\frac{D}{4} \equiv b^2\varepsilon \pmod{p}.$$

Condition (1) implies that $b^2\varepsilon$ is not 0 modulo p. Condition (3) implies that this quantity cannot be a square, as p is inert in $K = \mathbb{Q}(\sqrt{D})$. Since ε is a quadratic nonresidue modulo p, this equation has a solution modulo p.

Let $M \in M_{ns}^{\vec{\varepsilon}}(n,m)$ be the image of ω under the sought embedding and let $\tau \in \mathbb{C}$ such that $M \cdot \tau = \tau$ under the Möbius action. Then

$$M \cdot \tau = \frac{a\tau + b}{c\tau + d} = \tau \iff c\tau^2 + (d-a)\tau - b = 0,$$

so

$$\begin{aligned}\tau &= \frac{a-d \pm \sqrt{(d-a)^2 + 4bc}}{2c} = \frac{a-d \pm \sqrt{(d+a)^2 - 4(ad-bc)}}{2c} \\ &= \frac{a-d \pm \sqrt{D^2 - D(D-1)}}{2c} = \frac{a-d \pm \sqrt{D}}{2c},\end{aligned}$$

and τ can be chosen to lie in $\mathcal{H}$.

Appendix 5 in [Ser97] asserts that the corresponding point in the modular curve $X_{ns}^{\vec{\varepsilon}}(n,m)(\mathbb{C})$ under its identification with $\Gamma_{ns}^{\vec{\varepsilon}}(n,m)\backslash\mathcal{H}^*$, described in Section 4.4.1, actually lies in $X_{ns}^{\vec{\varepsilon}}(n,m)(H)$, where H is the ring class field attached to $\mathcal{O}$.

Using the modular parametrization described in Section 4.4.3 (Equation (4.5)) points can be obtained on $E(H)$. For examples, see Section 5 in [KP14].

Chapter 5
Stark-Heegner points attached to Cartan Non-split curves

This chapter will bring together the ideas of Sections 4.2, 4.3 and 4.4 setting up the computations required to find points on an elliptic curve E, conjecturally defined over narrow ring class fields of real quadratic extensions, in similar circumstances to those Kohen and Pacetti had in their construction of Heegner points. We will follow closely the path Darmon designed in his construction, which can be summarized as follows:

(1) Isolate a special prime p dividing the conductor N of the elliptic curve E.

(2) Construct a group to play the role of $\Gamma_0(N)$ in the classical case, which we denoted by $\Gamma_{p,M}$.

(3) Define a space of modular forms for $\Gamma_{p,M}$.

(4) Establish an isomorphism between a subspace of the space of classical weight 2 cusp forms for $\Gamma_0(N)$ and the space of weight 2 cusp forms for $\Gamma_{p,M}$.

(5) Define a system of integral measures.

(6) Define the double integrals using the measures above.

(7) Conjecture the existence of the semi-indefinite integral and relate it to the double integrals.

(8) Use the semi-indefinite integral to construct points on E.

Let K be a real quadratic field and E an elliptic curve defined over $\mathbb{Q}$ such that the sign of $L(E/K, s)$ is -1. We will limit ourselves to elliptic curves E with

bad reduction only at two odd primes, one of which is multiplicative and the other one additive. That is, $N = pq^2$, so p will play the same role as it did in Darmon's construction. (In particular, p is inert in K.) The case where q is split in the real quadratic field K was dealt with in [GM15] so we will deal with the case where q is inert. Fix K throughout the remaining of this chapter.

5.1 The group

The group $\Gamma_{p,M}$ was defined as the subgroup of $\mathbf{SL}_2(\mathbb{Z}[1/p])$ of matrices which are upper-triangular modulo M. Notice that $\Gamma_{p,M}$ can be seen as the determinant 1 elements of the tensor $M_0(N) \otimes \mathbb{Z}[1/p] = M_0(M)[1/p]$. The case of imaginary quadratic fields where p was split and q was inert used the group $\Gamma_{ns}^{\varepsilon}(q) \cap \Gamma_0(p)$, where ε was a quadratic nonresidue modulo q. The natural extension to the scenario we want to consider is to take the determinant 1 elements of the tensor product

$$(M_{ns}^{\varepsilon}(q) \cap M_0(p)) \otimes \mathbb{Z}[1/p] = M_{ns}^{\varepsilon}(q)[1/p],$$

where $M_{ns}^{\varepsilon}(q)[1/p]$ is the ring of 2×2 matrices with entries in $\mathbb{Z}[1/p]$ such that

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix} \equiv \begin{pmatrix} a & b \\ b\varepsilon & a \end{pmatrix} \pmod{q}.$$

When talking about the normalizer, we add the alternate condition

$$\begin{pmatrix} x & y \\ z & w \end{pmatrix} \equiv \begin{pmatrix} a & b \\ -b\varepsilon & -a \end{pmatrix} \pmod{q}.$$

This ring is called $M_{ns}^{+}(q)[1/p]$. The corresponding groups will be denoted $\Gamma_{ns}^{\varepsilon}(q)[1/p]$ and $\Gamma_{ns}^{+}(q)[1/p]$.

Both $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ and $\Gamma^{+}_{ns}(q)[1/p]$ are finite index subgroups of $\mathbf{SL}_2(\mathbb{Z}[1/p])$. Since the action of $\mathbf{SL}_2(\mathbb{Z}[1/p])$ on $\mathcal{H}_p$ and on $\mathcal{H}$ by Möbius transformations has dense orbits, the same is true for $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ and $\Gamma^{+}_{ns}(q)[1/p]$. Likewise, since the action of $\mathbf{SL}_2(\mathbb{Z}[1/p])$ on $\mathcal{H}_p \times \mathcal{H}$ is discrete, the same is true for $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ and $\Gamma^{+}_{ns}(q)[1/p]$.

There is a relationship between the different Cartan Non-split groups upon varying the quadratic nonresidue ε. We state the result in the next proposition.

Proposition 5.1. *Let ε and ε' be two quadratic nonresidues modulo q. There exists an inner automorphism of $\mathbf{SL}_2(\mathbb{Z}[1/p])$ which induces an isomorphism*

$$\Gamma^{\varepsilon}_{ns}(q)[1/p] \cong \Gamma^{\varepsilon'}_{ns}(q)[1/p].$$

Furthermore, it induces an isomorphism between the corresponding normalizers.

In order to prove it, we need a couple of lemmas.

Lemma 5.2. *For any quadratic nonresidue ε, the index $[\Gamma^{\varepsilon}_{ns}(q)[1/p] : \Gamma(q)[1/p]]$ is equal to $q+1$.*

Proof. This can be accomplished using the isomorphism

$$\begin{aligned} \Gamma^{\varepsilon}_{ns}(q)[1/p]/\Gamma(q)[1/p] &\longrightarrow (\mathbb{F}^{\times}_{q^2})_1 \\ \begin{pmatrix} a & b \\ b\varepsilon & a \end{pmatrix} &\longmapsto a + b\sqrt{\varepsilon}, \end{aligned}$$

where $(\mathbb{F}^{\times}_{q^2})_1$ is the set of norm 1 elements of $\mathbb{F}_{q^2} = \mathbb{F}_q(\sqrt{\varepsilon})$, which has q^2-1 elements. □

Lemma 5.3. *Let*

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z})$$

such that

$$ac \equiv bd\varepsilon \pmod{q}.$$

The inner automorphism of $\mathbf{SL}_2(\mathbb{Z}[1/p])$ *given by* $\gamma \mapsto M\gamma M^{-1}$ *restricted to* $\Gamma_{ns}^{\varepsilon}(q)[1/p]$ *yields an isomorphism into* $\Gamma_{ns}^{\varepsilon'}(q)[1/p]$*, where*

$$\varepsilon' \equiv \varepsilon(a^2 - b^2\varepsilon)^{-2} \pmod{q}.$$

Proof. For $\gamma \equiv \begin{pmatrix} x & y \\ y\varepsilon & x \end{pmatrix} \pmod{q}$, a quick computation reveals that

$$M\gamma M^{-1} \equiv \begin{pmatrix} (ad-bc)x - (ac-bd\varepsilon)y & (a^2 - b^2\varepsilon)y \\ -(c^2 - d^2\varepsilon)y & (ad-bc)x + (ac-bd\varepsilon)y \end{pmatrix} \pmod{q}.$$

Since $ac \equiv bd\varepsilon \pmod{q}$ and $ad - bc = 1$, the top-left entry and the bottom-right entry are both congruent to x modulo q.

Note that

$$\begin{aligned}(c^2 - d^2\varepsilon)(a^2 - b^2\varepsilon) &= c^2a^2 + d^2b^2\varepsilon^2 - \varepsilon(c^2b^2 + d^2a^2) \\ &= (a^2c^2 - 2acbd\varepsilon + b^2d^2\varepsilon^2) - \varepsilon(a^2d^2 - 2adbc + b^2c^2) \\ &= (ac - bd\varepsilon)^2 - \varepsilon(ad-bc)^2 \\ &\equiv -\varepsilon \pmod{q},\end{aligned}$$

so

$$-\frac{c^2-d^2\varepsilon}{a^2-b^2\varepsilon}=\frac{-(c^2-d^2\varepsilon)(a^2-b^2\varepsilon)}{(a^2-b^2\varepsilon)^2}\equiv\varepsilon(a^2-b^2\varepsilon)^{-2}\equiv\varepsilon'\pmod{q}.$$

Using this last congruence, we find that

$$(a^2-b^2\varepsilon)y\cdot\varepsilon'\equiv(a^2-b^2\varepsilon)y\cdot-\frac{c^2-d^2\varepsilon}{a^2-b^2\varepsilon}\equiv-(c^2-d^2\varepsilon)y\pmod{q},$$

so $M\gamma M^{-1}$ lies indeed in $\Gamma^{\varepsilon'}_{ns}(q)[1/p]$.

Let $\Gamma'\subseteq\Gamma^{\varepsilon'}_{ns}(q)[1/p]$ denote the image of $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ under the inner automorphism induced by M. Using Lemma 5.2 for ε and ε', we have that

$$\begin{aligned}
[\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma'] &= [\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma^{\varepsilon}_{ns}(q)[1/p]]\\
&= [\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma(q)[1/p]]\div[\Gamma^{\varepsilon}_{ns}(q)[1/p]:\Gamma(q)[1/p]]\\
&= [\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma(q)[1/p]]\div(q+1)\\
&= [\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma(q)[1/p]]\div[\Gamma^{\varepsilon'}_{ns}(q)[1/p]:\Gamma(q)[1/p]]\\
&= [\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma^{\varepsilon'}_{ns}(q)[1/p]],
\end{aligned}$$

but we also have

$$[\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma']=[\mathbf{SL}_2(\mathbb{Z}[1/p]):\Gamma^{\varepsilon}_{ns}(q)[1/p]]\cdot[\Gamma^{\varepsilon'}_{ns}(q)[1/p]:\Gamma'],$$

whence

$$[\Gamma^{\varepsilon'}_{ns}(q)[1/p]:\Gamma']=1.$$

This shows that $\Gamma'=\Gamma^{\varepsilon'}_{ns}(q)[1/p]$, as we wanted. □

Proof of Proposition 5.1. This statement is meaningful only when $q>3$, so assume this is the case. Since both ε and ε' are quadratic nonresidues, their quotient

is a quadratic residue. Let t be an integer such that $t^2 \equiv \varepsilon\varepsilon'^{-1} \pmod{q}$. Let us consider two cases:

- *Case 1:* t is a quadratic residue modulo q.

  Let a be an integer such that $a^2 \equiv t \pmod{q}$. Let M be a lift to $\mathbf{SL}_2(\mathbb{Z}[1/p])$ of the matrix

  $$\begin{pmatrix} a & 0 \\ 0 & t^{-1}a \end{pmatrix},$$

  which has determinant 1 modulo q.

- *Case 2:* t is a quadratic nonresidue modulo q.

  Consider the sum of Legendre symbols

  $$\sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right) = 0.$$

  Squaring it and reorganizing we obtain

  $$\begin{aligned} 0 &= \sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right)^2 + \sum_{a\in\mathbb{F}_q^\times} \sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right)\left(\frac{x-a}{q}\right) \\ &= (q-1) + \sum_{a\in\mathbb{F}_q^\times} \sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right)\left(\frac{x-a}{q}\right)\left(\frac{a^{-1}}{q}\right)^2 \\ &= (q-1) + \sum_{a\in\mathbb{F}_q^\times} \sum_{x\in\mathbb{F}_q} \left(\frac{xa^{-1}}{q}\right)\left(\frac{xa^{-1}-1}{q}\right) \\ &= (q-1) + (q-1)\sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right)\left(\frac{x-1}{q}\right), \end{aligned}$$

  whence

  $$\sum_{x\in\mathbb{F}_q} \left(\frac{x}{q}\right)\left(\frac{x-a}{q}\right) = -1$$

for any $a \in \mathbb{F}_q^\times$. In particular, it holds for $a = t$. This implies that there exists $t' \in \mathbb{F}_q - \{t\}$ such that t' is a quadratic nonresidue and $t' - t$ is a nonzero quadratic residue. Otherwise, for all quadratic nonresidues different from t, both Legendre symbols would be -1. These account for $(q-3)/2$ products equal to 1. For $x = t$ and $x = 0$ we obtain a zero product. This would mean that all the remaining $(q-1)/2$ products must be equal to -1, but all the remaining values for x are quadratic residues, so $x - t$ would have to be a quadratic nonresidue. This gives us a total of $(q-1)/2 + (q-3)/2 = q-2$ values of x for which $x-t$ is a quadratic nonresidue, producing a contradiction. Since ε is a quadratic nonresidue and $t'-t$ is a quadratic residue, $(t'-t)\varepsilon^{-1}$ is a quadratic nonresidue. This is, there is no integer k such that

$$k^2 \equiv (t'-t)\varepsilon^{-1} \pmod{q}, \qquad \text{or, equivalently,} \qquad t + k^2\varepsilon \equiv t' \pmod{q}.$$

This implies that the set

$$\left\{ t + 1^2\varepsilon,\, t + 2^2\varepsilon, \ldots,\, t + \left(\frac{q-1}{2}\right)^2 \varepsilon \right\}$$

must contain a quadratic residue, as the quadratic nonresidue t' is not in the set. Denote it by $x^2 \equiv t + y^2\varepsilon \pmod{q}$. Let M be a lift to $\mathbf{SL}_2(\mathbb{Z}[1/p])$ of the matrix

$$\begin{pmatrix} x & y \\ t^{-1}y\varepsilon & t^{-1}x \end{pmatrix},$$

which has determinant 1.

In both cases, the matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

has the property that

$$ac - bd\varepsilon \equiv \begin{cases} a \cdot 0 - 0 \cdot at^{-1}\varepsilon \\ x \cdot t^{-1}y\varepsilon - y \cdot t^{-1}x\varepsilon \end{cases} \equiv 0 \pmod{q},$$

so, by Lemma 5.3, M induces an isomorphism between the Cartan Non-split groups corresponding to

$$\varepsilon \quad \text{and} \quad \varepsilon(a^2 - b^2\varepsilon)^{-2} \equiv \varepsilon t^{-2} \equiv \varepsilon\varepsilon'\varepsilon^{-1} \equiv \varepsilon' \pmod{q}.$$

The last part of the proposition follows from the following fact: if H is a subgroup of G with normalizer N, then for every $g \in G$, gNg^{-1} is the normalizer of gHg^{-1}, as

$$n \in N_G(gHg^{-1}) \iff ngHg^{-1}n^{-1} = gHg^{-1} \iff g^{-1}ngHg^{-1}n^{-1}g = H$$
$$\iff g^{-1}ng \in N \iff n \in gNg^{-1}.$$

Let $G = \mathbf{SL}_2(\mathbb{Z}[1/p])$ and $H = \Gamma^{\varepsilon}_{ns}(q)[1/p]$ above. The result follows. □

5.1.1 Cusps

If two cusps in $\mathbb{P}^1(\mathbb{Q})$ are equivalent under the action of $\Gamma^{\varepsilon}_{ns}(q)$, they must also be equivalent under the action of $\Gamma^{\varepsilon}_{ns}(q)[1/p]$, as $\Gamma^{\varepsilon}_{ns}(q)[1/p] \supseteq \Gamma^{\varepsilon}_{ns}(q)$. However, the surplus of elements in the larger group allows for some inequivalent cusps in $\Gamma^{\varepsilon}_{ns}(q)$ to become equivalent in $\Gamma^{\varepsilon}_{ns}(q)[1/p]$. The same occurs with the normalizers. How this happens is summarized in the following proposition.

Proposition 5.4. *Let x, y and x', y' be two pairs of relatively prime integers. The cusps x/y and x'/y' are equivalent under the action of $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ or $\Gamma^{+}_{ns}(q)[1/p]$ if and only if the classes of $x^2 - y^2\varepsilon^{-1}$ and $x'^2 - y'^2\varepsilon^{-1}$ coincide in*

$$(\mathbb{Z}/q\mathbb{Z})^{\times}/\langle p^2\rangle \qquad or \qquad (\mathbb{Z}/q\mathbb{Z})^{\times}/\langle p^2, -p^2\rangle,$$

respectively.

In order to prove Proposition 5.4, we will make use of the following lemma.

Lemma 5.5. *Let a, b be two relatively prime integers. The cusp a/b is equivalent to ∞ under the action of $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ or $\Gamma^{+}_{ns}(q)[1/p]$ if and only if the class of $a^2 - b^2\varepsilon^{-1}$ is trivial in*

$$(\mathbb{Z}/q\mathbb{Z})^{\times}/\langle p^2\rangle \qquad or \qquad (\mathbb{Z}/q\mathbb{Z})^{\times}/\langle p^2, -p^2\rangle,$$

respectively.

Proof. Suppose that ∞ is equivalent to a/b under the action of $\Gamma^{\varepsilon}_{ns}(q)[1/p]$. This implies that there exists a matrix

$$M = \begin{pmatrix} x & z \\ y & w \end{pmatrix} \in \Gamma^{\varepsilon}_{ns}(q)[1/p],$$

such that

$$M(\infty) = \frac{a}{b}.$$

From $xw - yz = 1$ we can see that the ℓ-adic valuations of x and y ($\ell \neq p$) cannot both be positive. Since

$$\frac{a}{b} = M(\infty) = \frac{x \cdot 1 + z \cdot 0}{y \cdot 1 + w \cdot 0} = \frac{x}{y}$$

and a and b are relatively prime integers, we find that $a = \pm xp^s$ and $b = \pm yp^s$ for some integer s. Using $x \equiv w \pmod{q}$ and $y \equiv z\varepsilon \pmod{q}$, we conclude that

$$a^2 - b^2\varepsilon^{-1} = (x^2 - y^2\varepsilon^{-1})p^{2s} \equiv (xw - yz)p^{2s} \equiv p^{2s} \pmod{q},$$

which implies that the class of $a^2 - b^2\varepsilon^{-1}$ is trivial in

$$(\mathbb{Z}/q\mathbb{Z})^\times / \langle p^2 \rangle.$$

Conversely, suppose that $a^2 - b^2\varepsilon^{-1} \equiv p^{2s} \pmod{q}$ for some integer s. Let $\bar{b}$ be an integer such that $\bar{b}\varepsilon \equiv b \pmod{q}$, so $b^2\varepsilon^{-1} \equiv b\bar{b} \pmod{q}$. From $a^2 - b^2\varepsilon^{-1} \equiv p^{2s} \pmod{q}$, we find that $p^{2s} - (a^2 - b\bar{b})$ is divisible by q.

Let k_1 and k_2 be integers such that

$$ak_2 - bk_1 = \frac{p^{2s} - (a^2 - b\bar{b})}{q},$$

whose existence is guaranteed by the fact that a and b are relatively prime and the RHS is an integer. The matrix

$$M = \begin{pmatrix} a/p^s & (\bar{b} + qk_1)/p^s \\ b/p^s & (a + qk_2)/p^s \end{pmatrix}$$

has determinant

$$\frac{a(a + qk_2) - b(\bar{b} + qk_1)}{p^{2s}} = \frac{(a^2 - b\bar{b}) + (p^{2s} - (a^2 - b\bar{b}))}{p^{2s}} = 1.$$

Furthermore, $(\bar{b} + qk_1)/p^s \cdot \varepsilon \equiv b/p^s \pmod{q}$ and $a/p^s \equiv (a + qk_2)/p^s \pmod{q}$, implying that M lies in $\Gamma_{ns}^{\varepsilon}(q)[1/p]$. Finally,

$$M(\infty) = \frac{a/p^s}{b/p^s} = \frac{a}{b},$$

so the cusp a/b is equivalent to ∞ under the action of $\Gamma_{ns}^{\varepsilon}(q)[1/p]$.

The computations for the normalizer are completely analogous. □

Proof of Proposition 5.4. Let $x/y \in \mathbb{P}^1(\mathbb{Q})$ and let ρ denote the class of $x^2 - y^2\varepsilon^{-1}$ in $(\mathbb{Z}/q\mathbb{Z})^{\times}$. Let

$$M = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z})$$

be the matrix in Proposition 5.1 that induces an isomorphism between $\Gamma_{ns}^{\varepsilon}(q)[1/p]$ and $\Gamma_{ns}^{\varepsilon'}[1/p]$, where $\varepsilon' \equiv \varepsilon\rho^2 \pmod{q}$, this is,

$$M \equiv \begin{pmatrix} a & b \\ t^{-1}b\varepsilon & t^{-1}a \end{pmatrix} (\text{mod } q),\ a^2 - b^2\varepsilon \equiv t(\text{mod } q),\ t^2 \equiv \varepsilon\varepsilon'^{-1} \equiv \rho^{-2}(\text{mod } q).$$

We can choose t to be precisely ρ^{-1}. The action of M on the cusp x/y is given by

$$M\,(x/y) = \frac{a'x + b'y}{c'x + d'y}.$$

Consider

$$\begin{aligned}(a'x+b'y)^2-(c'x+d'y)^2\varepsilon'^{-1} &\equiv (a'x+b'y)^2-(c'x+d'y)^2\varepsilon^{-1}\rho^{-2} \pmod q\\ &\equiv (a^2-b^2\varepsilon)x^2+(b^2-a^2\varepsilon^{-1})y^2 \pmod q\\ &\equiv (a^2-b^2\varepsilon)(x^2-y^2\varepsilon^{-1}) \pmod q\\ &\equiv 1 \pmod q.\end{aligned}$$

By Lemma 5.5, the cusp $M(x/y)$ is equivalent to the cusp ∞ under the action of $\Gamma_{ns}^{\varepsilon'}(q)[1/p] = M\Gamma_{ns}^{\varepsilon}(q)[1/p]M^{-1}$ (and, a fortiori, by the action of its normalizer).

Let $x'/y' \in \mathbb{P}^1(\mathbb{Q})$ be another cusp and let ρ' denote the class of $x'^2 - y'^2/\varepsilon$ in $(\mathbb{Z}/q\mathbb{Z})^\times$. As before, we have

$$M(x'/y') = \frac{a'x'+b'y'}{c'x'+d'y'}$$

and, in order to test equivalence to ∞ we need to consider

$$\begin{aligned}(a'x'+b'y')^2-(c'x'+d'y')^2\varepsilon'^{-1} &\equiv (a^2-b^2\varepsilon)(x'^2-y'^2\varepsilon^{-1}) \pmod q\\ &\equiv t\rho' \pmod q.\end{aligned}$$

By Lemma 5.5, the cusp $M(x/'y')$ is equivalent to ∞ (and hence, to $M(x/y)$) under the action of $\Gamma_{ns}^{\varepsilon'}(q)[1/p]$ if and only if $t\rho'$ is in $\langle p^2\rangle$. By Proposition 5.1, this translates into the existence of $\gamma \in \Gamma_{ns}^{\varepsilon}(q)[1/p]$ such that

$$(M\gamma M^{-1})(M(x/y)) = M(x'/y') \iff \gamma(x/y) = x'/y'$$

if and only if $t\rho' \equiv \rho^{-1}\rho'$ lies in $\langle p^2 \rangle$, which is the same as asking for ρ and ρ' to lie in the same class in

$$(\mathbb{Z}/q\mathbb{Z})^\times / \langle p^2 \rangle.$$

As with Lemma 5.5, the computations for the normalizer are analogous. □

Now, we are in position to determine the number of inequivalent cusps under the action of the groups we are analyzing.

Corollary 5.6. *Let r be the size of the multiplicative subgroup of $(\mathbb{Z}/q\mathbb{Z})^\times$ generated by p^2. Let β be equal to 1 if $-1 \in \langle p^2 \rangle$ and 2 otherwise. Then*

$$\#\left(\Gamma_{ns}^{\varepsilon}(q)[1/p]\backslash\mathbb{P}^1(\mathbb{Q})\right) = \frac{q-1}{r} \qquad \text{and} \qquad \#\left(\Gamma_{ns}^{+}(q)[1/p]\backslash\mathbb{P}^1(\mathbb{Q})\right) = \frac{q-1}{\beta r}.$$

Proof. By Proposition 5.4, the cusps x/y and x'/y' are equivalent under the action of $\Gamma_{ns}^{\varepsilon}(q)[1/p]$ if and only if the classes of $x^2 - y^2\varepsilon^{-1}$ and $x'^2 - y'^2\varepsilon^{-1}$ in $(\mathbb{Z}/q\mathbb{Z})^\times/\langle p^2 \rangle$ coincide. In other words, the map

$$\begin{array}{ccc} \left(\Gamma_{ns}^{\varepsilon}(q)[1/p]\backslash\mathbb{P}^1(\mathbb{Q})\right) & \longrightarrow & (\mathbb{Z}/q\mathbb{Z})^\times/\langle p^2 \rangle \\ x/y & \longmapsto & x^2 - y^2\varepsilon^{-1} \end{array}$$

is a bijection.

Likewise, the cusps are equivalent under the action of $\Gamma_{ns}^{+}(q)[1/p]$ if an only if the classes coincide in $(\mathbb{Z}/q\mathbb{Z})^\times/\langle p^2, -p^2 \rangle$, yielding a bijection

$$\begin{array}{ccc} \left(\Gamma_{ns}^{+}(q)[1/p]\backslash\mathbb{P}^1(\mathbb{Q})\right) & \longrightarrow & (\mathbb{Z}/q\mathbb{Z})^\times/\langle p^2, -p^2 \rangle \\ x/y & \longmapsto & x^2 - y^2\varepsilon^{-1}. \end{array}$$

The first equality follows from

$$\#((\mathbb{Z}/q\mathbb{Z})^\times/\langle p^2\rangle) = \frac{\#((\mathbb{Z}/q\mathbb{Z})^\times)}{\#\langle p^2\rangle} = \frac{q-1}{r},$$

and the second one from

$$\#((\mathbb{Z}/q\mathbb{Z})^\times/\langle p^2,-p^2\rangle) = \frac{\#((\mathbb{Z}/q\mathbb{Z})^\times)}{\#\langle p^2,-p^2\rangle} = \frac{\#((\mathbb{Z}/q\mathbb{Z})^\times)}{\#\langle p^2\rangle} \div \frac{\#\langle p^2,-p^2\rangle}{\#\langle p^2\rangle} = \frac{q-1}{\beta r},$$

as the index of $\langle p^2\rangle$ in $\langle p^2,-p^2\rangle$ is precisely β. □

5.2 Modular Forms

Let Γ be $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ or $\Gamma^{+}_{ns}(q)[1/p]$. Denote by $\hat{\Gamma}$ the group $\Gamma^{\varepsilon}_{ns}(q,p)$ or $\Gamma^{+}_{ns}(q,p)$, respectively. Recall that $\mathcal{T} = \mathcal{T}_0 \cup \mathcal{T}_1$ is the Bruhat-Tits tree of $\mathbf{GL}_2(\mathbb{Q}_p)$, where $\mathcal{T}_0$ is the set of vertices and $\mathcal{T}_1$ is the set of unordered edges and $\mathcal{E}(\mathcal{T})$ is the set of ordered edges. Let $S_2(\mathcal{T},\Gamma)$ be the space of cusp forms of weight 2 for Γ, this is, all the functions

$$f\colon \mathcal{E}(\mathcal{T}) \times \mathcal{H} \longrightarrow \mathbb{C}$$

satisfying the three properties

(1) $f(\gamma e, \gamma\tau) = (c\tau+d)^2 f(e,\tau)$ for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma$.

(2) For each vertex $v \in T_0$ we have

$$\sum_{s(e)=v} f(e,\tau) = 0,$$

and for each ordered edge $e \in \mathcal{E}(\mathcal{T})$ we have $f(\bar{e},\tau) = -f(e,\tau)$.

(3) For each fixed oriented edge $e \in \mathcal{E}(\mathcal{T})$ the function $f_e(\tau)$ is a weight 2 cusp form for Γ_e, the stabilizer of e in Γ.

Proposition 5.7. *The stabilizers of v_0 and e_0 in $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ (resp. $\Gamma^{+}_{ns}(q)[1/p]$) are $\Gamma^{\varepsilon}_{ns}(q)$ and $\Gamma^{\varepsilon}_{ns}(q,p)$ (resp. $\Gamma^{+}_{ns}(q)$ and $\Gamma^{+}_{ns}(q,p)$), respectively.*

Proof. By Chapter 9, Exercise 1 in [Dar04], we have

$$\mathrm{Stab}_{\mathbf{SL}_2(\mathbb{Z}[1/p])}(v_0) = \mathbf{SL}_2(\mathbb{Z}) \qquad \text{and} \qquad \mathrm{Stab}_{\mathbf{SL}_2(\mathbb{Z}[1/p])}(e_0) = \Gamma_0(p).$$

From here, it's clear that

$$\mathrm{Stab}_{\Gamma}(v_0) = \mathbf{SL}_2(\mathbb{Z}) \cap \Gamma \qquad \text{and} \qquad \mathrm{Stab}_{\Gamma}(e_0) = \Gamma_0(p) \cap \Gamma,$$

whence the proposition follows. □

As in the case of classical Stark-Heegner points, for $f \in S_2(\mathcal{T}, \Gamma)$ we denote by f_0 the classical modular form f_{e_0} attached to the edge e_0. By the previous proposition, this is a modular form in $S_2(\hat{\Gamma})$.

We have the following lemma, which is analogous to Lemma 4.6.

Lemma 5.8. *The map*

$$\begin{aligned} S_2(\mathcal{T}, \Gamma) &\longrightarrow S_2(\hat{\Gamma}) \\ f &\longmapsto f_0 \end{aligned}$$

is injective. Furthermore, the image is $S_2(\hat{\Gamma})^{p-\mathrm{new}}$.

Proof. This follows almost exactly as Lemma 1.3 in [Dar01] and the comments before it. We will write the proof in the case of $\Gamma = \Gamma^{\varepsilon}_{ns}(q)[1/p]$ for simplicity of notation.

Let $\tilde{\Gamma} = M^{\varepsilon}_{ns}(q)[1/p]^{\times}_{+}$, which is, 2×2 matrices with entries in $\mathbb{Z}[1/p]$ satisfying the congruence condition modulo q from Section 5.1 with determinant p^{β}, $\beta \in \mathbb{Z}$.

Let $S_2(\mathcal{E}, \tilde{\Gamma})$ be the space of functions

$$f\colon \mathcal{E}(\mathcal{T}) \times \mathcal{H} \longrightarrow \mathbb{C}$$

satisfying the *two* properties

(1) $f(\gamma e, \gamma\tau) = (c\tau + d)^2 f(e, \tau)/\det(\gamma)$ for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \tilde{\Gamma}$.

(2) For each fixed oriented edge $e \in \mathcal{E}(\mathcal{T})$ the function $f_e(\tau)$ is a weight 2 cusp form for $\tilde{\Gamma}_e$, the stabilizer of e in $\tilde{\Gamma}$.

Note that (1) for γ with determinant 1 (i.e., $\gamma \in \Gamma$) gives (1) for any γ of even determinant, and (1) for γ with determinant p gives (1) for any γ of odd determinant.

Also, let $S_2(\mathcal{T}_0, \tilde{\Gamma})$ be the space of functions

$$f\colon \mathcal{T}_0 \times \mathcal{H} \longrightarrow \mathbb{C}$$

satisfying the *two* properties

(1) $f(\gamma v, \gamma\tau) = (c\tau + d)^2 f(v, \tau)/\det(\gamma)$ for all $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \tilde{\Gamma}$.

(2) For each fixed vertex $v \in \mathcal{T}_0$ the function $f_v(\tau)$ is a weight 2 cusp form for $\tilde{\Gamma}_v$, the stabilizer of v in $\tilde{\Gamma}$.

We say that a vertex is *even* if there is an even number of edges in the path connecting it to the root, and we say it is *odd* otherwise. The two orbits of the action of Γ on $\mathcal{T}_0$ are the even vertices and the odd vertices, precisely. However, the action of $\tilde{\Gamma}$ on $\mathcal{T}$ yields only one orbit for the vertices and one orbit for the oriented edges. If $\gamma \in \tilde{\Gamma}$ has determinant an even power of p, it preserves the Γ-orbit, whereas if it

has determinant an odd power of p it changes the Γ-orbit. Let $\alpha \in \tilde{\Gamma}$ such that α has determinant an odd power of p.

We have a map

$$\iota\colon S_2(\mathcal{T},\Gamma) \longrightarrow S_2(\mathcal{E},\tilde{\Gamma})$$

defined by

$$\iota(f)(e,\tau)d\tau = \begin{cases} f(e,\tau)d\tau & \text{if } s(e) \text{ is even} \\ f(\alpha e,\alpha\tau)d\alpha\tau & \text{if } s(e) \text{ is odd.} \end{cases}$$

Note that this definition is independent of α, as any other choice α' would also have as determinant an odd power of p, so $\alpha' = \gamma'\alpha$, where γ' has determinant an even power of p. The invariance of f for elements in Γ yields the required invariance for γ', which shows that $f(\alpha e,\alpha\tau)d\alpha\tau = f(\alpha' e,\alpha'\tau)d\alpha'\tau$. In particular, for the definition of $\iota(f)$, we may choose α with the further property of fixing the unordered edge e_0 (but not the ordered one).

The invariance for every element of $\tilde{\Gamma}$ follows from the following computations:

- If $s(e)$ is even and $\det(\gamma)$ is an even power of p we have

$$\iota(f)(\gamma e,\gamma\tau)d\gamma\tau = f(\gamma e,\gamma\tau)d\gamma\tau = f(e,\tau)d\tau = \iota(f)(e,\tau)d\tau.$$

- If $s(e)$ is even and $\det(\gamma)$ is an odd power of p we have

$$\iota(f)(\gamma e,\gamma\tau)d\gamma\tau = f(\alpha\gamma e,\alpha\gamma\tau)d\alpha\gamma\tau = f(e,\tau)d\tau = \iota(f)(e,\tau)d\tau,$$

as $s(\gamma e)$ is odd and $\alpha\gamma$ has determinant an even power of p.

- If $s(e)$ is odd and $\det(\gamma)$ is an even power of p we have

$$\iota(f)(\gamma e, \gamma\tau)d\gamma\tau = f(\alpha\gamma e, \alpha\gamma\tau)d\alpha\gamma\tau = \iota(f)(e,\tau)d\tau,$$

  as $s(\gamma e)$ and $s(e)$ are odd and $\alpha\gamma$ has determinant an odd power of p.

- If $s(e)$ is odd and $\det(\gamma)$ is an odd power of p we have

$$\iota(f)(\gamma e, \gamma\tau)d\gamma\tau = f(\gamma e, \gamma\tau)d\gamma\tau = \iota(f)(e,\tau)d\tau,$$

  as $s(e)$ is odd, $s(\gamma e)$ is even, and γ has determinant an odd power of p.

Moreover, this is an injection from $S_2(\mathcal{T},\Gamma) \hookrightarrow S_2(\mathcal{E},\tilde{\Gamma})$, since $\iota(f) = 0$ implies that $f(e,\tau) = 0$ for all edges such that $s(e)$ is even. From (2) in the definition of $S_2(\mathcal{T},\Gamma)$, for odd $s(e)$ we have $f(e,\tau) = -f(\bar{e},\tau) = 0$, as $s(\bar{e}) = t(e)$ is even. Whence, $f(e,\tau) = 0$ for all $e \in \mathcal{E}(\mathcal{T})$, making $f = 0$.

We also have two maps

$$\pi_s, \pi_t \colon S_2(\mathcal{E},\tilde{\Gamma}) \longrightarrow S_2(\mathcal{T}_0,\tilde{\Gamma})$$

given by

$$\pi_s(f)(v,\tau) = \sum_{s(e)=v} f(e,\tau) \qquad \text{and} \qquad \pi_t(f)(v,\tau) = \sum_{t(e)=v} f(e,\tau),$$

since any $\gamma \in \tilde{\Gamma}$ will map the set of edges with source (resp. target) v to the set of edges with source (resp. target) γv. The properties of $S_2(\mathcal{E},\tilde{\Gamma})$ translate into the properties for $S_2(\mathcal{T}_0,\tilde{\Gamma})$ this way.

Thus, we have the sequence

$$0 \longrightarrow S_2(\mathcal{T},\Gamma) \xrightarrow{\ \iota\ } S_2(\mathcal{E},\tilde{\Gamma}) \xrightarrow{\pi_s \oplus \pi_t} S_2(\mathcal{T}_0,\tilde{\Gamma}) \oplus S_2(\mathcal{T}_0,\tilde{\Gamma})$$

Recall that ι is injective, so in order to show exactness, we just need to show that $\ker(\pi_s \oplus \pi_t) = \mathrm{im}(\iota)$.

Let $f \in S_2(\mathcal{T},\Gamma)$. Then,

$$\pi_s(\iota(f))(v,\tau)d\tau = \sum_{s(e)=v} \iota(f)(e,\tau)d\tau = \begin{cases} \displaystyle\sum_{s(e)=v} f(e,\tau)d\tau & \text{if } v \text{ is even} \\ \displaystyle\sum_{s(e)=v} f(\alpha e,\alpha\tau)d\alpha\tau & \text{if } v \text{ is odd} \end{cases}$$

$$= \begin{cases} \left(\displaystyle\sum_{s(e)=v} f(e,\tau)\right) d\tau & \text{if } v \text{ is even} \\ \left(\displaystyle\sum_{s(e)=\alpha v} f(e,\alpha\tau)\right) d\alpha\tau & \text{if } v \text{ is odd} \end{cases}$$

$$= 0,$$

Where the last equality follows from (2) in the definition of $S_2(\mathcal{T},\Gamma)$. By the same token, and using the identity $f(e,\tau) = -f(\bar{e},\tau)$, we find that

$$\pi_t(\iota(f))(v,\tau) = 0,$$

implying that $\mathrm{im}(\iota) \subseteq \ker(\pi_s \oplus \pi_t)$.

Now, let $g \in \ker(\pi_s \oplus \pi_t)$. Let

$$f\colon \mathcal{E}(\mathcal{T}) \times \mathcal{H} \longrightarrow \mathbb{C}$$

defined by

$$g(e,\tau)d\tau = \begin{cases} f(e,\tau)d\tau & \text{if } s(e) \text{ is even} \\ f(\alpha^{-1}e, \alpha^{-1}\tau)d\alpha^{-1}\tau & \text{if } s(e) \text{ is odd.} \end{cases}$$

Since $g \in \ker(\pi_s \oplus \pi_t)$, we have that f satisfies condition (2) of the definition of $S_2(\mathcal{T}, \Gamma)$. The invariance of g for $\tilde{\Gamma}$ yields the required invariance of f for Γ. Lastly, it can readily be verified that $\iota(f) = g$. Hence, $\ker(\pi_s \oplus \pi_t) \subseteq \mathrm{im}(\iota)$, as required.

Note that $S_2(\mathcal{E}, \tilde{\Gamma}) \cong S_2(\Gamma^{\varepsilon}_{ns}(q,p))$. This, via the assignment $f \mapsto f_0$, which follows from Proposition 5.7. For each f_0 in the latter space, we can build f using the transitivity of the action of $\tilde{\Gamma}$ on $\mathcal{E}(\mathcal{T})$. Furthermore, this implies that one edge determines the behavior at all edges. In exactly the same way, we can see that $S_2(\mathcal{T}_0, \tilde{\Gamma}) \cong S_2(\Gamma^{\varepsilon}_{ns}(q))$, by taking the modular form attached to v_0, the root of $\mathcal{T}$.

The double coset operator given by $\Gamma^{\varepsilon}_{ns}(q,p) \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) \Gamma^{\varepsilon}_{ns}(q)$ is the natural trace map that symmetrizes a modular form in $S_2(\Gamma^{\varepsilon}_{ns}(q,p))$, yielding a modular form in $S_2(\Gamma^{\varepsilon}_{ns}(q))$, analogous to the level lowering operator in the classical case. The double coset operator given by $\Gamma^{\varepsilon}_{ns}(q,p)\alpha\Gamma^{\varepsilon}_{ns}(q)$ corresponds to the trace map that symmetrizes a modular form in $S_2(\Gamma^{\varepsilon}_{ns}(q,p))$ after applying the involution that arises slashing by α, analogous to the level lowering operator after applying the Atkin-Lehner involution. We denote the former by φ_s and the latter by φ_t. Thus, we obtain a map

$$\varphi_s \oplus \varphi_t \colon S_2(\Gamma^{\varepsilon}_{ns}(q,p)) \longrightarrow S_2(\Gamma^{\varepsilon}_{ns}(q)) \oplus S_2(\Gamma^{\varepsilon}_{ns}(q)).$$

The kernel of this map, as in the classical case, is the space of forms that are *new* at p, which is denoted by $S_2(\Gamma^{\varepsilon}_{ns}(q,p))^{p-\text{new}}$. By abuse of notation, let

$$\iota\colon S_2(\Gamma^{\varepsilon}_{ns}(q,p))^{p-\text{new}} \hookrightarrow S_2(\Gamma^{\varepsilon}_{ns}(q,p))$$

denote this inclusion.

This gives us an exact sequence

$$0 \longrightarrow S_2(\Gamma^{\varepsilon}_{ns}(q,p))^{p-\text{new}} \xrightarrow{\iota} S_2(\Gamma^{\varepsilon}_{ns}(q,p)) \xrightarrow{\varphi_s\oplus\varphi_t} S_2(\Gamma^{\varepsilon}_{ns}(q)) \oplus S_2(\Gamma^{\varepsilon}_{ns}(q)) \ ,$$

whence we derive the diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & S_2(\mathcal{T},\Gamma) & \xrightarrow{\iota} & S_2(\mathcal{E},\tilde{\Gamma}) & \xrightarrow{\pi_s\oplus\pi_t} & S_2(\mathcal{T}_0,\tilde{\Gamma}) \oplus S_2(\mathcal{T}_0,\tilde{\Gamma}) \\ & & \downarrow & & \downarrow & & \downarrow \\ & & S_2(\Gamma^{\varepsilon}_{ns}(q,p)) & \xrightarrow{\iota} & S_2(\Gamma^{\varepsilon}_{ns}(q,p)) & \xrightarrow{\varphi_s\oplus\varphi_t} & S_2(\Gamma^{\varepsilon}_{ns}(q)) \oplus S_2(\Gamma^{\varepsilon}_{ns}(q)), \end{array}$$

where the vertical maps are given by the assignment $f \mapsto f_0$. Commutativity of the left square follows trivially, as $f_0(\tau) = f(e_0,\tau)$ and $s(e_0) = v_0$ is even. Commutativity of the second square is a bit more subtle. The maps φ_s and φ_t can be expressed in a more concrete way by choosing representatives. Let γ_0, $\gamma_1, \ldots,$ γ_p be elements of $\Gamma^{\varepsilon}_{ns}(q)$ (the stabilizer of v_0) which map all the different oriented edges e such that $s(e) = v_0$ to e_0. These form a set of coset representatives

$$\Gamma^{\varepsilon}_{ns}(q) = \bigcup_{j=0}^{p} \Gamma^{\varepsilon}_{ns}(q,p)\gamma_j,$$

so

$$\varphi_s(f) = \sum_{j=0}^{p} f|_2[\gamma_j] \qquad \text{and} \qquad \varphi_t(f) = \sum_{j=0}^{p} f|_2[\alpha\gamma_j].$$

Let $f \in S_2(\mathcal{E}, \tilde{\Gamma})$ and let $j(\gamma, \tau)$ be the automorphy factor of γ at τ, i.e.,

$$\text{for} \quad g\colon \mathcal{H} \longrightarrow \mathbb{C}, \qquad g|_2[\gamma] = \det(\gamma) j(\gamma, \tau)^{-2} g(\gamma\tau).$$

Hence,

$$\begin{aligned} f(e_0, \tau)|_2[\gamma_j] &= \det(\gamma_j) j(\gamma_j, \tau)^{-2} f(e_0, \gamma_j\tau) = \det(\gamma_j) j(\gamma_j, \tau)^{-2} f(\gamma_j\gamma_j^{-1} e_0, \gamma_j\tau) \\ &= f(\gamma_j^{-1} e_0, \tau) \end{aligned}$$

and

$$\begin{aligned} f(e_0, \tau)|_2[\alpha\gamma_j] &= \det(\alpha\gamma_j) j(\alpha\gamma_j, \tau)^{-2} f(e_0, \alpha\gamma_j\tau) \\ &= \det(\alpha\gamma_j) j(\alpha\gamma_j, \tau)^{-2} f(\alpha\gamma_j\gamma_j^{-1}\alpha^{-1} e_0, \alpha\gamma_j\tau) = f(\gamma_j^{-1}\alpha^{-1} e_0, \tau). \end{aligned}$$

These two identities show that

$$\begin{aligned} \pi_s(f)_0(\tau) &= \sum_{s(e)=v_0} f(e, \tau) = \sum_{j=0}^{p} f(\gamma_j^{-1} e_0, \tau) = \sum_{j=0}^{p} f(e_0, \tau)|_2[\gamma_j] \\ &= \sum_{j=0}^{p} f_0(\tau)|_2[\gamma_j] = \varphi_s(f_0)(\tau) \end{aligned}$$

and

$$\begin{aligned} \pi_t(f)_0(\tau) &= \sum_{t(e)=v_0} f(e, \tau) = \sum_{j=0}^{p} f(\gamma_j^{-1}\alpha^{-1} e_0, \tau) = \sum_{j=0}^{p} f(e_0, \tau)|_2[\alpha\gamma_j] \\ &= \sum_{j=0}^{p} f_0(\tau)|_2[\alpha\gamma_j] = \varphi_t(f_0)(\tau), \end{aligned}$$

which establish the sought commutativity.

Since the diagram is commutative, for every $f \in S_2(\mathcal{T}, \Gamma)$ we have that

$$(\varphi_s \oplus \varphi_t)(\iota(f)_0) = ((\pi_s \oplus \pi_t)(\iota(f)))_0 = 0_0 = 0,$$

so $f_0 = \iota(f)_0 \in \ker(\varphi_s \oplus \varphi_t) = S_2(\Gamma^{\varepsilon}_{ns}(q,p))^{p-\text{new}}$. This means that we actually have the commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & S_2(\mathcal{T}, \Gamma) & \xrightarrow{\iota} & S_2(\mathcal{E}, \tilde{\Gamma}) & \xrightarrow{\pi_s \oplus \pi_t} & S_2(\mathcal{T}_0, \tilde{\Gamma}) \oplus S_2(\mathcal{T}_0, \tilde{\Gamma}) \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & S_2(\Gamma^{\varepsilon}_{ns}(q,p))^{p-\text{new}} & \xrightarrow{\iota} & S_2(\Gamma^{\varepsilon}_{ns}(q,p)) & \xrightarrow{\varphi_s \oplus \varphi_t} & S_2(\Gamma^{\varepsilon}_{ns}(q)) \oplus S_2(\Gamma^{\varepsilon}_{ns}(q)). \end{array}$$

The required isomorphism follows from the five-lemma. □

5.3 Measures, double Integrals and semi-indefinite Integrals

Starting with an elliptic curve E of conductor pq^2 we have a normalized weight 2 newform g associated to E. Theorem 4.21 associates to g an eigenform $f_0 \in S_2(\Gamma)$, where $\Gamma = \Gamma^{\varepsilon}_{ns}(q,p)$ or $\Gamma^{+}_{ns}(q,p)$ depending on the sign of the Atkin-Lehner involution at q, with the same eigenvalues as g away from q. Theorem 4.20 yields a multiplicative constant that shows we can assume f_0 to be a normalized rational eigenform, i.e., for every $\sigma_\ell \in \text{Gal}(\mathbb{Q}(\xi_q)/\mathbb{Q})$ we have $f_0|_2[A_\ell] = \sigma_\ell(f_0)$ (where A_ℓ is a matrix such that $\sigma_\ell(\infty) = A_\ell \cdot \infty$), its q-expansion has algebraic integers as coefficients and for every integer $m \geq 2$, f_0/m does not fulfil this last property. Since g is new at p (in fact, everywhere), f_0 will also be new at p. Finally, Lemma 5.8 associates to f_0 a modular form $f \in S_2(\mathcal{T}, \Gamma)$.

As in Section 4.3.3, in order to define a system of measures it suffices to define the values at the open sets U_e associated to each $e \in \mathcal{T}_1$. Let $x, y \in \mathbb{P}^1(\mathbb{Q})$ and c be

the Manin constant associated to E via f_0, as in Section 4.4.3. Let

$$\tilde{\mu}_f\{x \to y\}(U_e) = c \cdot 2\pi i \int_x^y f_e(z)dz.$$

Note that for $\gamma \in \Gamma$ we have

$$\begin{aligned}\tilde{\mu}_f\{\gamma x \to \gamma y\}(U_{\gamma e}) &= c \cdot 2\pi i \int_{\gamma x}^{\gamma y} f(\gamma e, z)dz \\ &= c \cdot 2\pi i \int_x^y f(e, \tau)d\tau = \tilde{\mu}_f\{x \to y\}(U_e), \end{aligned} \tag{5.1}$$

using the identity $f(\gamma e, \gamma\tau)d\gamma\tau = f(e, \tau)d\tau$.

Since the Abel-Jacobi map is not a rational map in this case, we need to average it following the recipe from Equation (4.4). The value $\tilde{\mu}_f\{x \to y\}(U_e)$ can be expressed as

$$\tilde{\mu}_f\{x \to y\}(U_e) = c \cdot 2\pi i \int_x^y f_e(z)dz = c \cdot 2\pi i \int_{\sigma_\ell(\infty)}^y f_e(z)dz - c \cdot 2\pi i \int_{\sigma_\ell(\infty)}^x f_e(z)dz,$$

so by averaging over all cusps we end up obtaining the measure

$$\tilde{\mu}'_f\{x \to y\}(U_e) = (q-1)\tilde{\mu}_f\{x \to y\}(U_e).$$

All these values can, again, be expressed in terms of the periods of f_0, which lie on a lattice $\Lambda_E \subseteq \mathbb{C}$. Denote by $\tilde{\Omega}_+$ and $\tilde{\Omega}_-$ the smallest positive real period and the smallest purely imaginary (with positive imaginary part) period of Λ_E, respectively. We know that Λ_E is generated either by $\tilde{\Omega}_+$ and $\tilde{\Omega}_-$ or contains the lattice generated

by these two periods with index 2 (see Chapter 2.8 in [Cre92]). Let

$$\Omega_{\pm} = \begin{cases} \tilde{\Omega}_{\pm} & \text{if } \Lambda_E = \langle \tilde{\Omega}_+, \tilde{\Omega}_- \rangle \\ \tilde{\Omega}_{\pm}/2 & \text{otherwise.} \end{cases}$$

Thus,

$$\tilde{\mu}_f\{x \to y\}(U_e) = \kappa_f^+\{x \to y\}(e) \cdot \Omega_+ + \kappa_f^-\{x \to y\}(e) \cdot \Omega_-,$$

where $\kappa_f^{\pm}\{x \to y\} \colon \mathcal{E}(\mathcal{T}) \longrightarrow \mathbb{Z}$. In order to account for the averaging of the cusps, we multiply this quantity by $q-1$. Choose a sign at infinity $w_\infty = \pm 1$ and define the system of integral distributions

$$\mu_f\{x \to y\}(U_e) = (q-1)\kappa_f^{w_\infty}\{x \to y\}(e).$$

The integrality of the distributions makes them p-adically bounded, so they are actually p-adic measures on $\mathbb{P}^1(\mathbb{Q}_p)$. The line integral from Section 4.3.1 provides us with the same double integral we had in Section 4.3.3. This is, attached to f, from τ_1 to τ_2 in $\mathcal{H}_p$, and from x to y in $\mathbb{P}^1(\mathbb{Q})$, we have

$$\int_{\tau_1}^{\tau_2} \int_x^y \omega_f = \int_{\mathbb{P}^1(\mathbb{Q}_p)} \log\left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{x \to y\}(t).$$

We also have the multiplicative counterpart, defined as

$$\prod_{\tau_1}^{\tau_2} \int_x^y \omega_f = \prod_{\mathbb{P}^1(\mathbb{Q}_p)} \left(\frac{t-\tau_2}{t-\tau_1}\right) d\mu_f\{x \to y\}(t) = \lim \prod_{\alpha} \left(\frac{t_\alpha - \tau_2}{t_\alpha - \tau_1}\right)^{\mu_f\{x \to y\}(U_\alpha)},$$

where the disjoint compact open sets U_α cover $\mathbb{P}^1(\mathbb{Q}_p)$ and the limit is taken over increasingly finer covers, with $t_\alpha \in U_\alpha$ is any sample point.

We also have an analogue of Lemma 4.8 in this scenario, exactly as expected.

Lemma 5.9. *For all $\tau_1, \tau_2, \tau_3 \in \mathcal{H}_p$, $x, y, z \in \mathbb{P}^1(\mathbb{Q})$ and $\gamma \in \Gamma$, the double integrals satisfy*

(1)

$$\oint_{\tau_1}^{\tau_3} \int_x^y \omega_f = \oint_{\tau_1}^{\tau_2} \int_x^y \omega_f \cdot \oint_{\tau_2}^{\tau_3} \int_x^y \omega_f \quad \textit{and} \quad \int_{\tau_1}^{\tau_3} \int_x^y \omega_f = \int_{\tau_1}^{\tau_2} \int_x^y \omega_f + \int_{\tau_2}^{\tau_3} \int_x^y \omega_f$$

(2)

$$\oint_{\tau_1}^{\tau_2} \int_x^z \omega_f = \oint_{\tau_1}^{\tau_2} \int_x^y \omega_f \cdot \oint_{\tau_1}^{\tau_2} \int_y^z \omega_f \quad \textit{and} \quad \int_{\tau_1}^{\tau_2} \int_x^z \omega_f = \int_{\tau_1}^{\tau_2} \int_x^y \omega_f + \int_{\tau_1}^{\tau_2} \int_y^z \omega_f$$

(3)

$$\oint_{\gamma\tau_1}^{\gamma\tau_2} \int_{\gamma x}^{\gamma y} \omega_f = \oint_{\tau_1}^{\tau_2} \int_x^y \omega_f \quad \textit{and} \quad \int_{\gamma\tau_1}^{\gamma\tau_2} \int_{\gamma x}^{\gamma y} \omega_f = \int_{\tau_1}^{\tau_2} \int_x^y \omega_f$$

Proof. Properties (1) and (2) follow formally from the definitions. Property (3) follows from Equation (5.1), as it implies

$$\mu_f\{\gamma x \to \gamma y\}(U_{\gamma e}) = \mu_f\{x \to y\}(U_e),$$

whence (3) is clear. □

Let $\Gamma' \subset \Gamma$ be the normal closure of the subset of Γ comprised of the all the commutators and the elements of Γ whose fixed points belong to $\mathbb{P}^1(\mathbb{Q})$. Denote by e_Γ the exponent of the group Γ/Γ'.

Proposition 5.10. *The exponent e_Γ is a divisor of $q+1$.*

Proof. Let

$$M = \begin{pmatrix} 1 & 0 \\ q & 1 \end{pmatrix},$$

which lies in $\Gamma(q)[1/p], \Gamma$ and $\mathbf{SL}_2(\mathbb{Z}[1/p])$. Let H denote the normal closure of M in $\mathbf{SL}_2(\mathbb{Z}[1/p])$. Theorem 1 in [Men67] establishes that H is precisely $\Gamma(q)[1/p]$. On the other hand, we have that H is generated by the set $\{AMA^{-1}\colon A \in \mathbf{SL}_2(\mathbb{Z}[1/p])\}$.

Let $r = A(0) \in \mathbb{P}^1(\mathbb{Q})$. Note that

$$AMA^{-1}(r) = AM(0) = A(0) = r,$$

as M stabilizes 0. This implies that $AMA^{-1} \in \Gamma'$, as it fixes the cusp r and AMA^{-1} belongs to Γ (it belonging to $\Gamma(q)[1/p] \subseteq \Gamma$). This, in turn, implies that $H \subseteq \Gamma'$, as the generating set of H is a subset of Γ'.

Now, the matrix $-I$, where I is the identity, fixes all of the cusps and is in Γ, so we actually have that $\pm\Gamma(q)[1/p] \subseteq \Gamma'$. From the multiplicativity of indices of nested subgroups we obtain the relation

$$[\Gamma : \Gamma(q)[1/p]] = [\Gamma : \Gamma'] \cdot [\Gamma' : \Gamma(q)[1/p]].$$

From Lemma 5.2 we have

$$\Gamma^{\varepsilon}_{ns}(q)[1/p]/\Gamma(q)[1/p] \cong \mathbb{Z}/(q+1)\mathbb{Z},$$

completing the proof. □

Let $\mathfrak{q}$ be Tate's p-adic period attached to E. In Section 4.3.5 we constructed cochains κ_τ, $\rho_{x,y}$, ρ_{τ_1,τ_2}, $\kappa^{\#}_\tau$ and c_τ. These were constructed in terms of the double integrals attached to the modular forms with which we were dealing back then. The same formal computations show that κ_τ is a cocycle and that, thanks to $\rho_{x,y}$ and ρ_{τ_1,τ_2}, it does not depend on the base point or on τ when seen in $H^2(\Gamma, \mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}})$.

Choose a cusp $x \in \mathbb{P}^1(\mathbb{Q})$. Before, $\Gamma = \Gamma_0(M)[1/p]$. If $\Gamma = \Gamma_{ns}^{\epsilon}(q)[1/p]$ or $\Gamma = \Gamma_{ns}^{+}(q)[1/p]$, we obtain the same exact sequence

$$0 \longrightarrow \mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}} \longrightarrow \mathcal{F} \longrightarrow \mathcal{M}_0 \longrightarrow 0,$$

where $\mathcal{F}$ is the group of $\mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}}$-valued functions on Γx and $\mathcal{M}_0$ is the group arising from restriction on $\mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}}$-valued modular symbols on $\Gamma x \times \Gamma x$. This induces a long exact sequence of cohomology groups which yields the connecting homomorphism

$$\delta\colon H^1(\Gamma, \mathcal{M}_0) \longrightarrow H^2(\Gamma, \mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}}).$$

Again, the same computations show that δ applied to c_τ is $\kappa_\tau^{\#}$. As a natural analogue of Conjecture 4.12, we formulate the following conjecture.

Conjecture 5.11. *The class of c_τ is trivial in $H^1(\Gamma, \mathcal{M}_0)$.*

As in Section 4.3.5, Conjecture 5.11 implies that c_τ is a coboundary, so we obtain a cochain $\tilde{\eta}_\tau$, unique up to multiplication by a 0-cocycle, such that

$$\int\!\!\!\!\!\!\diagup_{\tau}^{\gamma\tau} \int_y^z \omega_f = c_\tau(\gamma)\{y \to z\} = \tilde{\eta}_\tau\{\gamma^{-1}y \to \gamma^{-1}z\} \div \tilde{\eta}_\tau\{y \to z\}. \tag{5.2}$$

Lemma 5.12. *The map*

$$\begin{aligned} h'\colon \mathcal{M}_0^\Gamma &\longrightarrow \quad \mathrm{Hom}(\Gamma/\Gamma', \mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}}) \\ m &\longmapsto (\gamma\Gamma' \mapsto m\{x \to \gamma x\}) \end{aligned}$$

is a monomorphism.

Proof. See comments leading to Lemma 4.13. □

Lemma 5.10 implies that e_Γ is finite (and gives a bound for it). Since e_Γ is the exponent of Γ/Γ', Lemma 5.12 implies that any 0-cocyle is annihilated when raised to the e_Γ-th power. Then, the modular symbol $\tilde{\eta}_\tau^{e_\Gamma}$, which does not depend on the choice of $\tilde{\eta}_\tau$, yields a unique modular symbol, which we denote, momentarily, η_τ, such that

$$\left(\not\int_\tau^{\gamma\tau}\int_y^z \omega_f\right)^{e_\Gamma} = \eta_\tau\{\gamma^{-1}y \to \gamma^{-1}z\} \div \eta_\tau\{y \to z\}.$$

It is customary to use the notation

$$\not\int^\tau \int_y^z e_\Gamma\omega_f = \eta_\tau\{y \to z\}.$$

As in Section 4.3.5, we summarize all this in a conjecture. (Not a theorem, as it relies on Conjecture 5.11.)

Conjecture 5.13. *There exists a unique function*

$$\begin{array}{ccc} \mathcal{H}_p(\mathbb{C}_p)\times\Gamma x\times\Gamma x \longrightarrow & & \mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}} \\ (\tau,r,s) & \longmapsto & \not\int^\tau\int_r^s n\omega_f, \end{array}$$

such that, for all τ_1, $\tau_2 \in \mathcal{H}_p(\mathbb{C}_p)$, r, s, $t \in \Gamma x$ *and* $\gamma \in \Gamma$, *we have*

(1)

$$\not\int^\tau\int_r^s e_\Gamma\omega_f \times \not\int^\tau\int_s^t e_\Gamma\omega_f = \not\int^\tau\int_r^t e_\Gamma\omega_f$$

(2)

$$\not\int^{\tau_2}\int_r^s e_\Gamma\omega_f \div \not\int^{\tau_1}\int_r^s e_\Gamma\omega_f = \left(\not\int_{\tau_1}^{\tau_2}\int_r^s \omega_f\right)^{e_\Gamma}$$

(3)

$$\not\int^{\gamma\tau}\int_{\gamma r}^{\gamma s} e_\Gamma\omega_f = \not\int^\tau\int_r^s e_\Gamma\omega_f$$

5.4 The Stark-Heegner point

Let $\mathcal{O}_D \subseteq K$ be the order in K of discriminant D, with conductor relatively prime to N. Let $\omega_D = (D + \sqrt{D})/2$, so $\mathcal{O}_D = \mathbb{Z}[\omega_D]$.

Recall that in Section 4.4.5 we showed that if the discriminant of $\mathcal{O}_D$ was relatively prime to pq and q was inert in K we had an embedding

$$\mathcal{O}_D \hookrightarrow M_{ns}^{\varepsilon}(q) \subseteq M_{ns}^{\varepsilon}(q)[1/p],$$

so regard $\mathcal{O}_D$ as a subring of $M_{ns}^{\varepsilon}(q)[1/p]$.

Let τ be a fixed point of ω_D under the Möbius action, this is, if $\omega_D = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$\tau = \frac{a\tau + b}{c\tau + d} \qquad \text{or} \qquad c\tau^2 + (d-a)\tau - b = 0.$$

This is equivalent to saying that the column vector $\vec{\tau} = \begin{pmatrix} \tau \\ 1 \end{pmatrix}$ is an eigenvector of ω_D. Thus, if $r + s\omega_D \in \mathcal{O}_D$, we find that

$$(r + s\omega_D) \begin{pmatrix} \tau \\ 1 \end{pmatrix} = (r + s\omega) \begin{pmatrix} \tau \\ 1 \end{pmatrix},$$

where ω is the eigenvalue of ω_D associated to $\vec{\tau}$, and r and s are, by abuse of notation, scalar matrices on the LHS and integers on the RHS. Hence, τ is a fixed point of every element in $\mathcal{O}_D$.

Note that the discriminant of the quadratic equation defining τ is

$$\begin{aligned}(d-a)^2+4bc &= (d+a)^2-4(ad-bc) = (\omega_D+\bar{\omega}_D)^2-4\omega_D\bar{\omega}_D \\ &= \left(\frac{D+\sqrt{D}}{2}+\frac{D-\sqrt{D}}{2}\right)^2-4\left(\frac{D+\sqrt{D}}{2}\right)\left(\frac{D-\sqrt{D}}{2}\right) \\ &= D^2-(D^2-D) = D,\end{aligned}$$

so $\tau \in K$. Furthermore, since $K = \mathbb{Q}(\sqrt{D})$ and p is inert in K, $\sqrt{D}$ and, a posteriori τ, lie in $\mathbb{Q}_{p^2} - \mathbb{Q}_p$, where $\mathbb{Q}_{p^2}$ is the unramified quadratic extension of $\mathbb{Q}_p$. Combining these two, we find that $\tau \in \mathcal{H}_p \cap K$.

Let γ_τ be a generator of $(\mathcal{O}_D^\times)_1$, which is a rank one abelian group, contained in $M_{ns}^\varepsilon(q)[1/p]_1^\times = \Gamma_{ns}^\varepsilon(q)[1/p]$. If g is invariant under the Atkin-Lehner involution at q, the embedding is in $M_{ns}^+(q)[1/p]$ and we regard γ_τ as an element of $\Gamma_{ns}^+(q)[1/p]$. Denote by Γ the group $\Gamma_{ns}^\varepsilon(q)[1/p]$ or $\Gamma_{ns}^+(q)[1/p]$ accordingly.

Definition 5.14. Let $\tau \in \mathcal{H}_p \cap K$ as above and let $x \in \mathbb{P}^1(\mathbb{Q})$ be a cusp. The Stark-Heegner point associated to τ at the cusp x is the point $P_{\tau,x} \in E(\mathbb{C}_p)$ given by the Tate's uniformization corresponding to

$$\not\int_r^{\tau} \int_r^{\gamma_\tau r} e_\Gamma \omega_f \in \mathbb{C}_p^\times/\mathfrak{q}^{\mathbb{Z}}$$

for any $r \in \Gamma x$.

Note that this point is independent of the choice of cusp in Γx, as choosing two cusps $r, s \in \Gamma x$ yields

$$\begin{aligned}\fint^{\tau}\int_{r}^{\gamma_\tau r} e_\Gamma\omega_f \div \fint^{\tau}\int_{s}^{\gamma_\tau s} e_\Gamma\omega_f &= \fint^{\tau}\int_{r}^{s} e_\Gamma\omega_f \div \fint^{\tau}\int_{\gamma_\tau r}^{\gamma_\tau s} e_\Gamma\omega_f \\ &= \fint^{\tau}\int_{r}^{s} e_\Gamma\omega_f \div \fint^{\gamma_\tau^{-1}\tau}\int_{r}^{s} e_\Gamma\omega_f = 1,\end{aligned}$$

where the last equality comes from the fact that γ_τ stabilizes τ and the properties of Conjecture 5.13.

We further formulate the following conjecture regarding the algebraicity of the resulting points.

Conjecture 5.15. *Let $\tau \in \mathcal{H}_p \cap K$ as above and let $x \in \mathbb{P}^1(\mathbb{Q})$ be a cusp. Let H^+ be the narrow ring class field attached to $\mathcal{O}_D$. Then the point $P_{\tau,x}$ lies in $E(H^+)$.*

5.5 Setup for computations

Now that we defined the points, we would like to know how to actually compute them. [DP06] shows how to effectively compute Stark-Heegner points in the case of prime level. The techniques explained in Chapter 2 show how to compute the p-adic multiplicative double integrals to high accuracy in a general setup, but the techniques mentioned in Chapter 1 only allow for computation of the semi-indefinite integrals in the case of prime level. In [GM15], there is a method that works for composite levels, which again boils down to the effective computation of the double integrals to high accuracy.

5.5.1 The case $q = 3$.

Throughout this subsection, we will assume that the newform g is invariant under the usual Atkin-Lehner involution at q, so we will be working with the group

$\Gamma = \Gamma^+_{ns}(3)[1/p]$. By Corollary 5.6, all cusps are equivalent under the action of Γ (even without inverting p). This means that every cusp is in $\Gamma\infty$ and we can focus on only one indefinite integral, namely

$$\not\int^{\tau} \int_{\infty}^{\gamma_\tau \infty} e_\Gamma \omega_f. \tag{5.3}$$

Utilizing the continued fraction expansion of the rational number $\gamma_\tau\infty$, we obtain a (finite) sequence $\{p_j/q_j\}_{j=-1}^{k}$ of rational numbers. The matrices

$$M_j = \begin{pmatrix} (-1)^{j-1}p_j & p_{j-1} \\ (-1)^{j-1}q_j & q_{j-1} \end{pmatrix}$$

lie in $\mathbf{SL}_2(\mathbb{Z})$. The group $\Gamma^+_{ns}(3)$ has index 3 in $\mathbf{SL}_2(\mathbb{Z})$, so it is the disjoint union of its left-cosets

$$\mathbf{SL}_2(\mathbb{Z}) = \Gamma^+_{ns}(3) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bigcup \Gamma^+_{ns}(3) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \bigcup \Gamma^+_{ns}(3) \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix},$$

and we can write $M_j = \gamma_j r_j$, where $\gamma_j \in \Gamma^+_{ns}(3)$ and

$$r_j \in \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \right\}.$$

By properties (1) and (3) in Conjecture 5.13, we have

$$\begin{aligned} \not\int^{\tau} \int_{\infty}^{\gamma_\tau \infty} e_\gamma \omega_f &= \prod_{j=0}^{k} \not\int^{\tau} \int_{p_{j-1}/q_{j-1}}^{p_j/q_j} e_\Gamma \omega_f = \prod_{j=0}^{k} \not\int^{\tau} \int_{M_j(0)}^{M_j(\infty)} e_\Gamma \omega_f \\ &= \prod_{j=0}^{k} \not\int^{\tau} \int_{\gamma_j r_j(0)}^{\gamma_j r_j(\infty)} e_\Gamma \omega_f = \prod_{j=0}^{k} \not\int^{\gamma_j{}^{-1}\tau} \int_{r_j(0)}^{r_j(\infty)} e_\Gamma \omega_f, \end{aligned} \tag{5.4}$$

so we reduce the computation of (5.3) to that of computing

$$\fint_{}^{\tau'} \int_0^{\infty} e_\Gamma \omega_f, \quad \fint_{}^{\tau'} \int_0^{1} e_\Gamma \omega_f, \quad \fint_{}^{\tau'} \int_0^{1/2} e_\Gamma \omega_f, \tag{5.5}$$

for some $\tau' \in \mathcal{H}_p$. Note that in order to apply property (1), it is crucial that the cusps p_{j-1}/q_{j-1} and p_j/q_j be equivalent, which is guaranteed by the assumption $q = 3$ together with g having eigenvalue 1 at 3.

Consider the matrices

$$W_{0,\infty} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad W_{0,1} = \begin{pmatrix} -1 & 1 \\ -2 & 1 \end{pmatrix}, \qquad W_{0,1/2} = \begin{pmatrix} -2 & 1 \\ -5 & 2 \end{pmatrix},$$

which lie in $\Gamma_{ns}^+(3)$ and have the property that

$$W_{r,s}(r) = s \qquad \text{and} \qquad W_{r,s}(s) = r,$$

so, using the properties spelled out in Conjecture 5.13, we have the equality

$$\begin{aligned} \left(\fint_{}^{\tau'} \int_r^s e_\Gamma \omega_f \right)^2 &= \fint_{}^{\tau'} \int_r^s e_\Gamma \omega_f \div \fint_{}^{\tau'} \int_s^r e_\Gamma \omega_f = \fint_{}^{\tau'} \int_r^s e_\Gamma \omega_f \div \fint_{}^{\tau'} \int_{W_{r,s}^{-1}(r)}^{W_{r,s}^{-1}(s)} e_\Gamma \omega_f \\ &= \fint_{}^{\tau'} \int_r^s e_\Gamma \omega_f \div \fint_{}^{W_{r,s}\tau'} \int_r^s e_\Gamma \omega_f = \left(\fint_{W_{r,s}\tau'}^{\tau'} \int_r^s \omega_f \right)^{e_\Gamma}. \end{aligned} \tag{5.6}$$

Finally, in order to compute double integrals of the form

$$\fint_{\tau_1}^{\tau_2} \int_r^s,$$

we just need to follow what was explained at the end of Section 4.3.3.

5.5.2 More general values of q.

There are two obstructions in generalizing the procedure described in the previous subsection to larger values of q. On one hand, as mentioned right after Equation (5.5), we require the cusps p_{j-1}/q_{j-1} and p_j/q_j to be equivalent for every j. This obstruction can be salvaged by imposing a *nice* relationship between q and p. For example, Proposition 5.6 says that when p^2 generates all of the squares modulo q, all the cusps are equivalent under the action of $\Gamma_{ns}^+(q)[1/p]$.

On the other hand, we exhibited coset representatives r_j for which we found matrices $W_{r_j(0),r_j(\infty)} \in \Gamma_{ns}^+(q)[1/p]$ with the property that

$$W_{r_j(0),r_j(\infty)}(r_j(0)) = r_j(\infty) \qquad \text{and} \qquad W_{r_j(0),r_j(\infty)}(r_j(\infty)) = r_j(0).$$

A natural question to ask at this point is, when do these elements exist?

Proposition 5.16. *Let*

$$A = \begin{pmatrix} r_\infty & r_0 \\ s_\infty & s_0 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}).$$

Denote by r and s the cusps $A(\infty)$ and $A(0)$, respectively. Then, there exists a matrix $W_{r,s} \in \Gamma_{ns}^+(q)[1/p]$ whose action on $\mathbb{P}^1(\mathbb{Q})$ transposes r and s if and only if one of the following three conditions is met:

(i)

$$-\frac{s_0^2 - \varepsilon r_0^2}{s_\infty^2 - \varepsilon r_\infty^2} \in \langle p^2 \rangle$$

(ii)

$$q \mid r_\infty r_0 s_\infty s_0, \quad q \mid (r_\infty - s_0)(s_\infty - r_0) \quad \textit{and} \quad \frac{s_0^2 - \varepsilon r_0^2}{s_\infty^2 - \varepsilon r_\infty^2} \in \langle p^2 \rangle$$

(iii)

$$q \nmid r_\infty r_0 s_\infty s_0, \quad -\frac{r_0 s_0}{r_\infty s_\infty} \in \langle p^2 \rangle \quad \textit{and} \quad \varepsilon \equiv \pm \frac{s_\infty s_0}{r_\infty r_0} \pmod{q}.$$

Proof. Note that if $W = W_{r,s}$ exists, then we have

$$\big(WA\big)(0) = A(\infty) \quad \text{and} \quad \big(WA\big)(\infty) = A(0)$$

which is equivalent to

$$\big(A^{-1}WA\big)(0) = \infty \quad \text{and} \quad \big(A^{-1}WA\big)(\infty) = 0,$$

so the matrix $A^{-1}WA$ swaps 0 and ∞. The only matrices which swap 0 and ∞ have the form

$$\begin{pmatrix} 0 & * \\ * & 0 \end{pmatrix},$$

and if they lie in $\mathbf{SL}_2(\mathbb{Z}[1/p])$, then we have the equality

$$A^{-1}WA = \begin{pmatrix} 0 & \pm p^\alpha \\ \mp p^{-\alpha} & 0 \end{pmatrix}$$

for some $\alpha \in \mathbb{Z}$. Solving for W, we obtain

$$W = A \begin{pmatrix} 0 & \pm p^\alpha \\ \mp p^{-\alpha} & 0 \end{pmatrix} A^{-1} = \begin{pmatrix} \mp r_0 s_0 p^{-\alpha} \mp r_\infty s_\infty p^\alpha & \pm r_0^2 p^{-\alpha} \pm r_\infty^2 p^\alpha \\ \mp s_0^2 p^{-\alpha} \mp s_\infty^2 p^\alpha & \pm r_0 s_0 p^{-\alpha} \pm r_\infty s_\infty p^\alpha \end{pmatrix}.$$

Note that the trace of W is 0. If the diagonal entries are not divisible by q, then $W \in \Gamma_{ns}^+(q)[1/p]$ if and only if $W \in \Gamma_{ns}^+(q)[1/p] - \Gamma_{ns}^\varepsilon(q)[1/p]$, as q is odd. Thus, we

have that $W \in \Gamma^+_{ns}(q)[1/p]$ if and only if

$$\begin{aligned} \pm s_0^2 p^{-\alpha} \pm s_\infty^2 p^\alpha &\equiv (\pm r_0^2 p^{-\alpha} \pm r_\infty^2 p^\alpha)\varepsilon \pmod q \\ s_0^2 + s_\infty^2 p^{2\alpha} &\equiv (r_0^2 + r_\infty^2 p^{2\alpha})\varepsilon \pmod q \\ -\frac{s_0^2 - \varepsilon r_0^2}{s_\infty^2 - \varepsilon r_\infty^2} &\equiv p^{2\alpha} \pmod q, \end{aligned}$$

whence (i).

The diagonal entries are divisible by q if and only if

$$r_0 s_0 \equiv -r_\infty s_\infty p^{2\alpha} \pmod q. \tag{5.7}$$

If this is the case, then $W \in \Gamma^+_{ns}(q)[1/p]$ if and only if

$$\begin{aligned} \pm(\pm s_0^2 p^{-\alpha} \pm s_\infty^2 p^\alpha) &\equiv (\pm r_0^2 p^{-\alpha} \pm r_\infty^2 p^\alpha)\varepsilon \pmod q \\ \pm(s_0^2 + s_\infty^2 p^{2\alpha}) &\equiv (r_0^2 + r_\infty^2 p^{2\alpha})\varepsilon \pmod q. \end{aligned} \tag{5.8}$$

Now, we have two cases, depending on whether or not q divides $r_\infty r_0 s_\infty s_0$. If $q \mid r_\infty r_0 s_\infty s_0$, then Equation (5.7) and $r_\infty s_0 - r_0 s_\infty = 1$ imply $q \mid (r_\infty - s_0)(s_\infty - r_0)$.

If $r_\infty \equiv s_0 \equiv 0 \pmod q$, Equation (5.8) transforms into

$$\begin{aligned} \pm s_\infty^2 p^{2\alpha} &\equiv r_0^2 \varepsilon \pmod q \\ \pm\frac{\varepsilon r_0^2}{s_\infty^2} &\equiv p^{2\alpha} \pmod q. \end{aligned}$$

The RHS is a square, so $\pm\varepsilon$ has to be a square as well which is only possible if we have the negative sign, and the condition becomes equivalent to

$$\frac{s_0^2 - \varepsilon r_0^2}{s_\infty^2 - \varepsilon r_\infty^2} \equiv p^{2\alpha} \pmod q.$$

If $r_0 \equiv s_\infty \equiv 0 \pmod{q}$, Equation (5.8) transforms into

$$\pm s_0^2 \equiv r_\infty p^{2\alpha} \varepsilon \pmod{q}$$
$$\pm \frac{s_0^2}{\varepsilon r_\infty^2} \equiv p^{2\alpha} \pmod{q}.$$

Like before, we must choose the negative sign and again the condition translates into

$$\frac{s_0^2 - \varepsilon r_0^2}{s_\infty^2 - \varepsilon r_\infty^2} \equiv p^{2\alpha} \pmod{q},$$

whence (ii).

Finally, if $q \nmid r_\infty r_0 s_\infty s_0$, from Equation (5.7) we obtain

$$-\frac{r_0 s_0}{r_\infty s_\infty} \equiv p^{2\alpha} \pmod{q}.$$

Then, Equation (5.8) transforms into

$$\pm(s_0^2 - r_0 s_0 s_\infty / r_\infty) \equiv (r_0^2 - r_\infty r_0 s_0 / s_\infty)\varepsilon \pmod{q}.$$

Clearing denominators and factoring we obtain

$$\pm(r_\infty s_0 - r_0 s_\infty) s_0 s_\infty \equiv -(r_\infty s_0 - r_0 s_\infty) r_\infty r_0 \varepsilon \pmod{q},$$

whence

$$\varepsilon \equiv \pm \frac{s_\infty s_0}{r_\infty r_0} \pmod{q},$$

which is precisely (iii). □

Corollary 5.17. *Suppose that* $\langle p^2 \rangle = ((\mathbb{Z}/q\mathbb{Z})^\times)^2$. *Let*

$$A = \begin{pmatrix} r_\infty & r_0 \\ s_\infty & s_0 \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}).$$

Denote by r *and* s *the cusps* $A(\infty)$ *and* $A(0)$, *respectively. Then, there exists a matrix* $W_{r,s} \in \Gamma_{ns}^+(q)[1/p]$ *whose action on* $\mathbb{P}^1(\mathbb{Q})$ *transposes* r *and* s *if and only if one of the following three conditions is met:*

(i)

$$\left(\frac{\varepsilon - (s_\infty s_0 - \varepsilon r_\infty r_0)^2}{q}\right) = 1$$

(ii)

$$q \mid r_\infty r_0 s_\infty s_0, \quad q \mid (r_\infty - s_0)(s_\infty - r_0) \quad \textit{and} \quad \left(\frac{(s_\infty s_0 - \varepsilon r_\infty r_0)^2 - \varepsilon}{q}\right) = 1$$

(iii)

$$q \nmid r_\infty r_0 s_\infty s_0, \quad \left(\frac{-r_\infty r_0 s_\infty s_0}{q}\right) = 1 \quad \textit{and} \quad \varepsilon \equiv \pm \frac{s_\infty s_0}{r_\infty r_0} \pmod{q}.$$

Proof. Modulo squares, the elements

$$\frac{s_0^2 - \varepsilon r_0^2}{s_\infty^2 - \varepsilon r_\infty^2} \qquad \text{and} \qquad (s_\infty^2 - \varepsilon r_\infty^2)(s_0^2 - \varepsilon r_0^2)$$

are the same. Notice that

$$\begin{aligned}(s_\infty^2 - \varepsilon r_\infty^2)(s_0^2 - \varepsilon r_0^2) &= s_\infty^2 s_0^2 - \varepsilon r_0^2 s_\infty^2 - \varepsilon r_\infty^2 s_0^2 + \varepsilon^2 r_\infty^2 r_0^2 \\ &= s_\infty^2 s_0^2 + \varepsilon^2 r_\infty^2 r_0^2 - \varepsilon(r_\infty^2 s_0^2 + r_0^2 s_\infty^2) \\ &= (s_\infty^2 s_0^2 + 2 s_\infty s_0 \varepsilon r_\infty r_0 + \varepsilon^2 r_\infty^2 r_0^2) - \varepsilon(r_\infty^2 s_0^2 - 2 r_\infty s_0 r_0 s_\infty + r_0^2 s_\infty^2) \\ &= (s_\infty s_0 + \varepsilon r_\infty r_0)^2 - \varepsilon(r_\infty s_0 - r_0 s_\infty)^2 \\ &= (s_\infty s_0 + \varepsilon r_\infty r_0)^2 - \varepsilon.\end{aligned}$$

Likewise, modulo squares the elements

$$-\frac{r_0 s_0}{r_\infty s_\infty} \qquad \text{and} \qquad -r_\infty r_0 s_\infty s_0$$

are the same. The corollary follows directly from Proposition 5.16. □

According to Corollary 5.17, these matrices do not exist very often, as the conditions are very restrictive. Heuristically, half of the elements are squares, and (i) covers the majority of the cases, indicating that we expect roughly a proportion of just over $1/2$ of representatives to yield elements in $\Gamma_{ns}^+(q)[1/p]$ transposing the two cusps attached to them (the representatives).

In order to avoid this conundrum, we adapt an algorithm presented in [GM15]. Furthermore, this approach will allow us to dispose of the requirement that g be an eigenform with eigenvalue 1 for the Atkin-Lerner involution at q. Let us introduce some notation.

Let F be a number field and S a finite set of places of F, containing the archimedian ones. Let $\mathcal{O}_S$ denote the S-integers of F (the elements of F with non-negative valuation for every place not in S).

For an ideal $\mathcal{N}$ in $\mathcal{O}_S$ let

$$\Gamma_1(\mathcal{N}) = \left\{ \gamma \in \mathbf{SL}_2(\mathcal{O}_S) \colon \gamma \equiv \begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \pmod{\mathcal{N}} \right\}.$$

Lemma 5.18. *Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(\mathcal{N})$. Suppose that $c = u + ta$ for some unit $u \in \mathcal{O}_S^\times$ and some $t \in \mathcal{O}_S$. There exists $x \in \mathcal{O}_S$ such that*

$$\gamma = \begin{pmatrix} 1 & 0 \\ c + t(1-a) & 1 \end{pmatrix} \begin{pmatrix} 1 & -u^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ u(1-a) & 1 \end{pmatrix} \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

Proof. See Lemma 2.1 in [GM15], or with slightly different language, Lemma 2.2(b) in [BMS67]. □

Notice that since $a - 1, c \in \mathcal{N}$, the product in Lemma 5.18 consists of matrices in $\Gamma_1(\mathcal{N})$ that stabilize either 0 or ∞ when γ acts on $\mathbb{P}^1(\mathbb{Q})$. These matrices are precisely the elementary matrices of determinant 1. The following theorem allows us to remove the restriction imposed over a and c.

Theorem 5.19. *Let $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(\mathcal{N})$. Assuming GRH, the following algorithm terminates and computes an expression of γ as a product of elementary matrices of determinant 1 in $\Gamma_1(\mathcal{N})$.*

1. *Iterate over the elements λ in the ring of integers of F to find λ such that $a' = a + \lambda c$ generates a prime ideal and the reduction map*

$$\mathcal{O}_S^\times \longrightarrow (\mathcal{O}_S / a'\mathcal{O}_S)^\times$$

   is surjective.

2. *Set $\gamma' = \left(\begin{smallmatrix} 1 & \lambda \\ 0 & 1 \end{smallmatrix}\right) \gamma$ and let $\gamma' = \left(\begin{smallmatrix} a' & b' \\ c' & d' \end{smallmatrix}\right)$.*

3. Iterate over the elements $u \in \mathcal{O}_S^\times$ until finding

$$c' \equiv u \pmod{a'}.$$

4. Use Lemma 5.18 to find an expression of γ' as a product of elementary matrices.

Proof. See Theorem 2.3 in [GM15]. □

Corollary 5.20. *Assume GRH. Every matrix in $\Gamma_1(q^2)[1/p]$ can be expressed as a product of at most five elementary matrices in $\mathbf{SL}_2(\mathbb{Z}[1/p])$.*

Proof. It follows directly from Theorem 5.19 applied to $F = \mathbb{Q}$, $S = \{p, \infty\}$ and $\mathcal{N} = (q^2)$. □

We are interested in adapting Corollary 5.20 to find factorizations of a similar nature to matrices in $\Gamma_{ns}^{\varepsilon}(q)[1/p]$. We need a couple of lemmas for that.

Lemma 5.21. *Let*

$$\Gamma^\dagger(q)[1/p] = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(q)[1/p] : a \equiv 1 \pmod{q^2} \right\}$$

and let $A_q = \begin{pmatrix} q & 0 \\ 0 & 1 \end{pmatrix}$. The inner automorphism of $\mathbf{GL}_2(\mathbb{Z}[1/p])$ induced by A_q induces an isomorphism between $\Gamma_1(q^2)[1/p]$ and $\Gamma^\dagger(q)[1/p]$.

Proof. If

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{then} \quad A_q A A_q^{-1} = \begin{pmatrix} a & bq \\ c/q & d \end{pmatrix}.$$

Hence

$$\begin{aligned} A \in \Gamma_1(q^2)[1/p] &\iff c,\, a-1,\, d-1 \in q^2\mathbb{Z}[1/p] \text{ and } b \in \mathbb{Z}[1/p] \\ &\iff bq, c/q \in q\mathbb{Z}[1/p] \text{ and } a-1,\, d-1 \in q^2\mathbb{Z}[1/p] \\ &\iff A_q A A_q^{-1} \in \Gamma^\dagger(q)[1/p], \end{aligned}$$

showing the desired isomorphism. □

For a matrix $A \in \Gamma_1(q^2)$, denote by $\tilde{A} = A_q A A_q^{-1}$ its counterpart in $\Gamma^\dagger(q)[1/p]$.

Proposition 5.22. *Let $\tilde{A} \in \Gamma^\dagger(q)[1/p]$. Then there exists a factorization*

$$\tilde{A} = \tilde{U}_1 \tilde{L}_1 \tilde{U}_2 \tilde{L}_2 \tilde{U}_3,$$

where

$$\tilde{U}_i = \begin{pmatrix} 1 & x_i \\ 0 & 1 \end{pmatrix} \qquad \textit{and} \qquad \tilde{L}_i = \begin{pmatrix} 1 & 0 \\ y_i & 1 \end{pmatrix},$$

with x_i, $y_i \in q\mathbb{Z}[1/p]$.

Proof. By Corollary 5.20, there exist elementary matrices of determinant 1 U_i and L_i in $\Gamma_1(q^2)[1/p]$, with U_i upper triangular and L_i lower triangular, such that

$$A = U_1 L_1 U_2 L_2 U_3.$$

By Lemma 5.21, after conjugating by A_q, we obtain

$$\tilde{A} = A_q U_1 L_1 U_2 L_2 U_3 A_q^{-1} = \tilde{U}_1 \tilde{L}_1 \tilde{U}_2 \tilde{L}_2 \tilde{U}_3.$$

Since U_i is upper triangular, so is $\tilde{U}_i$ and it lies in $\Gamma(q)[1/p]$. Since L_i is lower triangular and lies in $\Gamma_1(q^2)[1/p]$, $\tilde{L}_i$ is lower triangular and lies in $\Gamma(q)[1/p]$. □

By definition, $\Gamma(q^2)[1/p] \subseteq \Gamma^\dagger(q)[1/p] \subseteq \Gamma(q)[1/p]$. The reduction map

$$\mathbf{SL}_2(\mathbb{Z}[1/p]) \longrightarrow \mathbf{SL}_2(\mathbb{Z}/n\mathbb{Z})$$

(where $\gcd(n,p) = 1$) is surjective (as the map with source $\mathbf{SL}_2(\mathbb{Z})$ is) and has as kernel $\Gamma(n)[1/p]$, so the index

$$[\mathbf{SL}_2(\mathbb{Z}[1/p]) : \Gamma(n)[1/p]] = \#(\mathbf{SL}_2(\mathbb{Z}/n\mathbb{Z})) = n^3 \prod_{\ell|n} \left(1 - \frac{1}{\ell^2}\right), \tag{5.9}$$

where the product is taken over the prime divisors ℓ of n. Substituting $n = q$ and $n = q^2$ in Equation (5.9), we obtain

$$[\Gamma(q)[1/p] : \Gamma(q^2)[1/p]] = \frac{q^6(1 - 1/q^2)}{q^3(1 - 1/q^2)} = q^3.$$

$\Gamma^\dagger(q)[1/p]$, lying strictly between these two groups, must have index q or q^2.

Lemma 5.23. $\Gamma^\dagger(q)[1/p] \trianglelefteq \Gamma(q)[1/p]$. *Moreover,*

$$\Gamma(q)[1/p]/\Gamma^\dagger(q)[1/p] \cong \mathbb{Z}/q\mathbb{Z}.$$

Proof. Consider the map

$$\begin{aligned} \pi_a : \Gamma(q)[1/p] &\longrightarrow (\mathbb{Z}/q^2\mathbb{Z})^\times \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\longmapsto \quad a. \end{aligned}$$

From

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} aA + bC & aB + bD \\ cA + dC & cB + dD \end{pmatrix},$$

when $b \equiv C \equiv 0 \pmod{q}$, $aA + bC \equiv aA \pmod{q^2}$. It follows that if $\gamma_1, \gamma_2 \in \Gamma(q)[1/p]$, then

$$\pi_a(\gamma_1)\pi_a(\gamma_2) = \pi_a(\gamma_1\gamma_2),$$

showing π_a is a homormorphism. By definition, $\ker(\pi_a) = \Gamma^\dagger(q)[1/p]$, showing it is a normal subgroup. The image of π_a is the subset of residues modulo q^2 which are 1 modulo q, which has q elements. From the first isomorphism theorem and the fact that the only group of order q is cyclic, the isomorphism sought holds. □

Lemma 5.24. *For any matrix $\gamma \in \Gamma(q)[1/p]$, $\gamma^q \in \Gamma^\dagger(q)[1/p]$.*

Proof. This is a trivial consequence of Lemma 5.23. □

Lemma 5.25. *There exists $M \in \mathbf{SL}_2(\mathbb{Z})$ such that $M(0)$ and $M(\infty)$ are equivalent under the action of Γ.*

Proof. If ∞ and 0 are already equivalent cusps (i.e., if $-\varepsilon \in \langle p^2 \rangle$) we can just let M be the identity.

By Proposition 5.4, if

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}_2(\mathbb{Z}),$$

it suffices to show that

$$a^2 - c^2\varepsilon^{-1} \equiv b^2 - d^2\varepsilon^{-1} \pmod{q}$$

in order to show that $M(\infty) = a/c$ and $M(0) = b/d$ are equivalent under the action of Γ. Let us show how to construct such M according to the following two cases:

- *Case 1:* $q \equiv \pm 1 \pmod{8}$.

In this case, 2 is a quadratic residue, so let ρ be a residue modulo q such that $\rho^2 \equiv 2^{-1} \pmod{q}$. Let $M \in \mathbf{SL}_2(\mathbb{Z})$ be such that

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} \rho & -\rho \\ \rho & \rho \end{pmatrix} \pmod{q},$$

which exists because the determinant is $\rho^2 + \rho^2 \equiv 2\rho^2 \equiv 1 \pmod{q}$. In this case,

$$a^2 - c^2\varepsilon^{-1} \equiv \rho^2(1 - \varepsilon^{-1}) \equiv b^2 - d^2\varepsilon^{-1} \pmod{q},$$

showing that M has the desired property.

- *Case 2:* $q \equiv \pm 3 \pmod{8}$.

  Now, 2 is a quadratic nonresidue, so 2ε is a quadratic residue. Let ρ be a residue modulo q such that $\rho^2 \equiv (2\varepsilon)^{-1} \pmod{q}$. Let $M \in \mathbf{SL}_2(\mathbb{Z})$ be such that

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} \rho & \rho \\ -\rho\varepsilon & \rho\varepsilon \end{pmatrix},$$

  which exists because the determinant is $\rho^2\varepsilon + \rho^2\varepsilon \equiv 2\varepsilon\rho^2 \equiv 1 \pmod{q}$. In this case,

$$a^2 - c^2\varepsilon^{-1} \equiv \rho^2(1 - \varepsilon) \equiv b^2 - d^2\varepsilon^{-1} \pmod{q},$$

  showing again that M has the desired property. □

Now we have everything we need at hand. Lemma 5.25 provides us with a matrix M in $\mathbf{SL}_2(\mathbb{Z})$ such that $M(0)$ and $M(\infty)$ are equivalent under the action of $\Gamma_{ns}^{\varepsilon}(q)[1/p]$. Let $\upsilon = M(\infty)$ and let $P_{\tau,\upsilon}$ be the Stark-Heegner point associated to τ

at υ. In order to compute this point, we need to compute the semi-indefinite integral

$$J_\tau = \not\int^\tau \int_\upsilon^{\gamma_\tau \upsilon} e_\Gamma \omega_f.$$

Proposition 5.26. *There exists a divisor d of $q(q+1)$ such that J_τ^d can be expressed as a product of double-integrals.*

Proof. Note that for any positive integer k we have

$$J_\tau^k = \left(\not\int^\tau \int_\upsilon^{\gamma_\tau \upsilon} e_\Gamma \omega_f \right)^k = \prod_{j=0}^{k-1} \not\int^\tau \int_{\gamma_\tau^j \upsilon}^{\gamma_\tau^{j+1} \upsilon} e_\Gamma \omega_f = \not\int^\tau \int_\upsilon^{\gamma_\tau^k \upsilon} e_\Gamma \omega_f,$$

where the second equality comes from the independence of base-point in the orbit of υ and the third equality from the properties of the semi-indefinite integral. Since $-I \in \Gamma_{ns}^\varepsilon(q)[1/p]$, we also have

$$\not\int^\tau \int_\upsilon^{-\gamma_\tau^k \upsilon} e_\Gamma \omega_f = \not\int^{-I\tau} \int_{-I\upsilon}^{-I\gamma_\tau^k \upsilon} e_\Gamma \omega_f = \not\int^\tau \int_\upsilon^{\gamma_\tau^k \upsilon} e_\Gamma \omega_f,$$

as $-I$ acts trivially on $\mathcal{H}_p$ and on $\mathbb{P}^1(\mathbb{Q})$.

By Lemma 5.2, since $\gamma_\tau \in \Gamma_{ns}^\varepsilon(q)[1/p]$, making $k = q+1$ above, makes γ_τ^k be in $\Gamma(q)[1/p]$, which is normal in $\mathbf{SL}_2(\mathbb{Z}[1/p])$ meaning that $M^{-1}\gamma_\tau^{q+1}M$ is also in $\Gamma(q)[1/p]$. By Lemma 5.24, if we further raise $M^{-1}\gamma_\tau^{q+1}M$ to the q-th power, we find that $M^{-1}\gamma_\tau^{q(q+1)}M \in \Gamma^\dagger(q)[1/p]$. Let d be the smallest positive integer such that either $M^{-1}\gamma_\tau^d M$ or $-M^{-1}\gamma_\tau^d M$ is in $\Gamma^\dagger(q)[1/p]$. It is clear that $d \mid q(q+1)$. Let $\gamma = \pm M^{-1}\gamma_\tau^d M \in \Gamma^\dagger(q)[1/p]$.

Proposition 5.22 yields matrices U_i and L_i in $\Gamma(q)[1/p]$ such that

$$\gamma = \tilde{U}_1 \tilde{L}_1 \tilde{U}_2 \tilde{L}_2 \tilde{U}_3,$$

where

$$U_i = \begin{pmatrix} 1 & x_i \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad L_i = \begin{pmatrix} 1 & 0 \\ y_i & 1 \end{pmatrix}.$$

For a matrix A, denote by $\bar{A} = MAM^{-1}$. Notice that if $A \in \Gamma(q)[1/p]$, then $\bar{A} \in \Gamma(q)[1/p]$, as it is normal. Thus, we obtain

$$\pm\gamma_\tau^d = \bar{\gamma} = \bar{U}_1\bar{L}_1\bar{U}_2\bar{L}_2\bar{U}_3,$$

where $\bar{U}_i,\ \bar{L}_i \in \Gamma(q)[1/p]$, $\bar{U}_iM(\infty) = M\tilde{U}_iM^{-1}M(\infty) = M\tilde{U}_i(\infty) = M(\infty)$ and $\bar{L}_iM(0) = M\tilde{L}_iM^{-1}M(0) = M\tilde{L}_i(0) = M(0)$. This is, $\bar{U}_i$ and $\bar{L}_i$ are matrices in $\Gamma^{\varepsilon}_{ns}(q)[1/p]$ which fix $\upsilon = M(\infty)$ and $\upsilon' = M(0)$, respectively.

To alleviate reading, denote by

$$\gamma_2 = \tilde{L}_1\tilde{U}_2\tilde{L}_2\tilde{U}_3, \quad \gamma_3 = \tilde{U}_2\tilde{L}_2\tilde{U}_3, \quad \gamma_4 = \tilde{L}_2\tilde{U}_3,$$

$$\tau_1 = \bar{U}_1^{-1}\tau, \quad \tau_2 = \bar{L}_1^{-1}\tau_1, \quad \tau_3 = \bar{U}_2^{-1}\tau_2, \quad \tau_4 = \bar{L}_2^{-1}\tau_3.$$

Using the properties of the semi-indefinite integral, following Equations (3.4) and (3.5) in [GM15], we have

$$\begin{aligned}
\oint^{\tau}\int_{\upsilon}^{\bar{\gamma}\upsilon} e_\Gamma\omega_f &= \oint^{\bar{U}_1\tau_1}\int_{\bar{U}_1\upsilon}^{\bar{U}_1\gamma_2\upsilon} e_\Gamma\omega_f = \oint^{\tau_1}\int_{\upsilon}^{\gamma_2\upsilon} e_\Gamma\omega_f \\
&= \oint^{\tau_1}\int_{\upsilon}^{\upsilon'} e_\Gamma\omega_f \times \oint^{\tau_1}\int_{\upsilon'}^{\gamma_2\upsilon} e_\Gamma\omega_f \\
&= \oint^{\bar{L}_1\tau_2}\int_{\bar{L}_1\upsilon'}^{\bar{L}_1\gamma_3\upsilon} e_\Gamma\omega_f \div \oint^{\tau_1}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \\
&= \oint^{\tau_2}\int_{\upsilon'}^{\gamma_3\upsilon} e_\Gamma\omega_f \div \oint^{\tau_1}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \\
&= \oint^{\tau_2}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \times \oint^{\tau_2}\int_{\upsilon}^{\gamma_3\upsilon} e_\Gamma\omega_f \div \oint^{\tau_1}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \\
&= \left(\oint_{\tau_1}^{\tau_2}\int_{\upsilon'}^{\upsilon} \omega_f\right)^{e_\Gamma} \times \oint^{\tau_2}\int_{\upsilon}^{\gamma_3\upsilon} e_\Gamma\omega_f
\end{aligned}$$

and exacly the same computation yields

$$\begin{aligned}
\oint^{\tau_2}\int_{\upsilon}^{\gamma_3\upsilon} e_\Gamma\omega_f &= \oint^{\bar{U}_2\tau_3}\int_{\bar{U}_3\upsilon}^{\bar{U}_3\gamma_4\upsilon} e_\Gamma\omega_f = \oint^{\tau_3}\int_{\upsilon}^{\gamma_4\upsilon} e_\Gamma\omega_f \\
&= \oint^{\tau_3}\int_{\upsilon}^{\upsilon'} e_\Gamma\omega_f \times \oint^{\tau_3}\int_{\upsilon'}^{\gamma_4\upsilon} e_\Gamma\omega_f \\
&= \oint^{\bar{L}_2\tau_4}\int_{\bar{L}_2\upsilon'}^{\bar{L}_2\bar{U}_3\upsilon} e_\Gamma\omega_f \div \oint^{\tau_3}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \\
&= \oint^{\tau_4}\int_{\upsilon'}^{\bar{U}_3\upsilon} e_\Gamma\omega_f \div \oint^{\tau_3}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \\
&= \oint^{\tau_4}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \times \oint^{\tau_4}\int_{\upsilon}^{\bar{U}_3\upsilon} e_\Gamma\omega_f \div \oint^{\tau_3}\int_{\upsilon'}^{\upsilon} e_\Gamma\omega_f \\
&= \left(\oint_{\tau_3}^{\tau_4}\int_{\upsilon'}^{\upsilon} \omega_f\right)^{e_\Gamma} \times \oint^{\tau_4}\int_{\bar{U}_3\upsilon}^{\bar{U}_3\upsilon} e_\Gamma\omega_f = \left(\oint_{\tau_3}^{\tau_4}\int_{\upsilon'}^{\upsilon} \omega_f\right)^{e_\Gamma}.
\end{aligned}$$

Putting both equations together, we obtain

$$J_\tau^d = \not\int_\upsilon^\tau \int_\upsilon^{\bar\gamma\upsilon} e_\Gamma \omega_f = \left(\not\int_{\tau_1}^{\tau_2} \int_{\upsilon'}^{\upsilon} \omega_f \times \not\int_{\tau_3}^{\tau_4} \int_{\upsilon'}^{\upsilon} \omega_f \right)^{e_\Gamma}. \tag{5.10}$$

□

Note the importance of choosing υ the way we chose it in Proposition 5.26, as the semi-indefinite integrals were split in such a way that we needed to evaluate them at the pair of cusps υ and υ', which are equivalent under the action of $\Gamma_{ns}^{\varepsilon}(q)[1/p]$ precisely by our choice of υ. Also, the factorization is heavily exploited, as the elements lie in the Cartan Non-split and they fix the cusps υ and υ'.

If there is a *nice* relationship between p and q, we can reduce even further the exponent d in Proposition 5.26 as follows.

Corollary 5.27. *Assume that $p^{q-1} \not\equiv 1 \pmod{q^2}$ and that $-\varepsilon \in \langle p^2 \rangle$. There exists a divisor d of $q+1$ such that J_τ^d can be expressed as a product of double-integrals.*

Proof. Let Let h' be the order of p modulo q^2. By Euler's Theorem, h' divides $q(q-1)$. By our assumption on p, h' does not divide $q-1$, implying $h' = qh$ for some h.

Note that $p^h \equiv 1 \pmod{q}$, as $p^h \equiv (p^h)^q \equiv p^{h'} \equiv 1 \pmod{q}$. This implies that for $k = 0, 1, \ldots, q-1$, the elements p^{hk} are all distinct modulo q^2 (the order is hq) and they are all congruent to 1 modulo q.

Let d be the smallest positive integer such that either γ_τ^d or $-\gamma_\tau^d$ lie in $\Gamma(q)[1/p]$, and denote this matrix by $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. By Lemma 5.2, $\gamma_\tau^{q+1} \in \Gamma(q)[1/p]$, so d is a

divisor of $q+1$. Let α be an integer such that $p^{-h\alpha} \equiv a \pmod{q^2}$. Then

$$\begin{pmatrix} p^{h\alpha} & 0 \\ 0 & p^{-h\alpha} \end{pmatrix} \gamma = \begin{pmatrix} ap^{h\alpha} & bp^{h\alpha} \\ cp^{-h\alpha} & dp^{-h\alpha} \end{pmatrix}$$

lies in $\Gamma^\dagger(q)[1/p]$. From here, we can continue as in Proposition 5.22, since the matrix we used to tweak γ fixes the equivalent cusps 0 and ∞. $\square$

Chapter 6
Further directions

The nature of this work is conjectural so a natural step from here is to perform computations corroborating the veracity of Conjecture 5.15, which would provide a supply of algebraic points on elliptic curves with additive reduction (of conductor pq^2) over ray class fields attached to real quadratic fields. We append a few partially computed examples in such a way that the *only* remaining task would be to compute the multiplicative double integrals. The code used to find the factorizations in Theorem 5.19 was written by Marc Masdeu and Xevi Guitart.

1. Let $D = 17$ and $E = 99a1$, with conductor $99 = 11 \cdot 3^2$, so $p = 11$ and $q = 3$, which are inert in $K = \mathbb{Q}(\sqrt{17})$. The eigenvalue of the newform corresponding to E at 3 is 1, and the only possible choice for ε is -1. The generator for $\mathcal{O}_{17}$ is $\omega_{17} = \frac{17+\sqrt{17}}{2}$, so in order to find an embedding we need a matrix with determinant 68 and trace 17, which lies in the Cartan. We can take the embedding given by

$$\omega_{17} \longmapsto \begin{pmatrix} 10 & -1 \\ -2 & 7 \end{pmatrix}.$$

The fundamental unit is given by $4+\sqrt{17}$, which has norm -1. Its square, $33+8\sqrt{17}$, generates $(\mathcal{O}_{17}^{\times})_1$.

$$33+8\sqrt{17} = -103+16\omega_{17} \longrightarrow \gamma_\tau = \begin{pmatrix} 57 & -16 \\ -32 & 9 \end{pmatrix},$$

whose fixed point is a root of the polynomial $-32\tau^2 - 48\tau + 16 = 0$, or, $2\tau^2 + 3\tau - 1 = 0$. Let $\tau = \dfrac{-3+\sqrt{17}}{4}$ be one such root.

We need to compute the integral

$$\oint^{\tau} \int_{\infty}^{\gamma_\tau \infty} e_\Gamma \omega_f = \oint^{\tau} \int_{\infty}^{-57/32} e_\Gamma \omega_f.$$

The continued fraction expansion yields convergents $-2/1$, $-7/4$, $-9/5$, $-16/9$ and $-57/32$, so we need to compute the product of the integrals

$$\oint^{\tau} \int_{\infty}^{-2} e_\Gamma \omega_f, \oint^{\tau} \int_{-2}^{-7/4} e_\Gamma \omega_f, \oint^{\tau} \int_{-7/4}^{-9/5} e_\Gamma \omega_f,$$

$$\oint^{\tau} \int_{-9/5}^{-16/9} e_\Gamma \omega_f, \oint^{\tau} \int_{-16/9}^{-57/32} e_\Gamma \omega_f.$$

These convergents give us the matrices

$$M_0 = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}, M_1 = \begin{pmatrix} -7 & -2 \\ 4 & 1 \end{pmatrix}, M_2 = \begin{pmatrix} 9 & -7 \\ -5 & 4 \end{pmatrix},$$

$$M_3 = \begin{pmatrix} -16 & -9 \\ 9 & 5 \end{pmatrix}, M_4 = \begin{pmatrix} 57 & -16 \\ -32 & 9 \end{pmatrix},$$

which yield representatives

$$r_0 = r_2 = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \qquad \text{and} \qquad r_1 = r_3 = r_4 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and

$$\gamma_0 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ \gamma_1 = \begin{pmatrix} -7 & -2 \\ 4 & 1 \end{pmatrix},\ \gamma_2 = \begin{pmatrix} 23 & -7 \\ -13 & 4 \end{pmatrix},$$

$$\gamma_3 = \begin{pmatrix} -16 & -9 \\ 9 & 5 \end{pmatrix},\ \gamma_4 = \begin{pmatrix} 57 & -16 \\ -32 & 9 \end{pmatrix}.$$

We let

$$\tau_0 = \gamma_0^{-1}\tau = \frac{-3-\sqrt{17}}{2},\ \tau_1 = \gamma_1^{-1}\tau = \frac{3-\sqrt{17}}{4},\ \tau_2 = \gamma_2^{-1}\tau = \frac{9-\sqrt{17}}{16},$$

$$\tau_3 = \gamma_3^{-1}\tau = \frac{3-\sqrt{17}}{2}, \tau_4 = \gamma_4^{-1}\tau = \frac{-3+\sqrt{17}}{4},$$

and

$$\tau_0' = W_{0,1/2}\tau_0 = \frac{9+\sqrt{17}}{32},\ \tau_1' = W_{0,\infty}\tau_1 = \frac{3+\sqrt{17}}{2},\ \tau_2' = W_{0,1/2}\tau_2 = \frac{9+\sqrt{17}}{16},$$

$$\tau_3' = W_{0,\infty}\tau_3 = \frac{3+\sqrt{17}}{4}, \tau_4' = W_{0,\infty}\tau_4 = \frac{-3-\sqrt{17}}{2}.$$

According to (5.4), we now need to compute the product of

$$\oint^{\tau_0} \int_0^{1/2} e_\Gamma \omega_f,\ \oint^{\tau_1} \int_0^{\infty} e_\Gamma \omega_f,\ \oint^{\tau_2} \int_0^{1/2} e_\Gamma \omega_f,$$

$$\oint^{\tau_3} \int_0^{\infty} e_\Gamma \omega_f,\ \oint^{\tau_4} \int_0^{\infty} e_\Gamma \omega_f.$$

Putting this together with (5.6), we obtain

$$\left(\not\int^{\tau}\int_{\infty}^{-57/32} e_\Gamma\omega_f\right)^2 = \left(\not\int_{\tau_0'}^{\tau_0}\int_0^{1/2}\omega_f \times \not\int_{\tau_1'}^{\tau_1}\int_0^{\infty}\omega_f \times \not\int_{\tau_2'}^{\tau_2}\int_0^{1/2}\omega_f \not\int_{\tau_3'}^{\tau_3}\int_0^{\infty}\omega_f \times \not\int_{\tau_4'}^{\tau_4}\int_0^{\infty}\omega_f\right)^{e_\Gamma}.$$

2. Let us use exactly the same setting as before, but let us apply the algorithm suggested for general values of q. The cusps 0 and ∞ are equivalent, so we take M to be the identity matrix. From the computations above, we have

$$\gamma_\tau = \begin{pmatrix} 57 & -16 \\ -32 & 9 \end{pmatrix},$$

and we can see that

$$\gamma = -\gamma_\tau^2 = \begin{pmatrix} -3761 & 1056 \\ 2112 & -593 \end{pmatrix} \in \Gamma(3)[1/11] \text{ and } \Gamma^\dagger(3)[1/11].$$

Proposition 5.22 yields the factorization

$$\gamma = \bar{U}_1\bar{L}_1\bar{U}_2\bar{L}_2\bar{U}_3,$$

where

$$\bar{U}_1 = \begin{pmatrix} 1 & -24/11 \\ 0 & 1 \end{pmatrix}, \quad \bar{L}_1 = \begin{pmatrix} 1 & 0 \\ 4902/11 & 1 \end{pmatrix}, \quad \bar{U}_2 = \begin{pmatrix} 1 & -3/11^3 \\ 0 & 1 \end{pmatrix},$$

$$\bar{L}_2 = \begin{pmatrix} 1 & 0 \\ -282\cdot 11^3 & 1 \end{pmatrix}, \quad \bar{U}_3 \begin{pmatrix} 1 & -45219/11^5 \\ 0 & 1 \end{pmatrix}.$$

Let

$$\tau_1 = \bar{U}_1^{-1}\tau = \frac{63}{44} + \frac{\sqrt{17}}{4}, \quad \tau_2 = \bar{L}_1^{-1}\tau_1 = \frac{-25690863 + 14641\sqrt{17}}{11411473384},$$

$$\tau_3 = \bar{U}_2^{-1}\tau_2 = \frac{39881499 + 14641\sqrt{17}}{11411473384}, \quad \tau_4 = \bar{L}_2^{-1}\tau_3 = -\frac{664029}{664204} + \frac{\sqrt{17}}{4}.$$

Equation (5.10) yields

$$\left(\fint^{\tau} \int_{\infty}^{-57/32} e_\Gamma \omega_f \right)^2 = \left(\fint_{\tau_1}^{\tau_2} \int_0^{\infty} \omega_f \times \fint_{\tau_3}^{\tau_4} \int_0^{\infty} \omega_f \right)^{e_\Gamma}.$$

3. Let $D = 5$ and $E = 147c1$, with conductor $147 = 3 \cdot 7^2$, so $p = 3$ and $q = 7$, which are inert in $K = \mathbb{Q}(\sqrt{5})$. Let $\varepsilon = -1$. The generator for $\mathcal{O}_5$ is $\omega_5 = \frac{5+\sqrt{5}}{2}$, so in order to find an embedding we need a matrix with determinant 5 and trace 5, which lies in the Cartan. We can take the embedding given by

$$\omega_5 \longmapsto \begin{pmatrix} -15 & -61 \\ 5 & 20 \end{pmatrix}.$$

The fundamental unit is given by $\frac{1+\sqrt{5}}{2}$, which has norm -1. Its square, $\frac{3+\sqrt{5}}{2}$, generates $(\mathcal{O}_5^\times)_1$.

$$\frac{1+\sqrt{5}}{2} = -1 + \omega_5 \longrightarrow \gamma_\tau = \begin{pmatrix} -16 & -61 \\ 5 & 19 \end{pmatrix},$$

whose fixed point is a root of the polynomial $5\tau^2 + 35\tau + 61 = 0$. Let $\tau = \frac{-35+\sqrt{17}}{10}$ be one such root. The cusps 0 and ∞ are equivalent, so we take M to be the

identity matrix. Let

$$\gamma = -\gamma_\tau^4 = \begin{pmatrix} 344 & 1281 \\ -105 & -391 \end{pmatrix} \in \Gamma(7)[1/3] \text{ and } \Gamma^\dagger(7)[1/3].$$

Proposition 5.22 yields the factorization

$$\gamma = \bar{U}_1 \bar{L}_1 \bar{U}_2 \bar{L}_2 \bar{U}_3,$$

where

$$\bar{U}_1 = \begin{pmatrix} 1 & -7/3 \\ 0 & 1 \end{pmatrix}, \quad \bar{L}_1 = \begin{pmatrix} 1 & 0 \\ -7/3 & 1 \end{pmatrix}, \quad \bar{U}_2 = \begin{pmatrix} 1 & 7/3^2 \\ 0 & 1 \end{pmatrix},$$

$$\bar{L}_2 = \begin{pmatrix} 1 & 0 \\ -12 \cdot 3^2 & 1 \end{pmatrix}, \quad \bar{U}_3 \begin{pmatrix} 1 & -301/3^4 \\ 0 & 1 \end{pmatrix}.$$

Let

$$\tau_1 = \bar{U}_1^{-1}\tau = -\frac{7}{6} + \frac{\sqrt{5}}{10}, \quad \tau_2 = \bar{L}_1^{-1}\tau_1 = \frac{1533 + 81\sqrt{5}}{2182},$$
$$\tau_3 = \bar{U}_2^{-1}\tau_2 = -\frac{1477}{19638} + \frac{81\sqrt{5}}{2182}, \quad \tau_4 = \bar{L}_2^{-1}\tau_3 = \frac{35}{162} + \frac{\sqrt{5}}{10}.$$

Equation (5.10) yields

$$\left(\fint^{\tau} \int_{\infty}^{-16/5} e_\Gamma \omega_f \right)^4 = \left(\fint_{\tau_1}^{\tau_2} \int_0^{\infty} \omega_f \times \fint_{\tau_3}^{\tau_4} \int_0^{\infty} \omega_f \right)^{e_\Gamma}.$$

References

[AL70] A. O. L. Atkin and J. Lehner. Hecke operators on $\Gamma_0(m)$. *Math. Ann.*, 185:134–160, 1970.

[BB12] Aurèlien Bajolet and Yuri Bilu. Computing integral points on $x_{ns}^+(p)$. *arXiv:1212.0665v1*, 2012.

[BCDT01] Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor. On the modularity of elliptic curves over $\mathbb{Q}$: wild 3-adic exercies. *J. Amer. Math. Soc.*, 14:843–939, 2001.

[BMS67] H. Bass, J. Milnor, and J.-P. Serre. Solution of the congruence subgroups problem for $\mathbf{SL_n}\,(\mathbf{n} \geq \mathbf{3})$ and $\mathbf{Sp_{2n}}\,(\mathbf{n} \geq \mathbf{2})$. *Inst. Hautes Études Sci. Publ. Math.*, 33:59–137, 1967.

[Che98] Imin Chen. The Jacobians of non-split Cartan modular curves. *Proc. London Math. Soc. (3)*, 77(1):1–38, 1998.

[CL05] Denis Charles and Kristin Lauter. Computing modular polynomials. *LMS J. Comput. Math.*, 8:195–204, 2005.

[Cox13] David A. Cox. *Primes of the Form x^2+ny^2: Fermat, Class Field Theory, and Complex Multiplication.* Pure and Applied Mathematics: A Wiley Series of Texts, Monographs, and Tracts. Wiley, 2 edition, 2013.

[Cre92] John E. Cremona. *Algorithms for Modular Elliptic Cruves.* Cambridge University Press, 2 edition, 1997, 1992.

[Dar01] Henri Darmon. Integration on $\mathcal{H}_p \times \mathcal{H}$ and arithmetic applications. *Annals of Mathematics*, pages 589–639, 2001.

[Dar04] Henri Darmon. *Rational points on modular elliptic curves*, volume 101 of *CBMS Regional Conference Series in Mathematics.* Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2004.

[DP06] Henri Darmon and Robert Pollack. Efficient calculation of Stark-Heegner points via overconvergent modular symbols. *Israel Journal of Mathematics*, 153(1):319–354, 2006.

[DS05] Fred Diamond and Jerry Shurman. *A First Course in Modular Forms*, volume 228 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2005.

[dSE00] Bart de Smit and Bas Edixhoven. Sur un résultat d'Imin Chen. *Math. Res. Lett.*, 7(2-3):147–153, 2000.

[Edi96] Bas Edixhoven. On a result of Imin Chen. *arXiv:alg-geom/9604008*, 1996.

[Elk98] Noam Elkies. Elliptic and modular curves over finite fields and related computational issues. *AMS/IP Stud. Adv. Math.*, 7:195–204, 1998.

[Fla89] Daniel E. Flath. *Introduction to Number Theory*. Wiley Classics Library. Wiley-Interscience, 1989.

[GH94] Phillip Griffiths and Joseph Harris. *Principles of Algebraic Geometry*. Wiley-Interscience, 1994.

[GM15] Xavier Guitart and Marc Masdeu. Elementary matrix decomposition and the computation of Darmon points with higher conductor. *Mathematics of Computation*, 84(292):875–893, 2015.

[Gro84] Benedict H. Gross. Heegner points on $X_0(N)$. In *Modular forms (Durham, 1983)*, Ellis Horwood Ser. Math. Appl.: Statist. Oper. Res., pages 87–105. Horwood, Chichester, 1984.

[GS93] Ralph Greenberg and Glenn Stevens. p-adic l-functions and p-adic periods of modular forms. *Invent. Math.*, 111(2):407–447, 1993.

[GZ86] Benedict H. Gross and Don B. Zagier. Heegner points and derivatives of L-series. *Invent. Math.*, 84(2):225–320, 1986.

[Har77] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1977.

[Kna92] Anthony W. Knapp. *Elliptic Curves*. Princeton University Press, 1992.

[KP14] Daniel Kohen and Ariel Pacetti. Heegner points on Cartan non-split curves. *arXiv:1403.7801v2*, 2014.

[Men67] J. Mennicke. On Ihara's modular group. *Invent. Math.*, 4:202–228, 1967.

[MTT86] B. Mazur, J. Tate, and J. Teitelbaum. On p-adic analogues of the conjectures of Birch and Swinnerton-Dyer. *Invent. Math.*, 84(1):1–48, 1986.

[RW14] Marusia Rebolledo and Christian Wuthrich. A moduli interpretation for the non-split cartan modular curve. *arXiv:1402.3498*, 2014.

[Ser70] Jean-Pierre Serre. Le problème des groupes de congruence pour $\mathbf{SL_2}$. *Ann. of Math.*, 92(2):489–527, 1970.

[Ser73] Jean-Pierre Serre. *A Course in Arithmetic*, volume 7 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1973.

[Ser97] Jean-Pierre Serre. *Lectures on the Mordell-Weil theorem*. Aspects of Mathematics. Friedr. Vieweg & Sohn, Braunschweig, third edition, 1997. Translated from the French and edited by Martin Brown from notes by Michel Waldschmidt, With a foreword by Brown and Serre.

[Sil86] Joseph H. Silverman. *The Arithmetic of Elliptic Curves*, volume 106 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2 edition, 2009,1986.

[Sil94] Joseph H. Silverman. *Advanced Topics in the Arithmetic of Elliptic Curves*, volume 151 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1994.

[ST92] Joseph H. Silverman and John Tate. *Rational points on elliptic curves*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1992.

[Wil95] Andrew Wiles. Modular Elliptic Curves and Fermat's Last Theorem. *Ann. of Math.*, 141:443–551, 1995.