Università di Pisa

Corso di Laurea Magistrale in Matematica

The Congruent Number Problem and ranks in quadratic twist families of elliptic curves

Tesi di Laurea Magistrale

Relatori:

Prof. Emmanuel Kowalski

Prof. Umberto Zannier

Candidato:

Francesco Naccarato

Anno Accademico 2021/2022

Table of Contents

1 Introduction

In the 3rd century BC, Diophantus observed that given a right triangle with integer side lengths $X < Y < Z$ (a so-called *Pythagorean triangle*), the numbers $Z^2 \pm 2XY$ are integer squares. Namely, they equal $(X \pm Y)^2$, so the numbers $(Y-X)^2,\ Z^2,\ (Y+X)^2$ are three integer squares in arithmetic progression, with common difference four times the area of the triangle.

We want to investigate these 3-term arithmetic progressions, relaxing the condition on the squares to being rational, while still keeping the difference integral. We have a bijection:

$$\left\{\begin{array}{l}\text{right triangles } T \text{ with rational}\\ \text{sides } X < Y < Z \text{ and integral area}\end{array}\right\} \overset{1:1}{\longleftrightarrow} \left\{\begin{array}{l}\text{triples } (r,s,t) \text{ of rational numbers}\\ \text{whose squares are in arithmetic}\\ \text{progression of integral difference}\end{array}\right\},$$

via the (clearly invertible) application $X = t - r,\ Y = t + r,\ Z = 2s$, which is well-defined since in this case the area of T is exactly the common difference of the progression, thanks to the normalization we have chosen for r, s and t:

$$\mathrm{Area}(T) = \frac{XY}{2} = \frac{(t+r)(t-r)}{2} = t^2 - s^2.$$

This phenomenon was observed and studied as early as the 10th century by Arab scholars, calling for the problem of determining which integers are the area of such a triangle (which we will call a *congruent triangle*), or, equivalently, the common difference of such an arithmetic progression. We refer to these integers as **congruent numbers** and to this question, which will be the main focus of this thesis, as the *Congruent Number Problem (CNP).*

Since the set of *rational* numbers that are the area of a right triangle with rational sides is the union of equivalence classes of $\mathbb{Q}^*/\mathbb{Q}^{*2}$ (we can rescale the sides by a rational factor), we can focus on finding squarefree congruent numbers (from now on the expression “congruent number” will possibly also mean “squarefree” depending on the context); hence, we already have two formulations for the CNP:

1. Determine all squarefree positive integers n such that there exists a right triangle of area n with rational sides;

2. Determine all squarefree positive integers n such that there exist three rational squares in arithmetic progression of common difference n.

We will find multiple other equivalent formulations, linking this apparently harmless problem to deep objects of modern Number Theory, such as elliptic curves and modular forms.

Before outlining the structure of this thesis, let us observe that, given rational numbers x, y, z with $x^2+y^2=z^2$, there exists (a unique) $l \in \mathbb{Q}^*$ such that $(X, Y, Z) = (lx, ly, lz)$ is a Pythagorean triple with $\gcd(X, Y, Z) = 1$ (such a Pythagorean triple is called *primitive*). It is an elementary and classical result, allegedly due to Euclid, that primitive Pythagorean triples are (up to switching X and Y if X is even) all of the form $X = u^2 - v^2,\ Y = 2uv,\ Z = u^2 + v^2$ as $u > v$ vary over pairs of coprime positive integers, one of which is even. So, by our previous reasoning, we obtain another characterization of congruent numbers: the (squarefree, positive integer) representatives of the $\mathbb{Q}^*/\mathbb{Q}^{*2}$-classes of numbers in the form $uv(u+v)(u-v)$ with $u, v \in \mathbb{Z}^+$. In short, we have a congruent number generating formula: let $u > v$ vary in the positive integers; then the congruent numbers are exactly those generated by factoring out the squares from the above expression:

$$n = uv(u+v)(u-v)/\square. \tag{1}$$

In Section 2 we will introduce elliptic curves, briefly discussing their group structure over number fields and finite fields, and then proceed to construct a quadratic twist family $\{E_n\}_{n\in(\mathbb{Z}\setminus\{0\})/\square}$, the so-called *congruent number family*, such that n is congruent if and only if E_n has a point of infinite order; more precisely, while stating our results we shall use the expression *congruent number family* to refer just to the subfamily for which n is positive, since a negative number cannot be congruent. Nevertheless, we defined the family with n as above since that is the expression for the index in generic quadratic twist families.

The problem of studying and finding nontorsion points on elliptic curves over $\mathbb{Q}$ (and, more generally, over number fields) is one of the richest and most active in modern Number Theory, so we will be able to attack the CNP with a wide variety of tools. These links between the CNP and very deep conjectures, along with evidence about the enormous growth, in terms of n, of the denominators of the sidelengths in congruent triangles of area n, hint at the arithmetic complexity hidden behind this problem.

We will utilize the L-function of an elliptic curve, which we will introduce in Section 3, to study its arithmetic. We will define the *analytic rank* of an elliptic curve and state the famous *Birch and Swinnerton-Dyer (BSD) Conjecture*, which will be crucial to understand a partial result towards the resolution of the CNP that we will prove in Sections 5 and 6. The behavior of the analytic rank in quadratic twist families, and especially in the congruent number family, will be studied under two aspects:

- Averages: we will obtain nontrivial bounds for the average analytic rank in quadratic twist families; ordering the curves by their height, we will present a novel and fairly simple proof that 3.25 is an upper bound, though weaker than the 1.5 obtained by Heath-Brown; ordering them by the height of their smallest rational point, we will follow an argument of Le Boudec that gives an unexpected lower bound of $1 + \epsilon$, reinforcing our belief that proving the *Minimalist Conjecture* (see Section 3) is very hard.

- Exceptional curves: we will prove a partial result towards the existence of infinitely many elliptic curves of fixed high analytic rank, conditional on the Generalized Riemann Hypothesis. We will then examine some heuristics and conjectures about the largest ranks appearing in these families.

In Section 4 the focus will be on modularity results connected to elliptic curves. After introducing the theory of modular forms and modular curves and developing the necessary tools of Hecke Theory, we will construct the L-function of a modular form and study some of its properties. We will show how the *Modularity Theorem*, which is, in its most general form, a groundbreaking result of Wiles, Breuil, Conrad, Diamond and Taylor, follows for the congruent number family from a classical theorem of Weil. We will end by defining modular forms of half-integer weight and outlining a deep correspondence, due to Shimura, that allows us to map such forms to forms with ordinary integer weight.

Section 5 will be devoted to the study of so-called *Heegner points* on modular and elliptic curves: we will describe this method of constructing rational points which appears to be unique in the modern toolset for explicitly solving Diophantine equations, and look at some of its deep consequences, including the construction of nontorsion points on some curves of the congruent number family.

In Section 6 we will turn our attention to the proofs of the two main results about congruent numbers presented in this thesis, due to Tunnell and Monsky respectively. For Tunnell's Theorem we will need a deep and technical result, due Waldspurger, relating the central values of twists of the L-function attached to a modular form to the coefficients of the L-function attached to its Shimura lifts. Monsky's Theorem will follow from a careful analysis, through the lens of Galois theory and elementary class field theory, of the points constructed in Section 5. We will conclude with a few remarks and ideas about what is left to prove in order to attain a complete understanding of this problem, including a generalization, due to Tian, Yuan and Zhang, of Monsky's result.

We end this introduction by stating the two main results, along with a simple corollary of the first, showing how the second one is a partial step towards a complete and unconditional characterization of congruent numbers.

Theorem 1.1 (Tunnell). *Let n be a squarefree positive integer and let (α, m) be $(2, n)$ if n is odd and $(4, \frac{n}{2})$ if n is even. If n is congruent, the number of integer solutions to $\alpha x^2 + y^2 + 8z^2 = m$ is twice that of integer solutions to $\alpha x^2 + y^2 + 32z^2 = m$. Moreover, if the BSD Conjecture holds, then the converse holds.*

Corollary 1.1.1. *Assuming the BSD Conjecture, positive integers congruent to* 5*,* 6 *or* 7 *modulo* 8 *are congruent.*

Proof. The two sets of solutions mentioned in the theorem are both empty in these cases, since the quadratic residues modulo 8 are $0, 1, 4$. An integer not in the form $4k$ and its squarefree part are in the same residue class modulo 8, so we are done by Theorem 1.1. □

Theorem 1.2 (Heegner, Monsky)**.** *In each residue class* 5*,* 6 *and* 7 *modulo* 8 *there are infinitely many congruent numbers with one odd prime factor and infinitely many with two odd prime factors.*

Notation

$\mu(n)$ will denote the Möbius function, defined as $\mu(n) = \begin{cases} 0 \text{ if } n \text{ is not squarefree,} \\ (-1)^{\omega(n)} \text{ else,} \end{cases}$ where $\omega(n)$ counts the prime factors of n;

$\sum_n{}' f(n) = \sum_n f(n)\mu^2(n)$ will denote a sum over squarefree positive integers;

$\mathrm{sqf}(n)$ will denote the *squarefree part* of n, the only squarefree integer d such that $\frac{n}{d}$ is an integer square;

The symbol (a, b) will denote the greatest common divisor of integers a and b, and $\phi(n)$ will denote the Euler totient function, counting the number of positive integers $d \leq n : (d, n) = 1$;

$\left(\dfrac{a}{m}\right)$ is the Jacobi symbol modulo $m = \prod_i p_i^{\alpha_i}$, defined as $\prod_i \left(\dfrac{a}{p_i}\right)^{\alpha_i}$, where $\left(\dfrac{a}{p}\right)$ is the usual Legendre symbol modulo a prime p, equating to 0 if $p \mid a$ and to 1 or -1 depending on a being a quadratic residue modulo p or not;

We refer to a group homomorphism $\chi : (\mathbb{Z}/N\mathbb{Z})^* \longrightarrow \mathbb{C}^*$ as a *Dirichlet character modulo* N. We can extend χ to a multiplicative map $\chi : \mathbb{Z} \longrightarrow \mathbb{C}$, defining it as 0 on integers d such that $(d, N) \neq 1$, and we will freely switch between these notations;

Given $f(t)$ a function of positive real variable, its Mellin transform is

$$g(s) = (\mathcal{M}f)(s) = \int_0^\infty f(t)t^s \frac{dt}{t},$$

for the values $s \in \mathbb{C}$ for which the integral converges;

Given $f : \mathbb{R}^n \longrightarrow \mathbb{C}$, its Fourier transform is

$$\hat{f}(y) = \int_{\mathbb{R}^n} f(x)e^{-2\pi i x \cdot y} dx,$$

for the values $y \in \mathbb{R}^n$ for which the integral converges;

The Gamma Function

$$\Gamma(s) = \int_0^\infty e^{-t} t^s \frac{dt}{t} = (\mathcal{M}e^{-t})(s)$$

is the only complex function that is analytic except for simple poles at $-\mathbb{N}$ and such that $\Gamma(n) = (n-1)!$;

Given a Riemann surface X, we denote the space of meromorphic functions $X \longrightarrow \mathbb{C}$ as $\mathbb{C}(X)$.

2 Congruent numbers and elliptic curves

In this section we go through the main aspects of the arithmetic and geometric theory of elliptic curves, focusing on the results that allow us to investigate their link with congruent numbers. The goal is to give a general overview of the subject, therefore we will not prove every result - especially if standard in the literature - and we refer to [38] for a more in-depth introduction.

2.1 Elliptic curves

An elliptic curve over a field K is a smooth, projective curve of genus 1 defined over K with a K-rational point. If the characteristic of K is different from 2, every elliptic curve has an affine model of the form

$$y^2 = f(x)$$

with $f \in K[x]$ having degree 3 and distinct roots, and any such equation defines the affine points of an elliptic curve. For our purposes, elliptic curves will generally be defined over $\mathbb{Q}$ or quadratic number fields, so we will be able to take f as $x^3 + ax + b,\ a, b \in K$, referring to the resulting equation as a *Weierstrass form, model* or *equation* for the elliptic curve. Requiring distinct roots for the polynomial in x is equivalent to asking for its discriminant to be nonzero. In the case of the Weierstrass model, this means $-16(4a^3 + 27b^2) \neq 0$; we call this quantity the *discriminant* of the model, and denote it by Δ_E.

Elliptic curves over $\mathbb{C}$ are compact Riemann surfaces, each biholomorphic to a complex torus via the so-called Weierstrass elliptic functions (see [38, VI]); a complex torus is a quotient of $\mathbb{C}$ by a lattice (in particular, as we will see in Section 5, $\mathbb{C}$-isomorphism classes of elliptic curves are in bijection with lattices in $\mathbb{C}$ up to $\mathrm{SL}(2, \mathbb{Z})$-equivalence via the *modular j-invariant*), so it inherits a group structure from $\mathbb{C}$, which gives a natural group structure on the corresponding elliptic curve, that consequently has the fundamental property that morphisms of curves are, up to translation, group homomorphisms. Let us denote the set of endomorphisms of a curve E, which is a ring when equipped with sum and composition, by $\mathrm{End}(E)$.

Definition 2.1. *An isogeny between elliptic curves E, E' is a nonzero group homomorphism $\phi : E \longrightarrow E'$. The degree $\deg \phi$ of an isogeny ϕ is the degree of its associated divisor as a morphism of curves (see [38, II,2]).*

Thus, $\mathrm{End}(E) = \{0\} \cup \{\text{isogenies } \ E \longrightarrow E\}$. Notice that, by definition, over $\mathbb{C}$

the degree of an isogeny ϕ is its degree as a holomorphic map of Riemann surfaces, and therefore just the cardinality of $\ker \phi$.

We immediately see that the set of points of order n in a complex elliptic curve is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^2$, since this is obvious for a complex torus. Moreover, multiplication-by-m maps for any $m \in \mathbb{Z}$ are group homomorphisms, or in other words, $\mathbb{Z} \hookrightarrow \text{End}(E)$. We will often denote the multiplication-by-m map as m, although the notation $[m]$ would be more precise. It is well known that any isogeny $\phi : E \longrightarrow E'$ possesses a unique *dual* isogeny $\hat{\phi}$ such that $\phi \circ \hat{\phi} = \hat{\phi} \circ \phi = [\deg \phi]$.

Definition 2.2. *If $\mathbb{Z} \subsetneq End(E)$ we say that E has complex multiplication.*

It turns out that over $\mathbb{C}$, if an elliptic curve has complex multiplication, then its endomorphism ring is an order R in the ring of integers $\mathcal{O}_K$ of a quadratic field K (i.e. a lattice in $\mathcal{O}_K$ which generates it over $\mathbb{Q}$). In that case, we say that E has complex multiplication by R, or that it is *CM by R*.

It is well known that the group structure on elliptic curves in Weierstrass form can be easily expressed geometrically: the identity element O is the point at infinity, and given $P_1, P_2 \in E$, their sum $P_1 + P_2$ is the symmetrical, with respect to the x axis, of the third intersection between E and the line between P and Q. In the degenerate cases where $P = O$ or $P = Q$ the lines to consider are, respectively, the vertical line containing Q and the tangent to E in P.

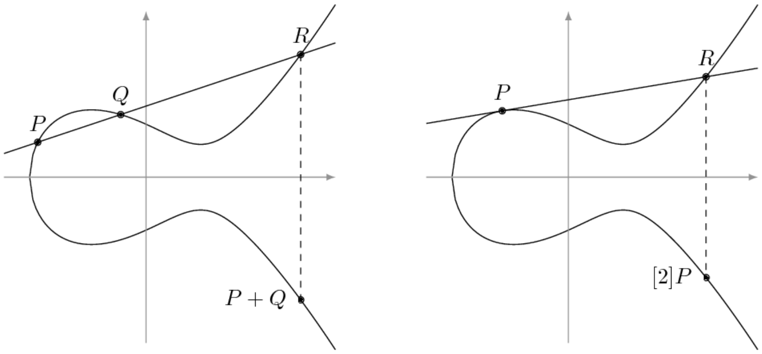

Figure 1: Group law on an elliptic curve

Example 2.3. • *Let us compute the 2-torsion of an elliptic curve over the complex numbers: $2P = O$ is equivalent to $P = -P$, that is, either $P = O$ or $y = 0$. This, since f has degree 3 and distinct roots, gives us 4 points in total, in accordance with the complex interpretation;*

- *Given $E : y^2 = x^3 + ax + b$, let us compute the double $2P$ of a point $P = (x_0, y_0) \in E$ (with $y_0 > 0$ without loss of generality): letting $m = \frac{3x_0^2+a}{2y_0}$ be the* *slope of the tangent in P, its equation is given by:*

$$y = mx - mx_0 + y_0.$$

To find $2P$ we can simply substitute into our Weierstrass equation and ask for the coefficient m^2 of the x^2 term to be 0: this yields $x_{2P} = m^2 - 2x_0$, which in turn gives $y_{2P} = y_0 - mx_0 - m^3$.

Projective varieties which are connected and possess a group structure compatible with their morphisms are called *abelian varieties.* A celebrated result by Weil, building upon the work [31] of Mordell (who did exactly the case of elliptic curves), tells us that, over any number field, the group of rational points of an abelian variety A, known as its Mordell-Weil group, is finitely generated:

$$A(K) \simeq \mathbb{Z}^{r(A)} \oplus A(K)_{\text{Tors}}. \tag{2}$$

In 1978, Mazur published a seminal paper [27] effectively proving the so-called *Torsion Conjecture* for rational elliptic curves, which gives a list of the 15 groups that appear as the torsion $E(\mathbb{Q})_{\text{Tors}}$ of the Mordell-Weil group of a rational elliptic curve. In 1996, Merel ([28]) proved the conjecture over number fields, that is, that over any number field K the size of the groups that can appear as the torsion subgroup of the rational points of elliptic curves over K is bounded in terms of $\deg(K)$.

Definition 2.4. *The quantity $r(A)$ in (2) is called the (algebraic) rank of A.*

The rank part is much less understood. For example, while the torsion of any elliptic curve over $\mathbb{Q}$ can be computed, there is no such algorithm proven to work unconditionally for computing the rank: deciding whether a given point is torsion or not is “easy”, but finding all nontorsion points appears to be hard. For instance, their arithmetic complexity, which is best defined via the *Néron-Tate height*, can be very large in terms of the coefficients, as we will see, and nontrivial upper bounds for the number of nontorsion points with complexity up to some B can currently be proven only under some hypotheses on the torsion (see for example [33]). In Section 5 we will also describe a highly nontrivial way to construct nontorsion points on some elliptic curves, which contributes to solidify our view that these points constitute a very complex and profound phenomenon.

Recently, the general belief about the behavior of the rank of rational elliptic curves has shifted from a claim of unboundedness towards one of global boundedness, thanks to the work of a few authors, including Elkies, Granville, Watkins, Park, Poonen and others ([47], [34]). If true, this phenomenon appears to be even harder to prove than the unboundedness claim, which for example has a proven analog in the function field case (see [45]). At the end of Section 3 we will briefly look at a heuristical argument for uniform boundedness in so-called *quadratic twist families*, which we will define soon and work a lot with.

2.2 Heights

Heights are an important tool for counting problems. The purpose of a height is to track the arithmetic complexity of a point, that is, of its coordinates. Hence, heights are defined as functions on projective spaces, in which varieties then embed, giving a natural generalization. We will just focus on the case of $\mathbb{P}^1(\mathbb{Q})$:

Definition 2.5. *The Weil height of a rational number* $x = \frac{a}{b}$, $\gcd(a,b) = 1$, *is* $h(x) = \log\max(|a|,|b|)$. *We refer to* $H(x) = \exp(h(x))$ *as the exponential Weil height.*

Given an elliptic curve $E : y^2 = x^3 + ax + b$ and a point $P = (x,y) \in E(\mathbb{Q})$, we define:

- $h(P) = h(x)$ its Weil, or naive, height (and the same for $H(P)$);
- $\hat{h}(P) = \lim_{n\to\infty} h(2^n P)4^{-n}$ its *Néron-Tate height*, also called *canonical height*;
- $H(E) = \max(4|a|^3, 27b^2)$ the height of the curve.

We will usually put a subscript on the heights to denote the curve to which they refer. It turns out (see [38, VIII,9]) that $\hat{h}$ is well-defined (the limit exists) and it satisfies some remarkable properties:

Proposition 2.6. 1. *$\hat{h}$ is a quadratic form on $E(\mathbb{Q})$, which is 0 exactly on its torsion, and is therefore positive definite on $E(\mathbb{Q})/E(\mathbb{Q})_{Tors}$;*

2. *$|h(P) - \hat{h}(P)| = O_E(1)$ where the constant is independent from the point;*

3. *if $E \overset{\phi}{\simeq} E'$, $\hat{h}_E(P) = \hat{h}_{E'}(\phi(P))$.*

The inequality in the second point of the above proposition makes counting points with respect the two heights almost indistinguishable: in the next section we will generalize it uniformly in certain families of curves. The third point basically tells us that the canonical height is not affected by coordinate changes, which means it is *intrinsic* to the variety, a very desirable property when working with geometric objects.

2.3 Elliptic curves over finite fields

We again refer to [38, V] for basic results about the algebraic geometry necessary to treat the topic.

We can look at elliptic curves over finite fields, say $\mathbb{F}_p$ with $p > 3$ for simplicity. Take a model $y^2 = x^3 + ax + b = f(x)$, $a, b \in \mathbb{Z}$ for E (we can always find it by starting with a Weierstrass model over $\mathbb{Q}$ and multiplying x and y by, respectively, the square and cube of the lcm of the denominators of the coefficients, which is easily seen to be an isomorphism over $\mathbb{Q}$); then, if $p \nmid \Delta_E$, E reduces to an elliptic curve $\tilde{E}/\mathbb{F}_p$ by reducing the coefficients, since the non-singularity is preserved. In this case, we say that the model has *good reduction* modulo p, whereas if $p \mid \Delta_E$, we say that the reduction is *bad*.

In this context, the set of rational points is clearly always finite: for any $x \in \mathbb{F}_p$, there are either 2 or 0 values of y (counting with multiplicity for the case $y = 0$) such that (x, y) is on the curve, depending on whether $x^3 + ax + b$ is a quadratic residue. Therefore, $\#E(\mathbb{F}_p) \leq 2p + 1$ (remembering the point at infinity). But, heuristically speaking, since the polynomial in x is squarefree in $\mathbb{F}_p[x]$, we would expect it to evaluate at a quadratic residue $\frac{p+1}{2}$ of the times, that is, we would expect around $p + 1$ rational points.

Another way, with a probabilistic flavour, to look at this is the following: we have:

$$\begin{aligned} \#E(\mathbb{F}_p) &= 1 + \#\{(x,y) \in \mathbb{F}_p^2 : y^2 = f(x)\} = \\ &= 1 + \sum_{x \in \mathbb{F}_p} \left(1 + \left(\frac{f(x)}{p}\right)\right) = p + 1 + \sum_{x \in \mathbb{F}_p} \left(\frac{f(x)}{p}\right). \end{aligned} \tag{3}$$

Since f should have no preference being a quadratic residue or not, the sum of Legendre symbols should behave like a random walk, therefore we would expect it to be of magnitude of the order of $\sqrt{p}$. The following well-known result makes this formal:

Theorem 2.7 (Hasse)**.** *In the above hypotheses,* $|\#E(\mathbb{F}_p) - (p+1)| < 2\sqrt{p}$.

Proof. Observe that the p-Frobenius map of $\mathbb{F}_p$ $F : (x, y) \mapsto (x^p, y^p)$ induces a degree p endomorphism of $\tilde{E}$ (now to be considered as the group of its points over an algebraic closure $\overline{\mathbb{F}_p}$ of $\mathbb{F}_p$), since the coefficients are fixed by it. By definition, $E(\mathbb{F}_p) = \ker(1 - F)$. It is easy to see that $1 - F$ is a separable map, so the size of the kernel equals its degree. Since the degree is a quadratic form on $\mathrm{End}(E)$, by the Cauchy-Schwartz inequality we get:

$$\begin{aligned} |\#E(\mathbb{F}_p) - p - 1| &= |\deg(1 - F) - \deg(F) - \deg(1)| \\ &\leq 2\sqrt{\deg(F)\deg(1)} = 2\sqrt{p}. \end{aligned}$$

□

We can look at this from another, slightly more insightful, point of view: given

$0 \neq \phi \in \mathrm{End}(E)$,

$$\deg(1-\phi) = (1-\phi)\widehat{(1-\phi)} = (1-\phi)(1-\hat{\phi}) = 1 - (\phi + \hat{\phi}) + \deg\phi. \tag{4}$$

Thus, $\mathrm{Tr}(\phi) = \phi + \hat{\phi} \in \mathbb{Z}$. If we define the *characteristic polynomial* of ϕ as $q_\phi(t) = t^2 - \mathrm{Tr}(\phi)t + \deg\phi$, we get that

$$q_\phi\left(\frac{m}{n}\right) = \frac{m^2 - mn\,\mathrm{Tr}(\phi) + n^2\deg\phi}{n^2} = \frac{\deg(m - n\phi)}{n^2} \geq 0,$$

by the same computation as in (4), so q_ϕ is nonnegative over $\mathbb{R}$ and hence, looking at its discriminant,

$$\mathrm{Tr}(\phi)^2 \leq 4\deg\phi \implies \mathrm{Tr}(\phi) \leq 2\sqrt{p}.$$

Applying this to $\phi = F$, by the first part of the above argument we get the Hasse bound $\#E(\mathbb{F}_p) = q_F(1) = p + 1 - \mathrm{Tr}(F)$. But we also obtain that there is a complex number $\alpha = \alpha(p, E)$, any of the roots of q_F, such that $\alpha + \bar{\alpha} = p + 1 - \#E(\mathbb{F}_p)$ and $|\alpha| = p^{\frac{1}{2}}$: this is known as the Riemann Hypothesis for our curve E. The reason for this resounding name is the following: given a curve C, let $N_r = N_r(C) = \#C(\mathbb{F}_{p^r})$; then we can encode the information about rational points on all finite fields of characteristic p with the *Zeta Function* of C:

$$Z_p(C, T) = \exp\left(\sum_{r=1}^{\infty} \frac{N_r}{r} T^r\right). \tag{5}$$

The name Zeta Function arises from the fact that for $C = \mathbb{P}^1$, since $N_r = p^r + 1$, $Z_p(C, T)$ can be written as $\exp(-\log(1-T) - \log(1-pT)) = \frac{1}{(1-T)(1-pT)}$ thanks to the Taylor expansion of $\log(1-x)$; therefore, substituting $T = p^{-s}$ and taking the product over all primes, we get:

$$\prod_p Z_p(\mathbb{P}^1, p^{-s}) = \zeta(s)\zeta(s-1). \tag{6}$$

The Zeta Function is the subject of the famous *Weil Conjectures*, which are now theorems. In the special case of genus 1 (actually also dimension 1, since the conjectures concern algebraic *varieties*), the rationality part of the conjecture states:

Theorem 2.8. *Let $E/\mathbb{F}_p$ be an elliptic curve. Then $Z_p(E, T)$ is a rational function of T: $Z_p(E, T)(1-T)(1-pT) = (1-\alpha T)(1-\bar{\alpha}T)$, where α is the same as above.*

It can be proven by noticing that, via the same methods generalized to $\mathbb{F}_{p^r}$, we get $\#E(\mathbb{F}_{p^r}) = p^r + 1 - \alpha^r - \bar{\alpha}^r$.
The statement that $|\alpha| = \sqrt{p}$ is another one of the Weil Conjectures, and is called Riemann Hypothesis because, if we again substitute $T = p^{-s}$ into (5), we get that the zeros of a Zeta Function have real part $\frac{1}{2}$.

Let us compute the number of rational points in a specific family of examples:

Example 2.9. *Let $E_n : y^2 = x^3 - n^2x$, which has discriminant $64n^6$, and let the prime $q \equiv 3 \pmod 4$ be coprime to n. Then $\#\tilde{E}(\mathbb{F}_q) = q + 1$.*

Indeed, other than the point at infinity and the three 2-torsion points, every other point will have $x \neq 0, \pm n$, so for these points there are $q - 3$ possible values for x. Since -1 is not a quadratic residue $\pmod q$ and $x^3 - n^2x$ is an odd function, it will evaluate to a quadratic residue at $x_0 \neq 0$ exactly when it will not do so at $-x_0$, and for any such x_0 there will be exactly two different values of y_0 such that (x_0, y_0) lies on the curve, since we already excluded the case with $y = 0$. Therefore, we get $\#\tilde{E}(\mathbb{F}_q) = 4 + 2\frac{q-3}{2} = q + 1$.

Observe that reducing a projective model $E : Y^2Z = X^3 + aXZ^2 + bZ^3$ modulo a good prime p gives a map $E(\mathbb{Q}) \longrightarrow \tilde{E}(\mathbb{F}_p)$, since we can lift every affine rational point to a projective point with integer coprime coordinates. The other result we need, whose proof requires the development of some of the theory of elliptic curves over local fields and will therefore be omitted, is the following:

Proposition 2.10. *Let $E : y^2 = x^3 + ax + b,\ a,\ b \in \mathbb{Z}$ be an elliptic curve, let $m \geq 1$ and let p be coprime with m and Δ_E. Then $E[m] \hookrightarrow \tilde{E}(\mathbb{F}_p)$.*

2.4 Quadratic twists

Let us take an elliptic curve $C : y^2 = f(x)$ with $f(x) = x^3 + ax + b \in \mathbb{Q}[x]$ and, given $d \in \mathbb{Q}^*$ squarefree, let us consider the curve $C_d : dy^2 = f(x)$. It is an elliptic curve since, over $\mathbb{Q}(\sqrt{d})$, it is isomorphic to C via $y \mapsto y\sqrt{d}$, which also gives:

$$C(\mathbb{Q}(\sqrt{d})) \simeq C_d(\mathbb{Q}). \tag{7}$$

Moreover, C_d has the Weierstrass form $C_d : y^2 = x^3 + ad^2x + bd^3$ obtained via $(x, y) \mapsto (x/d, y/d)$, since $(y/d)^2 = \frac{dy^2}{d^3} = \frac{f(x)}{d^3} = g(x/d)$ with $g(x) = x^3 + ad^2x + bd^3$. We say that C_d is the *quadratic twist* of C by d.

This construction is perfectly valid over any field K with $char(K) \neq 2$ and $d \in K^*$ squarefree. If K is a number field, we call $\{C_d\}_{d \in \mathcal{O}_K \setminus \{0\}/\square}$ the *quadratic twist family* of C. So over $\mathbb{Q}$ we are twisting by the squarefree integers. This easily implies that two twists of the same curve will never be isomorphic over $\mathbb{Q}$, a phenomenon that in general can happen, for instance over finite fields (whereas a curve will *never* be K-isomorphic to any of its other K-twists). We can then study the arithmetic of the curves in the family as d varies, and this turns out indeed to be very rich and complex: we will delve more into this in the next section, but already in the following subsection we will see, for instance, how these curves can have different ranks.

Let $E_1 = E$ be the rational elliptic curve defined by

$$E : y^2 = x^3 - x$$

and let $\{E_n\}$ be its quadratic twist family, for which we will from now on omit the subscript for the set in which the twisting parameter varies. So we have Weierstrass models $E_n : y^2 = x^3 - n^2x$.

Proposition 2.11. *The only torsion in $E_n(\mathbb{Q})$ is the 2-torsion; that is,*

$$E_n(\mathbb{Q})_{Tors} \simeq (\mathbb{Z}/2\mathbb{Z})^2 \ \forall n \geq 1.$$

Proof. On one hand, by Example 2.3, we know that $\{O, (0,0), (\pm n, 0)\}$ is the 2-torsion of E_n, isomorphic to $(\mathbb{Z}/2\mathbb{Z})^2$.
On the other hand, Proposition 2.10 implies that for all $m \geq 1$, $E_n(\mathbb{Q})[m] \hookrightarrow \tilde{E}(\mathbb{F}_p)$ for all sufficiently large p: so, by Example 2.9, $\#E_n(\mathbb{Q})[m] \mid \gcd(\{q+1\})$ as q varies over all sufficiently large primes $q \equiv 3 \pmod 4$, so clearly $\#E_n(\mathbb{Q})[m] \mid 4$ by the Chinese Remainder Theorem and Dirichlet's Theorem on primes in arithmetic progressions. But this means that $\#E(\mathbb{Q})_{\text{Tors}} \mid 4$ and the proposition follows. □

2.5 The congruent number family

Let n be a squarefree positive integer. We saw that if n is congruent, then the equations $u^2 \pm n = \square$ both have solutions in $\mathbb{Q}$. Multiplying them and putting $w = tv$ in the notation of the previous section, we get $u^4 - n^2 = w^2$. Finally, multiplying both sides by u^2 and putting $x = u^3$, $y = wu = tuv$, we get:

$$y^2 = x^3 - n^2x, \tag{8}$$

so we obtain a rational point on E_n. Since t, u, v are nonzero by the hypothesis on their squares - the only possibility would be that $t = 0$, in which case n would be a square, which would imply $n = 1$, but then $v = 2n = 2$ would not be a square - this point is not in its 2-torsion, so by Proposition 2.11 it is not a torsion point. We have proven one implication in the following:

Theorem 2.12. *Let n be a squarefree positive integer. Then n is congruent if and only if E_n has positive rank.*

For the other implication, let $P \in E_n$ be a nontorsion point and let $2P = (x, y)$, $y \neq 0$ so that, if $x = \frac{p}{q}$ with $\gcd(p, q) = 1$, then $\gcd(p, n) = 1$ and q is even by the duplication law described in Example 2.3. From (8) we get:

$$\frac{(p - qn)p(p + qn)}{q^3} = y^2$$

and since $(p, q) = 1$ the numerator and the denominator on the LHS are coprime, so they are both squares: in particular, $q = Q^2$. Moreover, since $\gcd(p, n) = 1$ we have

$\gcd(p, p \pm qn) = 1$ and, since q is even, p and qn have different parity, so $p \pm qn$ are also coprime. This means that the three factors in the numerator are also all squares, say T^2, U^2, V^2, which implies that

$$\left(\frac{p}{q} - n, \frac{p}{q}, \frac{p}{q} + n\right) = \left(\left(\frac{T}{Q}\right)^2, \left(\frac{U}{Q}\right)^2, \left(\frac{V}{Q}\right)^2\right) = (t^2, u^2, v^2)$$

is the required arithmetic progression.

We now have translated the CNP to the problem of determining which curves in the quadratic twist family of a specific elliptic curve have positive algebraic rank over $\mathbb{Q}$. In the next section we will define and focus more on the study of the *analytic* rank in such families, which, while conjectured to be equal to the algebraic rank, appears harder to control (with some exceptions). This will shed some light on what to expect from the solution of the CNP. Before delving into this, let us establish that:

Proposition 2.13. *E_1 has rank $0 : 1$ is not a congruent number.*

Proof. Take a nontorsion point on E_1 and double it as above to get (x, y) with $x = \frac{p}{q}$. Then again $q = Q^2,\ p = P^2$ are squares and so $(P^2 - Q^2)(P^2 + Q^2)$ is a square by the equation defining E_1. But, as Fermat notoriously proved by infinite descent, $X^4 - Y^4 = Z^2$ has no solutions in $(\mathbb{Z}^+)^3$: if the two factors are coprime, then $P^2 + Q^2$ is a square and we find a smaller solution thanks to the formula for primitive Pythagorean triples, otherwise P and Q are odd and both factors are twice a square, say $P^2 - Q^2 = 2R^2,\ P^2 + Q^2 = 2S^2$, but then $R^2 + S^2 = P^2$ is primitive and we can again perform the descent. □

Along with the following:

Example 2.14 (Zagier)**.** *$n = 157$ is a congruent number: the right triangle with catheti*

$$a = \frac{411340519227716149383203}{21666555693714761309610},\ b = \frac{6803298487826435051217540}{411340519227716149383203}$$

has area 157*,*

the proposition shows that there are congruent and noncongruent numbers, and, more generally, that different elliptic curves in a quadratic twist family can have different ranks.

Zagier's example was chosen, instead of presenting a claim easier to verify such as a congruent triangle for $n = 5$, because it has the interesting property that it is the *simplest* congruent triangle of area 157, in the sense of the number of digits involved. This shows that, hiding behind the CNP, there is a phenomenon of very high arithmetic complexity.

3 The analytic rank in quadratic twist families

3.1 The L-function of an elliptic curve

In this section we will work with elliptic curves $E/\mathbb{Q}$ given by a so-called *minimal model*, that is, an equation $y^2 = x^3 + ax + b,\ a, b \in \mathbb{Z}$ with Δ_E the smallest among the discriminants of integer Weierstrass models of E. This discriminant is also called *minimal.*

Given an elliptic curve $E/\mathbb{Q}$, set $a_p = a_p(E) = p+1-\#E(\mathbb{F}_p)$, where for primes of bad reduction the cardinality of the set of rational points is given by the first equality in (3). We will now introduce its L-function, formally defined through a so-called *Euler product,* whose complex-analytic properties (conjecturally) encode information about the elliptic curve. Recall the definition (5) of the Zeta Function relative to the primes of good reduction and extend it to those of bad reduction by setting $Z_p(E,T) = \frac{1-a_pT}{(1-T)(1-pT)}$.

Definition 3.1. *The L-function of an elliptic curve E is*

$$L(E,s) = \frac{\zeta(s)\zeta(s-1)}{\prod_p Z_p(E,p^{-s})}. \tag{9}$$

Observe that, by (6) and by Theorem 2.7, we have:

$$L(E,s) = \prod_p (1 - a_p p^{-s} + \epsilon(p)p^{1-2s})^{-1}, \tag{10}$$

where $\epsilon(p) = 0, 1$ depending on whether E has bad or good reduction modulo p. The choice of this formula will be more comprehensible after introducing the L-function of a modular form in Section 4, via the correspondence with normalized Hecke eigenforms. Nonetheless, while inside the realm of elliptic curves, we shall make a simple observation:

Remark 3.1.1. *If we could plug in $s = 1$ in* (9)*, we would get:*

$$L(E,1) = \prod_p \left(\frac{p - a_p + 1}{p}\right)^{-1} = \prod_p \frac{p}{|E(\mathbb{F}_p)|}. \tag{11}$$

This, albeit heuristically, gives some insight on the relationship between the behavior of $L(E,s)$ around $s = 1$ and the rank of E: if the latter is large, we expect a lot of rational points on the reduced curves, and hence $L(E,s)$ to rapidly go to 0 as s approaches 1.

We would like to look at $L(E,s)$ as a function of complex variable, but in order to do so we need information about its convergence. The Hasse bound allows us to prove:

Lemma 3.2. *The L-function of an elliptic curve converges absolutely for* $\operatorname{Re} s > \frac{3}{2}$.

Proof. If $\operatorname{Re} s > \frac{3}{2}$, the factors at bad primes have no poles because $|a_p p^{-s}| < \frac{2}{p} \leq 1$, while those at good primes have no poles since $|1 - a_p p^{-s} + p^{1-2s}| > 1 - 2p^{-1} - p^{-2} > 0$ for $p > 2$; but there are infinitely many factors, so now we need to look at the whole product: by the Riemann Hypothesis for the Zeta function of of the reduced curve, for a good reduction prime p we have $1 - a_p p^{-s} + p^{1-2s} = (1 - \alpha(p)p^{-s})(1 - \overline{\alpha(p)}p^{-s})$, so the product decomposes into two parts:

$$L(E,s) = \prod_p (1 - \alpha(p)p^{-s})^{-1} \prod_p (1 - \overline{\alpha(p)}p^{-s})^{-1}.$$

Putting $\operatorname{Re} s = 1.5 + \epsilon$, the absolute value of the denominator at p is at least $1 - p^{-1-\epsilon}$ for both parts, so by the geometric series formula they are both bounded by

$$\prod_p \sum_{k \in \mathbb{N}} (p^{-1-\epsilon})^k = \sum_{k \in \mathbb{N}} n^{-1-\epsilon},$$

which converges absolutely for $\epsilon > 0$. □

Let us remark that expanding this Euler product we obtain a Dirichlet series:

$$L(E,s) = \sum_{n \geq 1} a_n n^{-s}, \tag{12}$$

where one easily sees that the coefficients of p^{-s} are exactly the a_ps and, for primes of good reduction, $a_{p^2} = a_p^2 - 2p$.

Unfortunately, Lemma 3.2 is not enough to meaningfully look at $L(E,1)$, but as we will see in the case of congruent number curves,

Theorem 3.3. *$L(E,s)$ admits an analytic continuation to the whole complex plane, which satisfies the following functional equation for $\omega_E \in \{-1,1\}$:*

$$N_E^{\frac{s}{2}} (2\pi)^{-s} \Gamma(s) L(E,s) =: \Lambda(E,s) = \omega_E \Lambda(E, 2-s). \tag{13}$$

Here the sign of the functional equation ω_E is called the *root number* of E, while N_E is the *conductor* of E, an integer that we will rigorously define in Section 4 (but we can even take the one arising from this statement as its definition) which divides Δ_E and is divisible by every prime of bad reduction. This allows us to finally define:

Definition 3.4. *The analytic rank $r_{an}(E)$ of E is the order of vanishing of $L(E,s)$ at $s=1$. In other words, if we have a Taylor expansion $L(E,s)=\sum_{k\geq k_0} c_k(s-1)^k$ around 1, then $r_{an}=k_0$.*

Remark 3.4.1. *It is well known that $\Gamma(s)$ does not vanish on $\mathbb{C}$; therefore, taking derivatives of* (13) *and evaluating them at the symmetry point $s=1$ shows that the analytic rank is even if the root number is positive, and odd if the root number is negative.*

Conjecture 3.5 (Birch, Swinnerton-Dyer)**.** *For any elliptic curve $E/\mathbb{Q}$ we have $r(E)=r_{an}(E)$.*

This conjecture, especially in its original formulation

$$\prod_{p<X} \frac{p}{\#E(\mathbb{F}_p)} \sim c_E(\log X)^{-r(E)},$$

which was initially adopted and is equivalent to the modern one if we assume the convergence of the product in (11), is supported by overwhelming numerical evidence ([10]). Qualitatively it says, via the deep formalism of Dirichlet L-series and their analytic continuation, that rational points over finite fields relate to rational points over the global field as we expect, i.e. that the latter cannot "disappear" under the operation of reduction $\pmod p$ as p varies, and maybe even more deeply so, that $\mathbb{F}_p$-rational points "must come" from global rational points. The conjecture is currently proven only if r_{an} is less than 2, and the results involved are already incredibly deep: in Section 5 we will work with a fundamental construction for one of them, which proves that $r_{an}=1 \implies r\geq 1$.

Before working with these new tools we just defined, let us observe that, even if we have the analytic continuation, the equation in (11) is still a heuristic: there is no guarantee that the Euler product evaluated at 1 coincides with the value at 1 of the analytic continuation. The Euler product only makes sense for $\operatorname{Re} s > \frac{3}{2}$. In fact, in [13] Goldfeld proved:

Proposition 3.6. *If the product in* (11) *converges, then it equals $\frac{L(E,1)}{\sqrt{2}}$. Moreover, $L(E,s)$ satisfies the Riemann Hypothesis.*

This is surprising, but luckily does not affect anything in terms of order of convergence, so our heuristic is saved. By the second part of the proposition, it is very unlikely that we could make it formal.

3.2 Computations with $L(E_n,s)$

In this subsection, we will denote the standard scalar product on $\mathbb{R}^2$ as $x\cdot y$. Our goal is to prove Theorem 3.3 for our quadratic twist family $\{E_n\}$, and to determine their conductors and root numbers. The technique that we will use is classical, having as prototype the standard proof that the Riemann Zeta Function $\zeta(s)$ (which, in a way, is the "simplest" L-function) admits an analytic continuation that satisfies a certain functional equation: we manipulate the L-function in order to transform it in a so-called *Hecke L-function* of an imaginary quadratic field, an analog of ζ (in fact, with this step we completely forget about our starting elliptic curve, see (19)), which is then related to the Mellin transform of some kind of theta function, for example

$$\theta(t) = \sum_{n\in\mathbb{Z}} e^{-\pi t n^2} \tag{14}$$

in the case of ζ (we will use a slightly different definition for the theta function in the modular setting, see Section 4.2). The theta function satisfies a functional equation of the form

$$\theta(t) = t^{-\frac{1}{2}}\theta(t^{-1}), \tag{15}$$

where the branch of the square root is that for which $\sqrt{1} = 1$, which allows us to show that the Mellin transform is well-defined and to obtain the functional equation for the L-function. The functional equation of the theta-type functions are usually obtained through the Poisson summation formula:

Lemma 3.7. *Let $g : \mathbb{R}^n \longrightarrow \mathbb{C}$ be smooth, bounded and rapidly decreasing, i.e. $\lim_{|x|\to\infty} |x|^k g(x) = 0 \ \forall k$ (this ensures that the Fourier transform is defined on all $\mathbb{R}^n$). Then:*

$$\sum_{m\in\mathbb{Z}^n} g(m) = \sum_{m\in\mathbb{Z}^n} \hat{g}(m).$$

For an introduction to this technique and for proofs of the analytic tools involved see for example [21, II]. We will assume other elementary results in Fourier and complex analysis, including:

Lemma 3.8. *Assume that all functions to which we apply the Fourier transform are as in Lemma 3.7. We have:*

1. *if $g(x) = f(x+a)$ then $\hat{g}(y) = e^{2\pi i a\cdot y}\hat{f}(y)$;*
2. *if $g(x) = f(bx)$ then $\hat{g}(y) = b^{-n}\hat{f}(\frac{y}{b})$;*
3. *if $f(x) = e^{-\pi x\cdot x}$ then $\hat{f} = f$;*
4. *if $g = w\frac{\partial f}{\partial x}$ then $\hat{g}(y) = 2\pi i w\cdot y\hat{f}(y)$;*
5. *if $f : (0,\infty) \longrightarrow \mathbb{C}$ is e^{-ct} then $(\mathcal{M}f)(s) = c^{-s}\Gamma(s)$.*

Let w, for the remainder of this proof, be the vector $(1,i)$, and let $u=(u_1,u_2)\in\mathbb{R}^2\setminus\mathbb{Z}^2$. The idea is to express $L(E_n,s)$ using sums of the form

$$\sum_{m\in\mathbb{Z}^2}\frac{(m+u)\cdot w}{|m+u|^{2s}}.$$

First, observe that if $p\mid\Delta_{E_n}$ then either $p\mid n$ or n is odd and p is 2. In the first case E_n reduces to the cuspidal cubic $y^2=x^3$. Since $x^3=x^2x$, we easily see that x^3 is a quadratic residue $\pmod p$ $\iff$ x is, so we get that $E(\mathbb{F}_p)$ is given by the point at infinity, $(0,0)$ and then 2 points for each nonzero quadratic residue, yielding $a_p=0$. In the second case an immediate computation yields $\#E(\mathbb{F}_2)=3$, so by (10) we get:

$$L(E_n,s)=\prod_{p\nmid 2n}(1-a_pp^{-s}+p^{1-2s})^{-1}.$$

Consider n as fixed. We can say more on the complex numbers $\alpha_p=\alpha(E_n,p)$ other than the Riemann Hypothesis. Indeed, E_n has complex multiplication by the ring $\mathbb{Z}[i]$ since $\eta:(x,y)\mapsto(-x,iy)$ is an automorphism; let $p\equiv 1\pmod 4$ be of good reduction, $\tilde{E}_n$ be the reduction modulo p and F be the p-Frobenius. Recall that we saw in Section 2.3 that any endomorphism of $\tilde{E}_n$ is the root of a polynomial of degree 2 and integer coefficients, so $F\in\mathbb{Z}[\eta]$ since it is integral over $\mathbb{Z}$ and since $\mathbb{Q}(i)=\operatorname{Frac}(\mathbb{Z}[i])=\operatorname{Frac}(\mathbb{Z}[\eta])\subset\operatorname{Frac}(\mathbb{Z}[\eta,F])$, which has degree 2 by what we just said, and so equals $\operatorname{Frac}(\mathbb{Z}[\eta])$. Writing $F=a+b\eta$ and computing the degree, we get $p=a^2+(\eta+\hat{\eta})+b^2\deg\eta=a^2+b^2$, since η is an automorphism and $\eta^2=-1\Longrightarrow\hat{\eta}=-\eta$.

This, by the discussion in Section 2.3, tells us that, *up to signs,* $\alpha_p=a+bi$. But we know that $p+1-2\operatorname{Re}\alpha_p=\tilde{E}_n(\mathbb{F}_p)$, which has full 2-torsion, so a is odd since $p\equiv 1\pmod 4$. Similarly, we see that $\tilde{E}_n(\mathbb{F}_p)$ has a 4-torsion point P_4 for such p if and only if $\left(\dfrac{n}{p}\right)=1$: indeed, a simple computation via the duplication law shows that we must have $P_4=(\pm\sqrt{-1}n,(1+\sqrt{-1})\sqrt{\pm n})$. When this happens we have $8\mid\#E(\mathbb{F}_p)$, which gives us the sign of a. Writing p as $p=a^2+b^2$ where a,b have unspecified signs and parity, it is not too hard to see that these considerations can be summarized as follows:

$$\alpha_p=\begin{cases}\sqrt{-p} & \text{if } p\equiv 3 \pmod 4,\\ a+bi \text{ with signs s.t. } a+bi\equiv\left(\dfrac{n}{p}\right) \pmod{2+2i} & \text{if } p\equiv 1 \pmod 4,\end{cases}\tag{16}$$

where the $p\equiv 3\pmod 4$ case clearly follows from the Riemann Hypothesis for $\tilde{E}_n$ and Example 2.9.

For P above p, set $\alpha_P = \alpha_p$. This allows us to rewrite each Zeta Function $Z_p(E_n)$ as a product over prime ideals of $\mathbb{Z}[i]$ that divide p: as we know, p splits (into two prime ideals) in the Gaussian integers if and only if $p \equiv 1 \pmod 4$, so we get:

$$(1-T)(1-pT)Z_p(E_n, T) = \prod_{P|(p)} (1 - (\alpha_P T)^{\deg P}), \tag{17}$$

where $\deg P$ is the degree of inertia of P over (p).

Now we want to switch to a product over *elements* of $\mathbb{Z}[i]$, to better resemble an Euler product. The key is to define a multiplicative map $\tilde{\chi}_n$ on $\mathbb{Z}[i]$ such that for any prime ideal P, $\tilde{\chi}_n(x) = \alpha_P^{\deg P}$ for any generator x of P. Given x and $I = (x)$ an element and an ideal of $\mathbb{Z}[i]$, denote by $\mathbb{N}x = x\bar{x}$ and by $\mathbb{N}I = |\mathbb{N}x|$ their norms, which shall cause no confusion with the symbol denoting the natural numbers. Recall that they are multiplicative functions.

We know that prime ideals $P \nmid 2$ in $\mathbb{Z}[i]$ are either those above a rational prime $p \equiv 1 \pmod 4$ or generated by a rational prime $p \equiv 3 \pmod 4$, so the norm of a generator is precisely $\mathbb{N}(\alpha_P^{\deg P})$ by (16). Therefore, $\tilde{\chi}_n(x)$ will be of the form $x\chi'_n(x)$ with χ'_n being multiplicative and having modulus 1 (we are working with $p \nmid 2n$). Moreover, since the only point in which the definition of $\alpha_P(E_n)$ depends on n is in the $p \equiv 1 \pmod 4$ case, through the Legendre symbol $\left(\dfrac{n}{p}\right)$ in the congruence $\pmod{2i+2}$, we will have $\chi'_n(x) = \chi'_1(x)\left(\dfrac{n}{\mathbb{N}x}\right)$ with χ'_1 again multiplicative, since $\left(\dfrac{n}{p^2}\right) = \left(\dfrac{n}{p}\right)^2 = 1$. Notice that this, along with the definition of χ'_n, tells us that:

$$a_p(E_n) = a_p(E_1)\left(\frac{n}{p}\right). \tag{18}$$

At this point, (16) tells us that $\chi'_1(x) = j$ where $j \in \{1, -1, i, -i\}$ is the unique power of i such that $jx \equiv 1 \pmod{2i+2}$.

If we set $\tilde{\chi}_n(I) = \tilde{\chi}_n(x)$ for any generator x of I, our discussion and (17) give:

$$L(E_n, s) = \prod_{P \nmid 2n} \left(1 - \frac{\tilde{\chi}_n(P)}{(\mathbb{N}P)^s}\right)^{-1} = \sum_I \tilde{\chi}_n(I)(\mathbb{N}I)^{-s} = \frac{1}{4}\sum_{a,b\in\mathbb{Z}} \frac{(a+bi)\chi'_n(a+bi)}{(a^2+b^2)^s}, \tag{19}$$

since every such ideal has 4 generators, corresponding to multiplication by powers j of i. The second equality holds because $\tilde{\chi}_n$ and the norm are multiplicative and the Gaussian integers are a Dedekind domain.

Recalling the definition of χ'_n, we can see how it only depends on the class of $x \pmod{4n}$: clearly χ'_1 only depends on $x \pmod{2i+2}$, so we just need to see that

$\left(\dfrac{n}{\mathbb{N}x}\right)$ depends on $x \pmod{4n}$; if we had the Jacobi symbol with top and bottom switched that would be clear, but luckily we can use quadratic reciprocity: let $n = 2^\beta p_1 ... p_l$, $\beta \in \{0, 1\}$ and let $q_1, ..., q_k$ be the primes $\equiv 1 \pmod 4$ dividing $\mathbb{N}x$. Since all other primes appear with an even exponent, by quadratic reciprocity we get:

$$\left(\frac{n}{\mathbb{N}x}\right) = \left(\frac{2}{\mathbb{N}x}\right)^\beta \prod_{1\le i\le l, 1\le j\le k} \left(\frac{p_i}{q_j}\right) = \left(\frac{2}{\mathbb{N}x}\right)^\beta \prod_{i,j} \left(\frac{q_j}{p_i}\right) = \left(\frac{2}{\mathbb{N}x}\right)^\beta \left(\frac{\mathbb{N}x}{n_0}\right),$$

where $n_0 = n$ if n is odd and $\frac{n}{2}$ otherwise.
Since $x \equiv x' \pmod n \implies \mathbb{N}x \equiv \mathbb{N}x' \pmod n$ we just need to show that, if n is even and $x \equiv x' \pmod{4n}$, then $\left(\dfrac{2}{\mathbb{N}x}\right) = \left(\dfrac{2}{\mathbb{N}x'}\right)$. But this follows from the fact that $8 \mid 4n$ and $\left(\dfrac{2}{k}\right)$ is determined by $k \pmod 8$. Therefore, (19) becomes:

$$L(E_n, s) = \frac{1}{4} \sum_{0\le a,b<4n} \chi'_n(a+bi) \sum_{m\in\mathbb{Z}^2} \frac{a+bi+4nm\cdot w}{|(a,b)+4nm|^{2s}} =$$
$$= \frac{1}{4}(4n)^{1-2s} \sum_{0\le a,b<4n} \chi'_n(a+bi) \sum_{m\in\mathbb{Z}^2} \frac{(m+(\frac{a}{4n},\frac{b}{4n}))\cdot w}{|m+(\frac{a}{4n},\frac{b}{4n})|^{2s}}. \tag{20}$$

This is a form that we can relate to the Mellin transform of a 2-dimensional version of the theta function (14). Let:

$$\theta_u(t) = \sum_{m\in\mathbb{Z}^2} (m+u)\cdot w e^{-\pi t|m+u|^2}.$$

As for the theta function, we find a functional equation for $\theta_u(t)$ by means of the Poisson summation formula: let $g(x) = (x+u)\cdot we^{-\pi t|m+u|^2}$ (regarding t as fixed). Then, putting $f(x) = e^{-\pi|x|^2}$, we have:

$$g(x) = -\frac{1}{2\pi t} w \cdot \frac{\partial f(\sqrt{t}(x+u))}{\partial x},$$

and hence, by Lemma 3.8 1)-4), $\hat{g}(y) = -it^{-2} w\cdot ye^{2\pi i u\cdot y - \pi t^{-1}|y|^2}$. So if we define:

$$\theta^u(t) = \sum_{m\in\mathbb{Z}^2} m\cdot we^{2\pi i m\cdot u - \pi t|m|^2},$$

the Poisson summation formula gives us:

$$\theta_u(t) = -\frac{i}{t}\theta^u\left(\frac{1}{t}\right). \tag{21}$$

The functional equation allows us to prove that the integral

$$\int_0^\infty t^s \sum_{m\in\mathbb{Z}^2} (m+u)\cdot w e^{-\pi t|m+u|^2} \frac{dt}{t} \tag{22}$$

defining the Mellin transform of θ_u converges to an entire function of s: we have $|m+u|^2 > c > 0$ since $u \notin \mathbb{Z}^2$, so for t large enough we can upper bound the integrand with $e^{-\pi ct}$, whereas for small enough t, since the only term in θ^u with $|m| = 0$ vanishes due to the presence of $m \cdot w$, the functional equation allows us to upper bound the integrand with $e^{-\frac{\pi c}{t}}$, for some $0 < c < 1$. Therefore, we can interchange sum and integral in (22) to get:

$$(\mathcal{M}\theta_u)(s) = \sum_{m\in\mathbb{Z}^2} (m+u)\cdot w \mathcal{M} e^{-\pi t|m+u|^2} = \pi^{-s}\Gamma(s) \sum_{m\in\mathbb{Z}^2} \frac{(m+u)\cdot w}{|m+u|^{2s}}$$

thanks to Lemma 3.8 5). So (20) tells us that, for $\operatorname{Re} s > \frac{3}{2}$:

$$\pi^{-s}\Gamma(s)L(E_n, s) = \frac{1}{4}(4n)^{1-2s} \sum_{\substack{0\leq a,b<4n\\(a,b)\neq(0,0)}} \chi_n'(a+bi)(\mathcal{M}\theta_u)(s), \tag{23}$$

where $u = (\frac{a}{4n}, \frac{b}{4n}) \notin \mathbb{Z}^2$. Since we have a finite sum, an entire Mellin transform, and both $(4n)^{1-2s}$ and $\pi^s\Gamma(s)^{-1}$ are entire (see Remark 3.4.1), it follows that $L(E_n, s)$ admits an analytic continuation to the whole complex plane.

Moreover, the functional equation tells us that:

$$(\mathcal{M}\theta_u)(s) = -i\int_0^\infty t^{s-2}\theta^u\left(\frac{1}{t}\right)\frac{dt}{t} \overset{t\mapsto t^{-1}}{=} -i(\mathcal{M}\theta^u)(2-s).$$

Evaluating this Mellin transform (which is entire by the same argument as for that of θ_u) as we did above yields $\pi^{s-2}\Gamma(2-s)\sum_{m\in\mathbb{Z}^2} m\cdot w \exp\left(\frac{\pi i}{2n} m\cdot(a,b)|m|^{-2(2-s)}\right)$, so for $\operatorname{Re}(2-s) > \frac{3}{2}$ the expression in (23) equals:

$$-i(4n)^{1-2s}\pi^{s-2}\Gamma(2-s)\frac{1}{4}\sum_{m\in\mathbb{Z}^2}\frac{m\cdot w}{|m|^{2(2-s)}}\sum_{0\leq a,b<4n}\chi_n'(a+bi)e^{\frac{\pi i}{2n}m\cdot(a,b)}.$$

Calling the last sum $S_m(n)$ we have the following lemma, which is the application of classical results about Gauss sums (see Definition 4.28, and observe that by our discussion χ_n' is a Dirichlet character modulo $4n$) and whose proof can be found in [21, p.68]:

Lemma 3.9. *$S_m(n)$ is 0 if $m\cdot w \notin (1+i)\mathbb{Z}[i]$, while otherwise, setting $m\cdot w = (1+i)x$, we have:*

$$S_m = 4\chi_n'(x)\begin{cases}\left(\dfrac{-2}{n}\right)(1+i)n & \text{if } n \text{ is odd,}\\ \left(\dfrac{-1}{n_0}\right)(-1+i)n & \text{if } n = 2n_0 \text{ is even.}\end{cases}$$

Lemma 3.9 allows us to write (we focus just on the n odd case for simplicity):

$$\sum_{m\in\mathbb{Z}^2}\frac{m\cdot w}{|m|^{2(2-s)}}\sum_{0\le a,b<4n}\chi_n'(a+bi)e^{\frac{\pi i}{2n}m\cdot(a,b)}=$$
$$=(1+i)^2 2^s\left(\frac{-2}{n}\right)n\sum_{x\in\mathbb{Z}[i]}\tilde{\chi}_n(x)(\mathbb{N}x)^{s-2};$$

but by (20) the right-hand side is exactly $4i\left(\frac{-2}{n}\right)2^{s+1}nL(E_n,2-s)$, so for $\operatorname{Re}s<\frac{1}{2}$ the expression for $\pi^{-s}\Gamma(s)L(E_n,s)$ in (23) equals:

$$(4n)^{2-2s}\pi^{s-2}2^{s-1}\Gamma(2-s)\left(\frac{-2}{n}\right)L(E_n,2-s)=$$
$$=\left(\frac{-2}{n}\right)\pi^{s-2}\Gamma(2-s)(8n^2)^{1-s}L(E_n,2-s)$$

and so we find exactly the functional equation of (13) with $N_E=32n^2$ and $\omega_E=\left(\frac{-2}{n}\right)$. For the even n case, we have the same with $N_E=16n^2$ and $\omega_E=\left(\frac{-1}{n_0}\right)$. This tells us that the root number of E_n is 1 when $n\equiv 1,2,3 \pmod 8$ and -1 when $n\equiv 5,6,7 \pmod 8$.

By Remark 3.4.1 this means that, asymptotically, half of the congruent number curves have even analytic rank and half of them have odd analytic rank. This is obvious if we let n run over all positive integers, but it also holds under our definition with n only taking squarefree values, since standard asymptotics for squarefree numbers satisfying a congruence condition from Analytic Number Theory imply that squarefree integers are equidistributed among the six residue classes $\pmod 8$ above.

3.3 Average r_{an} I: the *explicit formula*

As we anticipated, the analytic rank is hard to control with current methods. Notice that if $\omega_E=-1$ then the analytic rank is odd, so in particular it is positive. This already has some important consequences, since the root number is not too hard to control, for example in quadratic twist families, as we just saw. The other tool that we have is the Guinand-Weil explicit formula (see [16], [48]), which requires us to assume the Riemann Hypothesis for $L(E,s)$ (not to be confused with the Riemann Hypothesis for the Zeta function of the reduction $\tilde{E}$ modulo a good prime, which is known!):

Proposition 3.10. *Suppose all zeroes ρ of $L(E,s)$ have real part* 1. *Let $g:\mathbb{R}\longrightarrow\mathbb{R}$ be even, piecewise C^1 and have compact support $K\subset[-1,1]$ and modulus bounded by*

1. *Let $\hat{g}$ be its Fourier transform and set $g_X(t) = g(t/\log X)$. Then, for any $X \geq 2$:*

$$r_{an}(E)\hat{g}(0) + \sum_{\tau = \operatorname{Im}\rho \neq 0} \hat{g}(\tau \log X) =$$

$$= g(0)\frac{\log N_E}{\log X} + \frac{2}{\log X}\left(U_1(E,X) + U_2(E,X)\right) + O\left(\frac{1}{\log X}\right), \tag{24}$$

where

$$U_k(X) = U_k(E,X) = -\sum_{p^k \leq X, p \geq 5} \frac{a_{p^k}}{p^k} g_X(k \log p) \log p.$$

Corollary 3.10.1. *Suppose $L(E,s)$ satisfies the Riemann Hypothesis. Then for $X \geq 2$ we have:*

$$r_{an} < \frac{\log N_E}{\log X} + \frac{2}{\log X}\left(U_1(X) + U_2(X)\right) + O\left(\frac{1}{\log X}\right). \tag{25}$$

Proof. Setting $g(t) = \max(0, 1 - |t|)$ we have $\hat{g}(t) = \left(\frac{\sin \pi t}{\pi t}\right)^2$ which is nonnegative; since both g and $\hat{g}$ evaluate to 1 for $t = 0$, we get (25) by neglecting the summation on the LHS of (24). □

Let us get rid of the U_2 term. As we observed in the previous section, computing the coefficients in (12) gives $a_{p^2} = a_p^2 - 2p$ for $p \nmid \Delta_E$ and $a_{p^2} = 1$ otherwise, so, bounding $g \leq 1$, we have:

$$|U_2(X)| \leq 2 \sum_{2 \leq p \leq \sqrt{x}} \frac{\log p}{p} = \log X + O(1) \tag{26}$$

by the Hasse Bound, where the second equality is a theorem of Mertens ([29]) (we will control other sums of this type via similar results akin to the *Prime Number Theorem*, without quoting all of them). Controlling U_1 is harder, for using the same bounds we get:

$$|U_1(X)| \leq 2 \sum_{2 \leq p \leq X} \frac{\log p}{\sqrt{p}} \gg \sqrt{X}. \tag{27}$$

It is already evident that it is hopeless to effectively use this formula for a *single* curve, since, even if we knew the implied constants in the error terms for some (necessarily small) X, we would still get a lower bound around $\log N_E$ (the actual result, with an implicit constant, is $\frac{1}{2}\frac{\log N_E}{\log \log N_E} + O\left(\frac{\log N_E}{(\log \log N_E)^2}\right)$, see [5]). The idea is to notice that, for some infinite *families* $\mathcal{F}$ of curves, letting the curve vary allows us to get enough cancellations inside $\sum_E U_1(E,X)$ to prove a finite upper bound for the average rank

$$\operatorname{avg} r_{an}(\mathcal{F}) \overset{\text{def}}{=} \limsup_{T \to \infty} \frac{1}{|\mathcal{F}(T)|} \sum_{E \in \mathcal{F}(T)} r_{an}(E), \tag{28}$$

where $\mathscr{F}(T) = \mathscr{F} \cap \{E/\mathbb{Q} : H(E) \leq T\}$; more specifically, for a quadratic twist family we can control the sum of the a_ps as the curve varies, and then proceed with the same technique as before to sum over p:

Theorem 3.11. *Let $\mathscr{F} = \{E^{(n)}\}_{n \in (\mathbb{Z}\setminus\{0\})/\square}$ be the quadratic twist family of some $E^{(1)}/\mathbb{Q}$. Then $\operatorname{avg} r_{an}(\mathscr{F}) \leq \frac{13}{4}$.*

Remark 3.11.1.

1. *We have to be careful about the indexing of the family: twisting by non-squarefree integers (and hence having infinitely many copies of the same twist) allows for a stronger bound, but it does not track the data in the correct way.*
2. *This result is weaker than the $\frac{3}{2}$ that Heath-Brown obtains in [20] (using nonconstant weights), but we shall produce an easier proof. In the same paper, Heath-Brown also studies the average analytic rank for all rational elliptic curves, bounding it by* 2.

Proof of Theorem 3.11. In Section 3.2 we saw that, for p coprime to $2n$,

$$a_p(E^{(n)}) = a_p(E^{(1)}) \left(\frac{n}{p}\right) \tag{29}$$

holds for the congruent number family (so with $n > 0$), but it holds (for primes of good reductions) in general: our proof can be easily extended to any quadratic twist family of a CM curve, while if the basis curve does not have complex multiplication then the Legendre symbol interpretation (3) shows that $a_p(E^{(n)}) = a_p(E^{(1)})$ or $-a_p(E^{(1)})$ depending on whether n is a square $\pmod p$.

Let us remark that Goldfeld improved the bound (26) by a constant, but valuable, factor (see [13]):

Lemma 3.12.

$$|U_2(E^{(n)})| \leq \frac{1}{4}\log X + O_E(\log\log n),$$

where we conveniently stated the result for a quadratic twist family because, by (29), all terms in $U_2(E^{(n)}, X)$ are equal to those in $U_2(E^{(1)}, X)$ except for those at primes of bad reduction, again since $a_{p^2} = a_p^2 - 2p$ and $\left(\frac{n}{p}\right)^2 = 1$. The contribution of primes of bad reduction accounts exactly for the error term.

Since the primes of bad reduction are those that divide $n\Delta_{E^{(1)}}$, their contribution in the sum ${\sum_{n\leq T}}' |U_1(E^{(n)}, X)|$ is $O\left(\frac{\log n}{\log\log n}\right)$ (each term being bounded by $\frac{2}{\sqrt{5}}\log 5$

by the Hasse bound) because a positive integer m has $O\left(\frac{\log m}{\log\log m}\right)$ prime factors; so, if $\log T = O(\log X)$, their contribution will vanish in the limit. Since this will be the case, we can disregard them, and we will use (29) for all primes for simplicity. This implies:

$$\sideset{}{'}\sum_{|n|\le T} |U_1(E^{(n)}, X)| = \left| \sideset{}{'}\sum_{|n|\le T} \sum_{5\le p\le X} \frac{g_X(\log p)\log p}{p} a_p(E^{(1)}) \left(\frac{n}{p}\right) \right| =$$

$$= \left| \sum_{5\le p\le X} \frac{g_X(\log p)\log p}{p} a_p(E^{(1)}) \sideset{}{'}\sum_{|n|\le T} \left(\frac{n}{p}\right) \right| \le 2 \sum_{2\le p\le X} \frac{\log p}{\sqrt{p}} \left| \sideset{}{'}\sum_{|n|\le T} \left(\frac{n}{p}\right) \right|. \qquad (30)$$

This leads us to study sums in the form:

$$\left| \sideset{}{'}\sum_{|n|\le T} \chi_p(n) \right|,$$

with χ_p a Dirichlet character modulo p. Without restricting to squarefree integers, improvements for prime modulus over the well-known Polya-Vinogradov inequality $\sum_{M\le n\le M+N} \chi_N(n) = O(\sqrt{N}\log N)$ give a bound of $O(Tp^{-\frac{1}{2}})$. Summing over squarefree integers only is not much harder: in [32, Lemma 2.3] Munsch showed that:

$$\left| \sideset{}{'}\sum_{n\le T} \chi_p(n) \right| \ll \sqrt{T}\log(T)^{\frac{3}{2}} p^{\frac{3}{16}}.$$

Since our sum is either 0 (if $p \equiv 3 \pmod 4$) or twice that considered by Munsch, substituting in (30) gives:

$$\sideset{}{'}\sum_{n\le T} |U_1(E^{(n)}, X)| \ll \sqrt{T}(\log T)^{\frac{3}{2}} \sum_{2\le p\le X} p^{-\frac{5}{16}} \log p \le C\sqrt{T}(\log T)^{\frac{3}{2}} X^{\frac{11}{16}}$$

for some absolute constant C and finally, combining this with our formula for the conductors, (25), Lemma 3.12 and (28), we obtain:

$$\operatorname{avg} r_{an}(\mathscr{F}) \le$$

$$\le \limsup_{T\to\infty} \frac{\pi^2}{12T}\Big(2 \sideset{}{'}\sum_{|n|\le T} \frac{\log|n|}{\log X} + \frac{12}{\pi^2} T \frac{|U_2(E^{(n)})|}{\log X} + C\sqrt{T}(\log T)^{\frac{3}{2}} X^{\frac{11}{16}}\Big),$$

and choosing $X = T^{\frac{8}{11}+\epsilon}$, $\epsilon > 0$, so that the third term vanishes, we get:

$$\operatorname{avg} r_{an}(\mathscr{F}) \le \frac{1}{2} + \lim_{T\to\infty} 2\frac{\log T}{\log T^{\frac{8}{11}+\epsilon}} = \frac{1}{2} + \frac{22}{8+11\epsilon} \Longrightarrow \operatorname{avg} r_{an}(\mathscr{F}) \le \frac{13}{4}, \qquad (31)$$

by letting $\epsilon \to 0$. □

Remark 3.12.1. *Technically, we have not used the definition in* (28) *precisely, since the curves in $\mathcal{F}$ with height less than or equal to T are not those with $n \leq T$, but those with $H(E^{(1)})n^6 \leq T$. But this just means rescaling $T \mapsto (T(H((E^{(1)}))^{-1})^{\frac{1}{6}}$, which clearly does not affect the limit.*

3.4 Average r_{an} II: ordering by smallest point

We just saw that, if we order the curves in a quadratic twist family by their height, we can bound the average analytic rank with a small constant. It is natural to ask what the real answer should be: extensive numerical evidence tells us that most elliptic curves, ordered this way, have *algebraic* rank either 0 or 1. Obviously we believe the BSD Conjecture to be true, so the same would hold for the analytic rank. In 1979 Goldfeld conjectured:

Conjecture 3.13 (Minimalist Conjecture)**.** *Let $\mathcal{F}$ be a quadratic twist family. Then, ordering the curves by height,* 50% *of them have (analytic) rank* 0 *and* 50% *have (analytic) rank* 1*.*

Here the density is to be taken as an asymptotic value: indeed, for most curves we know that at least $\frac{X^{\frac{1}{7}}}{\log X}$ of their twists of height up to X have rank at least 2 ([37]), an asymptotic which we can improve to $X^{\frac{1}{2}}$ for all curves if we consider the analytic rank ([14], [24]), and should probably equal $X^{\frac{3}{4}}$ ([26]). The name *Minimalist Conjecture* comes from the fact that it predicts what is, in light of the discussion at the end of Section 3.2, the simplest possible scenario for ranks in quadratic twist families. Notice that it implies that the average rank in such a family is $\frac{1}{2}$.

The reason for which elliptic curves with high (≥ 2) rank are sparse, and their distribution, are very poorly understood, especially for "high" ranks, say $r \geq 5$, for which only recently some heuristics, most notably [34], started to appear. The Minimalist Conjecture was subsequently extended to the family of all rational elliptic curves (up to isomorphism), seeing groundbreaking progress with the work [3] by Bhargava and Shankar, who in 2013 obtained an upper bound of 0.885 for the average rank. Unfortunately their methods do not generalize to quadratic twist families, which are too sparse. In the quadratic twist case, until recently the only finite bounds known were for specific families (such as $\frac{3}{2}$ for the congruent number family again by Heath-Brown, see [18]), but were conditional on the rank and the analytic rank having the same parity, the so-called Parity Conjecture. In the last year, Smith published two preprints ([40], [41]) proving, among other things, that 100% of the quadratic twists of an elliptic curve $E/\mathbb{Q}$ satisfying some technical conditions on the torsion have rank ≤ 1, and in fact that at least 50% have rank 0. We will now look into a somewhat "opposite" result, hinting at the fact that precisely capturing the distribution of ranks in these families may be a very challenging task.

Definition 3.14. *Given an elliptic curve $E/\mathbb{Q}$, denote:*

$$\eta_E = \begin{cases} \min_{P\in E(\mathbb{Q})\setminus E(\mathbb{Q})_{Tors}} \hat{h}(P) & \text{if } E(\mathbb{Q}) \setminus E(\mathbb{Q})_{Tors} \neq \emptyset, \\ \infty & \text{else.} \end{cases}$$

Moreover, given a family $\mathscr{F}$ of elliptic curves, its average analytic rank ordering the curves by η is:

$$\widetilde{\operatorname{avg}} r_{an}(\mathscr{F}) = \liminf_{T\to\infty} \frac{1}{|\mathscr{F}(T)|} \sum_{\eta_E \leq T} r_{an}(E).$$

Theorem 3.15 (Le Boudec)**.** *Let $\mathscr{F} = \{E^{(n)}\}_{n\in(\mathbb{Z}\setminus\{0\})/\square}$ be a quadratic twist family. Then:*

$$\widetilde{\operatorname{avg}} r_{an}(\mathscr{F}) > 1. \tag{32}$$

The meaning of this result is that, assuming BSD and just talking about "rank", if we neglect curves with rank 0, even though the *Minimalist Conjecture* philosophy (and Smith's results, for the right curves) tell us that we are left with curves which, taking the limit, have rank 1 with "probability 1", this fails to hold if we order these curves by the smallest height of its nontorsion rational points. The reason for a different behavior with respect to the curve height ordering is intuitively clear: having more than one nontrivial rational point, for a curve of height in a given range, makes it likely for the smallest of the heights to be lower than the typical height of a generator in rank 1 curves with height in the same range (just by an "independence" principle). Indeed, Le Boudec conjectures ([24, Conjecture 1]):

Conjecture 3.16. $\widetilde{\operatorname{avg}} r_{an}(\mathscr{F}) = \frac{3}{2}$.

Again talking about the algebraic rank r, this would simply mean that, looking at curves which have a nontorsion point ordered in this way, we again have a *Minimalist Conjecture*-type phenomenon, with half of them having no other points and half of them having another point. Interestingly, this has an easy explanation for r_{an}, since the equivalent statement would concern the parity of the root number (which, as we proved in the congruent number case, is given by a quadratic character modulo N_E), which would intuitively be equidistributed under our ordering, unless small rational points "preferred" some congruence classes for n.

Before proving the theorem, let us remark that, in addition to a negative root number implying an odd, and therefore positive, analytic rank, a positive root number also implies a positive (and even) analytic rank, *provided that the algebraic rank is positive.* This is a result of Kolyvagin ([22]), generalizing the work of Coates and Wiles ([8]), who proved the same statement for elliptic curves with complex multiplication. We state it without proof, since the methods involved go far beyond the topic of this thesis:

Theorem 3.17 (Coates-Wiles, Kolyvagin)**.** *Let $E/\mathbb{Q}$ be an elliptic curve. If $L(E,1) \neq 0$, then $E(\mathbb{Q})$ is finite.*

Proof of Theorem 3.15. For the sake of convenience, in this proof we will consider twists by non-necessarily squarefree integers, with no change in the notation. Let $E = E^{(1)} : y^2 = x^3 + ax + b,\ a, b \in \mathbb{Z}$ be our elliptic curve, and let us suppose that a and b divide $M = 12N_E$: this is not restrictive because $(x,y) \mapsto (xM^{-2}, yM^{-3})$ defines an isomorphism of E, and clearing denominators gives an equation with the required property. This is just a technical condition necessary to apply a lemma. Put $Q(u,v) = u(v^3 + au^2v + bu^3)$ and let:

- $A(X) = \{n \in \mathbb{Z} : |\mu(n)| = 1, |n| \leq X\}$;
- $B(Y) = \{n \in \mathbb{Z} : |\mu(n)| = 1, \eta_{E^{(n)}} \leq Y\}$;
- $$\Omega(Y) = \frac{1}{\#B(Y)} \sum_{\substack{n \in B(Y) \\ \omega_{E^{(n)}} = 1}} 1.$$

By Theorem 3.17, we just need to prove that:

$$\liminf_{Y \to \infty} \Omega(Y) > 0. \tag{33}$$

The proof of the theorem will follow from our ability to control three things:

1. the difference between the Néron-Tate and the Weil heights of a point on a curve in the twist family, *independently on the twist parameter*;
2. $\Omega(Y)$;
3. the number of integer lattice points (u,v) in a box such that $Q(u,v)$ is squarefree and $E^{(Q(u,v))}$ has root number 1.

We do this via the following lemmas:

Lemma 3.18.

$$\hat{h}_{E^{(n)}}(P) - h(P) = O_E(1).$$

Lemma 3.19.

$$\#B(Y) \ll Y^4.$$

Lemma 3.20. *Let $a, b \mid 12N_E$, let r_1, r_2 be integers and let*

$$S_Q(Z, r_1, r_2) = \sum_{\substack{|u|,|v| \le Z \\ (u,v) \equiv (r_1, r_2) \pmod M}} |\mu(Q(u,v))|.$$

Then there exists $c > 0$ such that $S_Q(Z, r_1, r_2) \ge cZ^2 + O(Z(\log Z)^{-\frac{1}{2}})$.

We omit the proof of the third lemma, which can be found in [14, Proposition 6], since their methods are outside the scope of this thesis.

The first lemma is easy: as we saw in (7), the map $P = (x, y) \mapsto (x, \sqrt{d}y) = Q$ defines an isomorphism between $E^{(d)}(\mathbb{Q})$ and $E(\mathbb{Q}(\sqrt{d}))$, so by Proposition 2.6 we have:

$$\hat{h}_{E^{(d)}}(P) = \hat{h}_E(Q) = h_E(Q) + O_E(1) = h_E(P) + O_E(1),$$

where the Weil height is the same because it only depends on the x coordinate.

For the second lemma, the key is to enlarge the set whose cardinality we are trying to bound, passing from twists to rational points on twists, in order to have an easy arithmetic way to compute the bound:

$$\begin{aligned} \#B(Y) \le & {\sum_{n\in\mathbb{Z}}}' \#\{P \in E^{(n)}(\mathbb{Q}) \setminus E^{(n)}(\mathbb{Q})_{\text{Tors}} : \hat{h}_{E^{(n)}}(P) \le \log Y\} \le \\ \le & {\sum_{n\in\mathbb{Z}}}' \#\{P \in E^{(n)}(\mathbb{Q}) \setminus E^{(n)}(\mathbb{Q})_{\text{Tors}} : h_{E^{(n)}}(P) \ll_E \log Y\}, \end{aligned} \tag{34}$$

where the second inequality follows from the first lemma. Notice that if we drop the requirement on the height, the set in the last summation is just the set of triples $(x, y, z) \in \mathbb{Z} \times (\mathbb{Z}^+)^2$ such that

$$dy^2 z = x^3 + axz^2 + bz^3 \tag{35}$$

and with $\gcd(x, y, z) = 1$ (just clearing denominators), where the positivity condition on y is ensured by the nontorsion assumption. Let $h = \gcd(d, z)$; clearly $h \mid x^3$, and since d is squarefree $h \mid x$. Let us write $d = hd_1, z = hz_1, x = hx_1$ with $\gcd(h, d_1) = 1$. Factoring out h^2 from both sides of (35) we get $d_1 y^2 z_1 = h(x_1^3 + ax_1 z_1^2 + bz_1^3)$. Since h is coprime with d_1 but also with y (otherwise if $1 \ne d_1 \mid x, y$ then $d_1 \mid z$ by (35), contradicting the fact that $\gcd(x, y, z) = 1$), we get $h \mid z_1$, giving:

$$d_1 y^2 z_2 = x_1^3 + ax_1 h^2 z_2^2 + bh^3 z_2^3 \tag{36}$$

for $z_1 = hz_2$. Let $k = \gcd(x_1, z_2)$; from $\gcd(y, xz) = \gcd(d_1, x_1) = 1$ it follows that $\gcd(k, d_1 y) = 1$, so by (36) we get $k^3 \mid z_2$, but clearly also $z_2 \mid x_1^3 \implies z_2 \mid k^3$, so $k^3 = z_2$, hence factoring it out and letting $x_1 = kx_2$ we finally get:

$$d_1 y^2 = x_2^3 + ax_2 h^2 k^4 + bh^3 k^6 = f(x). \tag{37}$$

Remembering that $d = d_1 h, x = hkx_2, z = h^2k^3$, we are counting quintuples $(d_1, x_2, y, h, k) \in (\mathbb{Z})^2 \times (\mathbb{Z}^+)^3$ satisfying (37) and such that:

$$\begin{cases} d_1 h \text{ is squarefree,} \\ \gcd(x_2, hk) = 1, \\ |x_2|, hk^2 \ll Y^2, \end{cases}$$

and then multiplying by 2 the result to account for our choice of positive y ($P \mapsto -P$ simply changes the sign of the y-coordinate). Notice that the requirement $|x_2|, hk \ll Y^2$ is equivalent to the one on the height in (34), since by construction $h_{E^{(d)}}(P) = h(\frac{x}{z}) = h(\frac{hkx_2}{h^2k^3}) = \log\max(|x_2|, hk^2)$. Hence, we get an upper bound:

$$\begin{aligned} B(Y) &\leq 2 \sum_{|x_2|, hk^2 \ll Y^2} \#\{(d_1, y) \in \mathbb{Z} \times \mathbb{Z}^+ : d_1 y^2 = f(x),\ |\mu(d_1)| = 1\} \\ &\leq 2 \sum_{|x_2|, hk^2 \ll Y^2} 1 \ll Y^4 \end{aligned}$$

since any nonzero integer can uniquely be written as the product of a squarefree integer and a square.

Going back to our goal of proving (33), Lemma 3.18 enables us to explicitly find a family of values of n such that $r(E^{(n)}) \geq 1$. $P = (vu^{-1}, u^{-2})$ is a point on $E^{(Q(u,v))}$: indeed, by the lemma, provided that $|\mu(Q(u,v))| = 1$ we get $\eta_{E^{(Q(u,v))}} \leq c_E \sqrt{\max(|u|, |v|)}$ for some $c_E > 0$, which tells us that if $|u|, |v|$ are sufficiently large in terms only of E, then P is nontorsion. Thus, there exists $C_E > 0$ such that

$$\eta_{E^{(Q(u,v))}} \leq Y$$

holds for any u, v with $|\mu(Q(u,v))| = 1$ and $|u|, |v| \leq C_E Y^2$, with at most finitely many exceptions depending only on E.

Setting $r_Q(n, Z) = \#\{(u, v) : |u|, |v| \leq Z,\ Q(u, v) = n\}$ and $C(Y) = \{n \in \mathbb{Z} : |\mu(n)| = 1,\ r_Q(n, C_E Y^2) \geq 1\}$, we get:

$$\#(B(Y) \setminus C(Y)) \ll 1$$

and hence, by Lemma 3.19,

$$\Omega(Y) \gg Y^{-4} \sum_{\substack{n \in C(Y) \\ \omega_{E^{(n)}} = 1}} 1.$$

We can use the Cauchy-Schwartz inequality to estimate the sum: notice that $\sum_{n \in \mathbb{Z}} r_Q(n, C_E Y^2) \mathbb{1}_{\omega_{E^{(n)}} = 1} = S_Q(C_E Y^2)$, and set $R_Q(Z) = \sum_{n \in \mathbb{Z}} r_Q(n, Z)^2$. Then by Lemmas 3.18 and 3.19:

$$\Omega(Y) \gg Y^{-4} \frac{S_Q(C_E Y^2)^2}{R_Q(C_E Y^2)} \gg \frac{Y^4}{R_Q(C_E Y^2)},$$

so we are left with the task of proving that $R_Q(C_E Y^2) \ll Y^4$, or, equivalently, that $R_Q(Z) \ll Z^2$. Being defined as a second moment, $R_Q(Z)$ is equal to

$$\#\{(u_1, u_2, u_3, u_4) : |u_i| \leq Z,\ Q(u_1, u_2) = Q(u_3, u_4)\}.$$

Notice that this set is in bijection with $\{(0,0,0,0)\} \cup \bigcup_{k \leq Z} \{(u_1, u_2, u_3, u_4) : |u_i| \leq Z/k,\ Q(u_1, u_2) = Q(u_3, u_4),\ \gcd(u_i) = 1\}$ and let V be the quartic surface in $\mathbb{P}^3(\mathbb{Q})$ defined by $Q(x_0, x_1) - Q(x_2, x_3)$. Then, accounting for an ambiguity of sign, we have:

$$R_Q(Z) = 1 + 2 \sum_{k \leq Z} \#\{\bar{x} \in V(\mathbb{Q}),\ H(\bar{x}) \leq Z/k\}$$

Let S be the finite set of lines contained in V and let $U = V \backslash \cup_{l \in S} l$. Then the condition $4a^3 + 27b^2 \neq 0$ yields that U is a smooth surface and hence, by [19, Theorem 10],

$$\#\{\bar{x} \in V(\mathbb{Q}) : H(x) \leq Z_0\} \ll_\epsilon Z_0^{2 - \frac{2}{9} + \epsilon}$$

and, since any line contributes with $\ll Z_0^2$ rational points of height bounded by Z_0, we finally get:

$$R_Q(Z) \ll \sum_{k \leq Z} \left(\frac{Z}{k}\right)^2 \ll Z^2,$$

which proves the theorem. □

3.5 Curves with high analytic rank

If we believe the Minimalist Conjecture (and we do), we should expect elliptic curves with rank > 1 to be quite rare. Assuming BSD, it is not too hard to numerically "check" this since, for a given curve, even if we do not know that the Euler product in (11) converges for $s = 1$, we can compute it over finitely many primes p, expand it as a Taylor series $\sum_{k \geq 0} c_k (s-1)^k$ and see that there is an index r such that c_r is far from 0 and $c_s \sim 0\ \forall s < r$. For example, the curve $E : y^2 + y = x^3 - 7x + 36$ is a good candidate for having analytic rank 4, because such a computation yields:

$$L(E, s) \approx (8.1737876581...)(s-1)^4 + O((s-1)^5),$$

where the coefficients of the lower order terms are $< 10^{-6}$. Therefore, taking a set of curves, for example the first B twists of a quadratic twist family, we can perform the above operation to a precision sufficient to convince ourselves that most of them indeed have analytic rank ≤ 1.

Nevertheless, we cannot prove that a *single* curve has analytic rank r for *any* $r > 3$: proving that $L^{(n)}(E, 1)$, the so-called *central value* of a derivative of the L-function, is nonzero (when it is) is not too hard with the aid of computational tools, but proving that it is 0 is a task that conceptually eludes any standard finitistic computation. As we will see in Section 5, even proving that some elliptic curves have analytic rank 3 is exceptionally hard and has deep consequences. For $r_{an} = 0, 1$ the task is easy since the root number is sufficient, while for $r_{an} = 2$ we need Theorem 3.17. For the algebraic rank the situation is a bit more encouraging, since a way to prove a lower bound is "simply" to find rational points, but, as we remarked in Section 2, bounding it from above is a very hard problem.

It is not too hard to prove that there are infinitely many elliptic curves over $\mathbb{Q}$ of rank 0, and the same result, while nontrivial, is known for rank 1 (see Section 6), but the question becomes open for any $r > 1$: for instance, the results in [2] fail just short of settling the $r = 2$ case, which is known only the assumption of the Parity Conjecture (see [7]). Notice that relaxing the condition to having rank $\geq r$ we obtain a far easier question, since, for instance, we have systematic ways to construct infinite families of curves having rank $\geq r$, with at most finitely many exception, for values of r up to 19. This is thanks to the fact that we can construct elliptic curves over $\mathbb{Q}(t)$ of rank up to 19 and to the *Specialization Theorem*:

Theorem 3.21 (Silverman). *Let $E/\mathbb{Q}(t)$ be an elliptic curve and let its specialization $E_{t_0}/\mathbb{Q}$ at $t = t_0$ be an elliptic curve, which happens with at most finitely many exceptions. Then $r(E) \leq r(E_{t_0})$ for all but finitely many values of t_0.*

The same task with $r_{an} \geq r$ is also doable for $r = 2$ and $r = 3$, as we will prove in our next theorem (for $r = 2$ we can already extract a proof, either from Le Boudec's result or from the current progress on the BSD Conjecture), but as we just said, even the case $r = 4$ appears out of reach. This is due to the aforementioned difficulty in lower-bounding the analytic rank. Motivated by this analysis, it is natural to ask the following question:

Question 1. *For which $r > 1$ do there exist infinitely many elliptic curves of (analytic) rank r?*

By what we said above, this question is much harder than the version with rank $\geq r$: in the case of the algebraic rank, the scenario that is difficult to exclude (for $r \gtrsim 20$) is the one in which there are elliptic curves with arbitrarily large rank, but only finitely many for any large rank (although we cannot currently rule out the unlikely scenario where there are $r_1 > r_2$ with finitely many curves having rank r_2 and infinitely many having rank r_1). We will focus on the analytic rank, proving a conditional partial result:

Theorem 3.22. *Assume the Riemann Hypothesis for the L-functions of the elliptic curves in the congruent number family. Then:*

1. *infinitely many congruent number curves E_n have even analytic rank $2 \leq r_{an}(E_n) \leq 10$;*

2. infinitely many congruent number curves E_n have odd analytic rank $3 \leq r_{an}(E_n) \leq 25$.

Remark 3.22.1. *Notice that, conditional on GRH, we either obtain an affirmative answer for our question when $r = 2, 3$ respectively, or we obtain infinitely many curves of analytic rank r for a fixed small $r > 3$, which would be an interesting result since we currently have no idea how to prove that a single curve has such an analytic rank.*

Proof of Theorem 3.22. Let us focus on the first statement. The second one will follow from the same techniques.

Consider the subfamily $\{E_n\}$ for $n = Q(u, v)$ already introduced in Le Boudec's proof of Theorem 3.15 (we again consider twists by all integers, obviously keeping in mind that our infinitude result concerns twists by numbers which do not differ by a rational square). In this case, $n = u(v^3 - vu^2) = uv(u + v)(u - v)$. Notice how this, modulo $\mathbb{Q}^{*2}$, is exactly our congruent number generating formula (1). We know that $P = \left(\frac{v}{u}, \frac{1}{u^2}\right)$ is a nontorsion rational point on $E_{Q(u,v)}$, so by Theorem 3.17 these curves have positive analytic rank. As we know, the root number is 1 when $n \equiv 1, 2, 3 \pmod 8$ (because its squarefree part is in its same residue class modulo 8, as we observed in the Introduction), which happens, for instance, if $u \equiv 6 \pmod 8$ and $v \equiv 1 \pmod 8$.

Set $P(u) = Q(u, 1)$ and let us consider the subfamily $\mathcal{F}_6^P = \{E_n\}$ with $n = P(u)$, $u \equiv 6 \pmod 8$. These curves have even analytic rank at least 2 by what we just said. If we "knew" that the numbers $n = u(u + 1)(u - 1)$, $u \equiv 6 \pmod 8$, had pairwise distinct squarefree parts (unlikely to be true), we would just need to prove that the average analytic rank of the family is less than 12. Luckily such a strong claim is not necessary:

Lemma 3.23. *Let D be a constant. The asymptotic density of the values of u for which* $\mathrm{sqf}(P(u)) < D$ *is* 0.

Proof of Lemma 3.23. Notice that, since our set of values of u has positive $\left(\frac{1}{8}\right)$ asymptotic density inside $\mathbb{N}$, it is equivalent to show that the asymptotic density of the specified set is 0 inside that of values of u or inside $\mathbb{N}$, so let us stick with $\mathbb{N}$. We show it even removing the congruence condition on u. Since u and $u^2 - 1$ are coprime, $\mathrm{sqf}(P(u)) = \mathrm{sqf}(u)\,\mathrm{sqf}(u^2 - 1)$. The values of u for which $\mathrm{sqf}(u) < D$, which form a superset of those we are interested in, are exactly those in the form $m\square$, $m < D$ squarefree. Clearly each of these finitely many sets has the asymptotic density (inside $\mathbb{N}$) of that of the squares, which is 0, so we get the desired claim. □

The lemma implies that we just need to strictly upper bound the average analytic rank of $\mathcal{F}_6^P$ by 12: if for all squarefree $d \geq D$ we had $r_{an}(E_d) \notin \{2, 4, 6, 8, 10\}$, then a proportion of density 1 of the curves in our family would have analytic rank at least 12 by the lemma, giving a lower bound of 12 for the average analytic rank.

Let us compute our upper bound using the explicit formula in its inequality form (25). Observe that by definition:

$$\operatorname{avg} r_{an}(\mathscr{F}_6^P) = \limsup_T \frac{8}{T} \sum_{\substack{u \le T \\ u \equiv 6 \pmod 8}} r_{an}(E_{P(u)}).$$

Let us treat the terms $\frac{\log N_E}{\log X}$, $\frac{2}{\log X} U_i(X)$, $i = 1, 2$ appearing in (25) separately. By Lemma 3.12, the third term is bounded by $\frac{1}{2}$ if $\log T = o(X)$, which will be the case (notice that this is exactly what we did in (31) with the $\frac{\log\log T}{\log X}$ term). Since $N_{E_n} = O(n^2)$ by our results of Section 3.2, we get $\frac{\log N_{E_{P(u)}}}{\log X} \sim 6\frac{\log u}{\log X}$, so the average of their sum is $\sim (\log X)^{-1}\frac{8}{T}\left(6\frac{T\log T}{8}\right) = 6\log T$, leaving us with a $6\frac{\log T}{\log X}$ factor.

Finally, we are left with U_1: working as in (30), again forgetting about the contributions at bad reduction primes, we have:

$$\sum_{u \equiv 6 \pmod 8} |U_1(E_{P(u)}, X)| = \left| \sum_u \sum_{5 \le p \le X} \frac{g_X(\log p)\log p}{p} a_p(E) \left(\frac{P(u)}{p}\right) \right| \le$$

$$\le \sum_{5 \le p \le X} \frac{g_X(\log p)\log p}{p} |a_p(E)| \left| \sum_u \left(\frac{P(u)}{p}\right) \right|.$$

Let us focus on the sum of Legendre symbols: since our values of p are coprime with 8, the Chinese Remainder Theorem and (3) imply that (let us keep the congruence condition on u implicit):

$$\left| \sum_u \left(\frac{P(u)}{p}\right) \right| = \left| \left\lfloor \frac{T}{8p} \right\rfloor a_p(E) + \sum_{x = p\left\lfloor \frac{T}{8p} \right\rfloor}^{\left\lfloor \frac{T-6}{8} \right\rfloor} \left(\frac{P(8x+6)}{p}\right) \right|, \tag{38}$$

since $y^2 = P(u)$ is precisely our Weierstrass equation for $E = E_1$. To estimate the remaining sum, let $f(x) = P(8x+6)$. Since the range of values over which x is running is shorter than $\frac{T}{8} - p(\frac{T}{8p} - 1) = p$, Weil's bounds on sum of characters (see [9, 1.3]) combined with a classical generalization of the Polya-Vinogradov inequality (see [1, 1.1]) imply, as Burgess states in [6, Abstract], that the modulus of the sum is $\ll \sqrt{p}\log p$, so we finally get:

$$\sum_u |U_1(E_{P(u)}, X)| \le \sum_{5 \le p \le X} \left(T\frac{\log p}{8p^2} a_p(E)^2 + c_0 \frac{\log^2 p}{\sqrt{p}} |a_p(E)| \right).$$

By Hasse's bound and Example 2.9:

$$\sum_{u} |U_1(E_{P(u)}, X)| \leq \sum_{\substack{5 \leq p \leq X \\ p \equiv 1 \pmod 4}} \left(\frac{T \log p}{2p} + 2c_0 \log^2 p \right) = \frac{T}{4} \log X + O(X \log X).$$

Now, to get our contribution for the average analytic rank, we need to multiply by $\frac{2}{\log X}$ and by $\frac{8}{T}$ for the average. Choosing $T = X^{1+\epsilon}$, $\epsilon > 0$ and letting $\epsilon \longrightarrow 0$, we get:

$$\operatorname{avg} r_{an}(\mathcal{F}_6^P) \leq \frac{1}{2} + 6(1+\epsilon) + 4 < 12.$$

For the second statement, the starting point is our knowledge, which we expressed in Section 3.1, that the BSD Conjecture is known for $r_{an} < 2$. Therefore, if we have a curve with two independent nontorsion points, it cannot have analytic rank ≤ 1. This allows us to apply the same idea as for the first statement: if we can find an infinite family of values $n = Q(u)$ for which the curves E_n have rank at least 2, root number -1 and unbounded squarefree parts for a density 1 subset, we can use the explicit formula to upper bound the average analytic rank in the family, hopefully getting a small bound.

In [17], Halbeisen and Hungerbühler prove that the curves E_n with $n = uv(u+2v)(2u+v)(u^2-v^2)(u^2+uv+v^2)$ have rank at least 2. Setting $v = 1$, define $Q(u) = u(u-1)(u+1)(u+2)(2u+1)(u^2+u+1)$ and $Q_1(u) = \frac{Q(u)}{(u^2-1)}$. Notice that, for $u \equiv 5 \pmod{16}$, we have $Q_1(u) \equiv 7 \pmod 8$ and $u^2 - 1 \equiv 8 \pmod{16}$. Therefore, $\operatorname{sqf}(Q(u)) \equiv 7 \cdot 2 \equiv 6 \pmod 8$. Moreover, for any u, the gcd between any two of the irreducible factors $q_i(u)$ of $Q(u)$ is in $\{1, 2, 3\}$, from which we have $\operatorname{sqf}(Q(u)) \geq 6^{-6} \prod_i \operatorname{sqf}(q_i(u))$ (we do not care about the suboptimality of the constant, just that it is absolute). This, by the same arguments as above (in particular Lemma 3.23 applied to $Q(u)$), means that strictly upper-bounding the average analytic rank of the family $\mathcal{F}_5^Q = \{E_n : n = Q(u),\ u \equiv 5 \pmod{16}\}$ by 27 implies the second statement of the theorem.

The upper bound is obtained in the exact same fashion of that for even analytic rank (in particular, notice how the application of the Chinese Remainder Theorem to sum over $\mathbb{F}_p$ is unchanged), with two differences:

- the expression for n now has degree 7 in our variable u, so the average of the log conductor terms for the same choice of $T = X^{1+\epsilon}$ gives us $14(1+\epsilon)$;
- in the RHS of (38), the first summation is not $\left\lfloor \frac{T}{8p} \right\rfloor a_p(E)$ since we are dealing with another polynomial, but the same bound [9, 1.3] mentioned above for

the sum over the whole $\mathbb{F}_p$ shows that we can bound that term with $6\left\lfloor \frac{T}{8p} \right\rfloor \sqrt{p}$. Therefore, our previous bound for the U_1 term gets multiplied by $\frac{6}{2} = 3$, leaving us with 12.

Since the bound for U_2 is unchanged, we obtain:

$$\operatorname{avg} r_{an}(\mathscr{F}_5^Q) \leq 14(1+\epsilon) + 12 + \frac{1}{2} < 27$$

as desired, which concludes the proof of Theorem 3.22. □

In the following discussion, which will be quite informal, assume the BSD Conjecture in order to compare results about algebraic and analytic rank.

It is natural to ask how "soft" Theorem 3.22 is, that is, by what margin the information loss in upper bounding the average analytic rank (which should be 2 and 3 in the families we considered) makes us incapable of ruling out scenarios which are false. Until not long ago, the predominant belief was that ranks are unbounded in quadratic twist families (see [47, Introduction]), so that one would reasonably expect infinitely many curves for each rank, and our theorem would not "allow" for any impossible scenario. But, as already mentioned in Section 2, the belief has now shifted.

Let us describe a heuristic for the ranks in the quadratic twist family of $E : y^2 = x^3+ax+b = f(x),\ f \in \mathbb{Z}[x]$ being bounded, with at most finitely many exceptions, by 8. If this heuristic is true, while Theorem 3.22 1) does not lose too much information, the second statement is quite weak.

The starting point is a refined form of the BSD Conjecture, which gives a formula for the $r-$th Taylor coefficient $\frac{L^{(r)}(E,1)}{r!}$ of the L-function of E, which is, according to the first part of the conjecture, the first nonzero coefficient. To (partially) understand it, we need to define two quantities. Note preliminarly that if we take the tensor $E(\mathbb{Q}) \otimes \mathbb{R} \simeq \mathbb{R}^r$, the torsion maps to 0, and we get an injection $E(\mathbb{Q})/E(\mathbb{Q})_{\text{Tors}} \simeq E(\mathbb{Q})/E(\mathbb{Q})_{\text{Tors}} \otimes 1 \hookrightarrow \mathbb{R}^r$. Moreover, on this image of the Mordell-Weil group modulo torsion we still have our positive definite quadratic form $\hat{h}$, which we can extend by linearity to $\mathbb{Q}^r$ and by continuity to $\mathbb{R}^r$. It is a classical result that, in the case of the canonical height, this procedure preserves positive definiteness.

Definition 3.24.

- *The regulator R_E of E is the volume with respect to $\hat{h}$ of a fundamental region for the lattice $E(\mathbb{Q})/E(\mathbb{Q})_{Tors}$.*
- *The real period of E is $\Omega_E = \int_{E(\mathbb{R})} |\frac{dx}{2y}|$.*

Observe that $\Omega_{E^{(d)}} = d^{-\frac{1}{2}}\Omega_E$. The refined form of the BSD conjecture says:

Conjecture 3.25 (Refined BSD)**.**

$$\frac{L^{(r)}(E,1)}{r!} = R_E\Omega_E \cdot (\textit{other factors}),$$

where the other factors, if we let the twist parameter $|d| \longrightarrow \infty$, are bounded from below by a positive constant. Let us talk about a twist $E^{(d)}$ of rank r. Assuming this conjecture, we would have $R_{E_d} \ll \sqrt{d}L^{(r)}(E_d,1)$.

Notice that the regulator, by definition, allows us to count rational points with canonical height up to B: the right asymptotic as $B \longrightarrow \infty$ is $\frac{B^{\frac{r}{2}}}{\sqrt{R_E}}$, since $\hat{h}$ is a quadratic form. Therefore, if we assume that $L^{(r)}(E_d,1) \ll d^{\epsilon}$, which the authors in [47] claim follows from a standard conjecture in Analytic Number Theory, we get:

$$\frac{B^{\frac{r}{2}}}{d^{\frac{1}{4}+\epsilon}} \ll \frac{B^{\frac{r}{2}}}{\sqrt{R_{E_d}}} \ll \#\{\text{nontorsion points of height up to } B \text{ in } E_d(\mathbb{Q})\}. \tag{39}$$

Let us introduce the notation $a \sim A$ to mean $A \leq a \leq 2A$ (so we let a vary in a so-called *dyadic interval*). In order to upper bound the RHS, we proceed like in the proof of Theorem 3.15, but starting from the model $y^2 = x^3 + ad^2x + bd^3$ seen in Section 2.4, to reduce ourselves to counting coprime integer solutions $(x,y,z) \in \mathbb{Z} \times (\mathbb{Z}^+)^2$ to $y^2z = x^3 + ad^2xz^2 + bd^3z^3$. We will then see, letting d vary in a dyadic interval, how a twist with rank too large gives a contradiction.

Let $\bar{f}(X,Z) = X^3 + aXZ^2 + bZ^3$: proceeding as in the paragraph following equation (35), we find that $z = z_1^3$ with $z_1 \mid x$, that clearly $\bar{f}(x,dx) \equiv 0 \pmod{y^2}$, and that these conditions give a solution. Therefore, setting $|d| \sim D$, $|x| \sim T$, $z \sim U/D$, $V = \max(T,U)$, we expect $y \approx \sqrt{DV^3/U}$ if $x^3 + ad^2x + bd^3$ does not have too much cancellation. Then, letting $N_D(T,U)$ be the number of solutions of the prescribed form with the variables free to vary in the given ranges, we expect:

$$N_D(T,U) \ll \sum_{d\sim D}\ \sum_{y\sim\sqrt{\frac{DV^3}{U}}}\ \sum_{z_1\sim(U/D)^{\frac{1}{3}}}\ \sum_{\substack{x\sim T,\ z_1|x\\ \bar{f}(x,dz_1^3)\equiv 0 \pmod{y^2}}} 1.$$

It is known that the density of solutions in $\mathbb{Z}^2$ to the congruence in the last summation is approximately equal to $\frac{\sigma_f(y^2)}{y^2}$, where $\sigma_g(t) = \#\{\text{roots of } g \text{ modulo } t\}$, so we get:

$$N_D(T,U) \ll \sum_{d\sim D}\ \sum_{y\sim\sqrt{\frac{DV^3}{U}}} \frac{\sigma_f(y^2)}{y^2} \sum_{z_1\sim(U/D)^{\frac{1}{3}}} \frac{T}{z_1} \ll TD\left(\frac{U}{DV^3}\right) \sum_{y\sim\sqrt{\frac{DV^3}{U}}} \sigma_f(y^2).$$

We can heuristically evaluate the last sum as follows: by the Chinese Remainder Theorem, if $y = \prod_{p|y} p^{\alpha_p}$, we have $\sigma_f(y^2) = \prod_{p|y}\sigma_f(p^{2\alpha_p})$. Now, by Hensel's Lemma, for any $p \nmid \Delta_E$ we have $\sigma_f(p^{2\alpha_p}) = \sigma_f(p)$. But $\sigma_f \leq 3$, so we have:

$$\sum_{y \sim \sqrt{\frac{DV^3}{U}}} \sigma_f(y^2) \ll \sum_{y \sim \sqrt{\frac{DV^3}{U}}} 3^{\omega(y)}. \tag{40}$$

The Erdős-Kac Theorem states that $Z(n) = \frac{\omega(n) - \log\log n}{\sqrt{\log\log n}}$, as a random variable on the space $\{n \in \mathbb{Z}^+ : n \leq Y\}$, follows a standard normal distribution $N(0,1)$. Using the claim in our dyadic interval we get:

$$\sum_{y \sim \sqrt{\frac{DV^3}{U}}} \sigma_f(y^2) \ll \frac{DV^3}{U} \left(\log\!\left(\frac{DV^3}{U} \right) \right)^{\log 3 + \frac{\log^2 3}{2}}$$

by the expectation of the exponential of a normal distribution. Therefore, setting $c = \log 3 + \frac{\log^2 3}{2}$, we have:

$$N_D(T,U) \ll TD \left(\sqrt{\frac{U}{DV^3}} \right) \left(\log\left(\frac{DV^3}{U} \right) \right)^c. \tag{41}$$

Now, we want to compare formulas for the number of points of height bounded by B in twists with $|d| \sim D$, so, neglecting the difference between canonical and naive height in the spirit of Proposition 2.6, we have to assume $|x|, z \leq e^B$ (and therefore set $T = U/D = e^B$ in (41)) and then sum dyadically (which gains a log factor). This gives a bound of $O(\sqrt{D} B^{c+1})$. In order to compare this with (39), we need to decide how large we can take B as a function of the approximate value D on the twist parameters. Similar arguments for heuristics on solutions to Pell equations work for (logarithmic) height up to $\sqrt{D}$, so Granville proposes the same value. Combining (39) and (41) we then get:

$$\frac{D^{\frac{r}{4}}}{D^{\frac{1}{4}+\epsilon}} \ll D^{\frac{1}{2}(c+1)},$$

which gives $r \leq 3 + 2(c+1) \sim 8.4$, as desired.

We remark that, in [47], the authors present a variant of the above argument that distinguishes three cases depending on the value of the average number of roots modulo primes $\eta_f \in \{1,2,3\}$ of f. They obtain a bound of 9 for curves with full rational 2-torsion and of 7 for curves such that $E(\mathbb{Q})[2] \simeq \mathbb{Z}/2\mathbb{Z}$, and this is due to the fact that their exponent in (41) for curves with full 2-torsion is not c but 2 (this does not appear to be justified). As they note, every elliptic curve with full rational 2-torsion is isogenous to, and therefore has the same rank of, one with $E(\mathbb{Q})[2] \simeq \mathbb{Z}/2\mathbb{Z}$, so their heuristic leads to an inconsistency. It is possible that, controlling the behavior of (40) for this last family via our Erdős-Kac argument, a bound bigger than 8 can be proven also in this last case, resolving the inconsistency.

4 Modular forms and modularity of elliptic curves

In this section we will develop part of the theory of modular forms, which are central objects in modern Number Theory. Modular forms, despite the complex-analytical nature of their definition, turn out to be related to the arithmetic of elliptic curves and even of the integers, encoding information about combinatorial identities and divisor functions, among other things. Since this theory is so vast, even focusing just on what serves our purposes, we refer to [11], [21] and [49] for a more complete introduction and for the proofs we will omit.

4.1 Congruence subgroups of the modular group

Let $\mathscr{H} := \{z \in \mathbb{C} : \operatorname{Im}(z) > 0\}$ and let $\Gamma(1) = \mathrm{SL}_2(\mathbb{Z})$. $\Gamma(1)$ is called the *modular group*. For any $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(1)$, we define: $\gamma z = \frac{az+b}{cz+d}$ (in what follows, when a matrix in $\Gamma(1)$ appears in an equation, a, b, c, d will denote its entries even if this is not specified, as long as this does not create confusion). The verification that $(\gamma_1\gamma_2)z = \gamma_1(\gamma_2 z)$, is left to the reader. Notice how, for such γ,

$$\operatorname{Im}(\gamma z) = |cz+d|^{-2} \operatorname{Im} z, \tag{42}$$

which tells us that what we just defined is an action of $\Gamma(1)$ on $\mathscr{H}$. Observe also that, given $\gamma \in \Gamma(1)$, the condition on the determinant implies that both a and d are coprime with b and c.

Definition 4.1. *Let $k \in \mathbb{Z}$. A modular form of weight k for $\Gamma(1)$ is a holomorphic function $f : \mathscr{H} \longrightarrow \mathbb{C}$ that stays bounded as $\operatorname{Im}(z) \to \infty$ and such that:*

$$f(\gamma z) = (cz+d)^k f(z) \ \forall \gamma \in \Gamma(1). \tag{43}$$

Denote $\mathrm{M}_k(1)$ the set of weight k modular forms for $\Gamma(1)$.

It is useful to introduce, for $\gamma \in \Gamma(1)$ as above, the modular *weight k operator* $[\gamma]_k$ on meromorphic functions, defined as:

$$f[\gamma]_k = (cz+d)^{-k} f(\gamma z), \tag{44}$$

which induces an action of $\Gamma(1)$ on $\mathbb{C}(\mathscr{H})$, for a simple computation shows:

Lemma 4.2. *Given $\gamma, \gamma' \in \Gamma(1)$, $[\gamma]_k[\gamma']_k = [\gamma\gamma']_k$.*

In this case, modular forms of weight k are exactly the fixed points of $[\gamma]_k$ on the subspace of holomorphic functions. Observe that a modular form is periodic, since plugging in $\gamma = T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ in (43) we get:

$$f(z+1) = f(z).$$

By the required holomorphy, this implies that a modular form has a Fourier expansion $f(z) = a_0 + a_1 q + a_2 q^2 + \dots$ with $q = e^{2\pi i z}$, where the fact that only terms with nonnegative index occur is equivalent to the boundedness condition.
This is due to the fact that $z \mapsto q$ is a holomorphic map from $\mathscr{H}$ to $\{z \in \mathbb{C} : 0 < |z| < 1\} = D \setminus \{0\}$ of period 1, so f induces a holomorphic map $\tilde{f} : D \to \mathbb{C}$ defined by $\tilde{f}(q) = f(z)$ and $\tilde{f}(0) = \lim_{\operatorname{Im} z \to \infty} f(z)$, which has the aforementioned Fourier expansion. We can also define the space of *modular functions* $\mathrm{M}_k^{\mathrm{mer}}(1)$ by only requiring the function f fixed by the modular operators to be meromorphic on $\mathscr{H}$ and have an expansion with only finitely many negative powers of q at infinity.

Definition 4.3. *A modular form f is called a cusp form if $a_0 = 0$. Denote $S_k(1)$ the set of weight k cusp forms for $\Gamma(1)$.*

The reason for this terminology is that the point at infinity in $\mathscr{H}$, or $0 \in D$, is called *cusp.* In general, a cusp is a point we need to add to compactify some quotients of $\mathscr{H}$: indeed, we can put a topology on $\mathscr{H}^* = \mathscr{H} \cup \mathbb{Q} \cup \{\infty\}$ that makes it compact (a fundamental system of neighbourhoods for $q \in \mathbb{Q}$ can be taken as $\{\{q\} \cup \{z \in \mathscr{H} : |z - q - \frac{i}{n}| < \frac{1}{n}\}\}_{n \geq 1}$, while for ∞ we can take $\{\operatorname{Im} z > n\}_{n \geq 1}$). The modular group naturally acts on $\mathscr{H}^*$ (set $\gamma\infty = \frac{a}{c}$) and, since one easily verifies that $\mathbb{Q} \cup \{\infty\}$ is an equivalence class under this action, we get:

Proposition 4.4. *With the above topology, let $Y(1) = \Gamma(1)\backslash\mathscr{H}$. Then $X(1) = \Gamma(1)\backslash\mathscr{H}^* = Y(1) \cup \{\infty\}$ is its compactification. $X(1)$ is called modular curve of level 1.*

With the quotient topology, $X(1)$ turns out to be a Riemann surface. It can be “seen” more explicitly by finding a fundamental domain for the action of the modular group: it is a classical result that $F = \{z \in \mathscr{H} : |\operatorname{Re} z| \leq \frac{1}{2}, |z| \geq 1\}$ is such a fundamental domain, in the sense that it is a closed subset intersecting each orbit and such that if it contains two elements of the same orbit then they both lie on its boundary. The fact that F is a fundamental domain can be used to prove the following:

Lemma 4.5. *The modular group is generated by the matrices $S = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.*

We outline a proof, without going in too much detail: take $\gamma \in \Gamma(1)$ and $z_1 \in F^\circ$. Let $z = \gamma z_1$. We want to prove that there exists $\beta \in \langle S, T \rangle$ such that $\beta z \in F$: by definition, this would imply $\gamma = \beta^{-1}$.

Suppose WLOG that z is in the region $\{|\operatorname{Re}\{\cdot\}| \leq \frac{1}{2}\}$ after moving it with a suitable power of T. If $|z| \geq 1$ we are done, otherwise observe that, by (42), $\operatorname{Im} Sz > \operatorname{Im} z$. Therefore, if we can prove that for any $z \in \mathscr{H}$ there are finitely many $\gamma \in \Gamma(1)$ (up to translations $\{T^n\}$) such that $\operatorname{Im} \gamma z < 1$, we would be done by repeating this process a finite number of times.

The required claim clearly follows from (42) and the two following facts:

1. given $\gamma_0 = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix} \in \Gamma(1)$, we have:
$$\{\gamma \in \Gamma(1) : (c, d) = (c_0, d_0)\} = \langle T \rangle \gamma_0;$$
2. for any $z \in \mathscr{H}$, the set $\{|cz + d|^2 : (c, d) \in \mathbb{Z}\}$ is discrete.

The first fact is an easy consequence of the Chinese Remainder Theorem, while the second is true because that set is the image of a discrete set ($\mathbb{Z}^2$) under a positive definite quadratic form.

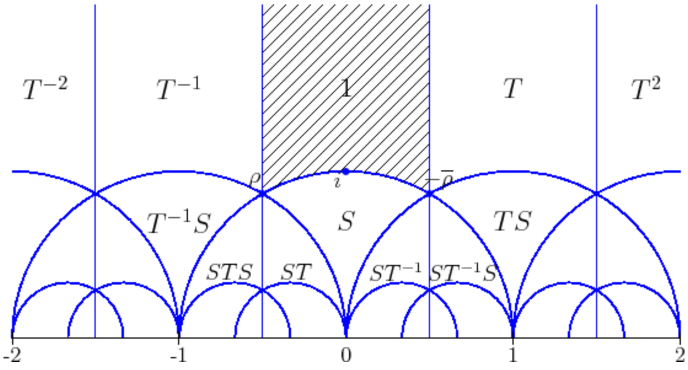

Figure 2: The standard fundamental domain for the action of the modular group

As the definition (43) suggests, we can define modular forms for other groups: let
$$\Gamma(N) = \{\gamma \in \mathrm{SL}_2(\mathbb{Z}) : \ a \equiv d \equiv 1 \pmod{N}, \ b \equiv c \equiv 0 \pmod{N}\},$$
in accordance with our definition of $\Gamma(1)$.

Definition 4.6. *A congruence subgroup is a subgroup $\Gamma \subset \Gamma(1)$ such that $\Gamma(N) \subset \Gamma$ for some N. In this case, we say that Γ has level N.*

Example 4.7. *Set*

$$\Gamma_0(N) = \{\gamma \in SL_2(\mathbb{Z}) : \; c \equiv 0 \pmod N\}$$

and

$$\Gamma_1(N) = \{\gamma \in SL_2(\mathbb{Z}) : \; a \equiv d \equiv 1 \pmod N, \; c \equiv 0 \pmod N\}.$$

These clearly are congruence subgroups of level N with $\Gamma_1(N) \subset \Gamma_0(N)$, and $\Gamma_1(N)$ is the kernel of the map $\Gamma_0(N) \longrightarrow \mathbb{Z}/N\mathbb{Z}, \; \gamma \mapsto d$, so it is normal in $\Gamma_0(N)$.

Notice that $\Gamma(N)$ is normal in $\Gamma(1)$ since it is the kernel of the projection homomorphism $\Gamma(1) \longrightarrow \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$. This also enables us to compute the index of $\Gamma(N)$ in the modular group:

$$[\Gamma(1) : \Gamma(N)] = N^3 \prod_{p|N}(1 - p^{-2}), \tag{45}$$

the computation of which can be found in [11, I,2]. A generic congruence subgroup is not normal, but its definition implies:

Lemma 4.8. *Let Γ be a congruence subgroup of level N and $\alpha \in \Gamma(1)$. Then $\Gamma' = \alpha^{-1}\Gamma\alpha$ is also a congruence subgroup of level N.*

Moreover, the kernel interpretation also tells us that the index of $\Gamma(N)$, and therefore of a congruence subgroup, in $\Gamma(1)$ is finite. We also clearly obtain a fundamental domain for the action of Γ as $\bigcup_{i=1}^{M} \gamma_i F$, where M is the index of Γ in $\Gamma(1)$ and the γ_i are coset representatives.

Given a congruence subgroup Γ, if we want to compactify the quotient $\Gamma\backslash\mathscr{H}$ we cannot just add the point at infinity, because the orbit $\{\Gamma\infty\}$ can intersect $\mathbb{Q}$ in a proper subset. Defining as above:

Definition 4.9. *The modular curve relative to Γ is $X(\Gamma) = \Gamma\backslash\mathscr{H}^*$,*

we again see that $X(\Gamma)$ is the compactification of $Y(\Gamma) = \Gamma\backslash\mathscr{H}$. Its set of cusps $X(\Gamma) \setminus Y(\Gamma)$ is finite: since every rational number is in the form $\alpha(\infty)$, $\alpha \in \Gamma(1)$, their number is at most that of right cosets $\Gamma\alpha$ in $\Gamma(1)$, which is equal to the index. For example, setting $\epsilon_\infty(\Gamma) = \{\text{cusps in } X(\Gamma)\}$, for $\Gamma = \Gamma(N)$ we have the following result, the proof of which can be found in [11, II,4]:

Lemma 4.10. *The action of $\Gamma(1)$ on $\epsilon_\infty(N) = \epsilon_\infty(\Gamma(N))$ comes from its action on $S_N = \{v = (a, b) \in (\mathbb{Z}/N\mathbb{Z})^2 : (a, b, N) = 1\}$ as row vectors: there is a surjection*

$U : S_N \twoheadrightarrow \epsilon_\infty(N)$ such that for any $v, v' \in S_N$ and $\gamma \in \Gamma(1)$, we have $v\gamma = v' \iff \gamma U(v) = U(v')$. Moreover, we have:

$$\#\epsilon_\infty(N) = \begin{cases} 1 & \textit{if } N = 1, \\ 3 & \textit{if } N = 2, \\ \frac{N^2}{2}\prod_{p|N}(1-\frac{1}{p^2}) & \textit{else.} \end{cases} \tag{46}$$

If $\Gamma \subset \Gamma'$ are congruence subgroups, we clearly have a projection map $X(\Gamma) \twoheadrightarrow X(\Gamma')$. This map fails to be a covering just at those points which are the images of the points in $\mathscr{H}^*$ such that $[\mathrm{Stab}_{\Gamma'}(z) : \mathrm{Stab}_\Gamma(z)] < [\Gamma' : \Gamma]$. Using the fact that the only points in $\mathscr{H}^*$ having nontrivial stabilizer, which are called *elliptic points,* for $\Gamma = \Gamma(1)$ are i and ζ_3 (where ζ_3 is a primitive cube root of unity), which easily follows from Lemma 4.5, it can be shown that he ramification points of the branched covering of modular curves as above form a finite set.

Let us look at the case $\Gamma = \Gamma(N),\ \Gamma' = \Gamma(M),\ M \mid N,\ N \neq M$. Let $N = rM$. The covering ramifies at the cusps: we may just verify this at $[\infty]_\Gamma$, since the cusps are permuted by $\Gamma(1)$; clearly a coset of representatives for $\mathrm{Stab}_\Gamma(z)$ in $\mathrm{Stab}_{\Gamma'}(z)$ is given by:

$$\{\begin{pmatrix} 1 & kM \\ 0 & 1 \end{pmatrix},\ 1 \leq k < r\},$$

so this index is r, which is smaller than $[\Gamma' : \Gamma] = [\Gamma(1) : \Gamma][\Gamma(1) : \Gamma']^{-1}$, thanks to (45). The interesting fact is that the cusps are the only ramification points, for we do not have elliptic points: indeed, by the above discussion, any elliptic point would be in the $\Gamma(1)$-orbit of i or ζ_3. Let us treat the first case: the other one is a similar computation. Since $\Gamma(1) = \langle S, T\rangle$ and S fixes i, an elliptic point for $\Gamma(N)$ is of the form $z = T^k i,\ k \in \mathbb{Z}$. But then we have $\gamma T^k i = T^k i \implies \gamma \in \mathrm{Stab}_{\Gamma(1)}(T^k i) = \{S, I, -S, -I\}$. Since the only one of those matrices that is in $\Gamma(N)$ is I, we are done.

We can finally define modular forms for a congruence subgroup:

Definition 4.11. *Let Γ be a congruence subgroup and k be an integer. A holomorphic function $f : \mathscr{H} \to \mathbb{C}$ is a modular form of weight k for Γ if*

$$f[\gamma]_k = f$$

for all $\gamma \in \Gamma$ and $f[\alpha]_k$ is bounded as $\operatorname{Im} z \to \infty$ for all $\alpha \in \Gamma(1)$. Denote $\mathrm{M}_k(\Gamma)$ the set of weight k modular forms for Γ.

The last condition amounts to saying that f is holomorphic at the cusps of the modular curve relative to Γ: indeed, observe that, if Γ has level N then $\begin{pmatrix} 1 & N \\ 0 & 1 \end{pmatrix} \in \Gamma$, so $f \in \mathrm{M}_k(\Gamma)$ has period N, so it induces $\tilde{f}(q_N) : D \longrightarrow \mathbb{C}$ with $q_N = e^{\frac{2\pi i z}{N}}$. But given $\alpha \in \Gamma(1)$, we have that $f[\alpha]_k$ is weight k modular with respect to $\alpha^{-1}\Gamma\alpha$, a congruence subgroup of level N by Lemma 4.8, because $f[\alpha]_k[\alpha^{-1}\gamma\alpha]_k = f[\gamma\alpha]_k = f[\gamma]_k[\alpha]_k = f[\alpha]_k$. Hence it has a Fourier expansion $\sum_{n \geq n_0} a_n q_N^n$, and the last condition in the definition is equivalent to asking that all functions $f[\alpha]_k$ expand as $a_0 + a_1 q_N + a_2 q_N^2 + ...$

Definition 4.12. *In the above notation, if $a_0 = 0 \; \forall \alpha \in \Gamma(1)$, we say that f is a weight k cusp form for Γ, and we denote their set as $S_k(\Gamma)$.*

Observe that, if k is odd, using (43) with $\gamma = -I$ gives $f(z) = -f(z)$, which means that the only modular form of odd weight for $\Gamma(1)$ is the zero function. In general, modular forms of a given weight for a congruence subgroup form a finite dimensional $\mathbb{C}$-vector space. In the case of $\Gamma(1)$, we have the dimension formulas:

Proposition 4.13.

$$\dim \mathrm{M}_k(\Gamma(1) = \begin{cases} 0 & \text{if } k < 0, \\ \lfloor \frac{k}{12} \rfloor & \text{if } k \geq 0 \text{ and } k \equiv 2 \pmod{12} \\ \lfloor \frac{k}{12} \rfloor + 1 & \text{else.} \end{cases} \tag{47}$$

As we will see, the modular forms that relate to elliptic curves have weight 2, but (47) implies that the only such function for $\Gamma(1)$ is identically 0, so we will have to look at modular forms with respect to congruence subgroups.

Finally, let us observe that $\mathrm{M}(\Gamma) = \oplus_{k \in \mathbb{Z}} \mathrm{M}_k(\Gamma)$ is naturally a graded $\mathbb{C}$-algebra with standard multiplication, since $(fg)[\gamma]_{k_1+k_2} = f(\gamma z) g(\gamma z)(cz+d)^{-k_1}(cz+d)^{-k_2} = f[\gamma]_{k_1} g[\gamma]_{k_2}$, and that if $\Gamma \subset \Gamma'$ we have $\mathrm{M}_k(\Gamma') \subset \mathrm{M}_k(\Gamma)$ and the same for cusp forms.

4.2 Constructing modular forms

The relation (43) characterizes modular forms as fixed points of the action of a group. When we have a finite group G acting on a vector space V, there is a general way to construct a fixed point of the action: just take some $v \in V$ and consider $u = \sum_{g \in G} gv$. Clearly $Gu = u$. Unfortunately, in our case congruence subgroups are not finite, but infinite sums still have a chance of making sense, since the vector space is that of holomorphic functions on $\mathscr{H}$.

We now construct so-called *Eisenstein series* using this technique, focusing on the case $\Gamma = \Gamma(1)$, but the same can be done for general congruence subgroups: we shall later sketch the construction for the case $\Gamma = \Gamma(N)$, which will be useful in Section 5.3.

Let $G = \Gamma(1)/\langle -I, T\rangle$, $\pi : \Gamma(1) \longrightarrow G$ and let k be even. Define:

$$E_k(z) = \sum_{g \in G} 1[g]_k(z) = \frac{1}{2} \sum_{\substack{c,d \in \mathbb{Z} \\ (c,d)=1}} (cz+d)^{-k},$$

where both the definition of the action $[\pi(\gamma)]_k := [\gamma]_k$ and the last equality make sense because of 1) in the proof of Lemma 4.5. Let also:

$$G_k(z) = \sum_{m \geq 1} m^{-k} E_k(z) = \sum_{(c,d) \in \mathbb{Z}^2 \setminus \{(0,0)\}} (cz+d)^{-k}. \tag{48}$$

E_k and G_k are called Eisenstein series of weight k. Notice that $G_k = 2\zeta(k)E_k$, and that we can take the same definitions as sums over $\mathbb{Z}^2$ for odd k, which would lead to term-by-term cancellations and consequently, if we can prove uniform convergence of the series on compact sets, their being identically 0. Uniform convergence on compact sets of (48) for integer $k \geq 3$ (and, more generally, for $\operatorname{Re} k > 2$) classically follows from the infinite sum for the Riemann Zeta Function being well defined for $\operatorname{Re} s > 1$:

Proposition 4.14. *The series in* (48) *converges totally, and therefore uniformly, on compact subsets $K \subset \mathcal{H}$ for $k \geq 3$. Hence, the Eisenstein series of weight $k \geq 3$ are holomorphic functions on $\mathcal{H}$.*

By our construction, this immediately implies that they are modular forms of weight k. Observe that G_k evaluates to $2\zeta(k)$ at the cusp ∞, because the part of the sum with $c \neq 0$ vanishes since we can interchange sum and limit by the uniform convergence, so the Eisenstein series are not cusp forms. Those of odd weight are identically 0 in accordance with what we observed above. It can be shown, by standard techniques to compute Fourier coefficients (see [49, 2.2]), that:

$$E_4(z) = 1 + 240 \sum_{n \geq 1} \sigma_3(n) q^n, \ E_6(z) = 1 - 504 \sum_{n \geq 1} \sigma_5(n) q^n, \tag{49}$$

where $\sigma_k(n) = \sum_{d|n} d^k$.

Observe that, by (47), $\mathrm{M}_k(1)$ is generated by E_k for $k = 4, 6, 8, 10$. Let us consider the case $k = 12 :$ E_4^3 and E_6^2 are nonzero elements of $\mathrm{M}_{12}(1)$, and they are linearly independent: otherwise, there exists $c \neq 0$ such that $\frac{E_6^3}{E_4^3} = cE_6$ is a nonzero modular form of weight 6, so $\frac{E_6}{E_4}$ is holomorphic, and therefore a nonzero modular form of weight 2, which is absurd. Therefore, $\{E_4^3, E_6^2\}$ is a basis of $\mathrm{M}_{12}(1)$. What if we are interested in a basis one of whose elements is a cusp form? We can simply take it to be a scalar multiple of $E_4^3 - E_6^2$: define the *modular discriminant* as:

$$\Delta(z) = \frac{E_4^3 - E_6^2}{1728}. \tag{50}$$

Since the Eisenstein series are holomorphic, Δ has no poles on $\mathscr{H}^*$. It has a zero at ∞, since the E_i evaluate to 1 there, which is easy by (49). This allows us to prove that, for $k > 2$, $\dim S_k(1) = \dim \mathrm{M}_k(1) - 1$: this is true with the LHS replaced by 0 for $k < 12$ and with $\dim \mathrm{M}_{k-12}(1)$ otherwise, by (47). But $f \mapsto f\Delta$ defines an isomorphism between $\mathrm{M}_{k-12}(1)$ and $S_k(1)$ by what we have just seen, and for $k < 12$ there are no cusp forms because the spaces are generated by an Eisenstein series.

Now let us examine the case of $\Gamma(N),\ N > 1$: we will construct a modular form E_k^s attached to each cusp $s \in \epsilon_\infty(N)$, with the property that, if $s_1, ..., s_m$ are the cusps, with m as in Lemma 4.10, then $E_k^{s_i}(s_j) \neq 0 \iff i = j$. Let us remark, even if we shall not need this fact, that one can use the Riemann-Roch Theorem to prove that $\dim \mathrm{M}_k(N) = \dim S_k(N) + \#\epsilon_\infty(N)$ (which is false for a generic congruence subgroup); therefore, as in the level 1 case, Eisenstein series together with $S_k(N)$ generate $\mathrm{M}_k(N)$.

We can repeat the above invariant vector construction, this time letting $G = \Gamma(N)/\langle T^N \rangle$, to define:

$$\begin{aligned} E_k^\infty(z) = \sum_{\gamma \in G} 1[\gamma]_k = \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1 \\ (c,d) \equiv (0,1) \pmod N}} (cz+d)^{-k} = \\ = 1 + \sum_{c \neq 0} \sum_{\substack{d \in \mathbb{Z} \\ \gcd(c,d)=1 \\ (c,d) \equiv (0,1) \pmod N}} (cz+d)^{-k}, \end{aligned}$$

(with the little abuse of notation of multiplying cosets and matrices) from which, as we did in the level 1 case, we see that this function is nonzero at $\infty \in X(N)$. Now, by Lemma 4.10 (and keeping its notation), letting WLOG $U : (0,1) \mapsto \infty$ (which implies $U((0, N-1)) = \infty$ by putting $\gamma = -I$ in the lemma) we can define, for any $v \in S_N,\ v = (0,1)\beta_v,\ \beta_v \in \Gamma(1)$:

$$E_k^v(z) = \sum_{\gamma \in G} 1[\beta_v \gamma]_k = \sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1 \\ (c,d) \equiv v \pmod N}} (cz+d)^{-k} \tag{51}$$

and $E_k^s = E_k^v$ for any $v \in U^{-1}(s)$, which is well defined: it suffices to show it for $s = \infty$, and if $v, v' \in U^{-1}(\infty)$, by Lemma 4.10 $\beta_{v'} \in \mathrm{Stab}_{\Gamma(1)}(\infty)\beta_v$, so we have (omitting the subscript k):

$$\begin{aligned} \sum_{\gamma \in G} 1[\beta_{v'}\gamma] = \sum_{\gamma \in G} 1[\beta_v T^N \gamma] = \sum_{\gamma \in G} 1[\beta_v T^N \gamma T^{-N} T^N] = \\ = \sum_{\gamma \in G} 1[\beta_v \gamma T^N] = \sum_{\gamma \in G} 1[\beta_v \gamma]. \end{aligned}$$

This automatically gives:

$$E_k^{\gamma(s)} = E_k^s[\gamma]_k. \tag{52}$$

Since

$$E_k^v(\infty) = \sum_{\substack{d=\pm 1\\ (0,d)\equiv v \pmod N}} d^{-k} + \lim_{\operatorname{Im} z\to\infty} \sum_{\substack{c\neq 0\\ \gcd(c,d)=1\\ (c,d)\equiv v \pmod N}} (cz+d)^{-k}$$

is 0 if $v \notin \{(0,1),(0,-1)\}$, which correspond to the cusp ∞, (52) proves the desired (non)vanishing properties, and in particular that this Eisenstein series are linearly independent.

Recall the theta function $\theta(t)$ that we introduced in (14), which now we consider with the change of variables $t = -2iz$. This immediately tells us that θ has period 1, so it is invariant under the modular action of T. (15) tells us that we have a functional equation:

$$\theta\left(-\frac{1}{4z}\right) = \sqrt{-2iz}\theta(z).$$

This theta function is famously linked to the problem of representing integers as sum of squares, because

$$\theta(z)^k = \sum_{n\in\mathbb{Z}} r(n,k)q^{n^2},$$

where $r(n,k) = \#\{m_1^2 + ... + m_k^2 = n\}$.

The functional equation implies that $\theta(\frac{z}{4z+1}) = \sqrt{4z+1}\theta(z)$, so for $F(z) = \theta(z)^4$ we have $F(\frac{z}{4z+1}) = (4z+1)^2F(z)$. Therefore, F satisfies (43) for $\gamma = \begin{pmatrix} 1 & 0\\ 4 & 1\end{pmatrix}$, but also for $\gamma = T$ and $\gamma = -I$. It turns out that the subgroup of $\Gamma(1)$ generated by this matrices is $\Gamma_0(4)$. This hints at a strategy to prove Lagrange's famous result that each natural number is the sum of 4 integer squares: if we knew a basis $f_1, ..., f_d$ of $\mathrm{M}_2(\Gamma_0(4))$, we could express F as a linear combination of the f_i after solving a $d \times d$ linear system for the first d Fourier coefficients, and then verify that all coefficients appearing in the sum defining F are nonzero. Observe also that θ^4 is a nonzero weight 2 modular form, albeit not for $\Gamma(1)$ ($\mathrm{M}_2(\Gamma_0(4))$ turns out to have dimension 2, see [49, 3.1]).

4.3 Hecke operators

Given a group G acting on a vector space V (denote the action $v[x]$) and a finite index subgroup H, we have the usual right and left actions of H on G that give partitions of G in right and left cosets Hx and xH, which can be interpreted as elements of the quotient sets $H\backslash G$ and G/H. But we can also consider *double cosets*, that is equivalence classes of the form HxH, which form the group $H\backslash G/H$ with the obvious multiplication, where one copy of H acts on the right and one on the left. Observe that, given $x \in G$, HxH is the union of right cosets of the form $Hxh,\ h \in H$. Under which condition can we say that this union can be taken as finite? Let $H' = x^{-1}Hx$ and $K = H \cap H'$. Then it turns out that we just need the index $[H : K]$ to be finite, for we have:

Lemma 4.15. *In the above notation and assumption, if we have $[H : K] = d$ with $H = \bigsqcup_{i=1}^{d} Ky_i,\ y_i \in H$, then $HxH = \bigsqcup_{i=1}^{d} Hxy_i$.*

Proof. given $h \in H$ we can write $h = ky_i,\ 1 \leq i \leq d$. But $k = x^{-1}h_1x$ for some $h_1 \in H$ by definition, hence an element of the form $h'xh$ with $h, h' \in H$ can be written as $h'xx^{-1}h_1xy_i = h'h_1xy_i \in Hxy_i$. We must show that the right cosets Hxy_i are disjoint: otherwise, there exists $1 \leq i, j \leq d,\ i \neq j$, and $h_1, h_2 \in H$ such that $h_1xy_i = h_2xy_j$, so $y_iy_j^{-1} = x^{-1}h_1^{-1}h_2x \in H'$, so $y_iy_j^{-1} \in K$ and therefore $y_i \in Ky_j$, contradicting the hypotheses. □

Keeping the above notation, under the hypotheses of Lemma 4.15 we can define the so-called *double coset operators*:

Definition 4.16. *The double coset operator associated with the double coset HxH is $[HxH] : V^H \longrightarrow V^H$, defined as $v[HxH] = \sum_{i=1}^{d} v[y_i]$.*

Proposition 4.17. *The operator $[HxH]$ on V^H is well defined: it does not depend on the choice of representatives of $K\backslash H$ nor on the choice of representative x for the double cosets HxH.*

Proof. we have $v[HxH][h] = \sum_{i=1}^{d} v[xy_ih] = v[HxHh] = v[HxH]$ by Lemma 4.15. Suppose that we use representatives $y_i' = k_iy_i$. Since $k_i = x^{-1}hx$ for some $h \in H$ by definition, we have $v[xk_iy_i] = v[hxy_i] = v[xy_i]$ since $v \in V^H$. Finally, if we choose a different representative x' for the double coset HxH then we will have $x' = hxh_1$, so by the first argument above and the fact that $v \in V^H$ we are done. □

This theory perfectly works for two subgroups H_1, H_2 acting right and left respectively on G, giving a double coset operator $H_1xH_2 : V^{H_1} \longrightarrow V^{H_2}$, under suitable finiteness hypotheses, but since we are interested in applying it with $H_1 = H_2 = \Gamma_1(N)$, this exposition is sufficient for us. For a complete introduction to the topic of abstract Hecke operators and Hecke algebras, see [11, VII,9].

Observe that, if we could extend the action of the modular group to $\mathrm{GL}_2^+(\mathbb{Q}) = \{\alpha \in \mathrm{GL}_2(\mathbb{Q}) : \det\alpha > 0\}$, we could apply our theory to $G = \mathrm{GL}_2^+(\mathbb{Q})$ and $H = \Gamma$

a congruence subgroup, because, even if $\Gamma(N)$ is not normal in $\mathrm{GL}_2^+(\mathbb{Q})$, Lemma 4.8 easily generalizes to $\alpha \in G$ by dropping the "level N" requirement, so H' as defined above is again a congruence subgroup (and the intersection of two congruence subgroups, say of levels N_1 and N_2, is clearly a congruence subgroup of level N_1N_2), which gives the required finiteness of indexes. For the rest of this section, let $\Gamma = \Gamma_1(N)$.

Before working with the practical version of these abstractly defined operators on spaces on modular forms, we need to extend the definition of the action (44) to $\mathrm{GL}_2^+(\mathbb{Q})$: for $\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{Q})$, set:

$$f[\alpha]_k(z) = (\det \alpha)^{\frac{k}{2}}(cz+d)^{-k} f(\alpha z),$$

where the action of G on $\mathscr{H}$ is defined as for $\Gamma(1)$ and is well defined by the positivity of the determinant. It is a routine exercise to verify that for $\alpha \in \mathrm{GL}_2^+(\mathbb{Q})$, the operator $[\alpha]_k$ induces an isomorphism $\mathrm{M}_k(\Gamma) \simeq \mathrm{M}_k(\alpha^{-1}\Gamma\alpha)$, $S_k(\Gamma) \simeq S_k(\alpha^{-1}\Gamma\alpha)$, since regularity at the cusps is preserved thanks to the fact that each coset of $\Gamma(1)$ in $\mathrm{GL}_2^+(\mathbb{Q})$ has an upper triangular representative.

For a prime p, let $\alpha_p = \begin{pmatrix} 1 & 0 \\ 0 & p \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{Q})$.

Definition 4.18. *Let p be a prime. The p-th Hecke operator on $\mathrm{M}_k(\Gamma)$ is defined as:*

$$T_p f = p^{\frac{k}{2}-1} \sum f[\Gamma\alpha_p\Gamma]. \tag{53}$$

Our goal is to show that these operators commute, preserve cusp forms (we need to prove that T_pf vanishes at the cusps if f does), and even more, that they preserve some important subspaces of cusp forms that we are going to introduce. In order to do so, we need to define another operator of the Hecke family.

Definition 4.19. *For a Dirichlet character $\chi \pmod N$, let $\mathrm{M}_k(N,\chi)$ (resp. $S_k(N,\chi)$) be the subspace of $\mathrm{M}_k(\Gamma)$ (resp. $S_k(\Gamma)$) such that:*

$$f[\gamma]_k = \chi(d) f \ \forall \gamma \in \Gamma_0(N),$$

where, as established at the beginning of the section, d is the bottom-right entry of γ.

Recall that Γ is normal in $\Gamma_0(N)$, with $\Gamma_0(N)/\Gamma \simeq (\mathbb{Z}/N\mathbb{Z})^*$. An easy computation shows that for any $\gamma_0 = \begin{pmatrix} a_0 & b_0 \\ c_0 & d_0 \end{pmatrix} \in \Gamma_0(N)$ we have $\gamma_0\Gamma = \{\gamma \in \Gamma_0(N) : d \equiv d_0 \pmod N\}$, so acting on $\mathrm{M}_k(\Gamma)$ with two matrices $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\gamma' = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$ in

$\Gamma_0(N)$ with $d \equiv d' \pmod N$ is the same, because we have $\gamma' = \gamma\gamma_1$ for some $\gamma_1 \in \Gamma$ and $f[\gamma]_k[\gamma_1]_k = f[\gamma\gamma_1\gamma^{-1}]_k[\gamma]_k = f[\gamma]_k$ by normality. Therefore, this action induces an action of $(\mathbb{Z}/N\mathbb{Z})^*$ on $S_k(\Gamma)$. This means that, if we define the *diamond operators*

$$\langle d\rangle : S_k(\Gamma) \longrightarrow S_k(\Gamma), \quad f \mapsto f[\gamma_0]_k \tag{54}$$

for any $\gamma_0 \in \Gamma_0(N)$ as above with $d_0 \equiv d \pmod N$, they are well defined for $(d, N) = 1$ (if f satisfies the hypotheses of Definition 4.12 then obviously $f[\gamma_0]_k$ does too), and they realize the aforementioned action of $(\mathbb{Z}/N\mathbb{Z})^*$. Hence, we obtain a faithful representation $(\mathbb{Z}/N\mathbb{Z})^* \longrightarrow \mathrm{GL}(S_k(\Gamma))$ since the diamond operators naturally form a group isomorphic to $(\mathbb{Z}/N\mathbb{Z})^*$ with the structure $\langle d_1\rangle\langle d_2\rangle = \langle d_1d_2\rangle$ which is induced by matrix multiplication (matrices $\gamma \in \Gamma_0(N)$ have $c \equiv 0 \pmod N$ and so the bottom right entry (mod N) of the product of two such matrices is the product of their bottom right entries (mod N)).

At this point, if we knew that the diamond operators were simultaneously diagonalizable, we would obtain the existence of a basis of common eigenvectors for Γ, which would consist of cusp forms f such that $\langle d\rangle f = \chi(d)f$ for some $\chi \in (\mathbb{Z}/N\mathbb{Z})^*$, by what we have just seen. A direct consequence of Maschke's Theorem is then that we obtain:

Proposition 4.20. *We have a decomposition*

$$S_k(\Gamma) = \oplus_{\chi\in(\mathbb{Z}/N\mathbb{Z})^*} S_k(N, \chi).$$

The simultaneous diagonalizability follows via classical linear algebra from the fact that diamond operators are the inverse of their adjoint with respect to a Hermitian, positive definite inner product on $S_k(\Gamma)$, the *Petersson inner product*. We refer the reader to [11, V,4] for a complete treatment of this fundamental construction, which we omit.

Proposition 4.21. *The weight k operators T_p map $S_k(\Gamma)$ to itself and commute with the weight k diamond operators $\langle d\rangle$. Therefore, T_p preserves $S_k(N, \chi)$ $\forall\chi \in (\mathbb{Z}/N\mathbb{Z})^*$ and we have $T_pT_q = T_qT_p$.*

Proof. We start by proving that Hecke and diamond operators formally commute. A tedious but straightforward computation shows:

$$\Gamma\alpha_p\Gamma = \{\alpha \in \mathrm{Mat}(2, \mathbb{Z}) : \alpha \equiv \begin{pmatrix} 1 & * \\ 0 & p \end{pmatrix} \pmod N,\ \det\alpha = p\}, \tag{55}$$

Since for any $\gamma_0 \in \Gamma_0(N)$ we have that $\gamma_0\alpha_p\gamma_0^{-1}$ is exactly in the above form, setting $\Gamma\alpha_p\Gamma = \bigsqcup_i \Gamma\beta_i$, we have:

$$\Gamma\alpha_p\Gamma = \Gamma\gamma_0\alpha_p\gamma_0^{-1}\Gamma = \gamma_0\gamma_0^{-1}\Gamma\gamma_0\alpha_p\gamma_0^{-1}\Gamma\gamma_0^{-1}\gamma_0 = \gamma_0\Gamma\alpha_p\Gamma\gamma_0^{-1} =$$
$$= \gamma_0\bigsqcup_i \Gamma\beta_i\gamma_0^{-1} = \gamma_0\bigsqcup_i \Gamma\gamma_0^{-1}\gamma_0\beta_i\gamma_0^{-1} = \bigsqcup_i \Gamma\gamma_0\beta_i\gamma_0^{-1},$$

where the third and last inequalities follow from the normality of Γ in $\Gamma_0(N)$. Comparing the right cosets decompositions gives $\bigsqcup_i \Gamma\gamma_0\beta_i = \bigsqcup_i \Gamma\beta_i\gamma_0$, although we may not have term-by-term equality. Still, taking γ_0 with bottom right entry $d_0 \equiv d \pmod N$, we get:

$$\langle d\rangle T_p f = \langle d\rangle \sum_i f[\beta_i]_k = \sum_i f[\gamma_0\beta_i]_k = \sum_i f[\beta_i\gamma_0]_k = T_p\langle d\rangle f.$$

It is a basic fact of linear algebra that two commuting operators preserve respective eigenspaces.

To prove that the Hecke operators fix $S_k(\Gamma)$ and that T_p and T_q commute, we must compute their action on cusp forms: we do so working with the Fourier expansion, recalling that the coefficients are called $a_n(f)$.

Lemma 4.22. *Let $\mathbf{1}_N$ be the trivial character* (mod N)*. Then for $f \in S_k(\Gamma)$ we have:*

$$a_n(T_p f) = a_{np}(f) + \mathbf{1}_N(p)p^{k-1}a_{n/p}(\langle p\rangle f). \tag{56}$$

Therefore, on the subspaces $S_k(N,\chi)$ we have:

$$a_n(T_p f) = a_{np}(f) + \chi(p)p^{k-1}a_{n/p}(f), \tag{57}$$

where $a_{n/p} = 0$ if $p \nmid n$.

Proof of Lemma 4.22. For $0 \le l \le p-1$, let $\alpha_p^l = \begin{pmatrix} 1 & l \\ 0 & p \end{pmatrix}$, so that $\alpha_p = \alpha_p^0$. Let us compute $\bigsqcup_{i=0}^{p-1} \Gamma\alpha_p^l$: a generic matrix in the l-th set has determinant p and is of the form $\begin{pmatrix} a & al+bp \\ c & cl+dp \end{pmatrix}$, $a, d \equiv 1 \pmod N$, $c \equiv 0 \pmod N$, so if $p \nmid a$, which is ensured if $p \mid N$, we get a generic matrix $\alpha \in \mathrm{Mat}(2,\mathbb{Z}) : \alpha \equiv \begin{pmatrix} 1 & * \\ 0 & p \end{pmatrix} \pmod N$, $\det\alpha = p$; but this is exactly the LHS in (55), so in this case we get a set of right coset representatives α_p^l for T_p.

If $p \nmid N$, it turns out that it suffices to complete the set of representatives with $\alpha_p^\infty = \begin{pmatrix} m & n \\ N & p \end{pmatrix}\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$ where m, n are such that $mp - nN = 1$, which is possible by

Bezout's Theorem, so that the first matrix is in $\Gamma_0(N)$ (and therefore acts as the p-th diamond operator). So for $p \mid N$ we just need to compute:

$$f[\alpha_p^l]_k(z) = p^{\frac{k}{2}} p^{-k} f\left(\frac{z+l}{p}\right) = p^{-\frac{k}{2}} \sum_{n=0}^{\infty} a_n(f) e^{\frac{2\pi i n(z+l)}{p}} = p^{-\frac{k}{2}} \sum_{n=0}^{\infty} a_n(f) q_p^n \zeta_p^{ln},$$

where $\zeta_p = e^{\frac{2\pi i}{p}}$. Summing over $0 \leq l \leq p-1$ then makes every term with $p \nmid n$ vanish, whereas the sum of roots of unity evaluates to p otherwise, giving:

$$f[\alpha]_k(z) = \sum_{p \mid n} a_n q_p^n = \sum_{n=0}^{\infty} a_{np} q^n,$$

which is exactly (56) since $\mathbf{1}_N(p) = 0$, remembering the factor $p^{\frac{k}{2}-1}$ in the definition of the Hecke operators. In the $p \nmid N$ case, the additional term $\alpha = \begin{pmatrix} m & n \\ N & p \end{pmatrix} \begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}$ acts via:

$$f[\alpha]_k(z) = (\langle p \rangle f)[\begin{pmatrix} p & 0 \\ 0 & 1 \end{pmatrix}]_k(z) = p^{\frac{k}{2}} (\langle p \rangle f)(pz),$$

so we get the additional term $p^{k-1} \sum_{n=0}^{\infty} a_n(\langle p \rangle f) q^{np}$, completing the proof of (56). The second part of the lemma follows directly from it since the $S_k(N, \chi)$ are precisely the spaces of eigenforms for the diamond operators. □

Going back to the proof of Proposition 4.21, (56) immediately gives that the Hecke operators fix the space of cusp forms. It is clearly sufficient to prove commutativity on the spaces $S_k(N, \chi)$ by Proposition 4.20, so computing $a_n(T_p(T_q f))$ using (57) we obtain:

$$a_{npq}(f) + \chi(q) q^{k-1} a_{np/q}(f) + \chi(p) p^{k-1} a_{nq/p}(f) + \chi(pq)(pq)^{k-1} a_{n/pq},$$

which is symmetric in p and q, so the claim follows. □

We conclude this section by observing that, if for $r \geq 2$ we set:

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

on $S_k(N, \chi)$, we get $T_{p^r} T_{q^s} = T_{q^s} T_{p^r}$ by induction with (56). So, after extending these operators to $S_k(\Gamma)$ by linearity thanks to Proposition 4.20, setting $T_n = \prod T_{p_i^{r_i}}$ where $n = \prod p_i^{r_i}$, we get that the operators T_n commute and that $T_{nm} = T_n T_m$ for $(m, n) = 1$.

4.4 L-functions and modularity

The definition of the Hecke operators T_n on $S_k(N,\chi)$ implies that their generating function has an Euler product:

$$\sum_{n=1}^{\infty} T_n n^{-s} = \prod_p (1 - T_p p^{-s} + \chi(p)p^{k-1-2s})^{-1}. \tag{58}$$

Notice how, for $k=2$, this resembles the Euler product (10) for the L-function of an elliptic curve.

Given a function $f(z) = \sum_{n\geq 0} a_n(f)q^n$, we construct its L-function as:

$$L_f(s) = \sum_{n=0}^{\infty} a_n(f)n^{-s}. \tag{59}$$

We look at the case $f \in \mathrm{M}_k(N,\chi)$. As usual, we do not know whether (or where) L_f defines a holomorphic function of complex variable s, but if we can prove that there is a constant $c = c(f)$ such that $|a_n(f)| = O(n^c)$ then we will know that $L_f(s)$ converges to the right of the vertical line $\operatorname{Re} s = c+1$. It turns out that the constant only depends on the weight of f.

Proposition 4.23. *1. If $f \in \mathrm{M}_k(\Gamma)$ then $|a_n(f)| = O(n^{k-1})$.*

2. If moreover f is a cusp form, $|a_n(f)| = O(n^{\frac{k}{2}})$.

Proof. We only prove the second claim, since it is what we will need in our exposition, but the first one follows from the same kind of computations applied to Eisenstein series and the fact that they form an orthogonal complement (with respect to the Petersson product) to the cusp forms.

Let $z = x+iy$ and $\phi(z) = f(z)y^{\frac{k}{2}}$. If we can show that $|\phi|$ is bounded by a constant C on $\mathscr{H}$ we are done, because:

$$|a_n(f)| = \left|\int_0^1 f(z)q^{-n}dx\right| \leq \int_0^1 |f(z)|e^{2\pi ny}dx \leq Cy^{-\frac{k}{2}}e^{2\pi ny},$$

so choosing $y = n^{-1}$ gives $|a_n(f)| \leq Ce^{2\pi}n^{\frac{k}{2}}$.

Now notice that ϕ is Γ-invariant: $\phi(\gamma z) = f(\gamma z)y^{\frac{k}{2}}((cz+d)^{-2})^{\frac{k}{2}} = f(z)y^{\frac{k}{2}}$ by (42). Therefore, it suffices to show that $|\phi|$ is bounded on a fundamental domain F_Γ for Γ. By what we saw in Section 4.1, this is of the form $\bigcup_{i=1}^M \gamma_i F$, with F a fundamental domain for $\Gamma(1)$ and the γ_i a set of coset representatives for $\Gamma\backslash\Gamma(1)$. By compactness of $F \cup \{\infty\}$, it suffices to show that in each of these translates, $|\phi|$

is bounded at ∞. As we will now see ϕ actually vanishes at ∞ in F, and this is sufficient since, by definition, cusp forms vanish at all cusps, that is, their translates by the whole modular group vanish at ∞, which will be precisely the condition that makes ϕ vanish at ∞; since $a_0(f) = 0$, we have:

$$|\phi(z)| \leq y^{\frac{k}{2}} \sum_{n=1}^{\infty} |a_n(f)| e^{-2\pi n y} \Longrightarrow \lim_{y \to \infty} |\phi(z)| = 0.$$

□

Now we know that $L_f(s)$ defines a holomorphic function for $\operatorname{Re} s > k$, and for $\operatorname{Re} s > \frac{k}{2} + 1$ if f is a cusp form. As we saw, it is always useful when dealing with L-functions to try to find an analytic continuation to the whole complex plane by means of (symmetry-based) functional equations. For $\mathrm{M}_k(N, \chi)$ we do so by means of the *Fricke involution*, the operator $[\omega_N]_k$ with $\omega_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{Q})$.

Define $\mathrm{M}_k^{\pm}(N, \chi) = \{f \in \mathrm{M}_k(N, \chi) : f[\omega_N]_k = \pm(-i)^k f\}$, and the same for $S_k(N, \chi)$. Then we have:

Theorem 4.24. *Let $f \in S_k(N, \chi)$. Then $L_f(s)$ has an analytic continuation to the whole complex plane. Moreover, if $f \in S_k^{\pm}(N, \chi)$, the function $\Lambda_f(s) = (\frac{\sqrt{N}}{2\pi})^s \Gamma(s) L_f(s)$ satisfies:*

$$\Lambda_f(s) = \pm \Lambda_f(k - s).$$

We start with a lemma:

Lemma 4.25. *1. If $f \in \mathrm{M}_k(N, \chi)$ (resp. $S_k(N, \chi)$) then $f[\omega_N]_k \in \mathrm{M}_k(N, \bar{\chi})$ (resp. $S_k(N, \bar{\chi})$) and the composition $[\omega_N]_k^2$ is the multiplication by $(-1)^k$;*

2. we have the decomposition $\mathrm{M}_k(N, \chi) = \mathrm{M}_k^+(N, \chi) \oplus \mathrm{M}_k^-(N, \chi)$ and the same for $S_k(N, \chi)$.

Proof. 1. Notice that for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$ we have:

$$f[\omega_N]_k[\gamma]_k = f[\omega_N \gamma \omega_N^{-1}]_k [\omega_N]_k = \bar{\chi}(d)[\omega_N]_k,$$

since $\omega_N \gamma \omega_N^{-1} = \begin{pmatrix} d & -c/N \\ -Nb & a \end{pmatrix}$ and $\chi(a) = \bar{\chi}(d)$ since $ad \equiv 1 \pmod{N}$. Moreover, $\omega_N^2 = -NI$, so its weight k action is the multiplication by $(N^2)^{\frac{k}{2}}(-N)^{-k} = (-1)^k$.

2. Define $f^{\pm} = \frac{1}{2}(f \pm i^k f[\omega_N]_k)$. Then clearly $f = f^+ + f^-$, and by definition $f^{\pm}[\omega_N]_k = \frac{1}{2}(f[\omega_N]_k \pm (-i)^k f) = (-i)^k f^{\pm}$. □

Proof of Theorem 4.24. Clearly $L_{c_1f_1+c_2f_2}(s)=c_1L_{f_1}(s)+c_2L_{f_2}(s)$ because the index on the LHS is a modular form with Fourier coefficients $c_1a_n(f_1)+c_2a_n(f_2)$, so we just need to prove the theorem for $f\in S_k^{\pm}(N,\chi)$ by Lemma 4.25. We work with the Mellin transform $g(s)=(\mathcal{M}\phi)(s)$ of the function $\phi(t)=f(it),\ t\in(0,\infty)$; for $\operatorname{Re}s>\frac{k}{2}+1$, we have:

$$g(s)=\int_0^\infty f(it)t^s dt=\int_0^\infty \sum_{n\geq 1} a_n(f)e^{-2\pi nt}t^{s-1}dt=$$
$$=\sum_{n\geq 1} a_n(f)\int_0^\infty e^{-2\pi nt}t^{s-1}dt=(2\pi)^{-s}\Gamma(s)L_f(s)$$

by Lemma 3.8. Therefore, in this region we have:

$$\begin{aligned}\Lambda_f(s)=&N^{\frac{s}{2}}\int_0^\infty f(it)t^{s-1}\frac{dt}{t}\overset{t\mapsto 1/Nt}{=}N^{-\frac{s}{2}}\int_0^\infty f\left(\frac{i}{Nt}\right)t^{-s}\frac{dt}{t}=\\ =N^{-\frac{s}{2}}&\int_0^\infty f\left(-\frac{1}{Nit}\right)t^{-s}\frac{dt}{t}=N^{\frac{k-s}{2}}\int_0^\infty f[\omega_N]_k(it)t^{k-s}\frac{dt}{t}=\\ &=\pm N^{\frac{k-s}{2}}\int_0^\infty f(it)t^{k-s}\frac{dt}{t}\overset{3.8}{=}\pm\Lambda_f(k-s),\end{aligned}$$

which, if we can prove that any of the integrals appearing above converges for $s\in\mathbb{C}$, would complete the proof. But, taking for example the third to last integral, we just need to prove that it converges with the smaller integration domain $(1,\infty)$, thanks to the transformation law of f under $t\mapsto -1/t$. But this follows directly from the fact that f goes to 0 exponentially as t goes to infinity since it is a cusp form, as seen in the proof of Proposition 4.23. □

Now we know that L-functions attached to modular forms behave very nicely: they admit an analytic continuation satisfying a functional equation. Moreover, the L-functions we already encountered that also satisfy these properties were those, studied in Section 3.2, of the elliptic curves congruent number family: by proving Theorem 3.3 in their case, without even mentioning modular forms, we proved that they satisfy the exact same type of functional equation.

This provokes a question:

Question 2. *Do all "nice" L-functions come from modular forms?*

The answer is, in a sense, on the affirmative:

Definition 4.26. *Given an L-function $L(s)=\sum_{n\geq 0}a_nn^{-s}$ and a positive integer N, we associate with L its Λ_N function as above: $\Lambda_N(s)=(\frac{\sqrt{N}}{2\pi})^s\Gamma(s)L(s)$. For a* *positive integer m and a character χ (mod m), we define the twists of L and Λ_N by χ as:*

$$L(\chi, s) = \sum_{n\geq 0} \chi(n) a_n n^{-s}, \quad \Lambda_N(\chi, s) = \left(\frac{m\sqrt{N}}{2\pi}\right)^s \Gamma(s) L(\chi, s).$$

Definition 4.27. *A set $\mathscr{P}$ of primes is called abundant if it has nonempty intersection with every arithmetic progression $\{a + nb,\ n \geq 0\}$ with $(a, b) = 1$.*

Definition 4.28. *Given a character χ (mod m), its Gauss sum is defined as $g(\chi) = \sum_{j=1}^{m} \chi(j) e^{\frac{2\pi i j}{m}}$.*

Theorem 4.29 (Weil)**.** *Let χ_0 be a character modulo $N \geq 1$. Let $f(z) = \sum_{n\geq 0} a_n q^n$ satisfy $|a_n| = O(n^c)$ for some $c \in \mathbb{R}$. Suppose that, for all the characters χ modulo an abundant set of primes $\mathscr{P} = \{p\}$, $\Lambda_N(\chi, s)$ has an analytic continuation bounded in any vertical strip and satisfying, for some $k \in \mathbb{N}$, the functional equation*

$$\Lambda_N(\chi, s) = C_\chi \Lambda_N(\bar{\chi}, k - s), \quad C_\chi = \pm\chi_0(p)\chi(-N)g(\chi)g(\bar{\chi})^{-1}.$$

Then $f \in \mathrm{M}_k^{\pm}(N, \chi_0)$; moreover, if we can take $c = k - 1 - \epsilon$ for some $\epsilon > 0$, then $f \in S_k^{\pm}(N, \chi)$.

Observe that what we did in Section 3.2 is to prove that the hypotheses of this theorem hold for $L(E_n, s)$ (or, to be precise, for the function f_{E_n} with Fourier coefficents the coefficients of $L(E_n, s)$) with $k = 2,\ N = N_{E_n},\ \chi_0 = 1$, but only for the trivial character χ. Indeed, boundedness in every vertical strip for Λ_f is clear from the paragraph after (23).

We say that a rational elliptic curve E is *modular* if its L-function is the same of that of a modular form f, and in this case we say that E *comes from* f. Therefore, to obtain *modularity* for the congruent number family, we need to repeat the proof in Section 3.2 but with $\tilde{\chi}_n(I)$ replaced with $\tilde{\chi}_n(I)\chi(\mathbb{N}I)$ for all χ modulo the primes not dividing N, which form an abundant set thanks to Dirichlet's Theorem. The technique is exactly the same, so we omit the lengthy computation, whose effect is just to multiply the constant C, which was the Legendre symbol $\left(\frac{-2}{n}\right)$ (say, for odd n), by $\chi(-N)g(\chi)g(\bar{\chi})^{-1}$. Moreover, by 3.2 we can take any $0 < \epsilon < \frac{3}{2}$ in the theorem, so E_n all come from cusp forms.

As we just said, for the congruent number curves, the conductor as we defined it in Section 3.1 is exactly the level of the modular form from which the curve arises via Weil's Theorem. The same technique we used can be extended to all elliptic curves with complex multiplication, proving that they arise from a form in $S_2^{\pm}(N, \mathbf{1}) = S_2^{\pm}(\Gamma_0(N))$: the additional endomorphisms are what allows one to write formulas like (16) for the values $\alpha_p(E)$, which in turn make it possible to treat the L-function as a Hecke L-function (see Section 3.2) and hence obtain the analytic continuation via the equality with the Mellin transform of the 2-dimensional theta functions.

In 2001, Breuil, Conrad, Diamond and Taylor completed a monumental achievement started by Taylor and Wiles, proving the Modularity Theorem:

Every rational elliptic curve E arises from a form in $S_2^{\pm}(\Gamma_0(N_E))$.

More than that, we know which cusp forms give rise to elliptic curves: they need to be eigenvetors for all the Hecke operators T_p, normalized with $a_1(f) = 1$ (these are called *eigenforms*). Observe that by (57) we have $a_1(T_p f) = a_p(f)$, so if $T_p f = \lambda_p f$ we get $\lambda_p a_1(f) = a_1(T_p f) = a_p(f)$, so the eigenvalue is $a_p(f)$. Finally, f needs to satisfy an additional condition to give rise to an elliptic curve: it needs not to come, in a sense which we do not make precise, from a modular form from the smaller space $S_2(\Gamma_1(Np^{-1}))$ for any $p \mid N$. Such a form in $S_k^{\pm}(N, \Gamma)$ is called *newform,* and we denote their space by $S_k^{\mathrm{new}}(N, \chi)$.

Therefore, (58) tells us that for a newform $f \in S_2^{\mathrm{new}}(\Gamma_0(N)) = S_2^{\mathrm{new}}(N, \mathbf{1})$ we have:

$$L_f(s) = \prod_p (1 - a_p p^{-s} + \mathbf{1}_N(p) p^{1-2s})^{-1}. \tag{60}$$

Therefore, the Modularity Theorem tells us that the values $a_p(E)$ are the eigenvalues of some linear operators on a finite dimensional $\mathbb{C}$-vector space. This is a form of *reciprocity:* see the beautiful Introduction in [11] for a general description of why this is.

4.5 Modular forms of half-integer weight

In what follows, we take the branch of the square root $\sqrt{z}$ such that $\sqrt{1} = 1$, which defines a holomorphic function on $\mathbb{C} \setminus \mathbb{R}^-$.

Recall that in Section 4.2 we saw that $F(z) = \theta^4(z)$ is a modular form for $\mathrm{M}_2(\Gamma_0(4))$, that is, $F(\gamma z) = (cz+d)^2 F(z)$ for $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(4)$. It is natural to ask what happens for other powers θ^k, which boils down to computing $\frac{\theta(\gamma z)}{\theta(z)}$, $\gamma \in \Gamma_0(4)$. Do we still obtain modular forms for this group? Obviously, by elementary properties of exponentiation the weight would have to be $\frac{k}{2}$, so we only have hopes for even k under our definitions. The following proposition, whose proof, albeit not deep, is long, computationally heavy and outside the aim of this thesis (see [21, III,4]), settles the question:

Proposition 4.30. *In the above situation, we have $\theta(\gamma z) = j(\gamma, z)\theta(z)$ for*

$$j(\gamma z) = \left(\frac{c}{d}\right)\epsilon_d^{-1}\sqrt{cz+d}, \quad \epsilon_d = \sqrt{\left(\frac{-1}{d}\right)}, \tag{61}$$

where for negative d the Legendre symbol is defined as $\left(\frac{c}{-d}\right)$ if $c > 0$ and as $-\left(\frac{c}{-d}\right)$ if $c < 0$ (and $\left(\frac{0}{d}\right) = 1$ if $|d| = 1$ and 0 otherwise).

Remark 4.30.1. *The automorphy factor $j(\gamma, z)$ satisfies $j(\gamma_1\gamma_2, z) = j(\gamma_1, \gamma_2 z)j(\gamma_2, z)$.*

Corollary 4.30.1. *For an even integer $k = 2l$, we have:*

$$\theta^k \in \mathrm{M}_l(4, \chi_{-1}^l),$$

where χ_{-1} is the character modulo 4 defined by $\chi_{-1}(n) = \left(\frac{-1}{n}\right)$.

Proof. We just need to verify the modularity relation for $\gamma \in \{-I, T, \begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}\}$. The first two are obvious; for the third, (61) gives:

$$\theta^k(\gamma z) = \left(\frac{0}{1}\right)^k \left(\frac{-1}{d}\right)(cz+d)^l\theta^k(z) = \chi_{-1}(d)(cz+d)^l\theta^k(z).$$

□

The proposition also gives us transformation laws for odd k, but in this case we do not have (for now) a theory that would allow us to gather information about representations of integers as sum of k squares ($k \geq 5$ being now trivial, but one could be interested in classifying which integers are the sum of three squares) or, more generally, as values of a quadratic form $A = \sum_{i,j=1}^k A_{ij}n_in_j$ in an odd number of variables, via the theta-like function $\sum_{\bar{n}} q^{\bar{n}A\bar{n}^t}$ for $\bar{n} = (n_1, ..., n_k)$.

Although the problem of studying representations of integers by quadratic forms was the original reason for which modular forms of half-integer weight $\frac{k}{2}$ were initially studied, we diverge from that line of inquiry, pursuing instead a deep link, expressed in the formalism of Hecke operators and L-functions, with modular forms of weight $k-1$. Before doing that, we need to set some definitions: indeed, it is not hard to show that defining the transformation law that a modular form of weight $\frac{k}{2}$ should satisfy by replacing $k \mapsto \frac{k}{2}$ in (43) leads to inconsistencies due to the sign choice for any congruence subgroup. Therefore, we focus on congruence subgroups $\Gamma \subset \Gamma_0(4)$ and require that *a weight $\frac{k}{2}$ modular form transforms as θ^k under the action of Γ:*

$$f(\gamma z) = j(\gamma, z)^k f(z) \ \forall \gamma \in \Gamma. \tag{62}$$

Defining the modular action as $f[\gamma]_{\frac{k}{2}}(z) = j(\gamma,z)^{-k}f(\gamma z)$, (62) is equivalent to asking that $f = f[\gamma]_{\frac{k}{2}}$. We are still not ready to give the precise definition, since we are yet to define what it means to be holomorphic/vanish at the cusps. In order to do so, we need to extend the action to $\mathrm{GL}_2^+(\mathbb{Q})$, while our definition of $j(\gamma,z)$ is only valid for $\Gamma_0(4)$. This makes it necessary to work with a 4-covering G of $\mathrm{GL}_2^+(\mathbb{Q})$.

Let $W = \{1,-1,i,-i\} \simeq \mathbb{Z}/4\mathbb{Z}$ be the group of fourth roots of unity. Define G as:

$$G = \{(\alpha,\phi) : \alpha \in \mathrm{GL}_2^+(\mathbb{Q}),\ \phi \in \mathbb{C}(\mathscr{H}) : \phi(z)^2 = w\frac{cz+d}{\sqrt{\det\alpha}}\} \text{ for some } w \in W\}.$$

Clearly, we have a natural projection $\pi : G \longrightarrow \mathrm{GL}_2^+(\mathbb{Q})$ given by the first component. Equip G with a product law $(\alpha_1,\phi_1(z))(\alpha_2,\phi_2(z)) = (\alpha_1\alpha_2, \phi_1(\alpha_2 z)\phi_2(z))$, whose associativity is immediate. Moreover,

$$\phi_1(\alpha_2 z)^2\phi_2(z)^2 = w_1w_2\frac{(c_1(\alpha_2 z)+d_1)(c_2z+d_2)}{\sqrt{\det\alpha_1\det\alpha_2}},$$

and the numerator is $c_1(a_2z+b_2)+d_1(c_2z+d_2) = (c_1a_2+d_1c_2)z+c_1b_2+d_1d_2$, which proves that the product is still in G as the left and bottom right entries of $\alpha_1\alpha_2$ are exactly $c_1a_2+d_1c_2$ and $c_1b_2+d_1d_2$. Clearly, the identity is $(I,1)$. Finally, a formal inverse of (α,ϕ) is $(\beta,\psi) = (\alpha^{-1}, \phi(\alpha^{-1}z)^{-1})$, which we just need to show belongs to G: setting $D = \det\alpha$, we have $\alpha^{-1} = D^{-1}\begin{pmatrix} d & -b \\ -c & a\end{pmatrix}$ and hence:

$$\psi^2 = w^{-1}\frac{\sqrt{D}}{c(\alpha^{-1}z)+d} = w^{-1}\frac{\sqrt{D}}{c\frac{dz-b}{-cz+a}+d} = w^{-1}(-cz+a)\sqrt{D^{-1}} = w^{-1}\frac{-\frac{c}{D}z+\frac{a}{D}}{\sqrt{D^{-1}}},$$

as desired.

Given $\eta = (\alpha,\phi)$ we can now define the weight $\frac{k}{2}$ modular operator, which defines an action of G thanks to Remark 4.30.1 and the definition of the product:

$$f[\eta]_{\frac{k}{2}} = f(\alpha z)\phi(z)^{-k}. \tag{63}$$

Let us work again with congruence subgroups $\Gamma \subset \Gamma_0(4)$; then, setting $\tilde{\Gamma} = \{(\gamma, j(\gamma,z))\}$ we obtain an injection $\tilde{\Gamma} \hookrightarrow G$. We now need to define the concept of holomorphicity/vanishing at the cusps $s \in \mathbb{Q}\cup\{\infty\}$:

- $s = \infty$: as in the integer weight case we take $\gamma = \pm T^h \in \Gamma,\ h \in \mathbb{Z}^+$, and since $j(\gamma,z) = 1$ in this case we know that f is invariant under $z \mapsto z+h$, and therefore has a Fourier expansion $\sum_{n\geq n_0} a_nq_h^n$ (remember that $f \in \mathbb{C}(\mathscr{H})$), and the conditions of meromorphicity/holomorphicity/vanishing are defined identically to the integer weight case;

- $s \in \mathbb{Q}$: take $\alpha \in \Gamma(1)$ such that $\alpha\infty = s$ and a lifting $\xi = (\alpha, \phi) \in G$. Let $\tilde{T}_\infty \subset G$ be the subgroup $\pi^{-1}(\langle T \rangle)$. We have $\xi^{-1}\tilde{\Gamma}\eta \subset \tilde{T}_\infty$, so by finitness of the index $\xi^{-1}\tilde{\Gamma}\xi = \langle(\pm T^h, w)\rangle$, $h > 0$ (note that w does not have to be 1 as if we were inside $\tilde{\Gamma}$). It is not hard to see that the generator does not depend on α but only on the cusp, so for a fixed point f of $[\tilde{\gamma}]_{\frac{k}{2}}$, $\gamma \in \tilde{\Gamma}$, setting $g = f[\xi]_{\frac{k}{2}}$, we have:

  $$g[\xi^{-1}\tilde{\gamma}\xi]_{\frac{k}{2}} = f[\tilde{\gamma}\xi]_{\frac{k}{2}} = g,$$

  so $g(z) = e^{\frac{\pi i j}{2}} g(z+h)$ with $j \in \{0, 1, 2, 3\}$, which means that $e^{-\frac{\pi i j z}{2h}}$ has period h and therefore that g expands as $\sum_{n \geq n_0} a_n q_h^{n+\frac{j}{4}}$. Defining $f(s)$ as $\lim_{z \to i\infty} g(z)$, we say that f is holomorphic at s if $a_n = 0$ $\forall n < 0$ and that it vanishes at s if $a_0 = 0$.

Definition 4.31. *Let $k \in \mathbb{Z}$ and let $\Gamma \subset \Gamma_0(4)$ be a congruence subgroup. We say that a function $f \in \mathbb{C}(\mathscr{H})$ which is fixed by $[\tilde{\gamma}]_{\frac{k}{2}}$ $\forall \tilde{\gamma} \in \tilde{\Gamma}$ is a modular form of weight $\frac{k}{2}$ for $\tilde{\Gamma}$ if it is holomorphic on $\mathscr{H}$ and at every cusp. Moreover, if f vanishes at every cusp, we call it a cusp form.*

We use the same notation of the integer case, without the need of repeating it, for the spaces of modular and cusp forms, and denote $\mathrm{M}_k(\tilde{\Gamma}_0(N), \chi)$, $S_k(\tilde{\Gamma}_0(N), \chi)$ the spaces of weight $\frac{k}{2}$ modular (cusp) forms f for $\tilde{\Gamma}_1(N)$ satisfying $f[\tilde{\gamma}]_{\frac{k}{2}} = \chi(d) f$ $\forall \tilde{\gamma} \in \tilde{\Gamma}_0(N)$, for N a multiple of 4. In these cases, the same argument via the faithful representation induced by the diamond operators shows:

$$\mathrm{M}_{\frac{k}{2}}(\tilde{\Gamma}_1(N)) = \bigoplus_{\chi} \mathrm{M}_{\frac{k}{2}}(\tilde{\Gamma}_0(N), \chi).$$

Observe that if χ is an odd character, that is if $\chi(-1) = -1$, then $\mathrm{M}_{\frac{k}{2}}(\tilde{\Gamma}_0(N), \chi) = 0$, as one sees by plugging $\tilde{\gamma} = (-I, 1)$ in the definition, getting $f = \chi(-1) f$. Interestingly, since $\chi = 1$ is the only even character modulo 4, this implies that $\mathrm{M}_{\frac{k}{2}}(\tilde{\Gamma}_0(4)) = \mathrm{M}_{\frac{k}{2}}(\tilde{\Gamma}_1(4))$.

4.6 The Shimura correspondence

In this section we will assume $4 \mid N$. We can define Hecke operators T_n for modular forms of half-integer weight, but we need to use the double coset operator interpretation for every n, giving a formula that also holds for the integer weight case, although we only used it for prime index.

Definition 4.32. *Let $n > 0$ and $\xi_n = \left(\begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix}, n^{\frac{1}{4}}\right)$. The n-th Hecke operator T_n on $\mathrm{M}_k(\tilde{\Gamma}_1(N))$ is defined as:*

$$T_n f = n^{\frac{k}{4}-1}[\tilde{\Gamma}_1(N)\xi_n\tilde{\Gamma}_1(N)]_{\frac{k}{2}}.$$

The remarkable fact is that, as a consequence of working with a lift of $\mathrm{GL}_2^+(\mathbb{Q})$, these operators are 0 if n is a nonsquare coprime to the level N:

Proposition 4.33. *If $(n, N) = 1$ and n is not the square of an integer, $T_n = 0$.*

Proof. Recall that, by Lemma 4.15,

$$T_n = n^{\frac{k}{4}-1}\sum_i [\xi_n \tilde{\gamma}_i]_{\frac{k}{2}}$$

for a set of right coset representatives $\tilde{\gamma}_i$ of $\tilde{\Gamma}'' = \xi_n^{-1}\tilde{\Gamma}_1(N)\xi_n \cap \tilde{\Gamma}_1(N)$ in $\tilde{\Gamma}_1(N)$.

The idea is to realize $\tilde{\Gamma}''$ as the lifting of the kernel of a map from $\Gamma' = \alpha_n^{-1}\Gamma_1(N)\alpha_n \cap \Gamma_1(N)$ into W (observe that $\pi(\xi_n) = \alpha_n$ by definition). There is an exact sequence:

$$0 \longrightarrow W \longrightarrow G \longrightarrow \mathrm{GL}_2^+(\mathbb{Q}) \longrightarrow 0$$

where the first map is $w \mapsto (I, w)$ and the second one the natural projection; given $\gamma \in \Gamma'$, putting $\gamma = \alpha_n^{-1}\gamma_1\alpha_n, \gamma_1 \in \Gamma_1(N)$, clearly $\tilde{\gamma}$ and $\xi_n^{-1}\tilde{\gamma}_1\xi_n$ both project to γ in $\mathrm{GL}_2^+(\mathbb{Q})$. Therefore, they differ by an element which we can identify in W thanks to the exact sequence, giving a map $\beta : \gamma \mapsto w$, which the product law of G and the definition of $j(\gamma, z)$ show is $\left(\dfrac{d}{n}\right)$. Let K be its kernel.

Notice that, with this definition $\tilde{\gamma} \in \tilde{\Gamma}'' \iff \gamma \in K$ follows immediately, because in this case $\tilde{\gamma} = \tilde{\gamma}(I, w)$. Now, under our hypotheses on n, $\tilde{\Gamma}''$ is a subgroup of $\tilde{\Gamma}'$ *of index* 2, because β has order 2. Decomposing $\tilde{\Gamma}' = \tilde{\Gamma}'' \cup \tilde{\Gamma}''\tilde{\tau}$ with $\tau = \alpha_n^{-1}\tau_1\alpha_n$ and $\tau_1 \in \Gamma_1(N)$, gives: $\tilde{\Gamma}_1(N) = \bigsqcup_i \tilde{\Gamma}''\tilde{\gamma}_i \cup \bigsqcup_i \tilde{\Gamma}''\tilde{\tau}\tilde{\gamma}_i$.

But

$$[\xi_n\tilde{\tau}\tilde{\gamma}_i]_{\frac{k}{2}} = [\xi_n\tilde{\tau}\xi_n^{-1}\xi_n\tilde{\gamma}_i]_{\frac{k}{2}} = [\tilde{\tau}_1(I, -1)\xi_n\tilde{\gamma}_i]_{\frac{k}{2}} = -[\xi_n\tilde{\gamma}_i]_{\frac{k}{2}},$$

where $\xi_n^{-1}\tilde{\tau}\xi_n = \tilde{\tau}_1(I, -1)$ follows from $\tilde{\tau}$ not being in $\tilde{\Gamma}'' \simeq \tilde{K}$, making all the terms in the first sum cancel out with those in the second one. □

The Hecke operators again commute and are multiplicative, so to describe them explicitly for $(n, N) = 1$ we just need to compute the T_{p^2} for $p \nmid N$. This is carried out exactly as we did in the integer weight case in Section 4.3, finding a set of right coset representatives and explicitly computing the modular action. We therefore omit the explicit computations for this case, which are longer and more tedious, and limit ourself to state:

Lemma 4.34. *Let* $4 \mid N$, χ *be a character* (mod N) *and* k *be an odd integer. For* $f \in \mathrm{M}_{\frac{k}{2}}(\tilde{\Gamma}_0(N), \chi)$ *and a prime* p *we have:*

$$a_n(T_{p^2}f) = a_{p^2n} + \chi(p)\left(\frac{(-1)^{\frac{k-1}{2}}n}{p}\right)p^{\frac{k-3}{2}}a_n + \chi(p^2)p^{k-2}a_{n/p^2}. \tag{64}$$

We are finally ready to formulate Shimura's correspondence, relating modular forms of half-integer weight $\frac{k}{2}$ to ordinary modular forms of weight $k-1$:

Theorem 4.35 (Shimura)**.** *Let* N *be a positive multiple of* 4, χ *a character* (mod N) *and* $k \geq 3$ *an odd integer. Let* $f(z) = \sum_{n\geq 0} a_n q^n \in S_{\frac{k}{2}}(\tilde{\Gamma}_0(N), \chi)$ *be an eigenform for* T_{p^2} *for all primes* p, *with corresponding eigenvalues* λ_p. *Then, setting*

$$\prod_p (1 - \lambda_p p^{-s} + \chi(p)^2 p^{k-2-2s})^{-1} = \sum_{n\geq 1} b_n n^{-s}, \tag{65}$$

we have $g = \sum_{n\geq 1} b_n q^n \in \mathrm{M}_{k-1}(N/2, \chi^2)$. *Moreover, if* $k \neq 3$, $g \in S_{k-1}(N/2, \chi^2)$.

Sketch of the proof. We limit ourselves to show how to reduce the proof to the verification, for a family of L-functions we are going to define, of the analytic continuation and functional equation hypotheses in Weil's Theorem, which is rather long and technical, and can be found in [36].

Recall how, in Section 4.5, the Euler product (58) for the Hecke operators allowed us express the L-function of an eigenform in $S_k(N, \chi) = S_k(\Gamma_0(N), \chi)$ as an Euler product itself ((60) is the trivial character case). In the half-integer weight case, we have a different situation since we can only let T_{p^2} act, so the eigenvector conditions coming from Lemma 4.34 only give information on the coefficients a_n, $p^2 \mid n$, on top of having a hard-to-control third term with a_{n/p^2}.

Luckily, having an eigenform for *all* T_{p^2}, as in the hypotheses of Shimura's result, allows us to deal with both problems simultaneously and to obtain substantial information on the L-function. Indeed, let t be an integer all of whose square factors divide N and set $\chi_t(m) = \chi(m)\left(\frac{(-1)^{\frac{k-1}{2}}t}{m}\right)$. Then, letting the T_{p^2} act on f, the relations in Lemma 4.34 give:

$$\sum_{n\geq 1} a_{n^2t} n^{-s} =$$
$$= a(t) \prod_p (1-\chi_t(p)p^{\frac{k-3}{2}-s})(1-\lambda_p p^{-s}+\chi(p^2)p^{k-2-2s})^{-1},$$

as can be seen from the same type of formal computations we did for an eigenvector of the ordinary Hecke operators. This means that

$$L_g(s) = \prod_p (1-\lambda_p p^{-s}+\chi(p)^2 p^{k-2-2s})^{-1} = \left(\sum_{n\geq 1} a_{n^2t} n^{-s}\right)\left(\sum_{n\geq 1} \chi_t(n) n^{\frac{k-3}{2}-s}\right) \tag{66}$$

for any t as above such that $a_t \neq 0$ (and there has to be such a t, because if $a_t = 0\ \forall t$ squarefree, then $a_t = 0\ \forall t$ by the eignevector condition, so $f = 0$ and f would not be an eigenvector).

At this point, if we prove that the RHS of (66) (or, to be precise, the twists of the associated function $\Lambda(s)$) satisfy the hypotheses of Weil's Theorem, with (see the notation in the statement): $c = \frac{k}{2}+1$, $\mathscr{P} = \{\text{primes not dividing } tN\}$, $\chi_0 = \chi^2$, $N = N_t$ and k replaced by $k-1$, we get precisely Shimura's result with the level replaced by $\gcd(N_t)$ for the values of t described after (66), since for odd $k > 3$ we have $\frac{k}{2}+1 < k-1$. What was shown (see [21, IV,4]) is that there is such a t such that $N_t = N/2$.

We only show that the RHS of (66) converges for $\operatorname{Re} s > \frac{k}{2}+1$: the factor on the right is controlled by replacing the character with 1 and converges for $\operatorname{Re} s > \frac{k-1}{2}$ by the same classical argument as for $\zeta(s)$; for the other sum, notice that the argument that gave $|a_n| = O(n^{\frac{k}{2}})$ for the coefficients of a weight k cusp form for $\Gamma_1(N)$ (Proposition 4.23) works in the half-integer weight case, so with our current notation $a_n = O(n^{\frac{k}{4}})$ and hence for any fixed t the sum converges for $\operatorname{Re} s > \frac{k}{2}+1$, as desired. □

5 Heegner points for congruent number curves

A goal for this section is to describe an ingenious construction, originally due to amateur mathematician Kurt Heegner, of rational points on elliptic curves, arising from their complex analytic nature mentioned in Section 2.1. These points, now known as Heegner points, turned out to be a critical tool in the proofs of deep conjectures, some of which have *a priori* nothing to do with elliptic curves or rational points, especially thanks to a stunning work [15] of Gross and Zagier. Nonetheless, we will mainly focus on the application of (a variant of) this construction to some of the curves in the congruent number family, in order to obtain Theorem 1.2.

In this section we use the variable z to denote a generic element in the upper-half plane $\mathscr{H}$, while adopting the letter τ for the case in which the element is imaginary quadratic.

5.1 Modular curves as moduli spaces

As we mentioned in Section 2.1, elliptic curves over $\mathbb{C}$ are realized as Riemann surfaces by taking quotients of $\mathbb{C}$ modulo lattices Λ, where the biholomorphisms are given in terms of meromorphic functions, called elliptic functions, that are Λ-periodic. In this way, elliptic functions parametrize points on elliptic curves like sine and cosine parametrize points on the circle.

Given a lattice Λ, the quotient $\mathbb{C}/\Lambda$ is a torus. This agrees with our notion of elliptic curves as varieties having dimension and genus equal to 1. The concept of holomorphic map of Riemann surfaces is very strict, and it turns out that a holomorphic map from a torus $\mathbb{C}/\Lambda$ to itself necessarily comes from the multiplication on $\mathbb{C}$ by some $\alpha \in \mathbb{C}^*$ such that $\alpha\Lambda \subset \Lambda$. Therefore, two tori $\mathbb{C}/\Lambda,\ \mathbb{C}/\Lambda'$ are biholomorphic if and only if there exists $\alpha \in \mathbb{C}^*$ such that $\Lambda = \alpha\Lambda'$. Taking a basis ω_1, ω_2 spanning Λ over $\mathbb{Z}$, this means that we can take $\omega_1 = 1$, and clearly also $\operatorname{Im}\omega_2 > 0$ up to changing its sign. Therefore, setting $\Lambda_z = \mathbb{Z} + z\mathbb{Z}$, we have a surjective map $\Omega : \mathscr{H} \longrightarrow \{\text{Elliptic curves over } \mathbb{C}\},\ z \mapsto E_z = \mathbb{C}/\Lambda_z$.

Lemma 5.1. *The elliptic curves having complex multiplication correspond via Ω to imaginary quadratic values $\tau \in \mathscr{H}$, with $End(E_\tau)$ being an order in the ring of integers $\mathcal{O}_K$ of $K = \mathbb{Q}(\tau)$.*

Proof. By what we said above, for an endomorphism ϕ we must have $0 \neq \alpha$ such that: $\begin{pmatrix} \alpha \\ \alpha\tau \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ \tau \end{pmatrix}$, which gives:

$$c + \tau d = (a + \tau b)\tau \implies b\tau^2 + (a - d)\tau - c = 0. \tag{67}$$

Now, if $a + \tau b = \alpha \notin \mathbb{Z}$, then $b \neq 0$ and (67) implies that τ is imaginary quadratic (recall that $\tau \in \mathscr{H}$), and since $\alpha = a + \tau b$ we have that $\mathrm{End}(E) \subset \mathbb{Q}(\tau)$ is integral over $\mathbb{Z}$, so it is an order in $K = \mathbb{Q}(\tau)$. □

Now, Ω is not injective, because two values $z, z' \in \mathscr{H}$ can give rise to the same lattice. Writing the required (and implied) relations:

$$\begin{cases} \begin{pmatrix} 1 \\ z \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ z' \end{pmatrix}, \\ \begin{pmatrix} 1 \\ z' \end{pmatrix} = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \begin{pmatrix} 1 \\ z \end{pmatrix} \end{cases}$$

and denoting as $\gamma, \gamma' \in \mathrm{Mat}_2(\mathbb{Z})$ the two matrices, we obtain $\gamma\gamma' = I \implies \gamma, \gamma' \in \pm\Gamma(1)$. But since they preserve $(1, \mathscr{H})$ the determinant must be positive, so we find the following remarkable fact:

The modular curve $Y(1)$ parametrizes elliptic curves over $\mathbb{C}$ up to isomorphism.

In this situation we say that the parametrizing object (or a compactification of it) is a *moduli space* for the parametrized set.

Let $g_2(z) = 60G_4(z),\ g_3(z) = 140G_6(z)$. The theory of elliptic functions shows that E_z is given by a Weierstrass model $y^2 = 4x^3 - g_2(z)x - g_3(z)$. We now introduce a function on the upper half plane that allows us to "classify" elliptic curves:

Definition 5.2. *The j-invariant is $j(z) = 1728\frac{g_2^3(z)}{\Delta(z)} = 1728\frac{E_4(z)^3}{E_4(z)^3 - E_6(z)^3}$.*

The j-invariant is a (weight 0, clearly) modular function for $\Gamma(1)$, in the sense of Section 4.1, for its Fourier expansion is $j(z) = \frac{1}{q} + 744 + ...$ (invariance under the modular group follows by that of the Eisenstein series). The fundamental property of the j-invariant is that it is a biholomorphism from $X(1)$ to $\mathbb{P}^1_{\mathbb{C}}$, which implies:

Proposition 5.3. *Two elliptic curves $E_{z_1}/\mathbb{C},\ E_{z_2}/\mathbb{C}$ are isomorphic over $\mathbb{C}$ if and only if $j(z_1) = j(z_2)$.*

Therefore, given an elliptic curve $E/\mathbb{C}$, we can attach to it its j-invariant $j(E)$, defined as $j(z)$ for any $z \in \mathscr{H}$ such that $E \simeq E_z$, and hence as $1728\frac{a^3}{a^3 - 27b^2}$ if E is given by a Weierstrass model $y^2 = 4x^3 - ax - b$. Given a complex number j_0, there exist an elliptic curve $E_0/\mathbb{Q}(j_0)$ with $j(E_0) = j_0$: if $j_0 \neq 0, 1728$ we can take

$$E_0 : y^2 = x^3 + \frac{3j_0}{1728 - j_0}x + \frac{2j_0}{1728 - j_0},$$

which one easily verifies has $j(E_0) = j_0,\ \Delta_{E_0} \neq 0$, and for $j_0 = 0, 1728$ we can take, respectively, $E_0 : y^2 = x^3 - 1$ and the congruent number curve E_1.

It is a classical result (see [11, I,5]) that $Y_0(N)$ is a moduli space for elliptic curves with a cyclic subgroup of order N:

$$\begin{aligned} &Y_0(N) = \{(E, C) : E/\mathbb{C},\ \mathbb{Z}/N\mathbb{Z} \simeq C \subset E\}/\sim, \text{ where} \\ &(E, C) \sim (E', C') \iff \exists \phi : E \xrightarrow{\sim} E' : \phi(C) = C'. \end{aligned} \tag{68}$$

Indeed, we can take representatives $(E, C) = (E_z, \langle \Lambda_z + \frac{1}{N} \rangle)$ for z in a fundamental domain for $\Gamma_0(N)$. Observe that, since the natural projection is an isogeny of degree N from E to E/C, we can map any such representative to a pair $E \longrightarrow E',\ E = E_z$ of elliptic curves with an isogeny of kernel $\simeq \mathbb{Z}/N\mathbb{Z}$, up to equivalence defined in terms of the obvious commutative diagram. Moreover, to any such pair of elliptic curves admitting such an isogeny we can associate a representative in the form (68) by taking a curve E_z isomorphic to E and the kernel of the isogeny as the group. These application are clearly inverses of one another, so we have obtained:

$$Y_0(N) = \{E \xrightarrow{\phi} E' : \ker \phi \simeq \mathbb{Z}/N\mathbb{Z}\}/\sim . \tag{69}$$

In Section 4 we saw, and proved in the case of the quadratic twist family $\{E_n\}$, that elliptic curves are modular, in the sense that their L-functions arise from cusp forms which are a normalized common eigenvector for the Hecke operators of weight 2 for $\Gamma_0(N),\ N = N_E$, and which transform well under the Fricke involution. In the seventies it was proven that, for a curve E of conductor N, this is equivalent to the existence of a nonconstant morphism $\phi_E : X_0(N) \longrightarrow E$ (see [35]). By definition, given a congruence subgroup Γ, we have $\mathbb{C}(X(\Gamma)) \simeq \mathrm{M}_0^{\mathrm{mer}}(\Gamma)$, so E is parametrized by modular functions.

5.2 Complex multiplication and Heegner points

Recall that the *ideal class group* of a number field K, which we denote by $\mathrm{Cl}(K)$, is defined as the quotient J_K/P_K of the (abelian) group of nonzero fractional ideals of the ring of integers $\mathcal{O}_K$ by its subgroup of principal ideals. The ideal class group serves the purpose of capturing how much unique factorization fails in $\mathcal{O}_K$: for example, it is trivial if and only if $\mathcal{O}_K$ is a UFD.

A very important result in Algebraic Number Theory is that $\mathrm{Cl}(K)$ is a finite group. Its order is the *class number* h_K of K. We borrow the following useful fact from Class Field Theory, relating $\mathrm{Cl}(K)$ to Galois extensions:

Proposition 5.4. *Let H_K be the maximal unramified abelian extension of K. Then $Gal(H_K/K) \simeq Cl(K)$.*

Remark 5.4.1. *Being unramified and being abelian are properties preserved under compositum of extensions, so the maximal such extension is well defined.*

We call H_K the Hilbert class field of K, and the isomorphism described in the proposition, known as *Artin isomorphism*, as $\sigma \mapsto \mathfrak{a}_\sigma, \mathfrak{a} \mapsto \sigma_{\mathfrak{a}}$.

Let us come back to our goal of constructing rational points on rational elliptic curves. Modular curves that are moduli spaces for elliptic curves with additional structure can be given algebraic models for which a point $(E_z, \text{structure})/\sim$ is defined over a number field L if and only E_z has a model defined over L and such that the structure is defined over L. We could use the map $\phi_E : X_0(N) \longrightarrow E$ described at the end of Section 5.1 to get rational points on E from those on the relative modular curve. But those are also not easy to find: for instance, Mazur's proof of the Torsion Conjecture mentioned in Section 2.1 boils down to proving that $X_1(N)$ has nontrivial rational points only for a few values of N.

There are a lot of rational points on $\mathscr{H}$, and we have the projection $\pi_0 : \mathscr{H} \longrightarrow \Gamma_0(N)\backslash\mathscr{H} \hookrightarrow X_0(N)$, but it is a transcendental map. Here a "miracle" occurs though: for imaginary quadratic values τ of the argument, we still get points which are rational on a finite extension of $\mathbb{Q}(\tau)$!

Definition 5.5. *A Heegner point is a class $\pi_0(\tau) = (E \longrightarrow E') \in X_0(N)$ such that E, E' have complex multiplication by $\mathcal{O}_{\mathbb{Q}(\tau)}$.*

Proposition 5.6. *Let $\tau \in \mathscr{H}$ be imaginary quadratic and let $x = \pi_0(\tau) \in X_0(N)$ be the associated Heegner point. Then x is defined over the Hilbert Class Field H_K of $K = \mathbb{Q}(\tau)$.*

Proof. The Main Theorem of complex multiplication implies that there is a bijection from elliptic curves with complex multiplication over an imaginary quadratic field and tori obtained via quotienting by elements of its class group, after embedding $K \hookrightarrow \mathbb{C}$:

$$\left\{\text{Elliptic curves CM by } \mathcal{O}_K, \text{ up to isomorphism}\right\} \overset{1:1}{\longleftrightarrow} \left\{\mathbb{C}/a,\ a \in \mathrm{Cl}(K)\right\}.$$

This is not just a pointwise correspondence, for we have an action, given by acting on the coefficients, of $G_{\bar{\mathbb{Q}}/\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$ on the left-hand set (and in general on the set of all elliptic curves with algebraic coefficients, since $\Delta_E \neq 0$ is an algebraic relation of the coefficients, so any automorphism preserves it). Indeed, a remarkable property of the Artin isomorphism is that this action transfers to the right-hand set as an action of $\mathrm{Cl}(K)$ via

$$(\mathbb{C}/\mathfrak{a})^{\sigma} = (\mathbb{C}/\mathfrak{a}\mathfrak{b}_{\sigma}^{-1}),\ \sigma \in G_{\bar{\mathbb{Q}}/\mathbb{Q}}. \tag{70}$$

By the discussion in Section 5.1 it suffices to show that $j(E_\tau)$ is defined over H_K. We show that it actually generates it.

Clearly for $\sigma \in G_{\bar{\mathbb{Q}}/\mathbb{Q}}$ we have $\mathrm{End}(E^\sigma) = \sigma^{-1}\mathrm{End}(E) \simeq \mathrm{End}(E)$; this means that, if E has complex multiplication by $\mathcal{O}_K$, so does E_σ. Moreover, since $j(E)$ is a rational function of the coefficients, we have

$$j(E^\sigma) = j(E)^\sigma, \tag{71}$$

which implies:

$$\mathrm{Gal}(\bar{\mathbb{Q}}/H_K) \overset{\text{Artin}}{=} \{\sigma \in G_{\bar{\mathbb{Q}}/\mathbb{Q}} : (\mathbb{C}/\mathfrak{a})^\sigma = \mathbb{C}/\mathfrak{a}\} \overset{\text{Main Th. CM}}{=}$$
$$= \{\sigma \in G_{\bar{\mathbb{Q}}/\mathbb{Q}} : E^\sigma = E\} \overset{(71)}{=} \{\sigma \in G_{\bar{\mathbb{Q}}/\mathbb{Q}} : j(E)^\sigma = j(E)\} = \mathrm{Gal}(\bar{\mathbb{Q}}/K(j(E))),$$

from which we get by elementary Galois Theory that $H_K = K(j(E))$. □

Definition 5.7. *An imaginary quadratic field K (or any of its generators $\tau \in \mathcal{H}$ over $\mathbb{Q}$) satisfies the Heegner hypothesis for N if p splits in K $\forall p \mid N$.*

If τ satisfies the Heegner hypothesis for $N = \prod_p p^{v_p(N)}$, $x = \pi(\tau)$ is a Heegner point. Indeed, by our correspondence, a necessary and sufficient condition for the existence of a Heegner point relative to K is that there exist fractional ideals $\mathfrak{a} \subset \mathfrak{b}$ such that the projection $\mathbb{C}/\mathfrak{a} \longrightarrow \mathbb{C}/\mathfrak{b}$ has kernel $\mathbb{Z}/N\mathbb{Z}$, or equivalently, there exists an ideal $\mathcal{N} = \mathfrak{a}\mathfrak{b}^{-1} \subset \mathcal{O}_K$ such that $\mathcal{O}_K/\mathcal{N} \simeq \mathbb{Z}/N\mathbb{Z}$. Under the Heegner hypothesis, we can take the above ideal as $\mathcal{N} = \prod_{p|N} \mathrm{p}^{v_p(N)},\ \mathrm{p}$ *above p.*

Remark 5.7.1. *It is not too hard to see that for any N there is a $\tau \in \mathcal{H}$, and in fact infinitely many, satisfying the Heegner hypothesis for N.*

Let us “bring” Heegner points on elliptic curves:

Lemma 5.8. *Let $E/\mathbb{Q}$ be an elliptic curve of conductor N and let ϕ_E be the associated modular parametrization. For an imaginary quadratic field K satisfying the Heegner hypothesis for N, let x_K be a Heegner point on $X_0(N)$ and let:*

$$y_K = \sum_{\sigma \in Gal(H_K/K)} \phi_E(x_k)^\sigma.$$

Then $y_K \in E(K)$.

Proof. We know that x_K, and therefore $\phi_E(x_K)$, are defined on H_K. The result follows from the Galois invariance of the sum defining y_K. □

We refer to y_K as the Heegner point on E relative to K, which we will also do for the following, $\mathbb{Q}$-rational, point:

Definition 5.9. *Given E and K as above, the rational Heegner point relative to K on E is $P_K = \overline{y_K} + y_K \in E(\mathbb{Q})$.*

Much of the importance of Heegner points relies on the following result, due to Gross and Zagier ([15]):

Theorem 5.10 (Gross-Zagier formula)**.** *Let $E/\mathbb{Q}$ be an elliptic curve satisfying $L(E,1) = 0$ and let $P_K \in E(\mathbb{Q})$ be a Heegner point. There exists a constant $0 \neq c(E,K)$ such that:*

$$\hat{h}(P_K) = c(E,K)L'(E,1).$$

Since a rational point has nonzero canonical height if and only if it is nontorsion, this formula implies, by Remark 5.7.1, that rational elliptic curves with analytic rank 1 have positive rank. In the early nineties, Kolyvagin managed to prove that if an elliptic curve has analytic rank 1, its Mordell-Weil group is generated by a single point (which is not necessarily a Heegner point, but it is if we do our discussion for $E(K)$), completing the proof of what is, currently, the best way to describe our progress towards the Birch and Swinnerton-Dyer Conjecture:

The BSD Conjecture holds for all elliptic curves with analytic rank less than 2.

A consequence of the Gross-Zagier formula is the existence of an elliptic curve with analytic rank at least 3: indeed, it is sufficient to exhibit an elliptic curve E with $\omega_E = -1$ and $L'(E,1) = 0$. By Theorem 5.10, the last condition is equivalent to having a torsion Heegner point. For a rational elliptic curve, both computing its root number and proving that one of its points is torsion are straightforward in nature (which one can easily see from the results we described in Section 2 and Section 4.4), and Gross and Zagier go on to verify the required properties for the curve $E : -139y^2 = x^3 + 10x^2 - 20x + 8$.

This is not simply a fancy result, because a few years earlier Goldfeld ([12]) had proved that the existence of a single elliptic curve with $L(E,s)$ having order of vanishing at least 3 at $s = 1$ implies a full solution to Gauss's Class Number Problem:

For $h \in \mathbb{Z}^+$, determine all imaginary quadratic fields of class number h

by showing that for any $\epsilon > 0$ there exists an absolute and computable constant $C(\epsilon)$ such that:

$$h_K > C(\epsilon)(\log d)^{1-\epsilon}$$

for any imaginary quadratic field K of discriminant $-d$.

5.3 Mock Heegner points

As explained at the start of Section 5.2, the idea behind Heegner points is to make use of the stunning fact that, for some imaginary quadratic values $\tau \in \mathscr{H}$, the modular functions for $\Gamma_0(N)$ that parametrize an elliptic curve $E/\mathbb{Q}$ of conductor N assume values defined over the Hilbert class field of $\mathbb{Q}(\tau)$. The construction we presented, while definitely deserving the "archetype" status because of relying on the simplest version of the Main Theorem of complex multiplication and admitting a Gross-Zagier formula, is not the only possible way to exploit this idea. Indeed, in order to construct a Heegner point $P_K \in E(\mathbb{Q})$, we needed to assume the Heegner hypothesis for K that primes of bad reduction for E split in K.

In what follows, we describe a slightly different construction for the congruent number curves, due to Monsky, where $p = 2$ (which clearly is a bad prime for them) ramifies in K. For this reason, the resulting points are called *mock* Heegner points. They are still defined on a finite extension of H_K, and we will be able to prove *directly,* without using any Gross-Zagier type result, that in some cases their trace under a suitable Galois action is nontorsion. The proof will be given in the next section, while now we focus on constructing the points.

Recall that in Section 4.2 we constructed Eisenstein series for $\Gamma(N)$ as:

$$\sum_{\substack{c,d \in \mathbb{Z} \\ \gcd(c,d)=1 \\ (c,d) \equiv v \pmod{N}}} (cz+d)^{-k}, \tag{72}$$

where $v \in S_N = \{v = (a,b) \in (\mathbb{Z}/N\mathbb{Z})^2 : (a,b,N) = 1\}$, with different values of v possibly giving the same function (this happens if and only if the respective entries are equal or have opposite signs, which implies the second statement in Lemma 4.10, but it is not important for us). It is not too hard to show that the series obtained by omitting the $\gcd(c,d) = 1$ requirement are still weight k modular forms for $\Gamma(N)$, as we proved for the ordinary Eisenstein series G_k. Moreover, we can extend the definition to all $v \in (\mathbb{Z}/N\mathbb{Z})^2 \setminus \{0\}$. The idea now is to bring the weight down to 2: as we know, in this case the sum does not converge absolutely, but we can deal with that by adding a correction term; define:

$$e_{v,N}(z) = \sum_{(c,d) \equiv v \pmod{N}} (cz+d)^{-2} - \sum_{(A,B) \neq (0,0)} (ANz+BN)^{-2}.$$

For the interested reader, we remark that $e_{v,N}$ is equal to the Weierstrass function of Λ_z evaluated at $\frac{1}{N}v$ and rescaled by a factor of N^{-2}. Indeed, the same standard technique used for the Weierstrass functions (see [38, VI,2]) shows that it converges totally on compact sets, so it converges absolutely and defines an element of $\mathbb{C}(\mathscr{H})$.

In order to show that the $e_{v,N}$ are in $\mathrm{M}_2(N)$ we just need to show that $s_N(z) = \sum_{(A,B)\neq(0,0)}(ANz + BN)^{-2}$ formally satisfies $s_N[\gamma]_2 = s_N$ for any $\gamma \in \Gamma(N)$. We have:

$$\begin{aligned} s_N(\gamma z) = (cz+d)^2 \sum_{(A,B)\neq(0,0)} &((aAN + cBN)z + bAN + dBN)^{-2} = \\ &= (cz+d)^2 s_N(z), \end{aligned}$$

where for the last equality we just need to show that, for any $a, d \equiv 1 \pmod N$, $b, c \equiv 0 \pmod N$, $ad - bc = 1$, the map $\mathscr{b} : \mathbb{Z}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{Z}^2 \setminus \{(0,0)\}$, $(A, B) \mapsto (aA + cB, bA + dB)$ is a bijection. Two pairs $(A, B), (A', B')$ give the same first coordinate in the image if and only if there exist $m \in \mathbb{Z} : A = A' + mc,\ B = B' - ma$, and analogously for the second coordinate. Both pairs of relations holding true at the same time would mean $mc = m_1 d,\ ma = m_1 b$, but this means that either $c = m_1 = m = 0$ or $\frac{d}{c} = \frac{b}{a} \iff ad - bc = 0$, which is absurd. Therefore, we obtain injectivity. Clearly the map is surjective *in each coordinate separately* by Bezout's Theorem, because $\gcd(a, c) = \gcd(b, d) = 1$ again by the determinant condition. To see that we have an actual bijection, fix $(j, k) \in \mathbb{Z}^2 \setminus \{(0,0)\}$ and take (A, B) such that $aA + cB = j$. By the same observation as in the injectivity argument, we just need to find $m \in \mathbb{Z}$ such that $bA + bcm + dB - adm = k$. But the expression on the left is just $bA + dB - m$, and we are done.

Set $N = 8$ in the above discussion, so that $e_v = e_{v,8}$, and define:

- $C_1 = e_{(0,4)} - e_{(4,0)},\ C_2 = e_{(4,4)} - e_{(4,0)},\ C_3 = C_1 - C_2,\ C_6 = e_{(2,4)} - e_{(2,0)};$
- $C_4 = e_{(1,0)} + e_{(1,4)} - e_{(5,0)} - e_{(5,4)},\ C_5 = e_{(1,6)} + e_{(3,6)} - e_{(5,6)} - e_{(7,6)}.$

It is not too hard to see that the C_is have no zeros in $\mathscr{H}$: those defined in the first line, since $\gcd(0, 2, 4, 8) = 2$, are actually modular under the action of $\Gamma(4)$ (and of $\Gamma(2)$ if we exclude C_6), by definition. Moreover, observe that by the modularity relation we just need to show that they are nonzero on any fundamental domain F for $\Gamma(4)$. From the values of the e_v at the cusps computed in [23, p.46], one easily sees that, for $v_1 \neq v_2$ appearing in the first line, if we write e_{v_1} and e_{v_2} as linear combinations of the basis $\{e_v\}_{v\in S_8}$ constructed from the Eisenstein series, there needs to be a basis element on one side that does not appear on the other side. Therefore, this e_vs are pairwise independent; but if one of the functions in the first line had a zero at $z_0 \in F$, the two meromorphic functions $e_{v_1}(z) - e_{v_1}(z_0),\ e_{v_2}(z) - e_{v_1}(z_0) \in \mathbb{C}(X(4))$ would both lie in $\{f \in \mathbb{C}(X(4)) : f(z_0) = 0\} = H^0(-z_0)$ (since the e_vs have no poles). It is well known that $X(4)$ has genus 1, so any divisor of positive degree has trivial H^1 as a consequence of the Riemann-Roch Theorem, so applying Riemann-Roch again we get $h^0(-z_0) = 1$, which gives the desired contradiction.

The same argument works for C_4 and C_5 because the e_vs appearing in their definition are pairwise linearly independent, since they come from elements in an Eisenstein basis for $\mathrm{M}_2(8)$, and because we see that $e_{(1,0)}+e_{(1,4)}$, $e_{(5,0)}+e_{(5,4)}$, $e_{(1,6)}+e_{(3,6)}$, $e_{(5,6)}+e_{(7,6)}$ all belong in $\mathrm{M}_2(4)$ since the matrices γ that transform one term of the addition into the other according to (52) form, along with the identity, a coset of $\Gamma(8)$ in $\Gamma(4)$.

Recall, by the discussion in Section 4.1, that $X(8) \longrightarrow X(4) \longrightarrow X(2)$ are ramified coverings, where the ramification points are the cusps $0, 1, \infty$ of $X(2)$. The above choice of the C_i is necessary for the next lemma, which follows from the values taken by the e_vs at the cusps. View each function on the lowest possible level modular curve; then:

Lemma 5.11. *1. C_3^2/C_1C_2 has a double zero at 0 and simple poles at $1, \infty$;*

2. C_6^2/C_1C_2 has no zeros or poles, and therefore is constant;

3. C_4C_5/C_1C_2 has double zeros on every cusp lying over 0 and simple poles on every cusp lying over $0, 1$.

Definition 5.12. *Set $X = \frac{i}{16\sqrt{2}}C_1C_2/C_4C_5$, $Y = \frac{1}{\sqrt{2}}(C_1 + C_2)/C_6$.*

These are modular functions of weight 0 and level 8 and 4 respectively, both holomorphic on $\mathscr{H}$, and with X also nonzero there, by our previous discussion. The importance of these functions is conveyed by the following result.

Proposition 5.13. *1. The modular functions X and Y satisfy $2Y^2 = X^4 + 1$.*

2. X and Y take values on $\mathbb{R}^+$ for $z \in i\mathbb{R}$, and we have $X(-\frac{2}{z}) = X(z)^{-1}$ and $Y(-\frac{2}{z}) = Y(z)X(z)^{-2}$.

Proof. 1) The ramification degree of $X(8) \longrightarrow X(2)$ at each cusp is 4 by (45) and (46), so by Lemma 5.11 1) and 3), C_3^2/C_1C_2 and X^4 have the same order at each cusp of $X(8)$ and are elsewhere holomorphic, so their ratio is holomorphic on $X(8)$, a compact Riemann surface, and therefore constant. Comparing the Fourier expansions at ∞ again thanks to [23] gives:

$$X^4 = C_3^2/4C_1C_2 = (C_1 - C_2)^2/4C_1C_2 \Longrightarrow X^4 + 1 = (C_1 + C_2)^2/4C_1C_1. \tag{73}$$

On the other hand, Lemma 5.11 2) implies $C_6^2/C_1C_2 = 4$ by comparing the Fourier coefficients at ∞. Therefore, (73) gives:

$$X^4 + 1 = (C_1 + C_2)^2/C_6^2 = 2Y^2.$$

2) Lemma 5.11 1), along with the fact that C_1, C_2 and C_3 do not vanish on $\mathscr{H}$, tells us that any meromorphic function on $C(2)$ having a double zero at 0 and simple poles at $1, \infty$, with no other zeros or poles, is a constant multiple of C_3^2/C_1C_2. Consider the function $F(z) = \frac{\Delta(\frac{z}{2})}{\Delta(z)} \in \mathbb{C}(C(2))$. Since the Eisenstein series E_k for $\Gamma(1)$ are a linear combination of the translates of the principal Eisenstein series $E_k^{(0,1)}$ for a congruence subgroup $\Gamma(N)$ (by the respective construction as invariant elements), and since the $\Gamma(1)$-translates of $E_k^{(0,1)}$ are Eisenstein series which vanish at all cusps of $X(N)$ except for $\Gamma(1)\infty$, we can use the expansion of the level 1 Eisenstein series to compute F at $\infty \in C(2)$. Since Δ has a simple zero at $i\infty \in \mathscr{H}^*$ and $\Delta(x+iy) = \rho_x e^{-2\pi y}$, $|\rho_x| = 1$ (see the proof of Proposition 4.23), we get that $F(z)$ has a simple pole at $\infty \in C(2)$. A similar computation for the cusp 0 using (51) and $\gamma = S = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ shows that F has a simple pole also at $1 \in C(2)$ and therefore, since Δ does not vanish in $\mathscr{H}$ and the maps between the modular curves $C(N)$ ramify only at the cusps, F must have a double zero at 0 and no other zero or pole. Hence, $F = \alpha C_3^2/C_1C_2$, and looking at the Fourier coefficients at ∞ gives $\alpha = 16$, so that $X^4 = \frac{1}{64}F$. Since

$$F\left(-\frac{2}{z}\right) = \frac{\Delta(-\frac{1}{z})}{\Delta(-\frac{2}{z})} = 2^{12}\frac{\Delta(z)}{\Delta(\frac{z}{2})} = 2^{12}F(z)^{-1},$$

we have $X(-\frac{2}{z})^4 = 2^6 F(z)^{-1} = X(z)^{-4}$. By the first claim,

$$2Y\left(-\frac{2}{z}\right)^2 = \frac{X(z)^4+1}{X(z)^4} = 2\frac{Y(z)^2}{X(z)^4},$$

therefore it suffices to check that both X and Y are positive on $i\mathbb{R}$: the key is to observe that the functions e_v satisfy $e_{(a,b)} = \overline{e_{(-a,b)}}$, which follows from substituting $c = \mathbb{Z}N + a$, $c = \mathbb{Z}N - a$ in (72) and from the coprimality of c, d. This means that all the C_i are real except for C_5 which is imaginary, which tells us that X and Y are real, and we finish again by looking at their Fourier coefficients at ∞. □

Let D be an odd squarefree positive integer, $\omega = i\sqrt{2D}$, $K = \mathbb{Q}(\omega)$ and let H be the Hilbert class field of K. Recall that we have an isomorphism $\mathfrak{a} \mapsto \sigma_{\mathfrak{a}}$ between $\mathrm{Cl}(K)$ and $\mathrm{Gal}(H/K)$. As remarked at the start of the argument, we cannot construct ordinary Heegner points for K (also we are working with $X(8)$ and not $X_0(8)$, but for that we could just take the projection $X(8) \longrightarrow X_0(8)$) since 2 ramifies in it, but a more general version of the correspondence in the proof of Proposition 5.6, a version of the so-called *Shimura reciprocity,* tells us that $x = X(\omega)$ and $y = Y(\omega)$ are defined in the *Ray class field* R_8 of K, an extension of degree 2 of H. Without diving into the details, and contenting ourselves with the standard construction of Heegner points, we simply limit to say that, studying the action of $\mathrm{Gal}(R_8/H)$ on x and y, Monsky ([30, p.49]) proves that:

Proposition 5.14. *1. we have $x \in H$, and $y \in H$ or $iy \in H$ depending on whether $D \equiv 1$ or $3 \pmod 4$;*

2. let $\mathfrak{a} \in Cl(K)$ and let $I \in \mathfrak{a}$ have odd norm (the ideal norm is extended multiplicatively to the group of nonzero fractional ideals); let $a\omega + b$ and d, with $8 \mid b$, $ad > 0$, generate I over $\mathbb{Z}$ and let $\beta = \frac{a\omega+b}{d}$. Then $\sigma_{\mathfrak{a}}^{-1}(x) = \left(\frac{2}{\mathbb{N}I}\right)X(\beta)$,

$$\text{and } \begin{cases} \sigma_{\mathfrak{a}}^{-1}(y) = \left(\frac{2}{\mathbb{N}I}\right)Y(\beta) & \text{if } D \equiv 1 \pmod 4, \\ \sigma_{\mathfrak{a}}^{-1}(iy) = i\left(\frac{-2}{\mathbb{N}I}\right)Y(\beta) & \text{if } D \equiv 3 \pmod 4. \end{cases}$$

Thanks to the first statement, the second one tells us that $X(\beta)$ is in H and $Y(\beta)$ is in H or $H(i)$. Therefore, since each ideal class contains an ideal with odd norm (we are quotienting by principal ideals), Proposition 5.14 allows us to attach to each ideal class $\mathfrak{a}$ of K a point

$$P_{\mathfrak{a}} = \begin{cases} (\sigma_{\mathfrak{a}}^{-1}(x), \sigma_{\mathfrak{a}}^{-1}(y)) & \text{if } D \equiv 1 \pmod 4, \\ (i\sigma_{\mathfrak{a}}^{-1}(x), \sigma_{\mathfrak{a}}^{-1}(iy)) & \text{if } D \equiv 3 \pmod 4. \end{cases} \tag{74}$$

whose coordinates are defined over H or $H(i)$ and satisfy $2Y^2 = X^4 + 1$.

Corollary 5.14.1. *The second part of the proposition implies that $P_{\mathfrak{a}}$ does not depend on the choice of the ideal I nor of its $\mathbb{Z}$-basis.*

Corollary 5.14.2. *If $D \equiv 1 \pmod 4$ and $\mathfrak{b}$ is an ideal class, we have $P_{\mathfrak{a}}^{\mathfrak{b}} = P_{\mathfrak{b}^{-1}\mathfrak{a}}$. If $D \equiv 3 \pmod 4$ and σ is an automorphism of $H(i)$ which fixes $K(i)$, we have $P_{\mathfrak{a}}^{\sigma} = P_{\mathfrak{b}^{-1}\mathfrak{a}}$, where $\mathfrak{b}$ is the ideal class such that $\sigma|_H = \sigma_{\mathfrak{b}}$.*

Proof. This follows directly from Proposition 5.14 and (70). □

5.4 A non-Weierstrass model for E_1

We managed to attach to each ideal class of K a point, defined over a degree 2 extension of $H = H_K$, on the affine curve $\tilde{C} : 2Y^2 = X^4 + 1$. Before resorting to the technique, analogous to that used in the case of ordinary Heegner points, of exploiting the Galois action to obtain points defined over $\mathbb{Q}$ or K, we are presented with the problem that $\tilde{C}$ does not "look like" an elliptic curve, in the sense that it is not defined by a Weierstrass equation. Moreover, if we take a projective model $2Y^2Z^2 = X^4 + Z^4$ of $\tilde{C}$ by adding the point at infinity $(0 : 1 : 0)$, this point turns out to be singular: all three partial derivatives $4X^3$, $4YZ^2$, $4Y^2Z - 4Z^3$ vanish there (while the same operation shows that $\tilde{C}$ as an affine curve is nonsingular).

Observe that $\tilde{C}$ has two unbounded real components, as its involution $(x,y) \mapsto (x,-y)$ clearly shows; in order to get a nonsingular projective model, it is necessary to blow-up the curve at $(0:1:0)$, which means to add two points at infinity, one for each real component: call this model C. Then C is a branched double cover of $\mathbb{P}^1_{\mathbb{C}}$, obtained via $\pi : (X:Y:Z) \mapsto (X:Z)$: let us calculate the genus g of C by applying the Riemann-Hurwitz Theorem to this map. Let $R = \sum_{P \in \mathbb{P}^1} (e_P - 1) \cdot P$ be the ramification divisor, with e_P the ramification degree of π at P. Then we have:

$$2g-2 = 2g(\mathbb{P}^1) - 2 + \deg(R) = -4 + \sum_{\pi^{-1}(P)=(x:0:z)} (2-1) = -4+4 \Longrightarrow g = 1,$$

where the last equalities hold because the ramification points, which obviously have ramification degree 2 since π has degree 2, are those having one preimage, that is $(\pm 1 : 0 : 1)$, $(\pm i : 0 : 1)$.

All of this combined tells us that C is a nonsingular projective curve of genus 1 defined over $\mathbb{C}$, and therefore that it is a complex elliptic curve. Moreover, since $(1,1) \in C(\mathbb{Q})$, it is an elliptic curve over $\mathbb{Q}$, with $O_C = (1,1)$ as its origin. Since $(x,y) \mapsto (x^{-1}, yx^{-2})$ is an involution *that preserves O_C*, it is a group endomorphism (by our discussion in Section 2.1) of order 2, and therefore it is the multiplication by -1. This tells us that the 2-torsion $C[2]$ is its set of fixed points, $\{(\pm 1, \pm 1)\}$ (where the signs can be different). Moreover, the regular maps $(x,y) \mapsto \begin{cases} (-x,-y) \\ (-x^{-1}, yx^{-2}), \\ (x^{-1}, -yx^{-2}) \end{cases}$ are involutions, so they must be translations, by the respective images $(-1,-1)$, $(-1,1)$, $(1,-1)$ of O_C. So for example we get:

$$(1,-1) - (x,y) = (x,-y), \quad (-1,1) - (x,y) = (-x,y). \tag{75}$$

It will be important to look at the point $P_{\mathfrak{m}} \in C(H(i))$ attached to the class $\mathfrak{m}$ of the maximal ideal $m = (\omega, 2)$ of $\mathcal{O}_K$. Notice that, in our notation, $P_{\mathcal{O}} = (x,y)$ or (ix, y) depending on whether $D \equiv 1$ or $3 \pmod 4$.

Lemma 5.15. *We have:*

$$P_{\mathfrak{m}} = \begin{cases} -P_{\mathcal{O}} & \text{if } D \equiv 1 \pmod 8, \\ -P_{\mathcal{O}} + (-1,-1) & \text{if } D \equiv 5 \pmod 8, \\ P_{\mathcal{O}} + (1,-1) & \text{if } D \equiv 3 \pmod 8, \\ P_{\mathcal{O}} + (-1,1) & \text{if } D \equiv 7 \pmod 8. \end{cases} \tag{76}$$

Proof. Denote $\xi_D = \left(\frac{2}{D}\right)$ and recall its dependence on the residue class of D modulo 8. Since $\omega = \sqrt{-2D}$, dividing by $\frac{D}{2}$ shows that $\mathfrak{m}$ is in the same ideal class of (ω, D), which has odd norm D. Therefore, Proposition 5.14 2), Corollary 5.14.1, along with Proposition 5.13 2) tell us that:

$$P_{\mathfrak{m}} = \begin{cases} (\xi_D X(\frac{\omega}{D}),\ \xi_D Y(\frac{\omega}{D})) \overset{\frac{\omega}{D}=-\frac{2}{\omega}}{=} (\xi_D x^{-1},\ \xi_D y x^{-2}) & \text{if } D \equiv 1 \pmod 4, \\ (i\xi_D X(\frac{\omega}{D}),\ i\left(\frac{-2}{D}\right) Y(\frac{\omega}{D})) \underset{-1 \not\equiv_D \square}{\overset{\frac{\omega}{D}=-\frac{2}{\omega}}{=}} (i\xi_D x,\ -\xi_D y x^{-2}) & \text{if } D \equiv 3 \pmod 4, \end{cases}$$

which proves the lemma thanks to above discussion. □

Lemma 5.16. *If $D \equiv 1 \pmod 4$, we have $P_{\mathfrak{a}^{-1}} = \overline{P_{\mathfrak{a}}}$.*

Proof. Since, if $I \in \mathfrak{a}$ has odd norm and $\mathbb{Z}$-basis $\{a\omega + b, d\}$, we can take $\bar{I}$ as an odd norm representative of $\mathfrak{a}^{-1}$ having $\mathbb{Z}$-basis $\{a\omega - b, d\}$, and since X and Y satisfy the functional equation $f(-\bar{z}) = \overline{f(z)}$ (because the C_i do, as a consequence of what we observed in the proof of Lemma 5.11), the claim follows by how the points are defined. □

We are ready to bring our discussion back to the congruent number curves: let $n \equiv 5, 6$ or $7 \pmod 8$, $L = \mathbb{Q}(\sqrt{n})$ and $V_n \subset C(L)$ be the subgroup of points P that are sent to $-P$ by the involution $\sqrt{n} \mapsto -\sqrt{n}$ of L. Clearly, $C[2] \subset V_n$. Let $E = E_1 : y^2 = x^3 - x$ be our basis congruent number curve.

Lemma 5.17. *There is a $\mathbb{Q}$-isomorphism $\phi : C \xrightarrow{\sim} E$. Moreover, the groups V_n and $E_n(\mathbb{Q})$ are isomorphic.*

Proof. Let $U = \left(\frac{X^2+1}{X^2-1}\right)^2 \in \mathbb{C}(C)$. Then:

$$U - 1 = \left(\frac{2X}{X^2-1}\right)^2,\ U + 1 = \left(\frac{2Y}{X^2-1}\right)^2,$$

so $U^3 - U = U(U-1)(U+1)$ is the square of $V = \frac{4XY(X^2+1)}{(X^2-1)^3}$. Therefore, $\psi : (X, Y) \mapsto (U, V)$ defines a morphism from C to E, whose kernel clearly is $C[2]$. This means that there is another morphism $\phi : C \longrightarrow E$ satisfying $\psi = 2\phi$, which is the required isomorphism.

Let us look at the image of V_n in E: by Example 2.3, for a point $(x, y) \in E(L)$ to be transformed into its negative $(x, -y)$ by the involution of L means that $x \in \mathbb{Q}$ and that y is of the form $q\sqrt{n}$, $q \in \mathbb{Q}$. But this set corresponds exactly to $E_n(\mathbb{Q})$ under the isomorphism (7). □

Corollary 5.17.1. *If we have $P \in V_n \setminus C[2]$, then n is a congruent number.*

6 Towards a classification of congruent numbers

In this final section, building on the material we developed throughout the thesis, we outline the proofs of Theorem 1.1 and Theorem 1.2 as stated in the Introduction. In particular, the theoretical foundations for the former are laid in Section 4, while the latter will follow by a careful analysis of the constructions presented in Section 5, along with a result in Gauss's genus theory of imaginary quadratic fields.

6.1 Tunnell's Theorem

We begin by outlining how to derive Theorem 1.1 from the results of Section 4. As we specified in the Introduction, a deep and technical result of Waldspurger is needed to translate questions about central values of twisted L-functions of weight $k-1$ modular forms to questions about coefficients of weight $\frac{k}{2}$ modular forms. Waldspurger's result, dating back to 1981, is very hard to even precisely state, since more than 5 pages of definitions are needed (see [46, VIII]). Therefore, we will limit ourselves to state what Waldspurger's result says (or, to be more precise, implies) in the case of the L-function $L(E,s)$ of the elliptic curve $E: y^2 = x^3 - x$. Before doing so, we follow the work [44] of Tunnell in computing preimages under the Shimura map to the weight 2 newform associated with E.

We begin with an observation: (58) and (65) immediately tell us that the Shimura correspondence commutes with the Hecke operators, in the sense that if $f \overset{Shim}{\longrightarrow} g$ and $T_{p^2} f = \lambda_p f$, then $T_p g = \lambda_p g$. Therefore, g is also a normalized eigenform. If we have a set $\{f_i\}_{i\in I} \subset S_{\frac{k}{2}}(\tilde{\Gamma}_0(N), \chi)$ of independent eigenforms for the operators T_{p^2}, then we can extend the Shimura map by linearity to their span (in this case the image of the Shimura correspondence will clearly not consist only of normalized forms).

Let

$$g(z) = q - 2q^5 - 3q^9 + 6q^{13} + 2q^{17} + ... \in S_2(\Gamma_0(32)) \tag{77}$$

be the level 32 newform such that $L_g(s) = L(E,s)$, where the coefficients can be computed by counting points on $\mathbb{F}_p$ (or, thanks to our work in Section 3.2, by doubling the odd integer x such that $x^2+y^2 = p$ has an integer solution and $x+iy \equiv 1 \pmod{2+2i}$), and then expanding the Euler product (10). In order to apply Waldspurger's result (which we are yet to state) we surely need to lift g to the space of half-integer weight forms, that is, to find an integer N and forms $\{f_i\}_{i\in I} \subset S_{\frac{3}{2}}(\tilde{\Gamma}_0(N), \chi)$ that map to g under the Shimura correspondence. Given a squarefree integer d, let us denote $\chi_d(\cdot) = \left(\frac{d}{\cdot}\right)$.

Notice that the character χ above has to be even, for at the end of Section 4.5 we have seen that $S_{\frac{k}{2}}(\tilde{\Gamma}_0(N), \chi) = 0$ for odd χ. But χ also needs to have modulus (it would be more appropriate to say *conductor*, a notion of smallest modulus) dividing N, and to satisfy $\chi^2 = 1$, by Shimura's Theorem. If we could take $N = 128$, it is not hard to show that the only possibilities for χ are the trivial character and χ_2. Under these hypotheses, Waldspurger's work tells us that:

Proposition 6.1. *Let $\{f_i\}_{i\in I} \subset S_{\frac{3}{2}}(\tilde{\Gamma}_0(128))$, $\{f_i'\}_{i\in I'} \subset S_{\frac{3}{2}}(\tilde{\Gamma}_0(128), \chi_2)$ be maximal sets of independent eigenforms for the operators T_{p^2} that map to g under the Shimura correspondence. Then there exists explicitly computable linear combinations $\sum a_n q^n = F = \sum_I c_i f_i$, $\sum a_n' q^n = F' = \sum_{I'} c_i f_i'$ and positive constants c, c' such that for any odd squarefree positive integer n we have:*

$$L_g(\chi_n, 1) = c a_n^2, \quad L_g(\chi_{2n}, 1) = c' a_n'^2.$$

Notice that $L_g(\chi_n, 1) = L(E_n, 1)$ by (18). Therefore, we are left with the task of proving that g does have nontrivial preimage in $S_{\frac{3}{2}}(\tilde{\Gamma}_0(128), \chi_2^e)$, $e = 1, 2$, and constructing it. For the sake of brevity, we will only cover the case $e = 2$, that is, that of the trivial character: the other one is similar. We will not describe how to compute the aforementioned linear combinations, limiting ourselves to using the coefficients obtained by Tunnell. The reason for which we have chosen $N = 128$ is that, as Tunnell proved, there is no $f \in S_{\frac{3}{2}}(\tilde{\Gamma}_0(64), \chi)$ mapping to g under the Shimura correspondence.

Proposition 6.2. *There is a maximal set $\{f_1, f_2\} \subset S_{\frac{3}{2}}(\tilde{\Gamma}_0(128))$ of independent eigenforms for all Hecke operators T_{p^2} that map to g under the Shimura correspondence.*

Proof. Our strategy is the following: we explicitly write $f_1, f_2 \in S_{\frac{3}{2}}(\tilde{\Gamma}_0(128))$, prove that they are eigenforms for all the Hecke operators T_{p^2}, and compute their eigenvalues for T_9 and T_{25}, showing that they coincide. We then infer by the Shimura correspondence that there are two eigenforms g_1, g_2 (but not necessarily cusp forms, since $k = 3$) of weight 2 and level at most 64 that have those eigenvalues for T_3, T_5, and use computations of Fourier expansions of eigenforms of small level and weight 2, done by Birch and others, to find that $g_1 = g_2 = g$. We will take as a given the fact, needed to prove that $\{f_1, f_2\}$ form a maximal independent set of Shimura preimages of g (and to prove that they are eigenforms for all the T_{p^2}), that $\dim_{\mathbb{C}} S_{\frac{3}{2}}(\tilde{\Gamma}_0(128)) = 3$.

Let:

$$f(z) = \sum_{m,n\in\mathbb{Z}} (-1)^n q^{(4m+1)^2+8n^2} = q - q^9 - q^{17} + 2q^{25} + \ldots$$

It is not too hard to show that $f(z) = (\theta(z) - \theta(4z))(\theta(32z) - \frac{1}{2}\theta(8z))$, which shows that $f \in \mathrm{M}_1(\tilde{\Gamma}_0(128), \chi_2)$, since $\theta(mz) \in S_{\frac{1}{2}}(\tilde{\Gamma}_0(4m), \chi_m)$ by Proposition 4.30. It can be shown that f is actually a cusp form (see [44]). Let $\theta_m(z) = \theta(mz)$. Define:

$$f_1 = f\theta_8, \quad f_2 = f(\theta_2 - \theta_8), \; f_3 = 2f\theta_{32} - f\theta_8.$$

Then $f_1, f_2, f_3 \in S_{\frac{3}{2}}(\tilde{\Gamma}_0(128))$ since $\chi_2^2 = 1$. From those of f we easily compute the first 27 coefficients of the Fourier expansion of f_1, $f_1 + f_2 = f\theta_2$ and f_3:

$$f_1(z) = q + q^9 - 4q^{17} - 3q^{25} + ..., \tag{78}$$

$$f_1(z) + f_2(z) = q + 2q^3 + q^9 - 2q^{11} - 4q^{17} - 2q^{19} - 3q^{25} + ..., \tag{79}$$

$$f_3(z) = q - 3q^9 + 5q^{25} + ... \tag{80}$$

Let us compute the action of T_9 and T_{25} on them: for T_9, by (64) we have:

$$\begin{aligned}
&a_1(T_9 f_1) = a_9(f_1) + \left(\frac{-1}{3}\right) a_1(f_1) = 1 - 1 = 0,\\
&a_2(T_9 f_1) = a_{18}(f_1) + a_2(f_1) = 0, \; a_3(T_9 f_1) = a_{27}(f_1) + a_3(f_1) = 0,\\
&a_1(T_9(f\theta_2)) = 1 - 1 = 0, \; a_2(T_9(f\theta_2)) = 0, \; a_3(T_9(f\theta_2)) = 0,\\
&a_1(T_9 f_3) = a_9(f_3) - a_1(f_3) = -3 - 1 = -4, \; a_2(T_9 f_3) = 0, \; a_3(T_9 f_3) = 0.
\end{aligned}$$

These expansions show that the forms f_1, f_2, f_3 are independent and therefore that they span $S_{\frac{3}{2}}(\tilde{\Gamma}_0(128))$. Writing $T_9 f_i, \; i = 1, 2, 3$ as a linear combination of f_1, f_2, f_3 then immediately shows that f_1 and f_2 are eigenvectors for T_9 of eigenvalue 0 and f_3 of eigenvalue -4. For T_{25} we only do the computation for a_1 to find the eigenvalue, but the procedure is the same and just requires to write out the first 75 coefficients. Again by (64) we have:

$$\begin{aligned}
&a_1(T_{25} f_1) = a_{25}(f_1) + \left(\frac{-1}{5}\right) a_1(f_1) = 1 - 3 = -2,\\
&a_1(T_{25}(f\theta_2)) = a_{25}(f\theta_2) + a_1(f\theta_2) = 1 - 3 = -2,\\
&a_1(T_{25} f_3) = a_{25}(f_3) + a_1(f_3) = 6,
\end{aligned}$$

and we find, by the same argument as above, that f_1, f_2 are eigenvectors for T_{25} of eigenvalue -2 and f_3 of eigenvalue 6. Also note from the respective expansions that $f_1, \; f_2, \; f_1 + f_2$ are pairwise independent.

As we can see by Table 3 in [4], g is the only eigenform for T_3 and T_5 having level at most 64 and respective eigenvalues $0, -2$, so, if we can prove that f_1, f_2 are eigenvectors for *all* the Hecke operators T_{p^2}, it will follow that it is the Shimura image of f_1 and f_2. Notice that, by the above discussion, the two-dimensional subspace generated by f_1, f_2 is orthogonal to that generated by f_3.

Let ψ be the quadratic character of modulus 4 and let $g_3(z) = \sum_{n \in \mathbb{Z}} \psi(n) n q^{n^2}$. Then the same techniques that can be used to prove Proposition 4.30 show that $g_3 \in S_{\frac{3}{2}}(\tilde{\Gamma}_0(128))$ (the behavior at the cusps is dealt with in the same way of that of the theta functions). Now, by (64) we find that g_3 is a common eigenvector of the T_{p^2} (with eigenvalues $1 + p\psi(p) + \frac{\overline{\psi(p)}}{p}$ if $p \neq 2$). But then, since their first three nonzero coefficients coincide, $g_3 = f_3$, so $\langle f_1, f_2 \rangle$ is stable under the action of the T_{p^2}: we just need to prove that f_1 and f_2 are individually eigenforms for all the Hecke operators.

The idea is now simple and neat: observe that the exponents appearing in the expansion of f are all congruent to 1 (mod 8), so all exponents appearing in the expansion of f_1 are congruent to 1 (mod 8) and those in the expansion of f_2 are congruent to 3 (mod 8) since $\theta_2(z) - \theta_8(z) = \sum_{d \text{ odd}} q^{2d^2}$. But the formulas in (64) preserve these conditions, because $p^2 \equiv 1$ or $4 \pmod 8$, so if $np^2 \equiv 1, 3 \pmod 8$ then $p \neq 2$ and $n \equiv 1, 3 \pmod 8$. Finally, $\{f_1, f_2\}$ is a maximal independent subset that maps to g under the Shimura correspondence because f_3 maps to a different eigenform since it has different eigenvalues. □

By examining the precise statement of Waldspurger's result, Tunnel computed that in our case the linear combination is just $F = f_1 + f_2 = f\theta_2$. For the sake of completeness, let us note that he also proved that the constant c in Theorem 6.1 in this case is $\frac{\Omega_E}{\sqrt{16n}}$. In the "even" case of Waldspurger's Theorem, where we twist by χ_2, he found that the linear combination is $F' = f\theta_4$ and that $c' = \frac{\Omega_E}{\sqrt{8n}}$.

Let $\mu(z) = \theta(z)(\theta(32z) - \frac{1}{2}\theta(8z))$. Since, for odd n, the coefficients $a_n(f\theta_2)$ and $a_n(f\theta_4)$ are respectively equal to $a_n(\mu\theta_2)$ and $a_n(\mu\theta_4)$, which expand as:

$$(\mu\theta_2)(z) = \sum_{x,y,z \in \mathbb{Z}} \left(q^{2x^2+y^2+32z^2} - \frac{1}{2} q^{2x^2+y^2+8z^2} \right),$$

$$(\mu\theta_4)(z) = \sum_{x,y,z \in \mathbb{Z}} \left(q^{4x^2+y^2+32z^2} - \frac{1}{2} q^{4x^2+y^2+8z^2} \right),$$

we obtain Theorem 1.1.

6.2 Monsky's Theorem

Let us keep the notation of Section 5. Our goal is to present Monsky's proof of Theorem 1.2. Notice that there are *a priori* six cases for the theorem. Since the techniques involved make it necessary to treat them somewhat separately, we cover the cases of one odd prime factor and only the case of $n \equiv 5 \pmod 8$ with two odd prime factors. Indeed, the other two cases require the choice of $D \equiv 3 \pmod 4$, which means that the points are defined over $H(i)$: controlling the additional action of $\mathrm{Gal}(H(i)/H)$, together with the presence of two odd prime factors makes the proofs quite long, although the techniques used to account for these facts are perfectly akin to those we use for the four cases we examine.

Let us denote by p_i a prime $p \equiv i \pmod 8$. Moreover, let us freely interchange between $\mathrm{Gal}(H/K)$ and $\mathrm{Cl}(K)$, e.g. speaking about the (Galois) action of an ideal class. As we will see, we will not let the whole $\mathrm{Cl}(K)$ act on the $P_{\mathfrak{a}}$, or in other words, we will not sum all the points constructed in Section 5, but we will restrict to those indexed by a particular subgroup (see for example the proof of Theorem 6.5 for the reason behind this).

Before picking up from where we left in the previous section, we need to take a small detour. In his investigation of binary quadratic forms, Gauss developed much of what is now known as the genus theory of imaginary quadratic fields. The fundamental fact of the theory is that, if $0 > \Delta \equiv 0, 1 \pmod 4$, then $\Gamma(1)$ acts on the set of quadratic forms $Q(x, y) = Ax^2 + Bxy + Cy^2$ with discriminant $\Delta = B^2 - 4AC$ by $Q^\gamma(x, y) = Q(ax + by, cx + dy)$, and the quotient has finite cardinality equal to the class number of the imaginary quadratic field of discriminant Δ (Gauss called the classes arising from this action *genera*). Another important fact is that if Γ is the largest extension of K that is abelian over $\mathbb{Q}$ (and not just over K as H) and unramified at all primes (the so-called *genus field* of K), then $\prod_{2 \neq q | D} K(\sqrt{q^*}) \subset \Gamma$, where $q^* = (-1)^{\frac{q-1}{2}} q$. We are interested in determining when $\mathrm{Cl}(K)^2$ has odd order.

Let $D = \prod_{j=1}^{l} q_j$ be squarefree and let J_2 be subgroup of $\mathrm{Cl}(K)$ consisting on the classes whose square is the principal class, which are called *ambiguous classes.* By Lagrange's Theorem, $\mathrm{Cl}(K)^2$ has odd order when the only ambiguous class it contains is trivial. Clearly the classes of the ideals generated by ω and each q_j are in J_2. Gauss showed that these classes are distinct and that they generate J_2, which therefore has cardinality 2^l. Since there is a surjective homomorphism $\mathrm{Cl}(K) \longrightarrow \mathrm{Cl}(K)^2$, $\mathfrak{a} \mapsto \mathfrak{a}^2$ with kernel J_2, $\mathrm{Cl}(K)^2$ has index 2^l. Since the q_j^*s are squares in H by the previous paragraph, the subgroup of $\mathrm{Cl}(K)$ that fixes $\sqrt{q_j^*}$ for all j contains $\mathrm{Cl}(K)^2$. Gauss showed that it has index 2^l, so it equals $\mathrm{Cl}(K)^2$.

Since, by Artin reciprocity (which generalizes quadratic reciprocity, see [25]), we know that a class $\mathfrak{a}$ fixes $\sqrt{q_j^*}$ if and only if $\mathbb{N}I$ is a quadratic residue $\pmod{q_j}$ for any ideal $I \in \mathfrak{a}$ with norm coprime to q_j, this allows us to understand when $\mathrm{Cl}(K)^2$ has odd order in a lot of cases:

Example 6.3. *If $D = p_3$ or p_5 we can take $(2, \omega)$ as an ideal of norm coprime to D, and since $\left(\frac{2}{D}\right) = -1$, $Cl(K)^2$ has odd order by the above discussion.*

We split our discussion in the cases $D \equiv 1, 3 \pmod 4$; the first one will yield the cases where $n \equiv 5 \pmod 8$ has 1 or 2 odd prime factors, while the second one will give the remaining two. In both cases, our n is going to be a divisor of $2D$, and we will study the point obtained by summing the $P_{\mathfrak{a}}$ relative to ideal classes in the subgroup $G_n \subset \mathrm{Cl}(K)$ that fixes $K(\sqrt{n})$:

$$P_n = \sum_{\mathfrak{a} \in G_n} P_{\mathfrak{a}}.$$

Case 1: $\mathbf{n} \equiv \mathbf{5} \pmod{\mathbf{8}}$.

Take $n_1 \equiv 2 \pmod 8$, $(n, n_1) = 1$ and let $2D = nn_1$, so that $D \equiv 1 \pmod 4$.

Lemma 6.4. *We have $\sqrt{n} \in H$ and $\mathfrak{m} \notin G_n$.*

Proof. 1) The extension of K generated by $\sqrt{n}$ is abelian, so we just need to check that it is unramified. Since $n \equiv 5 \pmod 8$ has an even number of prime factors $p \equiv 3 \pmod 4$, $K(\sqrt{n})$ is contained in the genus field Γ of K by the above discussion, and hence in H. For the second claim, it suffices to observe that, since 2 is inert from $\mathbb{Q}$ to $\mathbb{Q}(\sqrt{n})$, then m above 2 is inert from K to $K(\sqrt{n})$, so $\sigma_{\mathfrak{m}}(\sqrt{n}) = -\sqrt{n}$ by elementary Galois theory. □

Corollary 6.4.1. *We have $2P_n \in V_n$.*

Proof. By its definition, $P_n \in C(K(\sqrt{n}))$. So by Lemma 5.16, $P_n \in \mathbb{Q}(\sqrt{n})$. By Lemma 5.15 and Corollary 5.14.2 we have $(2P_{\mathfrak{a}})^{\mathfrak{m}} = -2P_{\mathfrak{a}}$, since the points $\neq P_{\mathcal{O}}$ appearing in the RHS of (76) are in the 2-torsion, so $2P_n^{\mathfrak{m}} = -2P_n^{\mathfrak{m}}$. But $\sqrt{n}^{\mathfrak{m}} = -\sqrt{n}$ by Lemma 6.4, so we obtain $2P_n \in V_n$ by definition of V_n. □

Theorem 6.5. *Let $n = D = p_5$. Then $2P_n \notin C[2]$, and n is a congruent number.*

Proof. Suppose on the contrary that $2P_n \in C[2]$, that is, $P_n \in C[4]$. We know that $P_n \in \mathbb{Q}(\sqrt{n})$. The homomorphism ψ described explicitly in the proof of Lemma 5.17 maps $C[4]$ to $E_1[2]$, so we see that we must have $P_n \in C[2]$, otherwise its image would have 0 y-coordinate and so the x coordinate of P_n would either be imaginary or belong to $\mathbb{Q}(\sqrt{2}) \setminus \mathbb{Q}$ (we rigorously prove this in the proof of Theorem 6.10). But then $P_n^{\mathfrak{m}} + P_n = (1, 1)$ as in the proof of Corollary 6.4.1. On the other hand,

$$P_n^{\mathfrak{m}} + P_n = \sum_{G_n} (P_{\mathfrak{a}}^{\mathfrak{m}} + P_{\mathfrak{a}}) \underset{6.4}{\overset{5.15}{=}} \sum_{G_n} (-1, -1) = (-1, -1),$$

because $G_n = \mathrm{Cl}(K)^2$ since $\mathrm{Cl}(K)^2 \subset G_n$ and they both have index 2 and hence odd order by Example 6.3, giving a contradiction. Therefore, n is a congruent number by Corollary 5.17.1. □

We now turn to the case of composite n, with the goal of constructing infinitely many $n \equiv 5 \pmod 8$ that are congruent. We start with an improvement on Corollary 6.4.1:

Lemma 6.6. *If D is composite, $P_n \in V_n$.*

Proof. Observe that if D is composite, the index of $\mathrm{Cl}(K)^2$ in $\mathrm{Cl}(K)$ is divisible by 4 (from the results of Gauss described in the discussion preceding Example 6.3). Therefore, since G_n clearly has index 2 in $\mathrm{Cl}(K)$, it has even order. Therefore, by Lemma 5.15 and Lemma 6.4 we get $P_n^{\mathfrak{m}} + P_n = (1, 1)$, which proves the lemma. □

Our goal is to prove the following result:

Theorem 6.7. *Let $n = p_1p_5$. Then, provided that $\left(\dfrac{p_1}{p_5}\right) = -1$, $P_n \notin C[2]$, and n is a congruent number.*

Before the actual proof, we need a preliminary result; let us now also define the point $Q = \sum_{\mathfrak{a}\in\mathrm{Cl}(K)^2} P_{\mathfrak{a}}$, and let $C_{\mathbb{R}}^{\pm}$ denote the two real connected components of C, the one with $y > 0$ and the one with $y < 0$.

Lemma 6.8. *If $\left(\dfrac{p_1}{p_5}\right) = -1$, then $P_n - (Q^{\mathfrak{m}} + Q) \in C_{\mathbb{R}}^{-}$.*

Proof. By elementary quadratic reciprocity, the hypothesis in the lemma implies that also $\left(\dfrac{p_5}{p_1}\right) = -1$, so since $\left(\dfrac{2}{p_1p_5}\right) = -1$ we get that $\mathrm{Cl}(K)^2$ has odd order by the discussion preceeding Example 6.3 (taking as ideals those above p_1, p_5). Therefore, $Q^m + Q = (-1, -1) \in C_{\mathbb{R}}^{-}$ by Lemma 5.15. Since -2 is a quadratic residue modulo p_1, $\mathfrak{m}$ fixes $\sqrt{p_1}$ again by the paragraph before Example 6.3, so the class $\mathfrak{a}$ of (ω, p_1) fixes $i\sqrt{2}$, but then it fixes $-i\sqrt{2}\frac{\omega}{2} = \sqrt{p_1p_5}$, that is it lies in G_n. Recall that under our hypotheses, $P_{\mathfrak{a}} = \left(\dfrac{2}{p_1}\right)(X(\frac{\omega}{p_1}), Y(\frac{\omega}{p_1}))$, so by Lemma 5.13 $P_{\mathfrak{a}} \in C_{\mathbb{R}}^{+}$. Now, since G_n has index 2 and a coset representative of the nontrivial class is $\mathfrak{m}$, Lemma 5.15 tells us that P_n lies on $C_{\mathbb{R}}^{+}$, so we get the desired claim. □

Proof of Theorem 6.7. Let $n' = p_5$ (not to be confused with $n_1 = 2$). Then we have $P_{n'} \in V_{n'}$ by Lemma 6.6. Since D has two prime factors, both G_n and $G_{n'}$ contain $\mathrm{Cl}(K)^2$ as a subgroup of index 2: take $\mathfrak{a} \in G_n \setminus \mathrm{Cl}(K)^2$; then we have:

$$Q + Q^{\mathfrak{a}} = P_n \text{ and } Q + Q^{\mathfrak{m}\mathfrak{a}} = P_{n'} \tag{81}$$

because $\mathfrak{m}\mathfrak{a} \in G_{n'}$ by Lemma 6.4. Therefore,

$$P_n + P_{n'} - 2Q \in C[2] \Longrightarrow P_n - 2Q \in V_{n'}. \tag{82}$$

The point Q is defined over H and fixed by $\mathrm{Cl}(K)^2$, so it is rational over $K(\sqrt{n}, \sqrt{n'})$; Lemma 5.16 then implies that it is defined over $\mathbb{Q}(\sqrt{n}, \sqrt{n'})$. Suppose the statement of the theorem is false. Then $P_n \in C[2] \subset V_{n'} \Longrightarrow 2Q \in V_{n'}$ by (82), so $2Q$ is defined over $K(\sqrt{n'})$, and hence $\mathbb{Q}(\sqrt{n'})$, by definition of $V_{n'}$. The extension of $\mathbb{Q}(\sqrt{n'})$ generated by the 2-division of $2Q$ cannot ramify at any odd prime, but $\mathbb{Q}(\sqrt{n}, \sqrt{n'})$ ramifies at p_1, so Q is actually defined on $\mathbb{Q}(\sqrt{n'})$. But then by (81) we have:

$$P_n = Q + Q^{\mathfrak{a}} = Q + Q^{\mathfrak{m}} \Longrightarrow P_n - (Q + Q^{\mathfrak{m}}) = (1,1) \in C_{\mathbb{R}}^{+},$$

which contradicts Lemma 6.8. As before, n is a congruent number by Corollary 5.17.1. □

Case 2: $\mathbf{n \equiv 6, 7}$ (mod $\mathbf{8}$).

Let $n_1 \equiv n - 5 \pmod 8$ and let us keep the notation for D, keeping in mind that in this case we will have $D \equiv 3 \pmod 4$. As a consequence of the definition of the points $P_{\mathfrak{a}}$, whose coordinates are multiplied by i compared to those for $D \equiv 1 \pmod 4$, we have a slight change of definitions.

Definition 6.9. *$G_n \subset Cl(K)$ is the subset of all classes that fix $i\sqrt{n}$, and $V_n \subset C(K(i\sqrt{n}))$ is the subgroup of points P sent to $-P$ by complex conjugation. Moreover, let Φ be a set of representatives for $Cl(K)/\langle\mathfrak{m}\rangle$. Then $P_{n,\Phi} \overset{def}{=} \sum_{\Phi} \epsilon_{\mathfrak{a}} P_{\mathfrak{a}}$, where $\epsilon_{\mathfrak{a}}$ is 1 if $\mathfrak{a} \in G_n$ and -1 otherwise.*

We then have analogous results to those given by Lemma 6.4, Corollary 6.4.1 and Lemma 6.6 for $P_{n,\Phi}$, with the only difference that in Lemma 6.4 $\sqrt{n}$ is replaced by $i\sqrt{n}$. We omit the proofs since they are completely identical, up to accounting, in Corollary 6.4.1 and Lemma 6.6, for the action of complex conjugation, a lift of which generates $\mathrm{Gal}(H(i)/H)$. See [30] for the detailed proofs.

Remark 6.9.1. *Since $\epsilon_{\mathfrak{ma}} = \epsilon_{\mathfrak{a}}$ and $P^{\mathfrak{ma}} - P^{\mathfrak{a}} \in C[2]$, changing the set of representatives Φ changes $P_{n,\Phi}$ by a 2-torsion element, so since our results are valid "modulo 2-torsion", we can meaningfully talk about P_n by taking an arbitrary Φ.*

Theorem 6.10. *Let $D = p_3$, $n = 2p_3$. Then $P_n \notin C[2]$, so n is a congruent number.*

Proof. We can explicitly describe $C[4]$ since it maps to $E[2]$ via the homomorphism ψ in the proof of 5.17: we need $V = 0$ or (U, V) to be the point at infinity, so we have:

$$C[4] = \{O_1, O_2, \left(0, \pm\frac{\sqrt{2}}{2}\right), (\pm i, \pm 1), (\pm 1, \pm 1), \left(\frac{\pm 1 \pm i}{\sqrt{2}}, 0\right)\}, \tag{83}$$

where O_1, O_2 are the points at infinity in C. Therefore, if we suppose by contradiction that P_n is in $C[4]$, then it is rational over $\mathbb{Q}(i, \sqrt{2})$, on which a generator of $\mathrm{Gal}(H(i)/H)$ restricts as the complex conjugation, and therefore:

$$\overline{P_n} - P_n = \sum_{\Phi}(-1,-1) = (-1,-1),$$

where Φ has odd order because, by our genus theory results, $\left(\dfrac{2}{p_3}\right) = -1$ implies that $\mathfrak{m} \notin \mathrm{Cl}(K)^2$ and that $\mathrm{Cl}(K)^2$ has index 2. Now, using (75), an easy computation gives $\bar{P} - P = (1,1)$ or $(1,-1)$ for any $P \in C[4]$ thanks to (83), which is the desired contradiction. □

Lemma 6.11. *Let $n = D = p_7$. Then $P_n \in C(H)$ and if $\mathfrak{a} \in Cl(K) \setminus Cl(K)^2$, then $P_n^{\mathfrak{a}} + P_n = (-1,1)$.*

Proof. This time we have $\left(\dfrac{2}{p_7}\right) = 1$, so $\mathfrak{m} \in \mathrm{Cl}(K)^2$ and hence Φ has odd cardinality. Then, proceeding as in the proof of Theorem 6.10 we find $\overline{P_n} - P_n = (1,1)$, which gives the first assertion. Let $\mathfrak{a}$ have even order $2s$. Then $\mathfrak{a}^s = \mathfrak{m}$, for it is a nontrivial ambiguous class, and $\mathfrak{m}$ is the only one since D is prime. If Φ_0 is a set of representatives for $\mathrm{Cl}(K)/\langle\mathfrak{a}\rangle$, then we can chose $\Phi = \bigcup_{i=1}^{s} \mathfrak{a}^i \Phi_0$ as a set of representatives for $\mathrm{Cl}(K)/\langle\mathfrak{m}\rangle$. Using Φ to define P_n and letting σ be a lift of $\mathfrak{a}$ to $H(i)$ that fixes i, we can compute the image of P_n:

$$\begin{aligned} P_n^{\sigma} &= \sum_{\mathfrak{b}\in\Phi} \epsilon_{\mathfrak{b}} P_{\mathfrak{a}^{-1}\mathfrak{b}} = \epsilon_{\mathfrak{a}} \sum_{\mathfrak{a}^{-1}\Phi} \epsilon_{\mathfrak{b}} P_{\mathfrak{b}} \overset{\mathrm{Cl}(K)^2=G_n}{\underset{\mathfrak{a}^s=\mathfrak{m}}{=}} \\ &\overset{\mathrm{Cl}(K)^2=G_n}{\underset{\mathfrak{a}^s=\mathfrak{m}}{=}} -\sum_{\Phi} P_{\mathfrak{b}} - \sum_{\Phi_0}(P_{\mathfrak{b}} - P_{\mathfrak{m}\mathfrak{b}}) \overset{5.15}{\underset{\#\Phi_0 \text{ is odd}}{=}} -P_n - (-1,1), \end{aligned}$$

and we get the desired claim. □

Theorem 6.12. *Let $n = D = p_7$. Then $2P_n \notin C[2]$, and n is a congruent number.*

Proof. Suppose that $P_{p_7} = P_n$ is a 4-torsion point: then, as we saw before, it is defined on $\mathbb{Q}(i, \sqrt{2})$, but by Lemma 6.11 it is defined on H, so it is defined on $\mathbb{Q}(\sqrt{2})$. Keeping the same notation as in Lemma 6.11, $\mathfrak{a} \notin \mathrm{Cl}(K)^2$ sends $\sqrt{2}$ to $-\sqrt{2}$, so it restricts to the nontrivial element τ of $\mathrm{Gal}(\mathbb{Q}(\sqrt{2})/\mathbb{Q})$, but one verifies that $P^{\tau} = -P$ or $-P + (1,-1)$ for a point $P \in C[4] \cap \mathbb{Q}(\sqrt{2})$: indeed, (83) tells us that acting with τ affects the two points at infinity (which are opposites of each other) switching them, and six finite points: the two with $X = 0$ and the four having $Y = 0$, but the latter are not in $\mathbb{Q}(\sqrt{2})$. Since the points P with $X = 0$ are permuted, (75) tells us that this means that P is sent to $(1,-1) - P$.

Since, by Lemma 6.11, $P_n^{\mathfrak{a}} = -P_n + (-1,1)$, this yields the desired contradiction. We once again conclude by Corollary 5.17.1. □

Since the Chinese Remainder Theorem and Dirichlet's Theorem on primes in arithmetic progressions imply that there are infinitely many pairs of primes satisfying the hypotheses of Theorem 6.7 (even by fixing one of them and finding infinitely many possibilities for the other), this result and Theorems 6.5, 6.10 and 6.12 imply the four cases of Theorem 1.2 that we wanted to prove.

6.3 Further developements

As we remarked in Section 3, two recent preprints by Alexander Smith prove that:

Theorem 6.13 (Smith). *In most of the quadratic twist families over $\mathbb{Q}$, at least 50% of the curves have rank 0, and 100% have rank at most 1.*

This immediately settles a longstanding conjecture on the behavior of "erratic" congruent numbers, i.e. those coming from a root number 1 twist:

Conjecture 6.14 (Density conjecture). *In the set of congruent numbers, those congruent to $5, 6$ or $7 \pmod 8$ form a density 1 subset.*

Smith's techniques are completely novel; he uses them to control a quantity called the 2^∞-*Selmer rank* of an elliptic curve. It is conjectured that this quantity is equal to the algebraic rank, but we only know that it is at least as large. It is possible to define the 2^k-Selmer rank $r_{2^k}(E)$ for any positive integer k, and these form a nonincreasing sequence. The idea of Smith is to prove that $\{r_{2^k}(E^{(d)})\}_{k\geq 1}$, as d varies, behaves like a time homogenous Markov chain, with 0 and 1 as the only absorbing points (with probability 1). Since the terms in such a sequence have the same parity and the 2^∞-Selmer rank is defined as the limit of the sequence, the result, already known, that $r_2(E^{(d)})$ has the same probability (depending on d) of being even or odd, yields Theorem 6.13.

In the other direction, that of proving that numbers in the $5, 6$ and 7 residue classes modulo 8 are indeed congruent numbers, there has been some substantial progress by Tian ([42]), generalizing Monsky's work: he managed to prove that in each of the classes there are infinitely many congruent numbers with *any* number of prime factors.

Tian considers the (mock) Heegner point

$$P = \begin{cases} \pi_0(\frac{\omega}{8}) & \text{if } n \equiv 5 \pmod 8, \\ \pi_0(\frac{\omega+2}{8}) & \text{else,} \end{cases}$$

(ω is as in Section 5: notice the similarity with the construction in (74)) which is again defined over $H(i)$ where H is the Hilbert class field of $K = \mathbb{Q}(\omega)$. Letting P_n be the same alternate sum as in Definition 6.9, we have that $P_n \in E(\mathbb{Q}\left(\sqrt{(-1)^{\frac{n-1}{2}}n}\right)) \simeq E_n(\mathbb{Q})$, and Tian proves:

Theorem 6.15. *Let* $n = \prod_{i=0}^{s} p_i$, $p_i \equiv 1 \pmod 8$ $\forall i > 1$, *and suppose that* $[Cl(K)[4] : Cl(K)[2]]$ *is* 2 *if* $p_0 \equiv \pm 1 \pmod 8$ *and* 1 *otherwise. Then:*

$$P_n \in 2^k E_n(\mathbb{Q}) \setminus 2^{k+1} E_n(\mathbb{Q}).$$

Since the only torsion in $E_n(\mathbb{Q})$ is the 2-torsion, Theorem 6.15 immediately implies that P_n is nontorsion, and since the condition is easily seen to be satisfied by infinitely many such n with a fixed number of prime factors, the original claim follows.

As we observed above, Tian's points are also different from the "canonical" Heegner points, coming from the same field K as Monsky's, where 2 ramifies. Nonetheless, Tian uses a generalization of the Gross-Zagier formula to prove his result by inducting on the number of primes.

In both Monsky's and Tian's works, it is fundamental to control the parity of $\mathrm{Cl}(K)^2$ to find when a mock Heegner point is nontorsion, and both authors need a lot of control on the multiplicative structure of n to do so. Joining forces with Yuan and Shou-Wu Zhang, Tian managed (see [43]) to push this method to what may well be its limit, obtaining a sufficient condition for $n = \prod p_i$ to be congruent depending just on properties of its Rédei matrix $\left(\left(\frac{p_i}{p_j}\right)\right)_{i,j}$. The authors stated their belief that this produces a positive density subset of integers in residue classes $5, 6, 7 \pmod 8$ that are congruent numbers, and Smith proved it in a (yet to be published) preprint ([39]) in 2016. To the best of our knowledge, the existence of a positive proportion of twists having rank 1 in a generic quadratic twist family is still an open question; let us remark that any ϵ improvement on Heath-Brown's upper bound of 1.5 for the average analytic rank in these families would yield a proportion of twists with analytic, and therefore also algebraic, rank equal to 1 of density at least $\frac{\epsilon}{2}$.

The reason for which proving that a number is congruent is so difficult is that it requires the *construction* of some highly nontrivial object, specifically a rational point, and our techniques to do so are very limited: this is a common phenomenon in all of modern mathematics. In this case, the difficulty stems from the fact that, as S.-W. Zhang remarks in [50], Heegner points are the only tool currently available to us for this scope. Moreover, even if we wanted to pursue Tunnell's path and solve the problem by proving BSD, the consensus about the conjecture is that its difficulty lies precisely in the lack of *constructions* of Heegner-type points that are nontorsion if $L(E, s)$ has a high order zero at $s = 1$.

7 Acknowledgments

I would like to thank Professor Emmanuel Kowalski for welcoming me as a foreign student in Zürich, for suggesting the topic of this thesis and for his patience and kindness in discussing all kinds of mathematical questions and ideas with me.

I would like to thank Professor Umberto Zannier for his guidance and constant availability over the last years, and for all the opportunities of mathematical growth that he enabled me to take.

I would like to thank Davide Lombardo for all the time he invested in carefully and thoroughly reading the manuscript, for his precious comments and suggestions, and for the useful and fruitful discussions about various parts of the thesis.

I would like to thank Nirvana Coppola for going through most of this work, for her advice on the exposition and for her sincere encouragements.

I would like to thank Paola Magrone for unfalteringly supporting and tolerating me during the last period of the writing of this thesis.

References

[1] P. Autissier, D. Bonolis, and Y. Lamzouri. The distribution of the maximum of partial sums of Kloosterman sums and other trace functions. *Compositio Mathematica*, 157(7): 1610–1651, 2021.

[2] M. Bhargava and W. Ho. On average sizes of Selmer groups and ranks in families of elliptic curves having marked points. preprint, https://arxiv.org/abs/2207.03309, 2022.

[3] M. Bhargava and A. Shankar. The average size of the 5-Selmer group of elliptic curves is 6, and the average rank is less than 1. preprint, https://arxiv.org/abs/1312.7859, 2013.

[4] B. Birch and W. Kuyk. Tables on elliptic curves. in modular functions of one variable iv. lecture notes in mathematics. *Springer*, 476: 81–144, 1979.

[5] A. Brumer. The average rank of elliptic curves i. *Inventiones mathematicae*, 109: 445–472, 1992.

[6] D. Burgess. On Dirichlet characters of polynomials. *Proceedings of the London Mathematical Society*, 3-13(1): 537–548, 1963.

[7] D. Byeon and K. Jeong. Infinitely many elliptic curves of rank exactly two. *Japan Academy Proceedings Series A: Mathematical Sciences*, 92A(5): 64–66, 2016.

[8] J. Coates and A. Wiles. On the conjecture of Birch and Swinnerton-Dyer. *Inventiones mathematicae*, 39: 223–251, 1977.

[9] T. Cochrane and C. Pinner. Using Stepanov's method for exponential sums involving rational functions. *Journal of Number Theory*, 116(2): 270–292, 2016.

[10] J. Cremona. Numerical evidence for the Birch and Swinnerton-Dyer conjecture, 2011. Talk at the BSD 50th Anniversary Conference, May 2011.

[11] F. Diamond and J. Shurman. *A first course in modular forms.* Reading, Mass. : Addison-Wesley, 2005.

[12] D. Goldfeld. Conjectures on elliptic curves over quadratic fields. *Lecture Notes in Math., Springer, Berlin*, 751: 108–118, 1979.

[13] D. Goldfeld. Sur les produits partiels eulériens attachés aux courbes elliptiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 294(14): 471–474, 1982.

[14] F. Gouvea and B. Mazur. The square-free sieve and the rank of elliptic curves. *Journal of the American Mathematical Society*, 4(1): 1–23, 1991.

[15] B. Gross and D. Zagier. Heegner points and derivatives of l-series. *Inventiones mathematicae*, 84: 225–320, 1986.

[16] A. Guinand. Some Fourier transforms in prime-number theory. *Quart. J. Math., Oxford*, 18: 53–64, 1942.

[17] L. Halbeisen and N. Hungerbühler. Heron triangles and their elliptic curves. *Journal of Number Theory*, 213: 232–253, 2020.

[18] D. Heath-Brown. The size of Selmer groups for the congruent number problem, ii. *Inventiones mathematicae*, 112(2): 331–370, 1994.

[19] D. Heath-Brown. The density of rational points on curves and surfaces. *Annals of Mathematics*, 155(2): 553–595, 2002.

[20] D. Heath-Brown. The average analytic rank of elliptic curves. *Duke Mathematical Journal*, 122(3): 591–623, 2004.

[21] N. Koblitz. *Introduction to Elliptic Curves and Modular Forms.* Springer, 1993.

[22] V. Kolyvagin. Euler systems, the Grothendieck festschrift, vol. ii, progr. math. *The Grothendieck Festschrift, Vol. II, Progr. Math.*, 87: 435–483, 1990.

[23] S. Lang. *Elliptic functions.* Springer, 1973.

[24] P. Le Boudec. Average rank in families of quadratic twists: a geometric point of view. *Mathematische Annalen*, 371: 695–705, 2018.

[25] F. Lemmermeyer. *Reciprocity laws: From Euler to Eisenstein.* Springer-Verlag, 2000.

[26] C. Li. Recent developments on quadratic twists of elliptic curves. *Proceedings of the International Consortium of Chinese Mathematicians 2017*, pages 381–399, 2020.

[27] B. Mazur. Rational isogenies of prime degree. *Inventiones Mathematicae*, 44(2): 129–162, 1978.

[28] L. Merel. Bornes pour la torsion des courbes elliptiques sur les corps de nombres. *Inventiones Mathematicae (in French)*, 124(1): 437–449, 1996.

[29] F. Mertens. Ein beitrag zur analytischen zahlentheorie. *J. reine angew. Math.*, 78: 46–62, 1874.

[30] P. Monsky. Mock Heegner points and congruent numbers. *Mathematische Zeitschrift*, 204: 45–68, 1990.

[31] L. Mordell. On the rational solutions of the indeterminate equations of the third and fourth degrees. *Proc. Camb. Phil. Soc.*, 21: 179–192, 1922.

[32] M. Munsch. Character sums over squarefree and squarefull numbers. *Archiv der Mathematik*, 102(6): 555–563, 2014.

[33] F. Naccarato. Counting rational points on elliptic curves with a rational 2-torsion point. *Rendiconti Lincei - Matematica e Applicazioni*, 32(3): 499–509, 2021.

[34] J. Park, B. Poonen, J. Voight, and M. Matchett Wood. A heuristic for boundedness of ranks of elliptic curves. *Journal of the European Mathematical Society*, 21(9): 2859–2903, 2019.

[35] G. Shimura. *Introduction to the arithmetic theory of automorphic functions.* Princeton University Press, 1971.

[36] G. Shimura. On modular forms of half integral weight. *Annals of Mathematics*, 97(3): 440–481, 1973.

[37] A. Silverberg. *The distribution of ranks in families of quadratic twists of elliptic curves.* Cambridge University Press, 2010.

[38] J. Silverman. *The arithmetic of elliptic curves.* Springer, 2009.

[39] A. Smith. The congruent numbers have positive natural density. preprint, https://arxiv.org/abs/1603.08479, 2016.

[40] A. Smith. The distribution of ℓ^∞-Selmer groups in degree ℓ-twist families i. preprint, https://arxiv.org/abs/2207.05674, 2022.

[41] A. Smith. The distribution of ℓ^∞-Selmer groups in degree ℓ-twist families ii. preprint, https://arxiv.org/abs/2207.05143, 2022.

[42] Y. Tian. Congruent numbers and Heegner points. *Cambridge Journal of Mathematics*, 2(1): 117–161, 2014.

[43] Y. Tian, X. Yuan, and S.-W. Zhang. Genus periods, genus points and congruent number problem. *Asian Journal of Mathematics*, 21(4): 721–774, 2017.

[44] J. Tunnell. A classical diophantine problem and modular forms of weight 3/2. *Inventiones mathematicae*, 72: 323–334, 1983.

[45] D. Ulmer. Elliptic curves with large rank over function fields. *Annals of Mathematics*, 155(1): 295–315, 2002.

[46] J.-L. Waldspurger. Sur les coefficients de Fourier des formes modulaires de poids demi-entier. *J. Math. pures et appl.*, 60: 374–484, 1981.

[47] M. Watkins, S. Donnelly, N. Elkies, T. Fisher, A. Granville, and N. Rogers. Ranks of quadratic twists of elliptic curves. *Publications mathématiques de Besançon*, 2: 63–98, 2014.

[48] A. Weil. Sur les formules explicites de la theorie des nombres. *Izv. Akad. Nauk. (ser. Math.)*, 36: 3–18, 1972.

[49] D. Zagier. Elliptic modular forms and their applications, 2008. Part of "The 1-2-3 of Modular Forms. Lectures at a Summer School in Nordfjordeid, Norway".

[50] S.-W. Zhang. Congruent numbers and Heegner points. *Colloquiasns, Colloquium De Giorgi 2010-2012*, 4: 61–68, 2012.