Applications of Chebotarev Density Theorems to Elliptic Curves

by

Joshua Swidinsky

B.Sc., University of Lethbridge, 2021

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Mathematics
Faculty of Science

SIMON FRASER UNIVERSITY
Spring 2024

Declaration of Committee

Name:	**Joshua Swidinsky**
Degree:	**Master of Science**
Thesis title:	**Applications of Chebotarev Density Theorems to Elliptic Curves**
Committee:	**Chair:** Michael Monagan Professor, Mathematics
	Imin Chen Supervisor Professor, Mathematics
	Stephen Choi Committee Member Professor, Mathematics
	Nils Bruin Examiner Professor, Mathematics

Abstract

A natural question to ask in the study of ℓ-adic and mod-ℓ representations attached to elliptic curves over $\mathbb{Q}$ is what conditions guarantee two such representations will be isomorphic. Related to this question is if we suppose two such representations are not isomorphic, does there exist a 'certificate' which proves they are not isomorphic? One way to show that two representations are not isomorphic is to show that they have different trace values on a group element. For an elliptic curve E over $\mathbb{Q}$ and a prime p of good reduction, one defines the trace of Frobenius $a_p(E)$, which is an integer independent of ℓ and arises as a trace value. In this thesis, we are interested in giving upper bounds for the smallest prime p such that $a_p(E) \neq a_p(E')$. Serre [30] gives a classical asymptotic bound, although the constants are quite large. Recent work of Mayle-Wang [18] provides an explicit bound with smaller constants. We expand on Mayle-Wang's work and further reduce the constants appearing in their result. Both methods use explicit forms of the Chebotarev density theorem which assume the generalized Riemann hypothesis.

Keywords: Elliptic curves; Chebotarev density; representation theory

Dedication

To my family.

Acknowledgements

I must begin by thanking my supervisor, Dr. Imin Chen. Thank you for your knowledge, guidance, and patience throughout my two years as your student.

Thanks to my family, for all your support and encouragement. I owe this all to you.

To all the friends from my undergraduate and graduate days, you have made my post-secondary experience the best it could be. I could not be more grateful.

Lastly, I must thank the SFU climbing community, the most welcoming group of people I've ever had the pleasure of meeting. You all kept me sane when the work would get overwhelming, and for that I thank you.

Table of Contents

List of Tables

Chapter 1

Introduction

A powerful tool in a modern mathematician's repertoire is representation theory, the study of relating an abstract group and its elements to linear transformations. This takes a group that may be complicated and difficult to work with and instead translates it into the language of matrices, which are well-understood objects. In number theory, often one is interested in studying the absolute Galois group of $\mathbb{Q}$, denoted $G_{\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$, that being the group of automorphisms of an algebraic closure of $\mathbb{Q}$, $\bar{\mathbb{Q}}$, fixing $\mathbb{Q}$, and the representations which arise from this Galois group. Such a representation is known as a Galois representation. These representations are, unfortunately, still rather difficult to work with, and so often one wishes for some geometric object upon which $G_{\mathbb{Q}}$ may act on. A simple example of such an object is an elliptic curve, that being a smooth plane curve of the form $y^2 = x^3 + ax + b$ with a special 'point at infinity'. These representations of elliptic curves appeared famously in Andrew Wiles' proof of Fermat's Last Theorem [36] (and with Richard Taylor in [35]), the famous conjecture there are no integer solutions $x, y, z \in \mathbb{Z}$ to the equation

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

when $n \geq 3$ and $xyz \neq 0$. There is also the famous Langlands' program, a series of conjectures relating Galois representations to automorphic forms (a topic we do not explore here; one can see [17] for Langlands' original article, or [6] for a more storied approach). A natural question arising from theses conjectures is if we are given two representations, under what conditions will the two representations be isomorphic? Or, as we shall explore in this thesis, if the representations are not isomorphic, can we find tighter upper bounds on a 'certificate' which shows, as representations, they are not isomorphic?

In 1983, Gerd Faltings [12] proved the Mordell conjecture which states (in our specialized case) that an elliptic curve has only finitely many points in $\mathbb{Q}$. Serre took ideas from this and developed a method in which one can validate that two representations are isomorphic (or not isomorphic) on some finite extension of the base field $\mathbb{Q}$. In particular, if we attach ℓ-adic representations to two different elliptic curves, we can check if these representations

are isomorphic over a finite extension of $\mathbb{Q}$. The method works by finding a prime p whose Frobenius element σ_p in the absolute Galois group gives a different trace value for the two representations. Serre [30] in 1981 proved, up to a constant, an upper bound on the smallest such prime depending only on knowledge of the primes of bad reduction of the two elliptic curves, in particular he proved it is of order $O((\log n)^2(\log\log n)^{12})$. More recently, Mayle-Wang [18] has given an explicit result that is much improved compared to Serre's asymptotic result. Both results rely on the celebrated Chebotarev density theorem, a major result in number theory that is a generalization of Dirichlet's primes in arithmetic progressions. In this thesis, we first examine Serre's [30] original argument, computing explicitly the constants in his asymptotic result, then look to improving Mayle-Wang [18] in most cases when we restrict to elliptic curves over the rationals.

The new tool we employ in our work is a particular description of a quotient of the *deviation group* of a group G, denoted $\delta(G)$. Most of this work follows the work done by Chênevert [9] in his thesis, and expands on work done by Serre. Much as Serre developed a method of proving a prime distinguished two representations of elliptic curves (given in detail in Chapter 4), the deviation group serves a similar purpose. Our contribution to this work is expanding on a remark given by Chênevert [9, pg. 114] in which he states that, in the case of 2-adic representations (that is, a representation whose codomain is $\mathrm{GL}_n(\mathbb{Z}_2)$), we can, in fact, replace $\delta(G)$ with a new set that serves the same purpose as the deviation group. Since a proof of this fact was never given we have given the details here in the context of Mayle-Wang's recent work.

We give a brief overview of this thesis. In Chapter 2, we provide a relatively self-contained overview of the material required for understanding the proofs appearing in subsequent chapters, including basic algebraic number theory (discriminants of number fields, ramification of primes), elliptic curves (basic definitions, torsion points, reduction modulo a prime), representations of elliptic curves, and the Chebotarev density theorem. Then, in Chapter 3, we investigate the discriminant of a particular number field, that being the rational numbers $\mathbb{Q}$ adjoined with torsion points of an elliptic curve. Chapter 4 sees us follow Serre's method as given in [30] but with more exposition given on his method as well as explicitly working out the constants appearing in his result. Lastly, in Chapter 5, we follow Mayle-Wang [18] and their method, but in more generality, considering instead the deviation group appearing in the work of Chênevert [9].

Chapter 2

Background

Here we give a treatment of the various topics required to understand our work in subsequent chapters. We begin with an outline of basic topics from algebraic number theory (including the notion of discriminants of number fields, ramification of primes, and various useful ideals), then move on to elliptic curves, galois representations (in particular ℓ-adic representations of elliptic curves), and an introduction to the Chebotarev density theorem.

2.1 Discriminants, Ramification, and Differents

There are three main objects we make use of throughout this thesis: the discriminant of a number field, the relative discriminant, and the (relative) different. We begin with the most basic of these three objects, that being the discriminant. For a number field K, we will fix the notation $n_K = [K : \mathbb{Q}]$. First, some preliminaries from algebraic number theory.

Definition 2.1.1. Let K be a number field. The ring of integers $\mathcal{O}_K$ is defined to be the integral closure of K, that is, the set of all elements $a \in K$ which satisfy a monic polynomial $f \in \mathbb{Z}[x]$.

The ring of integers will always be a free $\mathbb{Z}$-module, thus a basis set $E \subseteq \mathcal{O}_K$ exists. There is also the notion of an integral basis:

Definition 2.1.2. Let K be a number field, and $n = n_K$. The elements $\omega_1, \ldots, \omega_n \in \mathcal{O}_K$ are called an integral basis of $\mathcal{O}_K$ over $\mathbb{Z}$ if for any element $x \in \mathcal{O}_K$ can be written uniquely as $x = \alpha_1\omega_1 + \ldots + \alpha_n\omega_n$ for $\alpha_1, \ldots, \alpha_n \in \mathbb{Z}$.

Given a number field and its ring of integers, we define the discriminant in the following way:

Definition 2.1.3. Let $K/\mathbb{Q}$ be a number field with ring of integers $\mathcal{O}_K$, and $n = n_K$. Let $\omega_1, \ldots, \omega_n$ be an integral basis for $\mathcal{O}_K$. If $\sigma_1, \ldots, \sigma_n$ are embeddings of K into $\mathbb{C}$, then we define the discriminant of K, d_K, to be the square of the determinant of the matrix whose ij-entry is given by $\sigma_i(\omega_j)$.

It is worth mentioning that the discriminant of a number field does not depend on the choice of integral basis, for the determinant of the two integral bases will differ only by a unit. A classic result gives a lower bound on the possible size of a discriminant in terms of the size of the extension.

Theorem 2.1.4. *[19, Theorem 4.3] Let $K/\mathbb{Q}$ be a number field, and $n_K = [K : \mathbb{Q}]$. Then,*

$$|d_K|^{1/2} \geq \left(\frac{\pi}{4}\right)^{n_K/2} \frac{n_K^{n_K}}{n_K!}. \tag{2.1}$$

The quantity on the right-hand side of (2.1) is known as the Minkowski bound.

Similar to the discriminant, we can define the relative discriminant of a number field L/K when K is not $\mathbb{Q}$. This will instead be an ideal of $\mathcal{O}_K$, rather than an integer. As in the definition of Definition 2.1.3, if we are given an n-tuple of elements $(x_1, \ldots, x_n)$, not necessarily a basis, then the discriminant of these elements is the square of the determinant of the matrix whose elements are $\sigma_i(x_j)$ (see [25, Chapter 2.11]). We use this to define the relative discriminant.

Definition 2.1.5. [25, Chapter 13.2, Definition 3] Let $n = [L : K]$. The relative discriminant of L/K, $\mathfrak{d}_{L/K}$, is the ideal of $\mathcal{O}_K$ generated by the discriminant of elements $(x_1, \ldots, x_n)$ running over all possible bases $\{x_1, \ldots, x_n\}$ of L/K with each $x_i \in \mathcal{O}_L$.

If the base field is $K = \mathbb{Q}$, then $\mathfrak{d}_{L/K} = (d_L)$.

Our main use of the (relative or not) discriminant of a number field is to detect ramification of primes. In order to introduce ramification, we need to understand how ideals in the ring of integers can be decomposed into a product of prime ideals. By the Fundamental Theorem of Arithmetic, elements $a \in \mathbb{Z}$ can be written uniquely as a product of prime numbers. The question becomes, Does such a unique factorization hold for the ring of integers? The answer is yes for ideals of the ring of integers, but that for an arbitrary ideal of some ring, there need not be a unique factorization. Recall that a ring R is called a Dedekind domain if it is noetherian, integrally closed, and every nonzero prime ideal is maximal [22, Chapter I, Definition 3.2]. A very useful fact is that the ring of integers is a Dedekind domain.

Theorem 2.1.6. *The ring $\mathcal{O}_K$ is a Dedekind domain.*

Proof. See Neukirch [22] Chapter I, Theorem 3.1. □

We now give a decomposition of ideals of $\mathcal{O}_K$ into prime ideals.

Theorem 2.1.7. *Every ideal I of $\mathcal{O}_K$ admits a factorization $I = \mathfrak{p}_1 \ldots \mathfrak{p}_r$ into nonzero prime ideals $\mathfrak{p}_i \subseteq \mathcal{O}_K$ which is unique up to the order of the factors.*

Proof. See Neukirch [22] Chapter I, Theorem 3.3. □

Knowing this decomposition, we can state what it means for a prime to ramify. It turns out that only a select few primes enjoy ramification, and that determining these primes can be simple if one knows the discriminant $\mathfrak{d}_{L/K}$. First, a definition.

Definition 2.1.8. [19, Chapter 3] Let L/K be a separable extension, and $\mathcal{O}_L$, $\mathcal{O}_K$ the ring of integers of L, K respectively. Let $\mathfrak{p} \subseteq \mathcal{O}_K$ be a prime ideal of $\mathcal{O}_K$. Write

$$\mathfrak{p}\mathcal{O}_L = \mathfrak{P}_1^{e_1} \dots \mathfrak{P}_s^{e_s}$$

for $e_i \geq 1$, called the ramification index, and $\mathfrak{P}_j \subseteq \mathcal{O}_L$ prime ideals. If any $e_i > 1$, then $\mathfrak{p}$ is said to ramify in L.

In order to state the following theorem, we must generalize the concept of the division of integers into the division of ideals, that is, what it means to say one ideal divides another. The following definition accomplishes this.

Definition 2.1.9. [19, Chapter 3] Let $\mathfrak{p} \subseteq \mathcal{O}_K$ and $\mathfrak{P} \subseteq \mathcal{O}_L$ be prime ideals. We say $\mathfrak{P}$ divides $\mathfrak{p}$ if $\mathfrak{P}$ occurs in the factorization of $\mathfrak{p}\mathcal{O}_L$.

Theorem 2.1.10. *Let L be a finite extension of a number field K, and let $\mathcal{O}_L$, $\mathcal{O}_K$ be the ring of integers of L and K respectively. Assume $\mathcal{O}_L$ is a free $\mathcal{O}_K$-module. Then a prime $\mathfrak{p} \subseteq \mathcal{O}_K$ ramifies in L if and only if $\mathfrak{p}$ divides $\mathfrak{d}_{L/K}$.*

Proof. See Milne [19] Chapter 3, Theorem 3.35. □

In Chapter 3, we are interested in computing the discriminant of a particular number field. Rather than use Definition 2.1.3 or Definition 2.1.5 to compute the discriminant, we will instead find the primes which ramify, as per Theorem 2.1.10, which must then occur in the factorization of $\mathfrak{d}_{L/K}$.

We now state a few useful facts we shall employ about the relative discriminant. Before we can do this, however, we must first define the norm of an ideal. Since we have a factorization of an ideal into a product of prime ideals (Theorem 2.1.7), and we wish for the norm function to be a homomorphism between the ideals of $\mathcal{O}_L$ and the ideals of $\mathcal{O}_K$, it suffices that we define the norm for prime ideals.

Definition 2.1.11. [19, Chapter 4] Let L/K be a separable extension. Let $\mathfrak{P} \subseteq \mathcal{O}_L$ be a prime ideal, and define $\mathfrak{p} = \mathfrak{P} \cap \mathcal{O}_K$ and $f(\mathfrak{P}/\mathfrak{p}) = [\mathcal{O}_L/\mathfrak{P} : \mathcal{O}_K/\mathfrak{p}]$. We define the norm of $\mathfrak{P}$ as

$$\mathrm{N}_{L/K}(\mathfrak{P}) = \mathfrak{p}^{f(\mathfrak{P}/\mathfrak{p})}.$$

For completion, note that if $I \subseteq \mathcal{O}_L$ is any ideal with decomposition $I = \prod_{i=1}^{s} \mathfrak{P}_i$, then

$$\mathrm{N}_{L/K}(I) = \mathrm{N}_{L/K}\left(\prod_{i=1}^{s} \mathfrak{P}_i\right) = \prod_{i=1}^{s} \mathrm{N}_{L/K}(\mathfrak{P}_i)$$

which we can then compute using Definition 2.1.11.

We now state three facts related to the computation of the relative discriminant. The first fact we state helps us to compute the discriminant in a tower of fields.

Proposition 2.1.12. *For a tower of fields $K \subseteq L \subseteq M$, one has*

$$\mathfrak{d}_{M/K} = \mathfrak{d}_{L/K}^{[M:L]} N_{L/K}\left(\mathfrak{d}_{M/L}\right). \tag{2.2}$$

Proof. See Neukirch [22] Chapter III, Corollary 2.10. □

The second fact describes what happens if we have two linearly disjoint fields and we wish to consider the discriminant of their composite field.

Proposition 2.1.13. *Let L/K and L'/K be two finite Galois extensions. Assume $L \cap L' = K$. Then*

$$\mathfrak{d}_{LL'/K} = \mathfrak{d}_{L/K}^{[L':K]} \mathfrak{d}_{L'/K}^{[L:K]}.$$

Proof. See Neukirch [22] Proposition (2.11). □

The final tool we introduce is the different of a number field. Again we consider the situation of a separable extension L/K, where K is a number field and $[L : K] < \infty$. As the relative discriminant is used to detect ramification of prime ideals of $\mathcal{O}_K$, the relative different will likewise be used to detect ramification of prime ideals of $\mathcal{O}_L$ in L. To define the relative different, we first need to recall the definition of a fractional ideal.

Definition 2.1.14. [25, Chapter 7.1, Definition 1] Let A be an integral domain, K its associated field of fractions, and $M \subseteq K$ an A-module. M is a fractional ideal of A if there exists $a \in A$ such that $aM \subseteq A$. In addition, a nonzero fractional ideal N is invertible if there exists N' such that $N \cdot N' = A$.

The element a above can be thought of as clearing all denominators of elements in M.

Definition 2.1.15. [25, Chapter 13.2] Let L/K be a separable extension, and $M \subseteq L$. We define the complimentary set of M with respect to a Dedekind domain R to be the set

$$M^* = \{x \in L \mid \mathrm{Tr}_{L/K}(xM) \subseteq R\}.$$

A note on notation: for a ring R, we will define $R^\times$ to be the multiplicative group.

If we consider the complimentary set $\mathcal{O}_L^*$ with respect to $\mathcal{O}_L$, it turns out this will be a fractional ideal.

Proposition 2.1.16. *$\mathcal{O}_L^*$ is a fractional ideal of L with respect to $\mathcal{O}_L$.*

Proof. See Ribenboim [25] Chapter 13.2. □

Knowing now $\mathcal{O}_L^*$ is a fractional ideal, we can define the different of a number field.

Definition 2.1.17. [22, Chapter III, Definition 2.1] We define the different of a number field, $\mathfrak{D}_{L/K}$, to be the inverse of the fractional ideal $\mathcal{O}_L^*$, that is, $\mathfrak{D}_{L/K} = (\mathcal{O}_L^*)^{-1}$. $\mathfrak{D}_{L/K}$ is an ideal of $\mathcal{O}_L$.

The following relationship exists between the relative discriminant, the norm of a number field, and the different.

Proposition 2.1.18. *Let L/K be a number field. If $\mathfrak{d}_{L/K}$ denotes the discriminant of L/K, $N_{L/K}$ the norm, and $\mathfrak{D}_{L/K}$ the different, then the following relation exists between the discriminant and the different:*

$$\mathfrak{d}_{L/K} = N_{L/K}(\mathfrak{D}_{L/K}).$$

Proof. See Neukirch [22] Chapter III, Theorem 2.9 □

2.2 Elliptic Curves

The objects central to our study are elliptic curves. These curves are of such great importance in number theory that they appear in many places, including the famous proof of Fermat's Last Theorem. For our treatment of elliptic curves, we shall specify to the plane cubic form; for a more rigorous treatment (including background material on some basics of algebraic geometry, which we do not cover here), see Silverman [33]. What follows uses Silverman [33] as a guide.

We begin with a definition of an elliptic curve.

Definition 2.2.1. An elliptic curve E over a field K is a smooth (nonsingular) plane cubic equation of the form

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{2.3}$$

where $a_1, a_2, a_3, a_4, a_6 \in K$, and with an additional extra point $\mathcal{O}$ known as the point at infinity, from which there is an algebraic group law on the set of points of E. An elliptic curve written in the form above is said to be in Weierstrass form.

It is important to note that the set of points which on our elliptic curve lie in an algebraic closure of K, $\bar{K}$. The algebraic group law consists of an addition morphism $+ : E \times E \to E$, and a negation morphism $- : E \to E$, that then satisfy all the appropriate axioms so that $E(K)$ is a group, where $\mathcal{O}$ serves as the identity element.

There are a number of quantities one can define from the coefficients of an elliptic curve. The most important for our purposes is the discriminant, $\Delta(E)$, which is defined to be

$$\Delta(E) = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6, \tag{2.4}$$

where

$$b_2 = a_1^2 + 4a_2,$$

$$b_4 = 2a_4 + a_1a_3,$$

$$b_6 = a_3^2 + 4a_6,$$

and

$$b_8 = a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2.$$

We note that we can rephrase the condition of our elliptic curve being nonsingular by specifying $\Delta(E) \neq 0$ (see [33, Chapter III.1, Proposition 1.4(i)]). Similar to how the discriminant of a number field detects the ramification of primes, the discriminant of an elliptic curves detects whether or not a given curve (over a local field) will have good or bad reduction at a prime p. We will discuss what the reduction of an elliptic curve is later.

We can also define the j-invariant, given by

$$j = c_4^3/\Delta(E), \tag{2.5}$$

with

$$c_4 = b_2^2 - 24b_4. \tag{2.6}$$

The j-invariant describes an isomorphism class of elliptic curves, that is, over $\bar{K}$, two elliptic curves will be isomoprhic if and only if they have the same j-invariant (see Silverman [33, Chapter III.1, Proposition 1.4]).

We now consider the m-torsion group of an elliptic curve. First, recall that, together with $\mathcal{O}$, the points on an elliptic curve form a group under addition. So, if $P, Q \in E(\bar{K})$ are points on our elliptic curve (that is, $P = (a, b)$ for $a, b \in \bar{K}$, and a, b satisfy the equation of definition for E), then $P+Q$ defines a new point on E (see Silverman [33, Chapter III.2] for a precise outline of the group law). The point at infinity $\mathcal{O}$ acts as identity, so $P + \mathcal{O} = P$ for all $P \in E$. With this, we can define the multiplication-by-m map,

$$[m] : E \to E, \quad P \longmapsto \underbrace{P + \ldots + P}_{m \text{ times}}.$$

If $m < 0$, then $[m](P) = [-m](-P)$, where $-P$ is understood to be the inverse of the point P (the point such that $P - P = \mathcal{O}$), and $[0](P) = \mathcal{O}$. As in group theory, we can talk about the m-torsion of our elliptic curve, that being the subgroup

$$E[m] = \{P \in E(\bar{K}) \mid [m](P) = \mathcal{O}\}.$$

It turns out that $E[m]$ is rather easy to describe–in fact, it is isomorphic to two copies of $\mathbb{Z}/m\mathbb{Z}$.

Proposition 2.2.2. *Let E be an elliptic curve and $m \in \mathbb{Z}$ with $m \neq 0$. If $m \neq 0$ in K (that is, if either $Char(K) = 0$ or if $Char(K) = p$ then $p \nmid m$), then*

$$E[m] \cong \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

Proof. See Silverman [33] Chapter III.6 Corollary 6.4(b). □

Later, we will examine the field $\mathbb{Q}(E[\ell])$, where ℓ is prime. This field consists of the rational numbers together with points (x, y) that are annihilated by multipilcation-by-ℓ map.

We also define an isogeny between two elliptic curves.

Definition 2.2.3. [33, Chapter III.4] Let E, E' be two elliptic curves. An isogeny from E to E' is a morphism

$$\phi : E \to E'$$

satisfying $\phi(O_E) = O_{E'}$, where O_E, $O_{E'}$ denotes the point at infinity for E and E' respectively. Two elliptic curves are isogenous if there is an isogeny from E to E' such that $\phi(E) \neq \{O_{E'}\}$.

If the field K is finite, say $K = \mathbb{F}_q$, then we wish to count the number of solutions $(x, y) \in \mathbb{F}_q^2$ to the elliptic curve

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

with $a_1, a_2, a_3, a_4, a_6 \in \mathbb{F}_q$. Let $N_q(E)$ denote the number of these points on our elliptic curve over the field $\mathbb{F}_q$, and note that, since $\mathcal{O}$, the point at infinity, is always a solution, $N_q(E)$ will be one larger than the number of points $(x, y) \in \mathbb{F}_q^2$ satisfying the Weierstrass equation for E. The following theorem of Hasse describes an upper bound on the size of $N_q(E)$.

Theorem 2.2.4 (Hasse)**.** *Let E be an elliptic curve defined over a finite field $\mathbb{F}_q$. Then,*

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

Proof. See Silverman [33] Chapter V.1, Theorem 1.1. □

The usefulness of Hasse's Theorem, to us, comes in estimating a quantity known as the trace of Frobenius. For a prime q, we define

$$a_q(E) = 1 + q - N_q(E), \tag{2.7}$$

to be the trace of E. We see, using Hasse's Theorem, that $|a_q(E)| \leq 2q^{1/2}$. We will make use of this simple bound later.

The name *trace of Frobenius* is appropriate, as it arises from the qth Frobenius map $\phi_q : E \to E$ given by $(x, y) \to (x^q, y^q)$. Since the endomorphisms of E, $\mathrm{End}(E)$ (the isogenies from E to itself), forms a 2-dimentional $\mathbb{Z}$-module, each endomorphism has a quadratic characteristic polynomial associated with it, and therefore a trace. The trace, therefore, of ϕ_q in fact equals $a_q(E)$ (proofs of these facts appear in Silverman [33, Chapter V.2])

We can also consider an integral elliptic curve, that is, a curve whose coefficients a_i come from the ring of integers. In particular, for an elliptic curve $E(K)$, for K a local field complete with respect to a discrete valuation ν, and with ring of integers $\mathcal{O}_K$ and uniformizer π, given in Weierstrass form as

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

with the coefficients $a_i \in K$, we can perform a substitution $(x, y) \mapsto (u^{-2}x, u^{-3}y)$ and, if we choose u such that u is divisible by a large enough power of π, this substitution results in a Weierstrass equation whose coefficients now lie in $\mathcal{O}_K$ (we have, in a sense, cleared the denominators of the coefficients a_i). Thus, we will have $\nu(\Delta(E)) \geq 0$, and since this valuation is discrete, we can pick, among all such integral Weierstrass models, the one which minimizes $\nu(\Delta(E))$.

Definition 2.2.5. [33, Chapter VII.1] Let K be a local field, complete with respect to a discrete valuation ν, and let $\mathcal{O}_K$ be the ring of integers of K. Let E be an integral elliptic curve over $\mathcal{O}_K$. A Weierstrass equation for E is called a minimal Weierstrass equation for E at ν if $\nu(\Delta(E))$ is minimized. The minimal value of $\nu(\Delta(E))$ is called the valuation of the minimal discriminant of E at ν.

Recall that the j-invariant defined in (2.5) defines an isomorphism class of elliptic curves. Given two curves E and E' belonging to the same isomorphism class, it is possible for E and E' to have different reductions. Thus, we take the minimum Weierstrass equation to be the representative of each isomorphism class; this produces a well-defined way in which we can take the reduction of E. The following proposition proves such a minimal Weierstrass equation exists for any elliptic curve E.

Proposition 2.2.6. *Let K be a local field.*

a) Every elliptic curve E over K has a minimal Weierstrass equation.

b) A minimal Weierstrass equation is unique up to a change of coordinates

$$x = u^2x' + r, \qquad y = u^3y' + u^2sx' + t,$$

with $u \in \mathcal{O}_K^*$ (the group of units of $\mathcal{O}_K$) and $r, s, t \in \mathcal{O}_K$.

Proof. See Silverman [33] Chapter VII.1 Proposition 1.3. □

Having obtained a minimal Weierstrass equation for a given elliptic curve E, we now consider its reduction modulo π a uniformizer.

Definition 2.2.7. [33, Chapter VII.2] Let K and $\mathcal{O}_K$ be as above, and let π be a uniformizer for $\mathcal{O}_K$, that is, a generator for the unique maximal ideal of $\mathcal{O}_K$. Let $\tilde{t}$ denotes the reduction of t in $\mathcal{O}_K/\pi\mathcal{O}_K$ for any $t \in \mathcal{O}_K$. If E is given in minimal Weierstrass form, with $a_1, a_2, a_3, a_4, a_6 \in \mathcal{O}_K$, then the reduction of E modulo π, denoted $\tilde{E}$, is the curve obtainted by reducing its coefficients by π, that is,

$$\tilde{E} : y^2 + \tilde{a_1}xy + \tilde{a_3}y = x^3 + \tilde{a_2}x^2 + \tilde{a_4}x + \tilde{a_6}.$$

In the definition above, since every elliptic curve has a minimal Weierstrass model, and because we began with such a model, Proposition 2.2.6 guarantees this reduction is unique. Thus, there is no ambiguity when speaking about the reduction of an elliptic curve.

We now classify the reduction of an elliptic curve into one of three types.

Definition 2.2.8. [33, Chapter VII.5] Let E be an elliptic curve, and $\tilde{E}$ the reduction of E modulo π.

a) E has good (or stable) reduction at π if $\tilde{E}$ is nonsingular.

b) E has multiplicative (or semistable) reduction if $\tilde{E}$ has a node. The reduction is said to be split if the slopes of the tangent lines at the node are in $\mathcal{O}_K/\pi\mathcal{O}_K$, otherwise it is said to be nonsplit.

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

In cases (b) and (c), we say E has bad reduction at π.

We will say an elliptic curve E is semistable if it has good or multiplicative reduction.

In addition, there are two categories of ordinary reduction. We say an elliptic curves $\tilde{E}$ over a finite field $\mathbb{F}_p$ is ordinary if $\tilde{E}(\bar{\mathbb{F}}_p)$ has non-trivial p-torsion, and is supersingular otherwise. Then, E has good ordinary reduction at a prime p if $\tilde{E}$ is smooth and ordinary, and good supersingular reduction if $\tilde{E}$ is smooth and supersingular. We can define the height of the reduction of E at a prime of good reduction, h, as

$$h = \begin{cases} 1, & E \text{ has good ordinary reduction,} \\ 2, & E \text{ has good supersingular reduction .} \end{cases} \tag{2.8}$$

It is also worth noting how reduction behaves over finite extensions.

Proposition 2.2.9 (Semistable Reduction Theorem)**.** *Let E be an elliptic curve over a local field K.*

1. *Let K'/K be an unramified extension. Then the reduction type of E over K is the same as the reduction type of E over K'.*

2. *Let K'/K be a finite extension. If E has either good or multiplicative reduction over K, then it has the same reduction type over K'.*

3. *There exists a finite extension K'/K such that E has either good or split multiplicative reduction over K'.*

Proof. See Silverman [33] Chapter VII.5, Proposition 5.4. □

From part (c) above, we say that E has potentially good reduction if E has good reduction over the finite extension K', and it has potentially multiplicative reduction if E has multiplicative reduction over K'.

Our goal is to determine if a curve E will have good or bad reduction at π using only properties of E. We have the following proposition.

Proposition 2.2.10. *Let E be an elliptic curve over a local field K given in minimal Weierstrass equation.*

a) E has good reduction if and only if $\nu(\Delta(E)) = 0$, that is, $\Delta(E) \in \mathcal{O}_K^*$. In this case, $\tilde{E}$ is an elliptic curve over $\mathcal{O}_K/\pi\mathcal{O}_K$.

b) E has multiplicative reduction if and only if $\nu(\Delta(E)) > 0$ and $\nu(c_4) = 0$, that is, $\Delta(E) \in \pi\mathcal{O}_K$ and $c_4 \in \mathcal{O}_K^*$.

c) E has additive reduction if and only if $\nu(\Delta(E)) > 0$ and $\nu(c_4) > 0$, that is, $\Delta(E), c_4 \in \pi\mathcal{O}_K$.

Proof. See Silverman [33] Chapter VII.5 Proposition 5.1 □

The above proposition tells us the discriminant determines when an elliptic curve over a local field will have good reduction or bad reduction. In order to speak of reduction for an elliptic curve over $\mathbb{Q}$, we need to introduce a global minimal model for an elliptic curve. Let K be a number field, M_K the complete set of independent absolute values on K, M_K^0 the set of nonarchimedian absolute values. Let ord_ν be the normalized valuation of an absolute value $\nu \in M_K^0$, that is, it is a valuation such that $\mathrm{ord}_\nu(K^\times) = \mathbb{Z}$.

Definition 2.2.11. [33, Chapter VIII.8] The minimal discriminant of an elliptic curve E over K, denoted $\mathcal{D}_{E/K}$, is the integral ideal of K given by

$$\mathcal{D}_{E/K} = \prod_{\nu \in M_K^0} \mathfrak{p}_\nu^{\mathrm{ord}_\nu(\Delta_\nu)},$$

where $\mathfrak{p}_\nu$ is a prime ideal in $\mathcal{O}_K$ associated to ν, and Δ_ν is the discriminant of a minimal Weierstrass model for E over the local field K_ν, complete with respect to ν.

Our goal is to determine if there is a Weierstrass model that is simultaneously minimal for each absolute value $\nu \in M_K^0$. That leads to the definition of a global minimal Weierstrass equation.

Definition 2.2.12. [33, Chapter VIII.8] A global minimal Weierstrass equation for E over K is a Weierstrass equation

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

with $a_1, a_2, a_3, a_4, a_6 \in \mathcal{O}_K$ and such that the discriminant Δ_E satisfies $\mathcal{D}_{E/K} = (\Delta(E))$.

For elliptic curves over arbitrary fields K, there might not be a global minimal model (see Silverman [33, Chapter VIII.8, Example 8.5]), however for $\mathbb{Q}$ this is the case.

Proposition 2.2.13. *Let K be a number field. If K has class number one, then every elliptic curve E over K has a global minimal Weierstrass equation. In particular, this is true for $K = \mathbb{Q}$.*

Proof. See Silverman [33] Chapter VIII.8 Proposition 8.2 and Corollary 8.3. □

Thus, for an elliptic curve E over $\mathbb{Q}$, we can construct a global minimal model for E, then speak to its reduction modulo a prime q. In particular, if $\tilde{E}$ denotes the reduced curve, and we note that $\Delta(\tilde{E}) = \Delta(E) \pmod{q}$, then noting that a curve is nonsingular if and only if $\Delta(\tilde{E}) \neq 0$, which implies $q \not| \Delta(E)$. That gives the following corollary.

Corollary 2.2.14. *Let E be an elliptic curve over $\mathbb{Q}$. Let $\Delta(E)$ be the discriminant of an elliptic curve E in global minimal Weierstrass form. Then, E has good reduction at a prime q if and only if q does not divide $\Delta(E)$.*

From this theorem, we see that the set of primes for which E has bad reduction is finite. We will denote the set of primes for which E has bad reduction as $P(E)$.

2.3 Galois Representations

This section is dedicated to giving a brief description of two particular types of representations of elliptic curves: the ℓ-adic representation, and the mod ℓ representation. We begin by introducing these representations.

First, we recall the definition of a representation. We denote by $\mathrm{GL}_n(L)$ the group of invertible $n \times n$ matrices with coefficients in L.

Definition 2.3.1. [14, Definition 3.1] A representation of a group G over a field L is a homomorphism $\rho : G \to \mathrm{GL}_n(L)$. We say the representation is then of degree n.

Alternatively, we may define a representation of a group G over a field L as a homomorphism $\rho : G \to \mathrm{GL}(V)$ where V is an n-dimensional vector space over L and $\mathrm{GL}(V)$ denotes the group of linear automorphisms of V over L [29, Chapter 1]. The difference between the two definitions is that in the former definition, we have a specified basis for the vector space V over L.

We let $\bar{K}$ denote the algebraic closure of K, and write $G_K = \mathrm{Gal}(\bar{K}/K)$ to be the absolute Galois group of K, that is, all automorphisms $\sigma : \bar{K} \to \bar{K}$ for which $\sigma|_K = id_K$. In particular, we shall study later representations of the absolute Galois group of $\mathbb{Q}$, $G_{\mathbb{Q}} = \mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q})$.

We follow a similar treatment as Silverman [33, Chapter III.7]. As before, let E be an elliptic curve over K, and $E[\ell]$ the ℓ-torsion of E. If $\sigma \in G_K$ and $P \in E[\ell]$, then notice that

$$[\ell](\sigma(P)) = \sigma([\ell](P)) = \sigma(\mathcal{O}) = \mathcal{O}$$

and so G_K acts on the points in $E[\ell]$. This then gives us a representation

$$G_K \to \mathrm{Aut}(E[\ell]) \cong GL_2(\mathbb{Z}/\ell\mathbb{Z})$$

since $E[\ell]$ is a 2-dimensional vector space over $\mathbb{Z}/\ell\mathbb{Z}$ (Proposition 2.2.2). We can go further with this construction. Instead of considering a representation into $\mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z})$, we wish to instead construct a representation into $\mathrm{GL}_2(\mathbb{Z}_\ell)$, where $\mathbb{Z}_\ell$ denotes the ℓ-adic integers. We do this by 'gluing' together the mod ℓ representations given above by using an inverse limit.

We illustrate an inverse limit through an example, that being the construction of the ℓ-adic integers, following Neukirch [22, Chapter II.1]. Given a positive integer $k \in \mathbb{Z}$, we successively apply to k (and then all subsequent quotients) the division algorithm by dividing by ℓ, obtaining a system of equations

$$\begin{aligned} k &= a_0 + \ell k_1, \\ k_1 &= a_1 + \ell k_2, \\ &\vdots \\ k_{n-1} &= a_{n-1} + \ell k_n, \\ k_n &= a_n. \end{aligned}$$

Then, writing $k = \sum_{i=0}^{n} a_i \ell^i$ we obtain the ℓ-adic expansion of k. Allowing for negative integers, we turn our finite sum into a series $k = \sum_{i=0}^{\infty} a_i \ell^i$. This is an ℓ-adic integer, and the set of all such ℓ-adic integers is denoted by $\mathbb{Z}_\ell$. However, the motivation for a separate definition arrises when one attempts to show $\mathbb{Z}_\ell$ is a ring under the normal operations of addition and multiplication. Rather, we wish to view the ℓ-adic integers as a sequence of residue classes. If $k = \sum_{i=0}^{\infty} a_i \ell^i$ and $s_n = \sum_{i=0}^{n-1} a_i \ell^i \in \mathbb{Z}$, then we let $\bar{s_n} = s_n \pmod{\ell^n} \in$

$\mathbb{Z}/\ell^n\mathbb{Z}$. Each term in this sequence lies in a different ring, but we get the projection maps $\lambda_i : \mathbb{Z}/\ell^{i+1}\mathbb{Z} \to \mathbb{Z}/\ell^i\mathbb{Z}$ defined by $\bar{s} \to \bar{s} \pmod{\ell^i}$ so that $\lambda_i(\bar{s}_{i+1}) = \bar{s}_i$. This gives us the projections

$$\mathbb{Z}/\ell\mathbb{Z} \xleftarrow{\lambda_1} \mathbb{Z}/\ell^2\mathbb{Z} \xleftarrow{\lambda_2} \mathbb{Z}/\ell^3\mathbb{Z} \xleftarrow{\lambda_3} \ldots \xleftarrow{\lambda_n} \mathbb{Z}/\ell^{n+1}\mathbb{Z} \xleftarrow{\lambda_{n+1}} \ldots.$$

We then have

$$\varprojlim_n \mathbb{Z}/\ell^n\mathbb{Z} = \{(x_m)_{m=1}^{\infty} \mid x_m \in \mathbb{Z}/\ell^m\mathbb{Z} \text{ and } \lambda_m(x_{m+1}) = x_m\}.$$

This is the projective limit, which we denote by $\varprojlim_n \mathbb{Z}/\ell^n\mathbb{Z}$. The following proposition says that this construction does in fact define the ℓ-adic integers.

Proposition 2.3.2. *Associating to every ℓ-adic integer $k = \sum_{i=0}^{n} a_i\ell^i$ the sequence $(\bar{s}_n)_{n=1}^{\infty}$ as above gives a bijection $\mathbb{Z}_\ell \to \varprojlim_n \mathbb{Z}/\ell^n\mathbb{Z}$.*

Proof. See Neukirch [22] Chapter II, Proposition 1.3 □

An important detail in this construction is that this inverse limit was done with respect to the sequence $(\bar{s}_n)_{n=1}^{\infty}$, and the maps λ_i act in this 'inverse' direction.

We mimic the above construction of $\mathbb{Z}_\ell$, this time gluing together the groups $E[\ell^n]$. This produces an object known as the Tate module.

Definition 2.3.3. [33, Chapter III.7] Let E be an elliptic curve over a number field K and $\ell \in \mathbb{Z}$ a prime, with $\mathrm{char}(K) \neq \ell$. The ($\ell$-adic) Tate module of E is the group

$$T_\ell(E) = \varprojlim_n E[\ell^n],$$

where the inverse limit is taken with respect to the multiplication by ℓ maps

$$E[\ell^{n+1}] \to E[\ell^n], \quad P \mapsto \ell P.$$

We note that for the group $E[\ell^n]$, the points annihilated by the multiplication-by-ℓ^n map lie in $\bar{K}$. Notice, also, that since $P \in E[\ell^{n+1}]$, the point ℓP (that being the point P added to itself ℓ times) must lie in $E[\ell^n]$. This map, then, works in an "inverse" way as the maps λ_i did in the previous example.

When comparing the construction of the Tate module to the ℓ-adic integers, the group $E[\ell^n]$ corresponds to $\mathbb{Z}/\ell^n\mathbb{Z}$, whereas the maps λ_i now correspond to maps $E[\ell^{n+1}] \to E[\ell^n]$.

Notice that the action of the absolute Galois group G_K commutes with the reduction maps on each torsion group $E[\ell^n]$, so G_K also acts on $T_\ell(E)$. This action is $\mathbb{Z}_\ell$-linear as E is defined over K. Thus, we get a representation

$$G_K \to \mathrm{Aut}(T_\ell(E)) \cong \mathrm{GL}_2(\mathbb{Z}_\ell).$$

Since each $E[\ell^n]$ is a $\mathbb{Z}/\ell^n\mathbb{Z}$ module, when we take the inverse limit we see that $T_\ell(E)$ is a $\mathbb{Z}_\ell$-module. This leads to the following definition:

Definition 2.3.4. [33, Chapter III.7] The ℓ-adic representation of G_K attached to an elliptic E over a number field K is the homomorphism

$$\rho_{E,\ell} : G_K \to \mathrm{GL}_2(\mathbb{Z}_\ell).$$

We also have the representation

$$\bar{\rho}_{E,\ell} : G_K \to \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z}).$$

If p is a prime of good reduction, and σ_p denotes the Frobenius element at p of G_K (that is, the automorphism is defined as $\sigma_p(x) = x^p$), then we have (from Serre [30, pg. 188])

$$\operatorname{tr} \rho_{E,\ell}(\sigma_p) = a_p(E) \text{ and } \det \rho_{E,\ell}(\sigma_p) = p. \tag{2.9}$$

and

$$\operatorname{tr} \bar{\rho}_{E,\ell}(\sigma_p) \equiv a_p(E) \pmod{\ell} \text{ and } \det \bar{\rho}_{E,\ell}(\sigma_p) \equiv p \pmod{\ell}. \tag{2.10}$$

Proofs for (2.9) and (2.10) can be found in Silverman [33, Chapter V, Chapter VII].

Given a second elliptic curve E', with the analogous representations $\rho_{E',\ell}$ and $\bar{\rho}_{E',\ell}$, we wish to determine if $\rho_{E,\ell}$ and $\rho_{E',\ell}$ are isomorphic. By isomorphic representations, we shall mean:

Definition 2.3.5. [29, Chapter 1] Let $\rho_1 : G \to \mathrm{GL}(V_1)$ and $\rho_2 : G \to \mathrm{GL}(V_2)$ be two representations of a group G. We say ρ_1 is isomorphic to ρ_2 and write $\rho_1 \cong \rho_2$ if there exists $\phi : V_1 \to V_2$ such that ϕ is an isomorphism between V_1 and V_2 and is G-invariant, that is, $\phi \circ \rho_1(g) = \rho_2(g) \circ \phi$ for all $g \in G$.

The work we do in this thesis concerns explicit forms of the isogeny theorem, which states two elliptic curves are isogenous if and only if their Tate modules are isomorphic. More specifically, we have the following statement from Silverman (although the proofs are due to Tate [34] and Faltings [12]):

Theorem 2.3.6 (Isogeny Theorem)**.** *Let $\ell \neq \mathrm{char}(K)$. Denote by $\mathrm{Hom}_K(E, E')$ the group of isogenies from E to E' defined over K, and similarly the set $\mathrm{Hom}(T_\ell(E), T_\ell(E'))$ to be the group of $\mathbb{Z}_\ell$-linear maps from $T_\ell(E)$ to $T_\ell(E')$ that commute with the action of $G_{\bar{K}/K}$. Then, the natural map*

$$\mathrm{Hom}_K(E, E') \to \mathrm{Hom}(T_\ell(E), T_\ell(E'))$$

is an isomorphism in the following two cases:

a) K is a finite field (Tate [34]),

b) K is a number field (Faltings [12]).

In his seminal paper, Faltings [12] proved the following (which is taken from a translation of his paper found in [10]):

Theorem 2.3.7. *Let A and A' be abelian varieties over a number field K. Let N_v be the inertial degree of the place v, F_v a Frobenius element at v, and I_v the inertial group at v. Define*

$$L_v(s, A) = \det(1 - (N_v)^s \cdot F_v \mid T_\ell(A)^{I_v})^{-1},$$

and $L(s, A) = \prod_v L_v(s, A)$ the associated L-series of A, with the product running over all but finitely many places v of K. Then the following are equivalent:

i) A and A' are isogenous.

ii) $T_\ell(A) \bigotimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell \cong T_\ell(A') \bigotimes_{\mathbb{Z}_\ell} \mathbb{Q}_\ell$, where $T_\ell(A)$ and $T_\ell(A')$ are the Tate modules of A and A' respectively, and $\mathbb{Q}_\ell$ is the field of fractions of $\mathbb{Z}_\ell$.

iii) $L_v(s, A) = L_v(s, A')$ for all but finitely many places v of K.

iv) $L_v(s, A) = L_v(s, A')$ for all places v of K.

Proof. See Faltings [12, Corollary 2]. □

We make two quick notes. First, there is a definition for $L_v(s, A)$ when A has bad reduction (see Faltings [12]). Second, in the case of elliptic curves over $\mathbb{Q}$, Silverman [33, Chapter C.16] gives

$$L(s, E) = \prod_{p \notin P(E)} L_p(s, E)^{-1}$$

with

$$L_p(s, E) = 1 - a_p(E)p^{-s} + p^{1-2s}$$

for a prime p of good reduction. Thus, $L_p(s, E) = L_p(s, E')$ as a function of s if and only if $a_p(E) = a_p(E')$.

For elliptic curves, Serre gives the following proposition:

Proposition 2.3.8. *[27, pg. IV-15] Let E and E' be elliptic curves over a field K. The following are equivalent:*

a) The Galois representations $\rho_{E,\ell}$ and $\rho_{E',\ell}$ are isomorphic for all ℓ.

b) The Galois representations $\rho_{E,\ell}$ and $\rho_{E',\ell}$ are isomorphic for one ℓ.

c) We have $a_p(\tilde{E}) = a_p(\tilde{E'})$ for all p such that E, E' have good reduction at p (recalling that $\tilde{E}$ denotes the reduction of E at p).

d) For a set of places of K of density one, we have $a_p(\tilde{E}) = a_p(\tilde{E}')$.

Proof. See Serre [27, pg. IV-15]. □

From here, Serre [30] states two cases that we can consider in regards to the ℓ-adic representations:

a) If the ℓ-adic representations $\rho_{E,\ell}$ and $\rho_{E',\ell}$ are isomorphic for all primes ℓ, then the sets $P(E)$ and $P(E')$ are equal, and $a_p(E) = a_p(E')$ for all $p \notin P(E)$. This is exactly the statement of Proposition 2.3.8, in particular this comes from part (c) after noting that $a_p(E) = a_p(\tilde{E})$.

b) If there exists a prime ℓ_0 such that ρ_{E,ℓ_0} and ρ_{E',ℓ_0} are not isomorphic, then the set of primes p for which $a_p(E) \neq a_p(E')$ has positive density (see Definition 2.4.4 for a precise definition of density), in particular it is infinite. This is given by the negation of Proposition 2.3.8, in particular the negation of the statement (d), that being a set of primes for which $a_p(E) \neq a_p(E')$.

We shall be interested in studying the second case, that being the negation of Proposition 2.3.8. We will assume the representations are not isomorphic, so that the set of primes p such that $a_p(E) \neq a_p(E')$ is infinite. We then work to give an upper bound on the smallest prime in this set.

In relation to representations, we also make short use of the inertia group. We have the following two definitions.

Definition 2.3.9. [22, Definition 9.2] Let L/K be a finite Galois extension, with Galois group G. Let $\mathfrak{P} \subseteq \mathcal{O}_L$ be a prime ideal. Then, the subgroup $G_{\mathfrak{P}}$ of G defined by

$$G_{\mathfrak{P}} = \{\sigma \in G \mid \sigma(\mathfrak{P}) = \mathfrak{P}\}$$

is called the decomposition group of $\mathfrak{P}$ over K.

Definition 2.3.10. [22, Definition 9.5] Let $k(\mathfrak{P}) = \mathcal{O}_L/\mathfrak{P}$ and $k(\mathfrak{p}) = \mathcal{O}_K/\mathfrak{p}$. The inertia group of $\mathfrak{P}$ over K, $I_{\mathfrak{P}} \subseteq G_{\mathfrak{P}}$, is the kernel of the homomorphism

$$G_{\mathfrak{P}} \to \mathrm{Gal}(k(\mathfrak{P})/k(\mathfrak{p})).$$

2.4 Prime Number Theorem and Chebotarev Density Theorem

The results we give in this thesis rely on Chebotarev's Density Theorem. We begin first with a special case of Chebotarev's Density Theorem, the famous Prime Number Theorem.

First, we must establish what it means for two functions to be asymptotically equivalent.

Definition 2.4.1. [1, Chapter 3.2] Let f and g be two functions. We say f is asymptotically equivalent to g and write $f(x) \sim g(x)$ if $\lim_{x\to\infty} f(x)/g(x) = 1$.

With the definition of asymptotic equivalence, it is now straight-forward to give the Prime Number Theorem. First, we give the definition of three various prime counting functions.

Definition 2.4.2. [1, Chapter 4.2] We denote by $\pi(x)$ the prime counting function, given explicitly as

$$\pi(x) = \sum_{p\leq x} 1. \tag{2.11}$$

We denote by $\theta(x)$ and $\psi(x)$ the Chebyshev functions, with explicit formulas

$$\theta(x) = \sum_{p\leq x} \log p \tag{2.12}$$

and

$$\psi(x) = \sum_{k\in\mathbb{N}, p^k\leq x} \log p. \tag{2.13}$$

The Prime Number Theorem describes the asymptotic behavior of $\pi(x)$, $\theta(x)$, and $\psi(x)$.

Theorem 2.4.3 (Prime Number Theorem)**.** *Let $\pi(x)$, $\theta(x)$, and $\psi(x)$ be as above. One has*

$$\pi(x) \sim \mathrm{Li}(x) = \int_2^x \frac{dx}{\log x}$$

and

$$\theta(x) \sim x.$$

and

$$\psi(x) \sim x.$$

Proof. See [1, Chapter 13] or [23]. □

It is worth noting that all three asymptotics given above are equivalent; thus, it suffices to show the validity of one for all three to be true. Hadamard and de la Vallée Poussin proved the theorem, independently, in 1896; their proof required the use of the Riemann Zeta function and complex analysis.

We now introduce the Chebotarev Density Theorem. By density, we shall mean:

Definition 2.4.4. [22, Chapter 13, Definition 13.1] Let M be a set of prime ideals of K. The limit

$$d(M) = \lim_{s\to 1^+} \frac{\sum_{\mathfrak{p}\in M} \mathrm{N}_{L/K}(\mathfrak{p})^{-s}}{\sum_{\mathfrak{p}} \mathrm{N}_{L/K}(\mathfrak{p})^{-s}},$$

provided it exists, is called the Dirichlet density of M.

The Chebotarev Density Theorem makes an assertion about the *Artin symbol*. We define this next, following [20].

Let L/K be a finite extension of number fields, with galois group G. For every prime ideal $\mathfrak{P} \subseteq \mathcal{O}_L$, there is a Frobenius element $\sigma \in G$ such that $\sigma(\mathfrak{P}) = \mathfrak{P}$, and for every element $a \in \mathcal{O}_L$, we have

$$\phi_{\mathfrak{p}}(a) \equiv a^q \pmod{\mathfrak{P}}$$

with q equal to the number of elements in the residue field $\mathcal{O}_K/\mathfrak{p}$, with $\mathfrak{p} = \mathfrak{P} \cap \mathcal{O}_L$. When $\mathfrak{P}$ is unramified, the element σ is unique, and is denoted $(\frac{L/K}{\mathfrak{P}})$.

Definition 2.4.5. [20, pg. 8] Let L/K be a finite extension of number fields with galois group G, and $\mathfrak{p} \subseteq \mathcal{O}_K$ an unramified prime ideal of K. If $\mathfrak{P} \subseteq \mathcal{O}_L$ is a prime ideal above $\mathfrak{p}$, then the Artin symbol of $\mathfrak{P}$, $(\frac{L/K}{\mathfrak{P}})$, is equal to the Frobenius element $\sigma \in G$ as above. For the prime $\mathfrak{p}$ in K, the Artin symbol is defined as the conjugacy class

$$\left(\frac{L/K}{\mathfrak{p}}\right) = \left\{\left(\frac{L/K}{\mathfrak{P}'}\right) \mid \mathfrak{P}' \cap \mathcal{O}_L = \mathfrak{p}\right\} \subseteq G.$$

We define $P_{L/K}(\sigma)$, for $\sigma \in G$, to be the set

$$P_{L/K}(\sigma) = \left\{\mathfrak{p} \subseteq \mathcal{O}_K \mid \exists \mathfrak{P} \subseteq \mathcal{O}_L \text{ such that } \left(\frac{L/K}{\mathfrak{P}}\right) = \sigma\right\} \tag{2.14}$$

The Chebotarev Density Theorem describes the density of the set $P_{L/K}(\sigma)$.

Theorem 2.4.6 (Chebotarev Density Theorem)**.** *[22, Theorem 13.4] Let L/K be a Galois extension with Galois group G. Then, for every $\sigma \in G$, the set $P_{L/K}(\sigma)$ has positive density, and it is given by*

$$d(P_{L/K}(\sigma)) = \frac{|[\sigma]|}{|G|},$$

where

$$[\sigma] = \{\tau\sigma\tau^{-1} \mid \tau \in G\}.$$

It is not quite clear how Theorem 2.4.6, in its current form, is related to the Prime Number Theorem. To mould it into an analytic form, we define a new prime counting function, $\pi_C(x, L/K)$, for a conjugacy class C of the galois group of L/K, to be the function

$$\pi_C(x, L/K) = \left|\left\{\mathfrak{p} \mid \mathfrak{p} \text{ unramified in } L, \left(\frac{L/K}{\mathfrak{p}}\right) = C, \mathrm{N}_{K/\mathbb{Q}}\mathfrak{p} \leq x\right\}\right|$$

(see [16]).

Theorem 2.4.7 (Chebotarev Density Theorem)**.** *Let $\pi_C(x, L/K)$ be as above. Then,*

$$\pi_C(x, L/K) \sim \frac{|C|}{|G|}\,\mathrm{Li}(x).$$

In this form, it is much more clear how Chebotarev Density Theorem relates to the Prime Number Theorem. In particular, if we take $L = K = \mathbb{Q}$, then there is only one conjugacy class, of size one, and $|G| = 1$, so we obtain exactly the Prime Number Theorem.

Effective versions of Chebotarev's Density Theorem exist as well, which shall be our primary application. In essence, we shall be applying results in which the constants are explicitly computable in terms of the discriminants of L and K, as well as the size of each extension over $\mathbb{Q}$. The first of these was given by Lagarias and Odlyzko [16]. Their result (as well as ours) are conditional results, and rely on the validity of the famous Generalized Riemann Hypothesis (abbreviated GRH). We introduce briefly this conjecture.

The Riemann zeta function, $\zeta(s)$, for $s \in \mathbb{C}$, is defined to be the series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

(see [13, Chapter 1]). For $\operatorname{Re} s > 1$, the series is well-defined and converges. For $\operatorname{Re} s < 1$, we can analytically continue $\zeta(s)$ to the entirety of the complex plane, leaivng only a simple pole at $s = 1$ (see [13, Chapter 1, Theorem 1.2] for a detailed look at this continuation). An important point about this extension of the domain of $\zeta(s)$ to all of $\mathbb{C}$ is that we may now consider its zeros, that is, those values $s_0 \in \mathbb{C}$ such that $\zeta(s_0) = 0$. A class of zeros, known as the 'trivial zeros', occur at the negative even integers $\{-2, -4, -6, \ldots\}$; thus, $\zeta(-2k) = 0$ for $k \geq 1$ an integer (see [3, Chapter 1]). Other than the trivial zeros, $\zeta(s) \neq 0$ for $0 < \operatorname{Re} s < 1$; all other zeros occur within this region, known as the critical strip. The Riemann Hypothesis, arguably the most famous unsolved problem in mathematics, claims all zeros within the critical strip must lie on the line $\operatorname{Re} s = 1/2$.

Conjecture (Riemann Hypothesis)**.** If $s_0 \in \mathbb{C}$ is a nontrivial zero of $\zeta(s)$, then $\operatorname{Re} s_0 = 1/2$.

While this remains unproven, much work has gone into understanding further the distribution of zeros within the critical strip. For example, zero-free regions describe portions of the complex plane in which no zeros of $\zeta(s)$ exist, and extending these into the critical strip help understand the distribution of zeros further (see [13, Chapter 6] for a survey of some zero-free region results). Likewise, zero density estimates detect possible zeros in given regions in the critical strip (seee [13, Chapter 11]); of course, if the Riemann Hypothesis is true, then the density of zeros in the critical strip is zero outside of the line $\operatorname{Re} s = 1/2$. Lastly, numerical verification heights exist, that is, numbers H_0 such that for any $s \in \mathbb{C}$ that is a non-trivial zero satisfying $\operatorname{im} s \leq H_0$, then $\operatorname{Re} s = 1/2$ (see [3, Chapter 1, Table 2.2] for an overview of the history of verification heights). Many other equivalent statements exist for the Riemann Hypothesis (see [3, Chapter 10]).

As the Riemann Hypothesis describes the behavior of zeros in the critical strip for $\zeta(s)$, the Generalized Riemann Hypothesis concerns the distribution of zeros of L-functions. These are defined in terms of Dirichlet characters, for which we now give a definition:

Definition 2.4.8. [4, Chapter 12.2] A Dirichlet character $\chi(n)$ is a character of the group $(\mathbb{Z}/m\mathbb{Z})^*$ for $m \geq 1$ that is zero whenever $(n, m) \neq 1$. A Dirichlet character is multiplicative, so $\chi(ab) = \chi(a)\chi(b)$ whenever a, b are coprime, and is periodic.

Definition 2.4.9. [4, Chapter 12.2] Let χ be a character of $(\mathbb{Z}/m\mathbb{Z})^*$. The Dirichlet L-function associated to χ is the series

$$L(s, \chi) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}.$$

We note the similarity of $L(s, \chi)$ to $\zeta(s)$; in particular $\zeta(s)$ is an L-function with the trivial character $\chi(n) = 1$ for all $n \in \mathbb{Z}$. Thus, the statement of the Generalized Riemann Hypothesis should not be too surprising:

Conjecture (Generalized Riemann Hypothesis)**.** Let $L(s, \chi)$ be an L-function associated to the Dirichlet character χ. If $s_0 \in \mathbb{C}$ is a nontrivial zero of $L(s, \chi)$, then $\operatorname{Re} s_0 = 1/2$.

All the methods described regarding the zeta function–zero free regions, zero density estimates, height verifications–exist for L-functions as well, but are more difficult. It is worth noting as well that conditional results are often much stronger than unconditional results.

Definition 2.4.10. [22, Chapter VII, Definition 5.1] The Dedekind zeta function of a number field K is defined by the series

$$\zeta_K(s) = \sum_{\mathfrak{a}} \frac{1}{N_{K/\mathbb{Q}}(\mathfrak{a})}$$

where $\mathfrak{a}$ varies over the integral ideals of K.

Remark 2.4.11. The Extended Riemann Hypothesis (ERH) extends the Riemann Hypothesis to Dedekind zeta functions, as the Generalized Riemann Hypothesis extends the Riemann Hypothesis to L-functions.

We now state the first explicit form of Theorem 2.4.7.

Theorem 2.4.12. *[16, Theorem 1.1] There exists an effectively computable positive absolute constant c_1 such that if ERH holds for the Dedekind zeta function of L, then for every $x \geq 2$,*

$$|\pi_C(x, L/K) - \frac{|C|}{|G|} \operatorname{Li}(x)| \leq c_1 \left(\frac{|C|}{|G|} x^{1/2} \log(|d_L| x^{n_L}) + \log |d_L| \right).$$

An important corollary, one that we shall make use of, is finding an x_0 such that $\pi_C(x_0, L/K) > 0$.

Corollary 2.4.13. *[16, Corollary 1.2] There exists an effectively computable positive absolute constant c_2 such that if GRH holds for the Dedekind zeta function of $L \neq \mathbb{Q}$, then*

for every conjugacy class C of G there exists an unramified prime ideal $\mathfrak{p}$ of K such that $(\frac{L/K}{\mathfrak{p}}) = C$ and

$$N_{K/\mathbb{Q}}(\mathfrak{p}) \leq c_2(\log|d_L|)^2(\log\log|d_L|)^4.$$

If $L = \mathbb{Q}$, then $\mathfrak{p} = (2)$ is a solution.

The above is a non-nullity result about $\pi_C(x, L/K)$; it asserts the size of x we must take to ensure that $\pi_C(x, L/K)$ is nonzero, that is, there is some prime ideal whose Artin symbol hits C and with norm smaller than x. A simpler non-nullity result comes from [24].

Theorem 2.4.14. *There exists an absolute constant $c_1 > 0$ such that, for $C \neq \varnothing$ and assuming GRH, we have*

$$\pi_C(x, L/K) \geq 1$$

for all $x \geq 2$ such that $x \geq c_1(\log|d_L|)^2$.

Oesterlé [24] finds that $c_1 = 70$, although his proof was seemingly never published.

An improvement to Lagarias and Odlyzko is given by Bach-Sorenson [2]:

Theorem 2.4.15. *[2, Theorem 5.1] Assume GRH. Let $K/\mathbb{Q}$ be a Galois extension of number fields, with $K \neq \mathbb{Q}$. Let d_K denote the discriminant of K. Let n_K denote the degree of K. Let $C \subseteq \mathrm{Gal}(K/\mathbb{Q})$ be a nonempty subset closed under conjugation. Then, there is a prime p of $\mathbb{Q}$ unramified in K with $(\frac{K/\mathbb{Q}}{p}) \subseteq C$, of residue degree 1, satisfying*

$$p \leq (a\log|d_K| + bn_K + c)^2$$

for some triple (a, b, c) taken from Table 2.1. We may take $a = 4$, $b = 2.5$, and $c = 5$ to cover all cases of $\log|d_K|$ and $n_K = [K : \mathbb{Q}]$.

Remark 2.4.16. We note the values $a = 4$, $b = 2.5$, and $c = 5$ apply for all values of $n_K = [K : \mathbb{Q}]$ and $\log|d_K|$. For particular ranges of $\log|d_K|$ and $[K : \mathbb{Q}]$, we may take refined constants appearing in Table 2.1 coming from [2, Table 3].

We make mention of a fact we shall employ later. By Lemma 4.1.6, we may move the $\log|d_K|$ term into the term involving n_K, and vice versa we may move n_K into the $\log|d_K|$ term. This is at the expense of larger constants, however allows us to force either a or b to be 0. See Remark 4.1.5.

A corollary to the above is given in Mayle-Wang [18, Corollary 6] for quadratic extensions when we need to pick the prime p more specifically:

	$n = [K : \mathbb{Q}]$		
$\log \lvert d_K \rvert$	2	3-4	5-9
1-5	$(3.29, 1.48, 4.9)$	—	—
5-10	$(2.662, 0.75, 4.8)$	$(2.808, 0.58, 4.7)$	—
10-25	$(2.301, 0.52, 5)$	$(2.524, 0.45, 4.9)$	$(2.736, 0.35, 4.7)$
25-100	$(1.881, 0.34, 5.5)$	$(2.035, 0.27, 5.3)$	$(2.231, 0.21, 5.1)$
100-1000	$(1.446, 0.23, 6.8)$	$(1.527, 0.17, 6.4)$	$(1.629, 0.11, 6.1)$
1000-10000	$(1.125, 0.63, 10.9)$	$(1.148, 0.5, 10.2)$	$(1.178, 0.37, 9.5)$
10000-100000	$(1.032, 0.44, 20.2)$	$(1.038, 0.5, 18.7)$	$(1.046, 0.56, 17.3)$
100000+	$(1.008, -0.06, 47.7)$	$(1.01, -0.03, 41.9)$	$(1.012, 0, 37.8)$

	$n = [K : \mathbb{Q}]$		
$\log \lvert d_K \rvert$	10-14	15-49	50+
1-5	—	—	—
5-10	—	—	—
10-25	$(2.303, 0.19, 4.8)$	—	—
25-100	$(2.297, 0.19, 5)$	$(2.228, 0.1, 4.9)$	—
100-1000	$(1.667, 0.09, 6)$	$(1.745, 0.04, 5.8)$	$(1.755, 0, 5.7)$
1000-10000	$(1.189, 0.32, 9.2)$	$(1.212, 0.24, 8.8)$	$(1.257, 0, 7.3)$
10000-100000	$(1.049, 0.59, 16.8)$	$(1.054, 0.63, 16)$	$(1.095, 0, 8.2)$
100000+	$(1.012, 0, 37.8)$	$(1.014, 0.02, 35.9)$	$(1.017, 0.07, 31.8)$

Table 2.1: Values (a, b, c) appearing in Theorem 2.4.15 for a number field $K/\mathbb{Q}$, appearing in [2, Table 3], with degree n_K and d_K the discriminant. The line — is used for combinations of n and $\log \lvert d_K \rvert$ that are not possible by Minkowski's Theorem.

Corollary 2.4.17. *[18, Corollary 6] Assume GRH. Let $K/\mathbb{Q}$ be a Galois extension of number fields, with $K \neq \mathbb{Q}$. Let m be a positive integer, and set $\tilde{K} = K(\sqrt{m})$. Denote $d_{\tilde{K}}$ to be the absolute value of the discriminant of $\tilde{K}$. Let $n_{\tilde{K}}$ denote the degree of $\tilde{K}$. Let $C \subseteq \mathrm{Gal}(K/\mathbb{Q})$ be a nonempty subset that is closed under conjugation. Then there exists a prime number p not dividing m that is unramified in $K/\mathbb{Q}$ with $\left(\frac{K/\mathbb{Q}}{p}\right) \subseteq C$ and satisfying*

$$p \leq (\tilde{a} \log \lvert d_{\tilde{K}} \rvert + \tilde{b} n_{\tilde{K}} + \tilde{c})^2,$$

where $\tilde{a}, \tilde{b}, \tilde{c}$ are absolute constants that may be taken to be $4, 2.5$, and 5 respectively or may be taken to be the improved values given in [2, Table 3] associated with $\tilde{K}$.

Before proving the next proposition, we require a simple lemma regarding a particular inequality.

Lemma 2.4.18. *Let K be a Galois number field, d_K the absolute value of the discriminant, and n_K the degree of K. Then, for all $n_K \geq 74$,*

$$\exp\left(\frac{0.07}{0.24} n_K + \frac{24.5}{0.24}\right) \leq d_K.$$

Proof. To prove this lemma, we 'wedge' the Minkowski bound (2.1) between the left-hand and right-hand side. For simplicity, label the sequences

$$a_n = \exp\left(\frac{0.07}{0.24}n + \frac{24.5}{0.24}\right) \quad \text{and} \quad b_n = (\pi/4)^n (n^n/n!)^2.$$

By Theorem 2.1.4, we always have $b_{n_K} \leq d_K$. We claim $a_n \leq b_n$ for $n \geq 74$, which we prove by induction on n.

Base case: If $n = 74$, then computationally one can check $a_{74} \leq b_{74}$. We have $a_{74} \approx 5.102 \times 10^{53}$, and $b_{74} \approx 6.979 \times 10^{53}$.

Inductive step: Suppose $a_n \leq b_n$ for some $n \geq 74$. We have

$$\begin{aligned} a_{n+1} &= \exp\left(\frac{0.07}{0.24}(n+1) + \frac{24.5}{0.24}\right) \\ &= \exp\left(\frac{0.07}{0.24}n + \frac{24.5}{0.24}\right)\exp\left(\frac{0.07}{0.24}\right) \\ &\leq b_n \exp\left(\frac{0.07}{0.24}\right) \\ &\leq b_{n+1}. \end{aligned}$$

The final inequality requires some work to prove. First, we have

$$\begin{aligned} \exp\left(\frac{0.07}{0.24}\right) b_n \leq b_{n+1} &\iff \exp\left(\frac{0.07}{0.24}\right)(\pi/4)^n (n^n/n!)^2 \leq (\pi/4)^{n+1}((n+1)^{n+1}/(n+1)!)^2 \\ &\iff \exp\left(\frac{0.07}{0.24}\right) n^{2n} \leq \pi/4(n+1)^{2n} \\ &\iff \exp\left(\frac{0.07}{0.24}\right) 4/\pi \leq \frac{(n+1)^{2n}}{n^{2n}}. \end{aligned} \tag{2.15}$$

On the left-hand side, we have $\exp\left(\frac{0.07}{0.24}\right) 4/\pi \approx 1.705$. On the right-hand side, when $n = 74$, we have $\frac{(75)^{148}}{74^{148}} \approx 7.291$. Also, this is an increasing function, as

$$\frac{d}{dx}\left(\frac{x+1}{x}\right)^{2x} = \left(1 + \frac{1}{x}\right)^{2x}\left[2\log\left(1 + \frac{1}{x}\right) - \frac{2}{x+1}\right]$$

which is positive, so (2.15) holds (one could also see $\lim_{n\to\infty}(\frac{n+1}{n})^{2n} = e^2$). Thus the claim is proven. □

We give our own version of Theorem 2.4.15 and Corollary 2.4.17. The idea is to collapse Table 2.1 into a 1-dimensional table, removing the condition on $\log|d_K|$ so that each tuple is valid for a range of n_K. We do this by picking a "pivot" tuple for each column, for which all tuples appearing before the pivot are absorbed into a special constant p_0, and all tuples appearing after are checked to be smaller than the pivot tuple. By smaller, we shall mean

$n_{\tilde{K}} = [\tilde{K} : \mathbb{Q}]$	$(p_0, \bar{a}, \bar{b}, \bar{c})$
2	$(2111964, 1.125, 0, 12.16)$
3-4	$(2353401, 1.148, 0.5, 10.2)$
5-9	$(2676790, 1.178, 0.37, 9.5)$
10-14	$(2803146, 1.189, 0.32, 9.2)$
15-49	$(3072167, 1.212, 0.240, 8.80)$
50+	$(3100065, 1.257, 0, 7.3)$

Table 2.2: 4-tuples $(p_0, \bar{a}, \bar{b}, \bar{c})$ as appearing in Proposition 2.4.19.

a tuple (a, b, c) is smaller (or provides a stronger bound) than a tuple (a', b', c') (both from Table 2.1) if

$$(a \log |d_K| + bn_K + c)^2 \leq (a' \log |d_K| + b' n_K + c)^2$$

for all appropriate values of $\log |d_K|$ and n_K.

Proposition 2.4.19. *Assume GRH. Let $K/\mathbb{Q}$ be a Galois extension of number fields with $K \neq \mathbb{Q}$. Let m be a positive integer, and set $\tilde{K} = K(\sqrt{m})$. Denote $d_{\tilde{K}}$ to be the absolute value of the discriminant of $\tilde{K}$. Let $n_{\tilde{K}}$ denote the degree of $\tilde{K}$. Let $C \subseteq \mathrm{Gal}(K/\mathbb{Q})$ be a nonempty subset that is closed under conjugation. Then there exists a 4-tuple $(p_0, \bar{a}, \bar{b}, \bar{c})$ taken from Table 2.2 and a prime number p not dividing m that is unramified in $K/\mathbb{Q}$ with $\left(\frac{K/\mathbb{Q}}{p}\right) \subseteq C$ and satisfying*

$$p \leq \max\{p_0, (\bar{a} \log |d_{\tilde{K}}| + \bar{b} n_{\tilde{K}} + \bar{c})^2\}.$$

Proof. Apply Corollary 2.4.17, to get a prime p not dividing m that is unramified in $K/\mathbb{Q}$ and satisfies $p \leq (\tilde{a} \log |d_{\tilde{K}}| + \tilde{b} n_{\tilde{K}} + \tilde{c})^2$.

The remainder of the corollary is proved through an exhaustive case analysis through all possible ranges on $n_{\tilde{K}}$ appearing in Table 2.1.

We begin by examining the case $n_{\tilde{K}} = 2$ (the first column of Table 2.1). We choose the tuple $(1.125, 0.63, 10.9)$ to be our pivot. Note that, because $n_{\tilde{K}}$ is fixed in this case, we can absorb $\tilde{b}$ into $\tilde{c}$, giving us the tuple $(1.125, 0, 12.16)$. We then find p_0 as follows:

- If $1 \leq \log |d_{\tilde{K}}| \leq 5$, then

$$p \leq (3.29 \log |d_{\tilde{K}}| + 1.48 \cdot 2 + 4.9)^2 \leq (3.29 \cdot 5 + 7.86)^2 \leq 590.977.$$

- If $5 < \log |d_{\tilde{K}}| \leq 10$, then

$$p \leq (2.662 \log |d_{\tilde{K}}| + 0.75 \cdot 2 + 4.8)^2 \leq (2.662 \cdot 10 + 6.3)^2 \leq 1083.727.$$

- If $10 < \log|d_{\tilde{K}}| \leq 25$, then

$$p \leq (2.301 \log|d_{\tilde{K}}| + 0.52 \cdot 2 + 5)^2 \leq (2.301 \cdot 25 + 6.04)^2 \leq 4040.510.$$

- If $25 < \log|d_{\tilde{K}}| \leq 100$, then

$$p \leq (1.881 \log|d_{\tilde{K}}| + 0.34 \cdot 2 + 5.5)^2 \leq (1.881 \cdot 100 + 0.34 \cdot 2 + 5.5)^2 \leq 37744.719.$$

- If $100 < \log|d_{\tilde{K}}| \leq 1000$, then

$$p \leq (1.446 \log|d_{\tilde{K}}| + 0.23 \cdot 2 + 6.8)^2 = (1.446 \cdot 1000 + 7.26)^2 \leq 2111964.628.$$

Considering the maximum of all possible values above, we see $p \leq 2111964$. We take this to be our p_0 in the case $n_{\tilde{K}} = 2$.

It remains to verify that the two tuples appearing after our pivot point are smaller than our pivot tuple.

- If $10000 < \log|d_{\tilde{K}}| \leq 100000$, then

$$p \leq (1.032 \log|d_{\tilde{K}}| + 0.44 \cdot 2 + 20.2)^2 \leq (1.125 \log|d_{\tilde{K}}| + 12.16)^2$$

  where the last inequality holds provided $\log|d_{\tilde{K}}| \geq 95.914$.

- If $10000 < \log|d_{\tilde{K}}|$, then

$$p \leq (1.008 \log|d_{\tilde{K}}| - 0.06 \cdot 2 + 47.7)^2 \leq (1.125 \log|d_{\tilde{K}}| + 12.16)^2$$

  where the last inequality holds provided $\log|d_{\tilde{K}}| \geq 302.736$.

Thus, we see $p \leq (1.125 \log|d_{\tilde{K}}| + 12.16)^2$ in the above two cases. Therefore, we have $p \leq \max\{2111964, (1.125 \log|d_{\tilde{K}}| + 12.16)^2\}$, which corresponds to the tuple $(2111964, 1.125, 0, 12.16)$.

Next, suppose now $n_{\tilde{K}} = 3$ or $n_{\tilde{K}} = 4$, which corresponds to the second column of Table 2.1. We pick the tuple $(1.148, 0.5, 10.2)$ to be our pivot this time. We begin by noting that, if $\log|d_{\tilde{K}}| < 5$, then using the generic bound with $a = 4$, $b = 2.5$, and $c = 5$, we have

$$p \leq (4 \log|d_{\tilde{K}}| + 2.5 n_{\tilde{K}} + 5)^2 \leq 1225. \tag{2.16}$$

Similarly, we find p_0 as follows:

- If $5 \leq \log|d_{\tilde{K}}| \leq 10$, then

$$p \leq \begin{cases} 1191.631, & n_{\tilde{K}} = 3 \\ 1232.011, & n_{\tilde{K}} = 4 \end{cases}.$$

- If $10 < \log|d_{\tilde{K}}| \leq 25$, then

$$p \leq \begin{cases} 4809.423, & n_{\tilde{K}} = 3 \\ 4872.041, & n_{\tilde{K}} = 4 \end{cases}.$$

- If $25 < \log|d_{\tilde{K}}| \leq 100$, then

$$p \leq \begin{cases} 43936.352, & n_{\tilde{K}} = 3 \\ 44049.614, & n_{\tilde{K}} = 4 \end{cases}.$$

- If $100 < \log|d_{\tilde{K}}| \leq 1000$, then

$$p \leq \begin{cases} 2352879.888, & n_{\tilde{K}} = 3 \\ 2353401.446, & n_{\tilde{K}} = 4 \end{cases}.$$

Looking at the maximum of all ranges above, we see $p \leq 2353401$. We take this to be our p_0 for the range $3 \leq n_{\tilde{K}} \leq 4$. We now verify the two tuples which appear after our pivot are, in fact, smaller than our pivot.

- If $10000 < \log|d_{\tilde{K}}| \leq 100000$, then

$$p \leq (1.038 \log|d_{\tilde{K}}| + 0.5n_{\tilde{K}} + 18.7)^2 \leq (1.148 \log|d_{\tilde{K}}| + 0.5n_{\tilde{K}} + 10.2)^2$$

  where the last inequality holds for both $n_{\tilde{K}} = 3$ and $n_{\tilde{K}} = 4$ provided that $\log|d_{\tilde{K}}| \geq 77.273$.

- If $100000 < \log|d_{\tilde{K}}|$, then

$$p \leq (1.01 \log|d_{\tilde{K}}| - 0.03n_{\tilde{K}} + 41.9)^2 \leq (1.148 \log|d_{\tilde{K}}| + 0.5n_{\tilde{K}} + 10.2)^2$$

  where the last inequality holds for both $n_{\tilde{K}} = 3$ and $n_{\tilde{K}} = 4$ provided that $\log|d_{\tilde{K}}| \geq 218.189$.

Again, we have $p \leq \max\{2353401, (1.148 \log|d_{\tilde{K}}| + 0.5n_{\tilde{K}} + 10.2)^2\}$. Thus, the appropriate 4-tuple for the case $n_{\tilde{K}} = 3$ or $n_{\tilde{K}} = 4$ is $(2353401, 1.148, 0.5, 10.2)$.

The next three cases (when $5 \leq n_{\tilde{K}} \leq 9$, or $10 \leq n_{\tilde{K}} \leq 14$, or $15 \leq n_{\tilde{K}} \leq 49$) are proved in an identical manner to the above two cases, using code to verify the given 4-tuple. The pivots we used are, respectively, $(1.178, 0.37, 9.5)$ (for the case $5 \leq n_{\tilde{K}} \leq 9$), $(1.189, 0.32, 9.2)$ (for the case $10 \leq n_{\tilde{K}} \leq 14$), and $(1.212, 0.24, 8.8)$. The code which does this manual verification can be found in Appendix A.

The last case we establish explicitly. For both cases, we choose our pivot to be the tuple $(1.257, 0, 7.3)$. First, suppose $50 \leq n_{\tilde{K}} \leq 73$. This we, again, verify using the code, and we

get a 4-tuple $(3100065, 1.257, 0, 7.3)$. For the case $n_{\tilde{K}} \geq 74$, we note that if $100 \leq \log|d_{\tilde{K}}| \leq 1000$, then

$$p \leq (1.755 \log|d_{\tilde{K}}| + 5.7)^2 \leq 3100064.49.$$

If $1000 < \log|d_{\tilde{K}}| \leq 10000$, then

$$p \leq (1.257 \log|d_{\tilde{K}}| + 7.3)^2.$$

We must verify the above, our pivot, is larger than the remaining two cases.

- If $10000 < \log|d_{\tilde{K}}| \leq 100000$, then

  $$p \leq (1.095 \log|d_{\tilde{K}}| + 8.2)^2 \leq (1.257 \log|d_{\tilde{K}}| + 7.3)^2$$

  where the last inequality holds provided that $\log|d_{\tilde{K}}| \geq 5.556$.

- If $\log|d_{\tilde{K}}| > 100000$, then

  $$p \leq (1.017 \log|d_{\tilde{K}}| + 0.07 n_{\tilde{K}} + 31.8)^2 \leq (1.257 \log|d_{\tilde{K}}| + 7.3)^2,$$

  where the last inequality holds if

  $$\begin{aligned} 1.017 \log|d_{\tilde{K}}| + 0.07 n_{\tilde{K}} + 31.8 \leq 1.257 \log|d_{\tilde{K}}| + 7.3 &\iff 0.07 n_{\tilde{K}} + 24.5 \leq 0.24 \log|d_{\tilde{K}}| \\ &\iff \exp\left(\frac{0.07}{0.24} n_{\tilde{K}} + \frac{24.5}{0.24}\right) \leq d_{\tilde{K}}. \end{aligned} \tag{2.17}$$

  By Lemma 2.4.18, the last inequality of (2.17) holds for $n_{\tilde{K}} \geq 74$, thus $p \leq (1.257 \log|d_{\tilde{K}}| + 7.3)^2$.

Taking the maximum of all cases again, we see $p \leq \max\{3100064, (1.257 \log|d_{\tilde{K}}| + 7.3)^2\}$. This corresponds to the 4-tuple $(3100064, 1.257, 0, 7.3)$, the same tuple as at the beginning of this case. This completes our analysis. □

Remark 2.4.20. By appropriately scaling $\bar{a}, \bar{b}, \bar{c}$ appearing in each case to $\bar{A}, \bar{B}, \bar{C}$, we can make it so that $p_0 \leq (\bar{A} \log|d_{\tilde{K}}| + \bar{B} n_{\tilde{K}} + \bar{C})^2$ for all values of $\log|d_{\tilde{K}}|$ and $n_{\tilde{K}}$ in the appropriate range. This reduces the 4-tuple $(p_0, \bar{a}, \bar{b}, \bar{c})$ to a 3-tuple $(\bar{A}, \bar{B}, \bar{C})$. However, the scaling factor can be quite large for each case, and remove any potential advantage in smaller constants. For example, in the first case, we have the 4-tuple $(2111964, 1.125, 0, 12.16)$. We would then need to consider a scalar $k \in \mathbb{R}$ such that

$$2111964 \leq k(1.125 \log|d_{\tilde{K}}| + 12.16)^2.$$

Given that $\log|d_{\tilde{K}}| \geq 1$, we would need to take $k \geq 11966$. This would make $\bar{A} = 13461.75$ and $\bar{C} = 145506.56$. For the results presented in Chapter 4, the constants are already large enough to dominate any potential increase which may arise from such scaling factors; however, in Chapter 5, our constants are quite small, and such scaling factors would nullify our improvements.

Remark 2.4.21. The choice of which 3-tuple to pick as the pivot is entirely arbitrary. For example, for the case $n = 2$, we could have chosen the tuple $(1.881, 0.34, 5.5)$ coming from Table 2.1 to be our pivot, which would have resulted in a 4-tuple of $(4040, 1.881, 0, 6.18)$. When choosing a pivot, notice that taking a pivot corresponding to a larger range of $\log|d_{\tilde{K}}|$ results in a smaller $\bar{a}$ at the expense of a larger p_0. In our results where we make use of these constants (Chapter 5), we are required to take the final 4-tuple as our $n_{\tilde{K}}$ is large, so, $p_0 = 3100064$. For the other cases, we make the choice to push $\bar{a}$ to be as small as possible by making the pivot larger, but no larger than 3100064 (see Remark 5.3.1 for a concrete example of why smaller $\bar{a}$ is better for our purposes). For other applications where a smaller p_0 or $\bar{a}$ would be beneficial, one can reproduce Table 2.2 by choosing a different pivot.

Remark 2.4.22. In (2.16), we considered what happens when $\log|d_{\tilde{K}}| < 5$ (in the case when $n_{\tilde{K}} = 3$ or $n_{\tilde{K}} = 4$) even though in Table 2.1 this combination of $\log|d_{\tilde{K}}|$ and $n_{\tilde{K}}$ is claimed to not be possible. We applied Theorem 2.4.15 with the general values of a, b, c, those being $a = 4$, $b = 2.5$, and $c = 5$. We consider this 'impossible' case as there are some entries in Table 2.1 that do not give tuples for combinations of $n_{\tilde{K}}$ and $\log|d_{\tilde{K}}|$ that would, in fact, be possible. For example, [18, Remark 13] notes the maximal totally real subfield of $\mathbb{Q}(\zeta_7)$ has degree 3 and discriminant 49, yet Table 2.1 gives no tuple for a degree 3 field for which the logarithm of the discriminant is less than 5. Thus, in order to account for this case, we must use the generic bounds. This is incorporated into the code given in Appendix A, however given how large p_0 is in all cases, this is not sufficient to change our final result. However, if one were to pick different pivots, it is possible that p_0 would be small enough that this generic case will be larger than all tuples appearing before the pivot.

Chapter 3

Torsion Fields of Elliptic Curves

The goal of this chapter is to give an account of some facts related to the discriminant of torsion fields. In particular, if E is an elliptic curve over $\mathbb{Q}$ and ℓ a prime, we examine closely the field $\mathbb{Q}(E[\ell])$; specifically, we examine the relationship between the primes of bad reduction of E and the primes which ramify in $\mathbb{Q}(E[\ell])$. Our work in this section relies heavily on results of both Cali-Kraus [7] and Kraus [15].

We work to determine which primes ramify in the number field $\mathbb{Q}(E[\ell])$. As above, let E be an elliptic curve and ℓ a prime. Denote $L = \mathbb{Q}(E[\ell])$. We begin by stating a theorem due to Serre and Tate [32, Theorem 1] describing the primes which have good (or bad) reduction. Note the theorem is stated for general abelian varieties, which, in our case, can be thought of as being elliptic curves. The particular version we give appears in [18, Theorem 8].

Theorem 3.0.1. *Let $A/\mathbb{Q}$ be an abelian variety. For a prime number p, the following are equivalent:*

a) A has good reduction at p,

b) $\mathbb{Q}(A[m])/\mathbb{Q}$ is unramified at p for each positive integer m not divisible by p, and

c) $\mathbb{Q}(A[m])/\mathbb{Q}$ is unramified at p for infinitely many positive integers m, not divisible by p.

Proof. See [32, Theorem 1]. □

Theorem 3.0.1 gives us a clearer picture of what the discriminant of L shall look like. By Theorem 2.1.10, a prime p is ramified in L if and only p divides d_L; thus, the primes for which E has bad reduction should be ramified in L. This idea we shall expand upon and make more explicit in a different way.

Let us now turn our attention to the primes in $P(E)$, the primes where E has bad reduction. By the end of this section, we will prove a subset of the primes from $P(E)$ also ramify in L. To see this, however, we must state results of Kraus and Cali-Kraus.

The setup is as follows: let q and ℓ be primes, K a finite unramified extension of $\mathbb{Q}_q$, ν_q the q-adic valuation, and D an integer characterized by the following two equivalent properties:

a) the different of the extension $K(E[\ell])/K$ is the Dth power of the valuation (maximal) ideal of $K(E[\ell])$,

b) the discriminant ideal of the extension $K(E[\ell])/K$ is generated by $q^{nD/e}$, where n is the degree and e is the ramification index of the extension $K(E[\ell])/K$.

Before we begin our analysis, we must split into two cases, depending on whether or not $q = \ell$ or $q \neq \ell$. We begin with the former case, that being we assume $q = \ell$.

3.0.1 The case $q = \ell$

Our first goal is to replace the exponent appearing on q in the discriminant ideal. From here on, we shall replace q with ℓ (understanding that they are the same). Let K be a finite unramified extension of $\mathbb{Q}_\ell$ endowed with the valuation ν which extends that of $\mathbb{Q}_\ell$ (following Cali-Kraus).

Lemma 3.0.2. *Let ℓ be a prime. Let K be a finite extension of $\mathbb{Q}_\ell$ and E/K be an elliptic curve with potentially good reduction. Let n be the degree, and e be the ramification index, of the extension $K(E[\ell])/K$. Then $n/e \leq 48$.*

Proof. Let $L = K(E[\ell])$, G be the Galois group, and I be the inertia group of L/K. Recall that G/I is cyclic and isomorphic to the Galois group of the extension of residue fields $k(L)/k(K)$.

If E/K has good ordinary reduction or multiplicative reduction, then $\overline{\rho}_{E,\ell}(I)$ is conjugate to

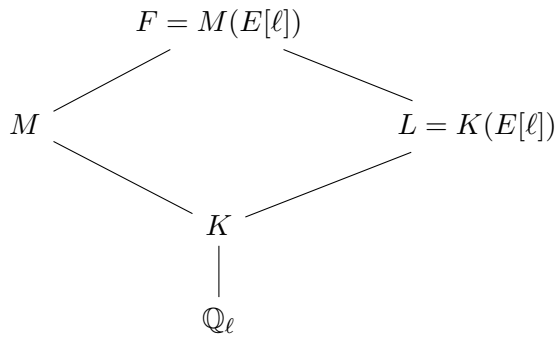
$$\begin{pmatrix} * & 0 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} * & * \\ 0 & 1 \end{pmatrix} \tag{3.1}$$

and has order $\ell - 1$ (see [28, Proposition 13]) so that $e = \#\overline{\rho}_{E,\ell}(I_{\mathfrak{L}}) = \ell - 1$. The image $\overline{\rho}_{E,\ell}(I)$ is normal in $\overline{\rho}_{E,\ell}(G)$. But the normalizer of (3.1) in $\mathrm{GL}_2(\mathbb{F}_\ell)$ is itself, so it follows that $n = e > 1$ and hence $n/e = 1$.

If E/K has good supersingular reduction, then $\overline{\rho}_{E,\ell}(I)$ is conjugate to a non-split Cartan subgroup of order $\ell^2 - 1$ (see [28, Proposition 12]) so that $e = \#\overline{\rho}_{E,\ell}(I_{\mathfrak{L}}) = \ell^2 - 1 > 1$. The image $\overline{\rho}_{E,\ell}(I)$ is normal in $\overline{\rho}_{E,\ell}(G)$. But the normalizer of C' in $\mathrm{GL}_2(\mathbb{F}_\ell)$ contains C' with index 2, so it follows that $n/e \leq 2$.

If E/K has additive reduction, after a finite extension M/K of degree dividing 24 that E/M has good or multiplicative reduction [28, §5.6]. Let n', e' be the degree, ramification

index of $M(E[\ell])/M$, respectively. Then $n/e \leq 24n'/e' \leq 48$ by the diagram:



□

Given this, we see both that, since $e > 1$, ℓ must be ramified in $K(E[\ell])/K$, and that the discriminant ideal of K is generated by at least ℓ^{48D}.

Thus, it remains to determine what D is. Let h denote the height of the reduction of E; recall from (2.8) that $h = 1$ if E has good ordinary reduction, and $h = 2$ if E has good supersingular reduction. If $h = 1$, then we let $j_{can}(\tilde{E})$ be the j-invariant of the canonical lifting of $\tilde{E}$ to E, such that if we reduce E modulo the maximal ideal of K, we get $\tilde{E}$. The following is a result of Kraus [15]:

Theorem 3.0.3. *[15, Theorem 2] Let ℓ be a prime. Suppose K is a finite extension of $\mathbb{Q}_\ell$, and E an elliptic curve which has good reduction on K. Denote by $\tilde{E}$ the elliptic curve obtained by reducing E modulo the maximal ideal of K.*

i) If $j(\tilde{E}) = 0$ (that is, $\nu_\ell(c_4) \geq 1$) and $h = 1$ (that is, $\ell \equiv 1 \pmod 3$), then

$$D = \begin{cases} \ell^2 - 2 & \text{if } \nu_\ell(c_4) = 1, \\ \ell - 2 & \text{if } \nu_\ell(c_4) \neq 1. \end{cases}$$

ii) If $j(\tilde{E}) = 1728$ (that is, $\nu_\ell(c_4) \geq 1$) and $h = 1$ (that is, $\ell \equiv 1 \pmod 4$), then

$$D = \begin{cases} \ell^2 - 2 & \text{if } \nu_\ell(c_6) = 1, \\ \ell - 2 & \text{if } \nu_\ell(c_6) \neq 1. \end{cases}$$

iii) If $j(\tilde{E}) \notin \{0, 1728\}$ and $h = 1$, then

$$D = \begin{cases} \ell^2 - 2 & \text{if } \nu_\ell(j - j_{can}(\tilde{E})) = 1, \\ \ell - 2 & \text{if } \nu_\ell(j - j_{can}(\tilde{E})) \neq 1. \end{cases}$$

iv) If $h = 2$,

$$D = \ell^2 - 2.$$

Applying both Lemma 3.0.2 and Theorem 3.0.3 in our situation, we see that ℓ ramifies in $\mathbb{Q}(E[\ell])/\mathbb{Q}$.

3.0.2 The case $q \neq \ell$

Suppose now that $q \neq \ell$. Like in the previous case, we must once again work to simplify the exponent appearing on q, that being $q^{nD/e}$. We have the following proposition coming from Darmon-Diamond-Taylor [11]:

Proposition 3.0.4. *[11, Proposition 2.12] Suppose that E has multiplicative reduction at q. Let $\delta : G_q/I_q \to \{\pm 1\}$ be the unique non-trivial unramified quadratic character of G_q if E has non-split multiplicative reduction, and let δ be the trivial character if E has split multiplicative reduction. Let $\epsilon_\ell : G_{\mathbb{Q}} : \mathbb{Z}_\ell^\times$ be the ℓ-adic cyclotomic character Then,*

$$\rho_{E,\ell} \mid_{G_q} \sim \begin{pmatrix} \epsilon_\ell & * \\ 0 & 1 \end{pmatrix} \otimes \delta,$$

and if $\ell \neq q$ then $m_q(\rho_{E,\ell}) = 1$.

We wish to give an analogous Lemma to Lemma 3.0.2 in the case $q \neq \ell$.

Lemma 3.0.5. *Let ℓ be an odd prime. Let K be a finite extension of $\mathbb{Q}_q$ and E/K be an elliptic curve with potentially multiplicative reduction. Let n be the degree, and e be the ramification index, of the extension $K(E[\ell])/K$, and $q \neq \ell$ a prime such that $K(E[\ell])/K$ is ramified at q. Then $n/e \leq 2(\ell - 1)$.*

Proof. As before, we consider the order of the image $\rho_{E,\ell}(G_q)/\rho_{E,\ell}(I_q)$. By Proposition 3.0.4, we have the image of $\rho_{E,\ell}(G_q)$ is exactly

$$\rho_{E,\ell} \mid_{G_q} \sim \begin{pmatrix} \epsilon_\ell & * \\ 0 & 1 \end{pmatrix} \otimes \delta.$$

Given there are at most $\ell(\ell-1)$ possibilities for ϵ_ℓ and $*$, and that δ can have at most order 2, we see $n \mid 2\ell(\ell-1)$, implying $n \leq 2\ell(\ell-1)$.

Now we consider $\rho_{E,\ell}(I_q)$. Note that $\epsilon_\ell(I_q) = 1$ since, in the case $q \neq \ell$, ϵ_ℓ is unramified at q. There are only two possibilities for $\rho_{E,\ell}(I_q)$: either it is trivial, or its image is given by matrices of the form

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix} \tag{3.2}$$

which would then have order ℓ. If the image is trivial, we have $e = 1$, which implies q does not ramify, a contradiction to our assumption that q is, in fact, ramified in the extension

$K(E[\ell])/K$. So, the image must, in fact, be given by matrices as in (3.2); thus $|\rho_{E,\ell}(I_q)| = \ell$. Lastly, when combined with the size of $\rho_{E,\ell}(G_q)$, we get the result $n/e \leq 2\ell(\ell-1)/\ell = 2(\ell-1)$. □

All that remains is to determine D; we use the following result of Cali-Kraus [7] to achieve this aim:

Theorem 3.0.6. *[7, Theorem 1] Let q, ℓ be primes such that $q \neq \ell$. Let K be a finite extension of $\mathbb{Q}_q$. Let $\nu_q(x)$ denote the valuation of x at q.*

1) Suppose $\nu_q(j) < 0$.

a) If E has reduction of multiplicative type on K, then

$$D = \begin{cases} 0 & \text{if } \ell \text{ divides } \nu_q(j), \\ \ell - 1 & \text{if } \ell \text{ does not divide } \nu_q(j). \end{cases}$$

b) If E has additive reduction on K, and $q \neq 2$, then

$$D = \begin{cases} 1 & \text{if } \ell \text{ divides } \nu_q(j) \text{ or } \ell = 2, \\ 2\ell - 1 & \text{if } \ell \text{ does not divide } \nu_q(j) \text{ and } \ell \neq 2, \end{cases}$$

otherwise if $q = 2$, then we are in one of two cases:

i) If $\nu_q(c_6) = 6$, then

$$D = \begin{cases} 2 & \text{if } \ell \text{ divides } \nu_q(j), \\ 3\ell - 1 & \text{if } \ell \text{ does not divide } \nu_q(j), \end{cases}$$

ii) else if $\nu_q(c_6) = 9$, then

$$D = \begin{cases} 3 & \text{if } \ell \text{ divides } \nu_q(j), \\ 4\ell - 1 & \text{if } \ell \text{ does not divide } \nu_q(j). \end{cases}$$

2) Suppose $\nu_q(j) \geq 0$, and E has addititve reduction on K. Assume $q \geq 5$. Let m be the denominator of $\nu_q(\Delta_E)/12$, with Δ_E the discriminant of E. We have

$$D = \begin{cases} m - 1 & \text{if } \ell \neq 2, \\ 1 & \text{if } \ell = 2 \text{ and } \nu_q(\Delta) \text{ is odd}, \\ 2 & \text{if } \ell = 2 \text{ and } \nu_q(\Delta) \text{ is even and not } 6, \\ 0 & \text{if } \ell = 2 \text{ and } \nu_q(\Delta) = 6. \end{cases}$$

3) In all remaining cases, D is given explicitly and satisfies $D \leq 68$.

If E has multiplicative reduction on K, then so long as ℓ does not divide $\nu_q(j)$ (for any $q \in P(E)$, the primes of bad reduction of E), then q will ramify. We have that $D \leq 2(4\ell - 1)$, which is linear in ℓ.

Chapter 4

The Method of Serre

This chapter is dedicated to describing the method Serre gives in [30] for finding an upper bound on the smallest prime p such that $a_p(E) \neq a_p(E')$. This result is not as refined as Mayle-Wang [18] (the results of which appear in Chapter 5), however we make a few modifications to the original theorem, including using Bach-Sorenson (Theorem 2.4.15) for our Chebotarev estimate, as well as a different bound on the discriminant. We begin by stating Serre's original result:

Theorem 4.0.1. *[30, Theorem 21] Let E and E' be two elliptic curves defined over $\mathbb{Q}$. Suppose that the set of primes p such that $a_p(E) \neq a_p(E')$ is infinite. Let S be a finite set of primes containing both $P(E)$ and $P(E')$; let $N_S = \prod_{\ell \in S} \ell$. Let $p = p(E, E', S)$ be the smallest prime number not belonging to S such that $a_p(E) \neq a_p(E')$. Under GRH, we have*

$$p \leq C_1 (\log N_S)^2 (\log \log 2N_S)^{12}, \tag{4.1}$$

where C_1 is an absolute constant.

An overview of the argument looks as follows: suppose we have two elliptic curves E and E', and let $P(E), P(E')$ denote the primes of bad reduction of E and E' respectively. Let p be the smallest prime such that $a_p(E) \neq a_p(E')$. We wish to pick a prime ℓ such that, modulo ℓ, the two traces are distinguished, that is, $a_p(E) \not\equiv a_p(E') \pmod{\ell}$. Once we have accomplished this, we can look at the field generated by the kernel of our mod ℓ representation, and apply a Chebotarev Density estimate to pick a prime q that will also satisfy $a_q(E) \not\equiv a_q(E') \pmod{\ell}$. By minimality, $p \leq q$, and so p will also satisfy the upper bound on q given to us by Chebotarev. Once we are in such a position, then some simple estimates and bookkeeping will give us our desired result.

We fix some notation: we let

$$P(L) = \{ q \in \mathbb{Z} \mid \ q \text{ prime and } q \text{ is unramified in } L \ \}$$

for a number field $L/\mathbb{Q}$, and for an elliptic curve E the analogous set

$$P(E) = \{q \in \mathbb{Z} \mid \ q \text{ prime and } E \text{ has bad reduction at } q \ \}.$$

For a set of primes P, we denote the product

$$N_P = \prod_{q \in P} q,$$

in particular we will consider the product

$$\mathfrak{N} = N_{P(E)} N_{P(E')}. \tag{4.2}$$

We give our modification of Theorem 4.0.1. Note that, asymptotically, it is the same as Theorem 4.0.1, however here we have explicitly worked out the constants and employed Proposition 2.4.19.

Theorem 4.0.2. *Let E and E' be two elliptic curves defined over $\mathbb{Q}$. Suppose that the set of primes p such that $a_p(E) \neq a_p(E')$ is infinite. Let p be the smallest prime number not belonging to $P(E) \cup P(E')$ such that $a_p(E) \neq a_p(E')$. Let S be a finite set of primes, such that for all primes q not belonging to S, both E and E' will have good reduction at q. Then,*

$$p \leq C_{12} (\log\log N_S)^{12} \left[\log N_S + \log\log\log N_S + \bar{c}\right]^2, \tag{4.3}$$

where

$$C_{12} = C_{11}(\log C_{10} + 1), \tag{4.4}$$

$$C_{11} = C_7 C_{10}^{12}, \tag{4.5}$$

$$C_{10} = C_9(\log(2 + \bar{c}) + 1), \tag{4.6}$$

$$C_9 = 2(\log C_8 + 2), \tag{4.7}$$

$$C_8 = C_4(2, 14) C_7, \tag{4.8}$$

$$C_7 = 6 C_6 C_5^{12} (1 + \log C_5)^2, \tag{4.9}$$

$$C_6 = 16 (\log 2)^2 \bar{a}^2, \tag{4.10}$$

$$C_5 = (C_3 \log 8 + 1/2), \tag{4.11}$$

and $C_4(r, s)$ and C_3 are as appears in Lemma 4.2.1 and Lemma 4.1.2 respectively, and $\bar{a}$, $\bar{c}$ can be taken to be 4 and 5 respectively by Theorem 2.4.15 or any appropriate value from Table 2.2.

Remark 4.0.3. Given our constants are taken from Proposition 2.4.19, one may wonder why we do not consider (4.3) with the maximum of p_0. In this case, the constant C_{12} is so large that (4.3) overpowers p_0. See Remark 4.2.2, where we have worked out C_{12} explicitly.

The work is split into two sections: first, some preliminary work detailing the tools we shall be employing in the main argument, then a section in which we give a proof of Theorem 4.0.2.

4.1 The Tools of Serre

We shall require two helpful lemmas of Serre [30], the latter of which will help us pick our special prime ℓ.

Lemma 4.1.1. *[30, Lemma 11] There exists an absolute constant $C_2 > 0$ such that $\sum_{q \leq x} \log q \geq C_2 x$ for all $x \geq 2$ (note that the summation runs over prime numbers $q \leq x$).*

Proof. By Theorem 2.4.3, we know $\theta(x) \sim x$. Thus, for large x, we have that

$$ax < \sum_{q \leq x} \log q < bx \tag{4.12}$$

for any constants $0 < a < 1 < b$. We can then scale a and b so that the inequality (4.12) holds for all $x \geq 2$. □

Lemma 4.1.2. *[30, Lemma 12] If $n \in \mathbb{Z}_{>0}$, there exists a prime number ℓ such that $n \not\equiv 0 \pmod{\ell}$ and $\ell \leq C_3 \log 2n$, where C_3 is an absolute constant.*

Proof. Take $C_3 = \sup(2/\log 2, 1/C_2)$. For the purpose of contradiction, suppose such a prime ℓ does not exist. Then, n would be divisible by all primes $\ell \leq x$, where $x = C_3 \log 2n$. Consider now the product $\prod_{\ell \leq x} \ell$; by our assumption, n is divisible by each prime ℓ in this product, and so we have $n \geq \prod_{\ell \leq x} \ell$, or $\log n \geq \sum_{\ell \leq x} \log \ell$. By our hypotheses, we have $x \geq C_3 \log 2 \geq 2$, and Lemma 4.1.1 gives us

$$\log n \geq \sum_{\ell \leq x} \log \ell \geq C_2 x = C_2 C_3 \log 2n \geq \log 2n,$$

but then $\log n \geq \log 2n$, which is a contradiction. □

We note that Serre [30] gives the optimal values of the two constants in Lemma 4.1.1 and 4.1.2 as $C_2 = 1/3 \log 2$ and $C_3 = 2/\log 2$. In addition, note that the prime ℓ is of order $\mathcal{O}(\log n)$.

Next, we give a brief description of how we shall distinguish the traces. Recall our goal is, given two elliptic curves, to find a prime $q \notin S$ so that $a_q(E) \neq a_q(E')$. Serre accomplishes this by considering a particular conjugacy class of the Galois group, and

showing that elements of this conjugacy class in fact are able to distinguish the trace elements. Then, applying an explicit Chebotarev estimate to that conjugacy class gives a prime q so that $a_q(E) \neq a_q(E')$. The construction is as follows. Recall the two-dimensional mod ℓ representations attached to E and E' as

$$\bar{\rho}_{E,\ell} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z}) \text{ and } \bar{\rho}_{E',\ell} : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z})$$

and define

$$\bar{\rho}_\ell : G_{\mathbb{Q}} \to \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z}) \times \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z}) \tag{4.13}$$

by $\bar{\rho}_\ell(x) = (\bar{\rho}_{E,\ell}(x), \bar{\rho}_{E',\ell}(x))$. Let $G_\ell = \bar{\rho}_\ell(G_{\mathbb{Q}}) \subset \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z}) \times \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z})$ be the image of the map $\bar{\rho}_\ell$. Define

$$C_\ell = \{(s, s') \in G_\ell \mid \mathrm{tr}(s) \neq \mathrm{tr}(s')\}. \tag{4.14}$$

We claim C_ℓ is nonempty. Let q be a prime number not belonging to $P(E) \cup P(E')$, and distinct from ℓ. The representation ρ_ℓ is unramified at q (see [11, Proposition 2.11]), so we can find the Frobenius element at q, σ_q, and get

$$\bar{\rho}_\ell(\sigma_q) = (\bar{\rho}_{E,\ell}(\sigma_q), \bar{\rho}_{E',\ell}(\sigma_q)).$$

This is an element of G_ℓ defined up to conjugation; let $\bar{\rho}_\ell(\sigma_q) = (s_q, s'_q)$. By (2.10), we have

$$\mathrm{tr}(s_q) \equiv a_q(E) \pmod{\ell} \text{ and } \mathrm{tr}(s'_q) \equiv a_q(E') \pmod{\ell}.$$

Then, (s_q, s'_q) belongs to C_ℓ if and only if

$$a_q(E) \not\equiv a_q(E') \pmod{\ell}. \tag{4.15}$$

According to the fact that $a_p(E) \not\equiv a_p(E') \pmod{\ell}$ and $\ell \neq p$, this condition is satisfied when q is exactly p. We therefore have $(s_p, s'_p) \in C_\ell$, which shows that C_ℓ is not empty.

Let H_ℓ be the intersection of G_ℓ with the group of homotheties (λ, λ), where λ runs through $(\mathbb{Z}/\ell\mathbb{Z})^*$. The conjugacy class C_ℓ is stable under multiplication by H_ℓ so that it is the inverse image of a non-empty subset C'_ℓ of $G'_\ell = G_\ell/H_\ell$.

Lemma 4.1.3. *[30, Lemma 13] With the setup as above, we have $|G'_\ell| < 2\ell^6$.*

Proof. Denote the kernel of the canonical homomorphism

$$G_\ell \to \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z}) \times \mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z}) \to \mathrm{PGL}_2(\mathbb{Z}/\ell\mathbb{Z}) \times \mathrm{PGL}_2(\mathbb{Z}/\ell\mathbb{Z}) \tag{4.16}$$

by $\tilde{H}_\ell$, which is the intersection of G_ℓ with the group of (λ, μ), where λ and μ run through $(\mathbb{Z}/\ell\mathbb{Z})^*$. We see $\tilde{H}_\ell$ contains H_ℓ. On the other hand, if $(s, s') \in G_\ell$, we have $\det(s) = \det(s')$ from (2.10). It follows that if $(\lambda, \mu) \in \tilde{H}_\ell$, we have $\lambda^2 = \mu^2$, thus $\lambda = \pm\mu$; this shows that

$(\tilde{H}_\ell : H_\ell) = 1$ or 2. From here,

$$|G'_\ell| = |G_\ell/H_\ell| \leq 2|G_\ell/\tilde{H}_\ell| \leq 2|\operatorname{PGL}_2(\mathbb{Z}/\ell\mathbb{Z})|^2 \leq 2(\ell^3 - \ell)^2 \leq 2\ell^6,$$

which gives the lemma. □

The last lemma we state comes from Serre [30, Theorem 6], however we employ a different Chebotarev estimate to the one Serre originally used.

Lemma 4.1.4. *Let L be a Galois extension of $\mathbb{Q}$ of degree $n_L < \infty$, and S a finite set of primes such that $L/\mathbb{Q}$ is unramified outside S. Assume GRH. For any conjugacy class C of $G = \operatorname{Gal}(L/\mathbb{Q})$, there exists a prime number $q \notin S$ such that $\sigma_q \in C$ and*

$$q \leq \left[2\bar{a}\left(\log|d_L| + n_L \log\left(N_{S\setminus P(L)}\right)\right) + \bar{c}\right]^2 \tag{4.17}$$

where $\bar{a}, \bar{c}$ can be taken to be 4 and 5 respectively, or can be chosen from Table 2.2, assuming, also, that $\bar{b} = 0$.

Proof. Let $P(L)$ be the set of primes that are ramified in L. Given that all primes outside S must not ramify in L, we have $P(L) \subseteq S$. We consider now two cases:

i) Suppose $P(L) = S$. Apply Proposition 2.4.19 to L and the conjugacy class C, so that we get a prime $p \notin S$ such that

$$p \leq (\bar{a}\log|d_L| + \bar{c})^2.$$

Noting that (4.17) is larger than the above bound, we conclude the result in this case.

ii) Suppose $P(L) \subsetneq S$. Let D be the product of the primes $q \in S \setminus P(L)$, and let F be the unique quadratic field such that $d_F = D$ if D is odd, and $d_F = 2D$ if D is even; we have $F = \mathbb{Q}(\sqrt{\pm D})$ in the first case and $F = \mathbb{Q}(\sqrt{\pm D/2})$ in the second case, the sign being chosen such that $\pm D \equiv 1 \pmod 4$ and $\pm D/2 \equiv 3 \pmod 4$ respectively. As L is not ramified at any prime factors of d_F (d_L and d_F are coprime), the fields L and F are linearly disjoint (see [21, Theorem 4.26]). Their compound $L' = L.F$ is Galois, with Galois group $G' = G \times \{\pm 1\}$. As d_L and d_F are relatively prime, Proposition 2.1.13 gives

$$d_{L'} = (d_F)^{n_L}(d_L)^2. \tag{4.18}$$

In particular, $P(L')$ is equal to S; the extension $L'/\mathbb{Q}$ is ramified at all the prime numbers q belonging to S. Apply Proposition 2.4.19 to $L'/\mathbb{Q}$ and to $C \times \{\pm q\}$ to get a prime $p \notin S$ with $\sigma_p \in C$ and

$$p \leq (\bar{a}\log|d_{L'}| + \bar{c})^2. \tag{4.19}$$

According to (4.18), we have

$$\log|d_{L'}| = n_L \log|d_F| + 2\log|d_L|. \tag{4.20}$$

We have the simple bound $d_F \leq D^2$, and so

$$n_L \log|d_F| \leq 2n_L \log|D| \leq 2n_L \log\left(\prod_{q\in S\setminus P(L)} q\right) = 2n_L \log N_{S\setminus P(L)}. \tag{4.21}$$

We deduce therefore

$$\log|d_{L'}| \leq 2n_L \log N_{S\setminus P(L)} + 2\log|d_L| \tag{4.22}$$

which, when combining with (4.19), we get

$$p \leq (\bar{a}\log|d_{L'}| + \bar{c})^2 \leq \left[2\bar{a}\left(\log|d_L| + n_L \log\left(\prod_{q\in S\setminus P(L)} q\right)\right) + \bar{c}\right]^2,$$

giving the desired bound

□

Remark 4.1.5. One may wonder what allows us to assume $\bar{b} = 0$ in Lemma 4.1.4 given that no such stipulation is made in Proposition 2.4.19. Let us suppose we take a 4-tuple from Table 2.2 (which, recall, is a condensed version of Table 2.1) $(p_0, \bar{a}, \bar{b}, \bar{c})$. For most cases appearing, $\bar{b} \neq 0$. However, in each case when $\bar{b}$ is nonzero, we are given an explicit range $n_{lower} \leq n_K \leq n_{upper}$ for which the 4-tuple is valid (the final case, when $n \geq 50$, we are given $\bar{b} = 0$). Then, we can take

$$(\bar{a}\log|d_K| + \bar{b}n_K + \bar{c})^2 \leq (\bar{a}\log|d_K| + \bar{b}n_{upper} + \bar{c})^2,$$

and so our 4-tuple becomes $(p_0, \bar{a}, 0, \bar{b}n_{upper} + \bar{c})$, thus eliminating $\bar{b}$.

We may also subsume the term $\bar{b}n_K$ into $\bar{a}$. By Lemma 4.1.6 (appearing below), we have $(\frac{1}{2}\log 3)n_K \leq \log|d_K|$, thus $n_K \leq \frac{2}{\log 3}\log|d_K|$. Therefore,

$$(\bar{a}\log|d_K| + \bar{b}n_K + \bar{c})^2 \leq \left(\left(\bar{a} + \frac{2\bar{b}}{\log 3}\right)\log|d_K| + \bar{c}\right).$$

and so the 4-tuple becomes $(p_0, \bar{a} + \frac{2\bar{b}}{\log 3}, 0, \bar{c})$. One may then consider which of the two gives a smaller bound given the situation.

Lastly, we give a bound on the discriminant of a number field $K/\mathbb{Q}$. This is given in Serre [30, Proposition 6], but we use the version found in [18, Lemma 7]. We note that the

radical of an integer a, $\mathrm{rad}(a)$, is defined to be the product of all prime divisors of a, that is, $\mathrm{rad}(a) = \prod_{q|a} q$.

Lemma 4.1.6. *[18, Lemma 7] If $K/\mathbb{Q}$ is a nontrivial finite Galois extension, then*

$$\left(\frac{1}{2}\log 3\right)[K:\mathbb{Q}] \leq \log d_K \leq ([K:\mathbb{Q}]-1)\log \mathrm{rad}(d_K) + [K:\mathbb{Q}]\log([K:\mathbb{Q}]),$$

where d_K is the absolute value of the discriminant of K.

Proof. The left-hand side inequality follows from Theorem 2.1.4 (see [30, pg. 139]). For the right-hand side, let $\mathfrak{D}_{K/\mathbb{Q}} \subseteq \mathcal{O}_K$ denote the different ideal, and note by Proposotion 2.1.18 we have

$$d_K = N_{K/\mathbb{Q}}(\mathfrak{D}_{K/\mathbb{Q}}) = \prod_{q|d_K} q^{\nu_q(N_{K/\mathbb{Q}}(\mathfrak{D}_{K/\mathbb{Q}}))}.$$

Taking logarithms of both sides, we get

$$\log d_K = \sum_{q|d_K} \nu_q(N_{K/\mathbb{Q}}(\mathfrak{D}_{K/\mathbb{Q}}))\log q = \sum_{q|d_K}\sum_{\mathfrak{q}|q} f_{\mathfrak{q}}\nu_{\mathfrak{q}}(\mathfrak{D}_{K/\mathbb{Q}})\log q. \tag{4.23}$$

For each prime ideal $\mathfrak{q} \subseteq \mathcal{O}_K$ lying above q, we have that

$$\nu_{\mathfrak{q}}(\mathfrak{D}_{K/\mathbb{Q}}) = e_{\mathfrak{q}} - 1 + s_{\mathfrak{q}}$$

for some integer $s_{\mathfrak{q}}$ satisfying $0 \leq s_{\mathfrak{q}} \leq \nu_{\mathfrak{q}}(e_{\mathfrak{q}})$. Thus,

$$\sum_{\mathfrak{q}|q} f_{\mathfrak{q}}\nu_{\mathfrak{q}}(\mathfrak{D}_{K/\mathbb{Q}}) = \sum_{\mathfrak{q}|q} f_{\mathfrak{q}}(e_{\mathfrak{q}}-1) + \sum_{\mathfrak{q}|q} f_{\mathfrak{q}}s_{\mathfrak{q}} \leq [K:\mathbb{Q}] - 1 + \sum_{\mathfrak{q}|q} f_{\mathfrak{q}}\nu_{\mathfrak{q}}(e_{\mathfrak{q}}). \tag{4.24}$$

Since $K/\mathbb{Q}$ is Galois, $e_{\mathfrak{q}}$ divides $[K:\mathbb{Q}]$ (see [19, Theorem 3.34]). Thus, $\nu_q(e_{\mathfrak{q}}) \leq \nu_q([K:\mathbb{Q}])$. Hence,

$$\sum_{\mathfrak{q}|q} f_{\mathfrak{q}}\nu_{\mathfrak{q}}(e_{\mathfrak{q}}) = \sum_{\mathfrak{q}|q} f_{\mathfrak{q}}e_{\mathfrak{q}}\nu_q(e_{\mathfrak{q}}) \leq \nu_q([K:\mathbb{Q}])\sum_{\mathfrak{q}|q} f_{\mathfrak{q}}e_{\mathfrak{q}} = \nu_q([K:\mathbb{Q}])[K:\mathbb{Q}]. \tag{4.25}$$

Lastly,

$$\sum_{q|d_K} \nu_q([K:\mathbb{Q}])[K:\mathbb{Q}]\log q \leq [K:\mathbb{Q}]\log[K:\mathbb{Q}]. \tag{4.26}$$

Applying all the above to (4.23), we obtain

$$\log d_K \leq ([K:\mathbb{Q}]-1)\sum_{q|d_K}\log q + [K:\mathbb{Q}]\log[K:\mathbb{Q}].$$

The result is then obtained by replacing $\sum_{q|d_K}\log q$ with $\log \mathrm{rad}(d_K)$. □

Remark 4.1.7. We could also study what happens if we use Lemma 4.1.6 in the proof of Lemma 4.1.4, in particular applying the result of Lemma 4.1.6 to (4.20). We would have:

$$\begin{aligned}\log|d_{L'}| &= n_L \log|d_F| + 2\log|d_L| \\ &\leq n_L((n_F-1)\log\operatorname{rad} d_F + n_F \log n_F) + 2((n_L-1)\log\operatorname{rad} d_L + n_L \log n_L) \\ &= n_L(\log\operatorname{rad} d_F + 2\log 2) + 2((n_L-1)\log\operatorname{rad} d_L + n_L\log n_L) \\ &\leq n_L(\log\operatorname{rad} d_F + 2\log 2 + 2\log\operatorname{rad} d_L + 2\log n_L) \\ &= n_L\left(2\log\operatorname{rad} d_L + \log\left(\prod_{q\in S\setminus P(L)} q\right) + 2\log(2n_L)\right)\end{aligned} \tag{4.27}$$

We could then compare (4.27) with (4.22) to see which gives the better bound on $\log|d_{L'}|$. However, given the scope of this chapter is to expand on Serre's original argument, we choose to use Lemma 4.1.4 in its current form.

4.2 Improvement

With the preliminary work now taken care of, we work to prove Theorem 4.0.2. We begin with a proof of a simple lemma we will make use of.

Lemma 4.2.1. *Let $r, s > 0$, with $r > 1$ and satisfying $s > (1-\frac{1}{r})\log(2)$. Then, for any real $x \geq 2$, we have*

$$x^{1/r} \leq C_4(r,s)\frac{x}{(\log x)^s} \tag{4.28}$$

where $C_4(r,s) = \max\left\{2^{\frac{1-r}{r}}(\log(2))^s, (\frac{sr}{r-1})^s e^{-s}\right\}$.

Proof. Beginning with (4.28), we have

$$x^{1/r} \leq C_4(r,s)\frac{x}{(\log x)^s} \iff x^{\frac{1-r}{r}}(\log x)^s \leq C_4(r,s). \tag{4.29}$$

First, notice if $r = 1$, then (4.29) becomes $(\log x)^s \leq C_4(1,s)$, but the left is unbounded as x approaches infinity, so no such constant exists to satisfy the inequality.

Let $f(x) = x^{\frac{1-r}{r}}(\log x)^s$. Since $C_4(r,s)$ must obey the above inequality, taking $C_4(r,s)$ to be the maximum of the function $f(x)$ shall satisfy (4.28). To that end, we find

$$\frac{d}{dx}f(x) = -\frac{x^{\frac{1-r}{r}-1}(\log(x))^{s-1}\left((r-1)\log(x) - sr\right)}{r}.$$

There are two critical values: $x = 1$ and $x = e^{\frac{sr}{r-1}}$ (note that, if $s = 1$, then $x = 1$ is not a critical value). Since $x \geq 2$, we do not consider the first critical value $x = 1$. Since r and s

satisfy $s > (1 - \frac{1}{r})\log(2)$, we notice

$$s > (1 - \frac{1}{r})\log(2) \iff \frac{sr}{r-1} > \log(2) \iff e^{\frac{sr}{r-1}} > 2$$

so the second critical point is to the right of $x = 2$. Thus, we take

$$C_4(r,s) = \max\{f(2), f(e^{\frac{sr}{r-1}})\} = \max\left\{2^{\frac{1-r}{r}}(\log(2))^s, \left(\frac{sr}{r-1}\right)^s e^{-s}\right\}$$

as desired. □

Now we prove Theorem 4.0.2.

Proof of Theorem 4.0.2. Let p be the smallest prime such that $a_p(E) \neq a_p(E')$. Let $n = |a_p(E) - a_p(E')| > 0$. By Lemma 4.1.2, there exists a prime ℓ such that $n \not\equiv 0 \pmod{\ell}$, and $\ell \leq C_3 \log 2n$; in particular, we know $a_p(E) \not\equiv a_p(E') \pmod{\ell}$. In addition, by Theorem 2.2.4 (Hasse's Theorem),

$$\begin{aligned} \ell &\leq C_3 \log(2|a_p(E) - a_p(E')|) \\ &\leq C_3 \log(2(|a_p(E)| + |a_p(E')|) \\ &\leq C_3 \log(8p^{1/2}) \\ &\leq C_5 \log p \end{aligned} \tag{4.30}$$

where $C_5 = (C_3 \log 8 + 1/2)$.

Let L be the subfield of $\bar{\mathbb{Q}}$ fixed by the kernel of the homomorphism $\mathrm{Gal}(\bar{\mathbb{Q}}/\mathbb{Q}) \to G_\ell \to G'_\ell$, where we recall that G_ℓ is the image of the map ρ_ℓ defined in (4.13), and $G'_\ell = G_\ell/H_\ell$. The extension L is galois, with galois group G'_ℓ. It is not ramified outside of the set $S_\ell = P(E) \cup P(E') \cup \{\ell\}$. Apply Lemma 4.1.4 to this extension and to a conjugacy class of G'_ℓ containing C'_ℓ to get a prime $q \notin S_\ell$ and $\sigma_q \in C'_\ell$ satisfying

$$\begin{aligned} q &\leq [2\bar{a}(\log|d_L| + n_L \log(N_{S_\ell \setminus P(L)})) + \bar{c}]^2 \\ &\leq [2\bar{a}((n_L - 1)\log \mathrm{rad}(d_L) + n_L \log(n_L) + n_L \log(N_{S_\ell \setminus P(L)})) + \bar{c}]^2 \\ &\leq [2\bar{a} n_L(\log N_S + \log(n_L)) + \bar{c}]^2 \\ &\leq [2\bar{a}(2\ell^6)(\log N_S + \log(2\ell^6)) + \bar{c}]^2 \\ &\leq [4\bar{a}\ell^6(\log N_S + \log 2 + 6\log \ell) + \bar{c}]^2 \\ &\leq 16(\log 2)^2 \bar{a}^2 \ell^{12}[\log N_S + 6\log \ell + \bar{c}]^2 \\ &= C_6 \ell^{12}[\log N_S + 6\log \ell + \bar{c}]^2 \end{aligned} \tag{4.31}$$

where we have used Lemma 4.1.6 on the second line, $|G_\ell| \leq 2\ell^6$ coming from Lemma 4.1.3, and $C_6 = 16(\log 2)^2\bar{a}^2$. Since p is the smallest prime such that $a_p(E) \neq a_p(E')$, and $a_q(E) \neq a_q(E')$, we have $p \leq q \leq C_6 \ell^{12}[\log N_S + 6\log \ell + \bar{c}]^2$.

We now work to remove ℓ. Recall, by (4.30), we have $\ell \le C_5 \log p$. Thus,

$$\begin{aligned} p &\le C_6(C_5 \log p)^{12}[\log N_S + 6\log(C_5 \log p) + \bar{c}]^2 \\ &\le C_7(\log p)^{12}[\log N_S + \log\log p + \bar{c}]^2 \end{aligned} \tag{4.32}$$

where $C_7 = 6C_6C_5^{12}(1+\log C_5)^2$.

Divide both sides of (4.32) by $(\log p)^{14}$ to get

$$\begin{aligned} \frac{p}{(\log p)^{14}} &\le C_7\left[\frac{\log N_S + \log\log p + \bar{c}}{\log p}\right]^2 \\ &= C_7\left[\frac{\log N_S + \bar{c}}{\log p} + \frac{\log\log p}{\log p}\right]^2 \\ &\le C_7\left[\log N_S + \bar{c} + 1\right]^2 \end{aligned} \tag{4.33}$$

where we have used the estimates

$$\frac{\log N_S + \bar{c}}{\log p} \le \log N_S + \bar{c} \quad \text{and} \quad \frac{\log\log p}{\log p} \le 1.$$

Apply Lemma 4.2.1 with $r = 2$ and $s = 14$ (notice $14 > \frac{\log(2)}{2}$) to get

$$p^{1/2} \le C_4(2,14)\frac{p}{(\log p)^{14}} \le C_8\left[\log N_S + \bar{c} + 1\right]^2$$

with $C_8 = C_4(2,14)C_7$. In particular,

$$\begin{aligned} \log p &\le 2\left(\log C_8 + 2\log(\log N_S + \bar{c} + 1)\right) \\ &\le 2(\log C_8 + 2)\log(\log N_S + \bar{c} + 1) \\ &= C_9 \log(\log N_S + \bar{c} + 1) \\ &\le C_9 \log((1 + \bar{c} + 1)\log N_S) \\ &= C_9(\log(2+\bar{c}) + \log\log N_S) \\ &\le C_9(\log(2+\bar{c}) + 1)\log\log N_S) \\ &= C_{10}\log\log N_S \end{aligned} \tag{4.34}$$

with $C_9 = 2(\log C_8 + 2)$ and $C_{10} = C_9(\log(2+\bar{c}) + 1)$. Lastly, combining (4.34) with (4.32) yields

$$\begin{aligned} p &\le C_7(\log p)^{12}\left[\log N_S + \log\log p + \bar{c}\right]^2 \\ &\le C_7C_{10}^{12}(\log\log N_S)^{12}\left[\log N_S + \log(C_{10}\log\log N_S) + \bar{c}\right]^2 \\ &\le C_{11}(\log\log N_S)^{12}\left[\log N_S + (\log C_{10} + 1)\log\log\log N_S + \bar{c}\right]^2 \\ &\le C_{11}(\log C_{10} + 1)(\log\log N_S)^{12}\left[\log N_S + \log\log\log N_S + \bar{c}\right]^2 \\ &= C_{12}(\log\log N_S)^{12}\left[\log N_S + \log\log\log N_S + \bar{c}\right]^2 \end{aligned} \tag{4.35}$$

with $C_{11} = C_7 C_{10}^{12}$ and $C_{12} = C_{11}(\log C_{10} + 1)$. □

Remark 4.2.2. Explicitly computing the constant C_{12} reveals it is quite large. Using $\bar{a} = 4$ and $\bar{c} = 5$, we find

$$C_{12} \approx 2.79 \times 10^{45}.$$

The size of C_{12} makes any practical use of Theorem 4.0.2 difficult. We will find the results of Chapter 5 are a vast improvement.

Chapter 5

The Method of Mayle-Wang

Recent work of Mayle-Wang [18] has given an explicit result on the smallest prime which achieves $a_p(E) \neq a_p(E')$. The constants are quite small, and, like Serre, depend on only knowledge of the primes of bad reduction of the two elliptic curves E and E'. We recall from (4.2) that $\mathfrak{N} = N_{P(E)}N_{P(E')}$ for two elliptic curves E and E'.

Theorem 5.0.1. *[18, Theorem 2] Assume GRH. Let E,E′ be two elliptic curves over $\mathbb{Q}$ without complex multiplication. Suppose E and E′ are not $\mathbb{Q}$-isogenous. Then there exists a prime p of good reduction for E and E′ such that $a_p(E) \neq a_p(E')$ and satisfying the inequality*

$$p \leq (482 \log \operatorname{rad}(2\mathfrak{N}) + 2880)^2. \tag{5.1}$$

In this chapter, we work to give a generalized result of Theorem 5.0.1 which will follow the same structure as Mayle-Wang, but where we choose to leave unevaluated a critical bound on a particular Galois group. Then, we shall use this generalization to give an improvement on the constants appearing in Theorem 5.0.1 in the case when the mod 2 representations are either isomorphic and irreducible, or not isomorphic.

Let G be a group, and suppose ρ_1 and ρ_2 are two non-isomorphic representations of G. The set $\delta(G)$ (described precisely in Definition 5.1.1) is known as the deviation group, and it is a finite set containing a subset which acts as a certificate that ρ_1 and ρ_2 are not isomorphic. The set $\varphi(G)$ (defined in Proposition 5.1.6) is a subset of a very explicit semi-direct product, and so estimating $|\varphi(G)|$ is much easier. We now give the general result:

Proposition 5.0.2. *Assume GRH. Let G be a group. Let E,E′ be two elliptic curves over $\mathbb{Q}$. Suppose E and E′ are not $\mathbb{Q}$-isogenous. Let $\delta(G)$ be the deviation group of G with respect to the ℓ-adic representations $\rho_{E,2}$ and $\rho_{E',2}$. Then there exists a prime p of good reduction for E and E′ such that $a_p(E) \neq a_p(E')$, and a 4-tuple $(p_0, \bar{a}, \bar{b}, \bar{c})$ from Table 2.2 (chosen such that the degree is equal to $2|\delta(G)|$) such that*

$$p \leq \max\{p_0, (\bar{a}((2|\delta(G)| - 1)\log \operatorname{rad}(2\mathfrak{N}) + 2|\delta(G)|\log(2|\delta(G)|)) + 2|\delta(G)|\bar{b} + \bar{c})^2\}. \tag{5.2}$$

Furthermore, if E and E' are such that their mod 2 representations are isomorphic and absolutely irreducible, then we may replace $|\delta(G)|$ with $|\varphi(G)|$.

We provide a proof in Section 5.3. Improvements to the size of the group $\delta(G)$ appearing above shall lead to further improvements in the constants.

What follows is our main result, in which we have employed Proposition 5.0.2 with an estimate on the size of $|\varphi(G)|$ to give an improvement on Theorem 5.0.1 when the representations are isomorphic and irreducible, or not isomorphic.

Theorem 5.0.3. *Assume GRH. Let E, E' be two elliptic curves over $\mathbb{Q}$. Suppose E and E' are not $\mathbb{Q}$-isogenous. Assume the mod 2 representations $\bar{\rho}_{E,2}$ and $\bar{\rho}_{E',2}$ are not isomorphic, or if they are isomorphic that they are absolutely irreducible. Then there exists a prime p of good reduction for E and E' such that $a_p(E) \neq a_p(E')$ and satisfying the inequality*

$$p \leq \max\{3100065, (120 \log \operatorname{rad}(2\mathfrak{N}) + 559)^2\}. \tag{5.3}$$

A proof is given in Section 5.3.

Remark 5.0.4. Mayle-Wang, in Theorem 5.0.1, include a hypothesis that the elliptic curves E and E' be without complex multiplication; in Proposition 5.0.2 and Theorem 5.0.3, we have dropped this assumption. The original result of Serre in Chapter 4 does not require such a hypothesis. All we require here is that the set of primes for which $a_p(E) \neq a_p(E')$ is infinite, which is satisfied once we assume the two elliptic curves are not $\mathbb{Q}$-isogenous, a consequence of Faltings' Theorem (Theorem 2.3.7).

5.1 The Deviation Group $\delta(G)$

As in Chapter 4, we required a way to determine when two representations are not isomorphic, by considering their trace value at a group element. The method, in this chapter, differs from the previous strategy. Here, we wish to construct a finite group, called the *deviation group*, denoted $\delta(G)$, from which we can find a finite subset that will determine if the two representations are isomorphic or not.

Our treatment of the deviation group will follow the exposition given in Ignasi's thesis [26]. We note that Ignasi's exposition is, itself, taken from Chênevert's thesis [9], whose work follows the work of Serre [31] (the propositions and lemmas which appear here, with the exception of Lemma 5.1.8, can also be found in in [9, Chapter 5]). Let G be a group, and L a local field (a finite extension of $\mathbb{Q}_\ell$, for ℓ prime) with ring of integers $\mathcal{O}_\lambda$, maximal ideal λ, and residue field $k = \mathcal{O}_\lambda/\lambda\mathcal{O}_\lambda$. We let π be a uniformizer, so $\lambda = \pi\mathcal{O}_\lambda$. Let $\rho_1, \rho_2 : G \to \mathrm{GL}_n(\mathcal{O}_\lambda)$ be two λ-adic representations. We begin by extending the map $\rho_1 \times \rho_2 : G \to \mathrm{GL}_n(\mathcal{O}_\lambda) \times \mathrm{GL}_n(\mathcal{O}_\lambda)$ from G to the group ring $\mathcal{O}_\lambda[G]$. Recall that the group ring $\mathcal{O}_\lambda[G]$ is a $\mathcal{O}_\lambda$-module with a basis being the elements of G; explicitly, it can be written

as

$$\mathcal{O}_\lambda[G] = \left\{ \sum a_i g_i \mid a_i \in \mathcal{O}_\lambda, g_i \in G, \text{ with only finitely many } a_i \text{ nonzero} \right\}.$$

We define the map $\rho : \mathcal{O}_\lambda[G] \to M_n(\mathcal{O}_\lambda) \oplus M_n(\mathcal{O}_\lambda)$, where $M_n(\mathcal{O}_\lambda)$ is defined to be the set of all matrices with entries in $\mathcal{O}_\lambda$, to be

$$\rho\left(\sum a_i g_i\right) = \left(\sum a_i \rho_1(g_i), \sum a_i \rho_2(g_i)\right).$$

Note that the image need not be contained in $\mathrm{GL}_n(\mathcal{O}_\lambda) \times \mathrm{GL}_n(\mathcal{O}_\lambda)$, as, in general, $\mathrm{GL}_n(\mathcal{O}_\lambda)$ need not be closed under taking $\mathcal{O}_\lambda$-linear combinations.

Let M be the full image of ρ inside $M_n(\mathcal{O}_\lambda) \oplus M_n(\mathcal{O}_\lambda)$, and consider the composition map $\delta : G \xrightarrow{\rho} M^\times \to (M/\lambda M)^\times$.

Definition 5.1.1. [26, Definition 2.1.1] The image $\delta(G)$ of G inside $(M/\lambda M)^\times$ is called the *deviation group* of the pair of representations ρ_1, ρ_2.

Remark 5.1.2. Since M is a subalgebra of $R = M_n(\mathcal{O}_\lambda) \times M_n(\mathcal{O}_\lambda)$, it might be tempting to think $\delta(G)$ is a subgroup of $(R/\lambda R)^\times = \mathrm{GL}_2(k) \times \mathrm{GL}_2(k)$ but this may not be the case. See the remark after [26, Definition 2.1.1].

The deviation group turns out to be finite, as described by the following proposition.

Proposition 5.1.3. *[26, Proposition 2.1.2] The group $\delta(G)$ is finite, and in particular we have $|\delta(G)| \leq |k|^{2n^2}$.*

Proof. M is a submodule of the free $\mathcal{O}_\lambda$-module $M_n(\mathcal{O}_\lambda) \oplus M_n(\mathcal{O}_\lambda)$. Since $\mathcal{O}_\lambda$ is a local ring, M is free and is of rank r, where r satisfies

$$r \leq \mathrm{rank}(M_n(\mathcal{O}_\lambda) \oplus M_n(\mathcal{O}_\lambda)) = 2n^2.$$

Given M is a $\mathcal{O}_\lambda$-module, $M/\lambda M$ is a k-algebra of dimension r. Hence,

$$|\delta(G)| \leq |(M/\lambda M)^\times| \leq |k|^r \leq |k|^{2n^2}$$

as claimed. □

Remark 5.1.4. A similar bound on $|\delta(G)|$ is employed by Mayle-Wang in their proof of Theorem 5.0.1, although they do not explicitly mention the deviation group (they implicitly work with it). See the proof in [18, Theorem 2].

Let us turn our attention now to the practical use of $\delta(G)$, that being its ability to help us determine when two representations are isomorphic.

Proposition 5.1.5. *[26, Proposition 2.1.3] Let $\Sigma \subseteq G$ be a subset that surjects onto $\delta(G)$. Then, $\rho_1 \sim \rho_2$ if and only if $\mathrm{tr}(\rho_1(g)) = \mathrm{tr}(\rho_2(g))$ for all $g \in \Sigma$.*

Proof. ($\Longrightarrow$) If $\rho_1 \sim \rho_2$, then it must be the case that $\mathrm{tr}(\rho_1(g)) = \mathrm{tr}(\rho_2(g))$.

($\Longleftarrow$) Suppose $\mathrm{tr}(\rho_1(g)) = \mathrm{tr}(\rho_2(g))$ for all $g \in \Sigma$, but that $\rho_1 \not\sim \rho_2$. Then, there is some $g_0 \in G$ such that $\mathrm{tr}(\rho_1(g_0)) \neq \mathrm{tr}(\rho_2(g_0))$. Since this is not an equality in $\mathcal{O}_\lambda$, it implies there exists an integer $\alpha \geq 1$ such that

$$\mathrm{tr}(\rho_1(g_0)) \equiv \mathrm{tr}(\rho_2(g_0)) \pmod{\lambda^\alpha} \quad \text{and} \quad \mathrm{tr}(\rho_1(g_0)) \not\equiv \mathrm{tr}(\rho_2(g_0)) \pmod{\lambda^{\alpha+1}}.$$

Letting π be our uniformizer so that $\lambda = \pi\mathcal{O}_\lambda$ (as above), define the map

$$\begin{aligned} \tilde{\phi} : G &\to \mathcal{O}_\lambda \\ g &\mapsto \pi^{-\alpha}(\mathrm{tr}(\rho_2(g)) - \mathrm{tr}(\rho_1(g))). \end{aligned} \tag{5.4}$$

Our objective now is to descend $\tilde{\phi}$ to a new map Φ whose domain is $\delta(G)$ instead of G, and whose codomain is k instead of $\mathcal{O}_\lambda$. Since Σ surjects into $\delta(G)$, we will be able to find an element $g' \in \Sigma$ such that $\Phi(\delta(g'))$ is not in λM, and hence, the traces restricted to the set Σ will also be nonequal, which is what we want to prove.

We extend the map $\tilde{\phi}$ to an $\mathcal{O}_\lambda$-linear map

$$\phi : M \to \mathcal{O}_\lambda \tag{5.5}$$

which satisfies the following commutative diagram:

$$\begin{array}{ccc} G & \xrightarrow{\tilde{\phi}} & \mathcal{O}_\lambda \\ {\scriptstyle i}\big\downarrow & & \big\uparrow{\scriptstyle \phi} \\ \mathcal{O}_\lambda[G] & \xrightarrow{\rho} & M \end{array}$$

We claim that $\phi(M) \not\subseteq \lambda$, that is, ϕ is not the zero map. Given ϕ satisfies the above diagram, notice that $\phi(\rho(i(g_0))) = \pi^{-\alpha}(\mathrm{tr}(\rho_2(g_0)) - \mathrm{tr}(\rho_1(g_0))) \notin \lambda$ since $\mathrm{tr}(\rho_2(g_0)) - \mathrm{tr}(\rho_1(g_0)) \not\equiv 0 \pmod{\lambda^{\alpha+1}}$ by our assumption above.

Now considering quotients, the map ϕ descends to a nonzero k-linear map $M/\lambda M \to k$, hence, since $\delta(G) \subseteq (M/\lambda M)^\times$, to a function

$$\Phi : \delta(G) \to k. \tag{5.6}$$

Notice that $\rho(G)$ is a basis for $M = \rho(\mathcal{O}_\lambda[G])$ since G is a basis for $\mathcal{O}_\lambda[G]$ with coefficients in $\mathcal{O}_\lambda$. Since δ is given by the composition of ρ with the canonical homomorphism from $M^\times \to (M/\lambda M)^\times \subseteq M/\lambda M$, we see that $\delta(G)$ must span $M/\lambda M$. Thus, Φ is a nonzero map into k.

Given that Σ surjects onto $\delta(G)$, there exists a $g' \in \Sigma$ such that $\Phi(\delta(g')) \neq 0$, that is,

$$\phi(\rho(g')) = \pi^{-\alpha}(\mathrm{tr}(\rho_2(g')) - \mathrm{tr}(\rho_1(g'))) \notin \lambda.$$

In particular, $\mathrm{tr}(\rho_1(g')) \neq \mathrm{tr}(\rho_2(g'))$ which contradicts our initial assumption. $\square$

Before we introduce the next proposition, some exposition is need (following [26]). We assume now the representations $\rho_1, \rho_2 : G \to \mathrm{GL}_n(\mathcal{O}_\lambda)$ are not isomorphic, that is, they are not conjugate in $\mathrm{GL}_n(\mathcal{O}_\lambda)$, but that the residual representations $\bar{\rho}_1$ and $\bar{\rho}_2$ obtained from ρ_1 and ρ_2 by reduction modulo λ are isomorphic (see Definition 2.3.4 for this notation applied to the case of ℓ-adic representations). We then have an equality $\bar{\rho}_1 = P\bar{\rho}_2 P^{-1}$ for some matrix $P \in M_n(k)$.

Define β to be the largest integer such that ρ_1 and ρ_2 are conjugated modulo λ^β, that is, there is a matrix $P \in \mathrm{GL}_n(\mathcal{O}_\lambda)$ such that $\rho_1 \equiv P\rho_2 P^{-1} \pmod{\lambda^\beta}$; we then have $\beta \geq 1$, since $\bar{\rho}_1 \cong \bar{\rho}_2$. In addition, we showed above (in the proof of Proposition 5.1.5) that there is an integer $\alpha \geq 1$ such that $\mathrm{tr}(\rho_1) \equiv \mathrm{tr}(\rho_2) \pmod{\lambda^\alpha}$ and $\mathrm{tr}(\rho_1) \not\equiv \mathrm{tr}(\rho_2) \pmod{\lambda^{\alpha+1}}$; in particular, ρ_1 and ρ_2 are not conjugate modulo $\lambda^{\alpha+1}$, so $\beta \leq \alpha$. Given that ρ_1 and ρ_2 are conjugate modulo λ^β but not conjugate modulo $\lambda^{\beta+1}$, if we replace ρ_2 with a conjugate we may assume $\rho_1 \equiv \rho_2 \pmod{\lambda^\beta}$ but $\rho_1 \not\equiv \rho_2 \pmod{\lambda^{\beta+1}}$.

Hence, for any $g \in G$, we have

$$\rho_2(g) - \rho_1(g) \equiv 0 \pmod{\lambda^\beta} \Rightarrow \rho_2(g) - \rho_1(g) = \theta_g \pi^\beta \tag{5.7}$$

for some $\theta_g \in M_n(\mathcal{O}_\lambda)$ and π a uniformizer of λ. Rearranging, we get an equation for $\rho_2(g)$ of the form

$$\rho_2(g) = (I_n + \theta_g \pi^\beta \rho_1(g)^{-1})\rho_1(g). \tag{5.8}$$

Let $\theta : G \to M_n(\mathcal{O}_\lambda)$ be the map $g \to \theta_g \rho_1(g)^{-1}$, and notice that (5.8) becomes

$$\rho_2(g) = (I_n + \pi^\beta \theta(g))\rho_1(g). \tag{5.9}$$

Proposition 5.1.6. *[26, Proposition 2.2.1] Let $\rho_1, \rho_2 : G \to \mathrm{GL}_n(\mathcal{O}_\lambda)$ be representations that are not isomorphic, and suppose $\bar{\rho}_1, \bar{\rho}_2 : G \to \mathrm{GL}_n(k)$ are isomorphic. Let β be the largest integer such that ρ_1 and ρ_2 are conjugate modulo β, and as above, assume ρ_2 has been replaced by a conjugate such that $\rho_1 \equiv \rho_2 \pmod{\lambda^\beta}$. Let*

$$\begin{aligned} \varphi : G &\to M_n(k) \rtimes \mathrm{GL}_n(k) \\ g &\mapsto (\theta(g) \pmod{\lambda}, \rho_1(g) \pmod{\lambda}) \end{aligned} \tag{5.10}$$

where the semidirect product is with respect to the action of $\mathrm{GL}_n(k)$ on $M_n(k)$ by conjugation, that is multiplication is given by

$$(A, B) \cdot (C, D) = (A + BCB^{-1}, BD).$$

Then φ is a group homomorphism which factors through the deviation group $\delta(G)$.

Proof. First, let us show φ is a group homomorphism. That is, given $g, h \in G$, we must show

$$\begin{aligned}\varphi(gh) &= (\theta(gh) \pmod{\lambda}, \rho_1(gh) \pmod{\lambda}) \\ &= (\theta(g) + \rho_1(g)\theta(h)\rho_1(g)^{-1} \pmod{\lambda}, \rho_1(g)\rho_1(h) \pmod{\lambda}) \\ &= \varphi(g)\varphi(h),\end{aligned}$$

where the operation on $M_n(k)$ is addition. Since ρ_1 is a homomorphism, we know $\rho_1(gh) = \rho_1(g)\rho_1(h)$, and so the second component splits appropriately. Hence, it suffices that we show the first component splits, that is,

$$\varphi(gh)_1 = \theta(g) + \rho_1(g)\theta(h)\rho_1(g)^{-1} \pmod{\lambda}.$$

By (5.8), we have

$$\rho_2(g) = (I_n + \theta_g \pi^\beta \rho_1(g)^{-1})\rho_1(g) = (I_n + \pi^\beta \theta(g))\rho_1(g) \tag{5.11}$$

and

$$\rho_2(h) = (I_n + \theta_h \pi^\beta \rho_1(h)^{-1})\rho_1(h) = (I_n + \pi^\beta \theta(h))\rho_1(h) \tag{5.12}$$

and

$$\rho_2(gh) = (I_n + \theta_{gh} \pi^\beta \rho_1(gh)^{-1})\rho_1(gh) = (I_n + \pi^\beta \theta(gh))\rho_1(gh). \tag{5.13}$$

Using the right-hand sides of (5.11) and (5.12), we have

$$\begin{aligned}\rho_2(gh) &= \rho_2(g)\rho_2(h) \\ &= (I_n + \pi^\beta \theta(g))\rho_1(g)(I_n + \pi^\beta \theta(h))\rho_1(h) \\ &= \rho_1(g)\rho_1(h) + \pi^\beta(\theta(g)\rho_1(g)\rho_1(h) + \rho_1(g)\theta(h)\rho_1(h)) + \pi^{2\beta}\theta(g)\rho_1(g)\theta(h)\rho_1(h).\end{aligned} \tag{5.14}$$

Equating the right-hand side of (5.14) with the right-hand side of (5.13), then multiplying by $\rho_1(gh)^{-1}$ on the right and by $\pi^{-\beta}$, we obtain an equation for $\theta(gh)$:

$$\theta(gh) = \theta(g) + \rho_1(g)\theta(h)\rho_1(g)^{-1} + \pi^\beta \theta(g)\rho_1(g)\theta(h)\rho_1(g)^{-1}.$$

Since $\beta \geq 1$, reducing modulo $\lambda = \pi\mathcal{O}_\lambda$, we obtain the desired equality

$$\varphi(gh)_1 = \theta(gh) \pmod{\lambda} = \theta(g) + \rho_1(g)\theta(h)\rho_1(g)^{-1} \pmod{\lambda}.$$

This shows that φ is a group homomorphism. Now let us show φ factors through $\delta(G)$, i.e. let us show $\ker(\delta) \subseteq \ker(\varphi)$. Let $g \in \ker(\delta)$. Since $(\rho_1 \times \rho_2)(g) = \rho(g) \in I_n + \lambda M$, by the definition of M there exists a subset $\{a_h\}_{h \in G} \subseteq \mathcal{O}_\lambda$ with $a_h = 0$ for almost all $h \in G$

such that

$$\rho(g) = I_n + \pi \sum_{h \in G} a_h \rho(h).$$

Since this is a cartesian product $\rho_1 \times \rho_2(g)$, the equation above breaks down into a pair of equations

$$\rho_i(g) = I_n + \pi \sum_{h \in G} a_h \rho_i(h). \tag{5.15}$$

For $i = 1$, this implies $\rho_1(g) \equiv I_n \pmod{\lambda}$ (since $\lambda = \pi \mathcal{O}_\lambda$). This gives that the second component of $\varphi(g)$ is the identity element in $\mathrm{GL}_n(k)$. Moreover, for $i = 2$, we equate the right-hand side of (5.11) with the right-hand side of (5.15) and find that

$$\begin{aligned}
\rho_1(g) + \pi^\beta \theta(g) \rho_1(g) &= I_n + \pi \sum_{h \in G} a_h \rho_2(h) \\
&= I_n + \pi \sum_{h \in G} a_h (\rho_1(h) + \pi^\beta \theta(h) \rho_1(h)) \\
&= I_n + \pi \sum_{h \in G} a_h \rho_1(h) + \pi^{\beta+1} \sum_{h \in G} a_h \theta(h) \rho_1(h).
\end{aligned}$$

Subtracting $\rho_1(g) = I_n + \pi \sum_{h \in G} a_h \rho_1(h)$ and multiplying by $\pi^{-\beta}$, we see

$$\theta(g) \rho_1(g) = \pi \sum_{h \in G} a_h \theta(h) \rho_1(h)$$

which, upon multiplying by $\rho_1(g)^{-1}$ on the right becomes

$$\theta(g) = \pi \sum_{h \in G} a_h \theta(h) \rho_1(h g^{-1}).$$

Reducing modulo λ, we see

$$\theta(g) = \pi \sum_{h \in G} a_h \theta(h) \rho_1(h g^{-1}) \equiv 0 \pmod{\lambda}.$$

Therefore, the first component of $\varphi(g)$ is the identity element in $M_n(k)$. Thus, we have

$$\varphi(g) = (\theta(g) \pmod{\lambda}, \rho_1(g) \pmod{\lambda}) = (0, I_n)$$

hence $g \in \ker(\varphi)$. □

Remark 5.1.7. The homomorphism $\delta(G) \twoheadrightarrow \varphi(G)$ may not be injective. See [26, Remark 2.2.2].

Lastly, we prove a general lemma regarding determinants of matrices that we shall employ later.

Lemma 5.1.8. *[26, Lemma 2.2.3] Let R be a discrete valuation ring with uniformizer π, and F its field of fractions. For any $A \in \mathrm{GL}_n(F)$,*

$$\det(I_n + \pi A) = 1 + \pi \operatorname{tr}(A) + O(\pi^2).$$

Proof. Considering characteristic polynomials and working in F, we have

$$\begin{aligned}
\det(I_n + \pi A) &= \det(\pi(\pi^{-1} I_n + A)) \\
&= \pi^n \det(\pi^{-1} I_n + A) \\
&= \pi^n \left(\frac{1}{\pi^n} + \frac{1}{\pi^{n-1}} \operatorname{tr}(A) + \ldots + \det(A) \right) \\
&= 1 + \pi \operatorname{tr}(A) + \ldots + \pi^n \det(A) \\
&= 1 + \pi \operatorname{tr}(A) + O(\pi^2).
\end{aligned}$$

□

It can be difficult to compute the exact size of $\delta(G)$, or find a tighter upper bound for it. We will, in the following section, work to replace $\delta(G)$ with $\varphi(G)$ in the case of 2-adic representations. The codomain of φ is easily understood, and hence a bound for $|\varphi(G)|$ is easily computable. This is what allows us to prove Theorem 5.0.3.

5.2 The Tools of Mayle-Wang

The methodology of Mayle-Wang relies on the following proposition that is due to Serre (a proof of which can be found in [5, Theorem 4.7]). The proposition which follows is a refined version of Serre's original argument due to Mayle-Wang [18, Proposition 12] in which we have reworked the statement and proof to follow the work and notation done in Section 5.1. We note that the statement is similar to that of Proposition 5.1.5: in the previous proposition, we could decide on if two representations are isomorphic by considering the traces on a finite set; here, we show that if the representations are not isomorphic, then their traces must disagree on some finite set. While the proofs are very similar, the advantage of the following proposition is that it is in a form that we may easily apply Chebotarev to.

Proposition 5.2.1. *[18, Proposition 12] Let n a positive integer. Let G be a group and $\rho_1, \rho_2 : G \to \mathrm{GL}_n(\mathcal{O}_\lambda)$ be a group homomorphism, and $\delta(G)$ the deviation group of G with respect to the two representations ρ_1 and ρ_2. Suppose that there exists an element $g \in G$ such that $\operatorname{tr} \rho_1(g) \neq \rho_2(g)$. Then there exists a subset $C \subseteq \delta(G)$ for which*

1. *the set C is non-empty and closed under conjugation by $\delta(G)$, and*
2. *if the image in $\delta(G)$ of an element $g \in G$ belongs to C, then $\operatorname{tr} \rho_1(g) \neq \operatorname{tr} \rho_2(g)$.*

Proof. Let $R := M_n(\mathcal{O}_\lambda) \times M_n(\mathcal{O}_\lambda)$. Let M denote the $\mathcal{O}_\lambda$-subalgebra of R generated by the image of G under the product map

$$\rho_1 \times \rho_2 : G \to \mathrm{GL}_n(\mathcal{O}_\lambda) \times \mathrm{GL}_n(\mathcal{O}_\lambda).$$

Recall that $\delta(G)$ is the image of G under $\rho_1 \times \rho_2$ in $M/\lambda M$.

Let α be the largest nonnegative integer such that for each $g \in G$, one has that

$$\mathrm{tr}(\rho_1(g)) \equiv \mathrm{tr}(\rho_2(g)) \pmod{\lambda^\alpha}.$$

As M is a $\mathcal{O}_\lambda$-subalgebra generated by the image of G under $\rho_1 \times \rho_2$, it follows that the congruence $\mathrm{tr}\, x_1 \equiv \mathrm{tr}\, x_2 \pmod{\lambda^\alpha}$ holds for each pair $(x_1, x_2) \in M$. We obtain the $\mathcal{O}_\lambda$-module homomorphism $\phi : M \to \mathcal{O}_\lambda$ given by

$$\phi(x_1, x_2) = \pi^{-\alpha}(\mathrm{tr}(x_2) - \mathrm{tr}(x_1)).$$

Since $\phi(\lambda M) \subseteq \lambda\mathcal{O}_\lambda$, we may consider the induced $\mathcal{O}_\lambda/\lambda\mathcal{O}_\lambda$-module homomorphism $\bar{\phi} : M/\lambda M \to \mathcal{O}_\lambda/\lambda\mathcal{O}_\lambda$.

Let $C = \delta(G) \setminus \ker \bar{\phi}$ be the set of elements in $\delta(G)$ whose image under $\bar{\phi}$ in $M/\lambda M$ all are nonzero. From the definition of α and the linearity of the trace map, there exists $g_0 \in G$ such that

$$\mathrm{tr}(\rho_1(g_0)) \not\equiv \mathrm{tr}(\rho_2(g_0)) \pmod{\lambda^{\alpha+1}}.$$

Note that the image of $(\rho_1 \times \rho_2)(g_0)$ in $\delta(G)$ is contained in C, so C is nonempty. Also, C is closed under conjugation since the trace map is invariant under conjugation.

Finally, suppose that $g \in G$ is such that the image of g in $\delta(G)$ is contained in C. Then, $\phi(\rho_1 \times \rho_2(g)) \notin \lambda\mathcal{O}_\lambda$, and in particular $\mathrm{tr}\, \rho_1(g) \neq \mathrm{tr}\, \rho_2(g)$. □

We now give an analogous version of Proposition 5.2.1 in the case where the mod 2 representations are isomorphic and absolutely irreducible. This allows us to replace $\delta(G)$ in Proposition 5.2.1 with $\varphi(G)$ from Proposition 5.1.6, a set which is easier to estimate the size of. The idea to replace $\delta(G)$ with $\varphi(G)$ comes from Chênevert [9, pg. 114], in which he gives a remark that, in the 2-adic case, Serre [31] implies that $\delta(G) \cong \varphi(G)$. However, in a conversation with Chênevert, Serre mentions he might not have proven the map $\delta(G) \to \varphi(G)$ was an isomorphism, but, in an unpublished letter to Tate, that the α in the proof of Proposition 5.1.5 is equal to the β coming from the construction of the function φ. We show, in the 2-adic case, that $\alpha = \beta$, and that we can replace $\delta(G)$ in Proposition 5.2.1 with $\varphi(G)$ and get the same conclusion, that is, there is a subset $C \subseteq \varphi(G)$ that is a conjugacy class, and if $g \in G$ is such that $\varphi(g) \in C$, then $\mathrm{tr}\, \rho_1(g) \neq \mathrm{tr}\, \rho_2(g)$.

In order to prove this special case, we require a theorem of Carayol [8].

Theorem 5.2.2. *[8, Theorem 1] Let A be a local ring, R an A-algebra, and let $\rho_1, \rho_2 : R \to M_n(A)$ be two representations of R of the same dimension n. Suppose that the residual representation $\bar{\rho} : R \otimes_A F \to M_n(F)$, where F is the residue field of A, is absolutely irreducible. Suppose that the traces for ρ_1 and ρ_2 are the same for every $r \in R$. Then, ρ_1 and ρ_2 are isomorphic as representations, that is, there exists a matrix $Q \in \mathrm{GL}_n(A)$ such that $\rho_1(r) = Q\rho_2(r)Q^{-1}$ for all $r \in R$.*

We now prove the special case.

Proposition 5.2.3. *Let n be a positive integer. Let G be a group and $\rho_1, \rho_2 : G \to \mathrm{GL}_n(\mathbb{Z}_2)$ be a group homomorphism, and suppose the mod 2 representations $\bar{\rho}_1, \bar{\rho}_2$ are isomorphic and absolutely irreducible. Suppose that there exists an element $g \in G$ such that $\operatorname{tr}\rho_1(g) \neq \operatorname{tr}\rho_2(g)$. Then there exists a subset $C \subseteq \varphi(G)$ for which*

1. *the set C is non-empty and closed under conjugation by $\varphi(G)$, and*
2. *if the image in $\varphi(G)$ of an element $g \in G$ belongs to C, then $\operatorname{tr}\rho_1(g) \neq \operatorname{tr}\rho_2(g)$.*

Proof. Our setup begins, as it did, in Section 5.1. Let α be the largest nonnegative integer such that for each $g \in G$, we have

$$\operatorname{tr}(\rho_1(g)) \equiv \operatorname{tr}(\rho_2(g)) \pmod{2^\alpha} \quad \text{and} \quad \operatorname{tr}(\rho_1(g)) \not\equiv \operatorname{tr}(\rho_2(g)) \pmod{2^{\alpha+1}}.$$

In addition, we let β be the largest integer such that ρ_1 and ρ_2 are conjugated modulo λ^β, that is, there is a matrix $P \in \mathrm{GL}_n(\mathcal{O}_\lambda)$ such that $\rho_1 \equiv P\rho_2P^{-1} \pmod{\lambda^\beta}$. As demonstrated before, we have $\beta \leq \alpha$. Also, given that ρ_1 and ρ_2 are conjugate modulo λ^β but not conjugate modulo $\lambda^{\beta+1}$, if we replace ρ_2 with a conjugate $P\rho_2P^{-1}$ for $P \in \mathrm{GL}_n(\mathbb{Z}_2)$, we may assume

$$\rho_1 \equiv P\rho_2P^{-1} \pmod{2^\beta} \text{ and } \rho_1 \not\equiv P\rho_2P^{-1} \pmod{2^{\beta+1}}. \tag{5.16}$$

This implies $P\rho_2(g)P^{-1} - \rho_1(g) \equiv 0 \pmod{2^\beta}$ for any $g \in G$. In particular, we get $P\rho_2(g)P^{-1} - \rho_1(g) = \theta_g 2^\beta$ for some $\theta_g \in M_n(\mathbb{Z}_2)$, which we can write as

$$\theta_g = \frac{P\rho_2(g)P^{-1} - \rho_1(g)}{2^\beta}. \tag{5.17}$$

In particular, note that

$$\operatorname{tr}(\theta_g) = 2^{-\beta}(\operatorname{tr}(P\rho_2(g)P^{-1}) - \operatorname{tr}(\rho_1(g))) = 2^{-\beta}(\operatorname{tr}(\rho_2(g)) - \operatorname{tr}(\rho_1(g))) \tag{5.18}$$

by the invariance of trace under conjugation.

We now show $\alpha = \beta$. Extend the maps ρ_1, ρ_2 to the group ring $\mathbb{Z}/2^\alpha\mathbb{Z}[G]$ by $\rho_i(\sum a_j g_j) = \sum a_j \rho_i(g_j)$, for $i = 1, 2$ and $a_j \in \mathbb{Z}/2^\alpha\mathbb{Z}$ and $g_j \in G$. Then, notice that

$$\begin{aligned} \operatorname{tr}(\rho_1(\textstyle\sum a_j g_j)) \pmod{2^\alpha} &\equiv \operatorname{tr}(\textstyle\sum a_j \rho_1(g_j)) \pmod{2^\alpha} \\ &\equiv \textstyle\sum a_j \operatorname{tr}(\rho_1(g_j)) \pmod{2^\alpha} \\ &\equiv \textstyle\sum a_j \operatorname{tr}(\rho_2(g_j)) \pmod{2^\alpha} \\ &\equiv \operatorname{tr}(\rho_2(\textstyle\sum a_j g_j)) \pmod{2^\alpha}. \end{aligned} \tag{5.19}$$

Since we satisfy the hypotheses of Theorem 5.2.2 with $A = \mathbb{Z}/2^\alpha\mathbb{Z}$ and $R = \mathbb{Z}/2^\alpha\mathbb{Z}[G]$, we can find a matrix $Q \in \mathrm{GL}_n(\mathbb{Z}/2^\alpha\mathbb{Z})$ such that $\rho_1(g) \equiv Q\rho_2(g)Q^{-1} \pmod{2^\alpha}$ for all $g \in G$. However, β is the largest integer such that ρ_1 and ρ_2 are conjugate modulo 2^β, so $\alpha \leq \beta$, implying $\alpha = \beta$.

Recall, from (5.10), the map $\varphi : G \to M_n(\mathbb{F}_2) \rtimes \mathrm{GL}_n(\mathbb{F}_2)$ is defined by

$$\varphi(g) = (\theta(g) \pmod 2, \rho_1(g) \pmod 2))) = ([\theta_g \rho_1(g)^{-1}]_2, [\rho_1(g)]_2). \tag{5.20}$$

We note our use of the notation $[N]_2$, for $N \in M_n(\mathbb{Z}_2)$, to denote the residue class of N taken modulo 2.

Define the map $\phi' : M_n(\mathbb{F}_2) \rtimes \mathrm{GL}_n(\mathbb{F}_2) \to \mathbb{F}_2$ by

$$\phi'((A, B)) = \operatorname{tr}(AB) \tag{5.21}$$

where the product of matrices is taken to be the action of $\mathrm{GL}_n(\mathbb{F}_2)$ on $M_n(\mathbb{F}_2)$, and the trace is considered to be modulo 2. By (5.20), notice

$$\begin{aligned} \phi'(\varphi(g)) &= \operatorname{tr}([\theta_g \rho_1(g)^{-1}]_2 [\rho_1(g)]_2) \\ &= \operatorname{tr}([\theta_g \rho_1(g)^{-1} \rho_1(g)]_2) \\ &= \operatorname{tr}([\theta_g]_2) \\ &= [\operatorname{tr}(\theta_g)]_2 \\ &= [2^{-\alpha}(\operatorname{tr}(\rho_2(g)) - \operatorname{tr}(\rho_1(g)))]_2 \end{aligned} \tag{5.22}$$

where we have used (5.18) above (with β replaced with α, since $\alpha = \beta$) and the fact that, for a matrix $N \in M_n(\mathbb{Z}_2)$ with $\operatorname{tr}(N) = \sum_{i=1}^n a_{ii}$ for entries $a_{ii} \in \mathbb{Z}_2$ along the diagonal, we have

$$\begin{aligned} \operatorname{tr}([N]_2) &= \sum_{i=1}^n [a_{ii}]_2 \\ &= \left[\sum_{i=1}^n a_{ii}\right]_2 \\ &= [\operatorname{tr}(N)]_2 \end{aligned} \tag{5.23}$$

which shows the final equality (noting our use of $[x]_2$ in (5.23) to denote the residue class of a 2-adic integer $x \in \mathbb{Z}_2$).

Let C be the set of elements in $\varphi(G)$ that take a nonzero value under the map ϕ'. From the definition of α and the linearity of the trace map, there exists $g_0 \in G$ such that

$$\operatorname{tr}(\rho_1(g_0)) \not\equiv \operatorname{tr}(\rho_2(g_0)) \pmod{2^{\alpha+1}}.$$

Note that the image of g_0 in $\varphi(G)$ is inside C, so C is nonempty. In addition, let $\phi(h) \in \phi(G)$ for some $h \in G$; then, given φ is a homomorphism (Proposition 5.1.6) and by (5.22) and the invariance under conjugation of the trace map,

$$\begin{aligned}
\phi'(\varphi(h)\varphi(g)\varphi(h)^{-1}) &= \phi'(\varphi(hgh^{-1})) \\
&= [2^{-\alpha}(\operatorname{tr}(\rho_2(hgh^{-1})) - \operatorname{tr}(\rho_1(hgh^{-1})))]_2 \\
&= [2^{-\alpha}(\operatorname{tr}(\rho_2(h)\rho_2(g)\rho_2(h)^{-1}) - \operatorname{tr}(\rho_1(h)\rho_1(g)\rho_1(h)^{-1}))]_2 \\
&= [2^{-\alpha}(\operatorname{tr}(\rho_2(g)) - \operatorname{tr}(\rho_1(g)))]_2 \\
&= \phi'(\varphi(g)) \\
&\neq 0,
\end{aligned} \tag{5.24}$$

so C is closed under conjugation. Finally, suppose that $g \in G$ is such that the image of g in $\varphi(G)$ is contained in C. Then, $\phi'(\varphi(g)) \neq 0$, in particular $\operatorname{tr}\rho_1(g) \neq \operatorname{tr}\rho_2(g)$. □

5.3 Improvement

We now give a proof for Proposition 5.0.2. Then, using the improvements brought about by the deviation group and Proposition 5.1.6, we give an improvement of the result of Mayle-Wang.

For a prime ℓ and an elliptic curve E, we define

$$\mathbb{Q}(E[\ell^\infty]) = \bigcup_{k=1}^{\infty} \mathbb{Q}(E[\ell^k]).$$

We begin with the proof of Proposition 5.0.2. We note that the work which follows is the same as that of Mayle-Wang [18], except for our use of Proposition 2.4.19 and Table 2.2.

Proof of Proposition 5.0.2. Let $\mathcal{O}_\lambda = \mathbb{Z}_2$. Let $A = E \times E'$ and apply Proposition 5.2.1 with $\ell = 2$, $n = 2$, $G = \operatorname{Gal}(\mathbb{Q}(A[2^\infty]/\mathbb{Q}))$, and the 2-adic representations $\rho_1 = \rho_{E,2}$ and $\rho_2 = \rho_{E',2}$. By Theorem 2.3.7, since the two curves E and E' are not $\mathbb{Q}$-isogenous, ρ_1 and ρ_2 are not isomorphic; therefore, by Proposition 2.3.8, there is some prime p such that $a_p(E) \neq a_p(E')$.

By Proposition 5.2.1, there exists a conjugacy class $C \subseteq \delta(G)$ obeying the conclusion. Let K be a subfield of $\mathbb{Q}(A[2^\infty])$ for which $\mathrm{Gal}(K/\mathbb{Q}) = \delta(G)$, which exists by the fundamental theorem of infinite Galois theory.

Let $m = \mathrm{rad}(N_{P(E)}N_{P(E')})$. Apply Proposition 2.4.19 to $\tilde{K} = K(\sqrt{m})$ and C as appearing in Proposition 5.2.1 to obtain a prime p not dividing m such that $\left(\frac{K/\mathbb{Q}}{p}\right) \subseteq C$ which satisfies

$$p \leq \max\{p_0, (\bar{a}\log|d_{\tilde{K}}| + \bar{b}[\tilde{K}:\mathbb{Q}] + \bar{c})^2\} = \max\{p_0, (\bar{a}\log|d_{\tilde{K}}| + 2|\delta(G)|\bar{b} + \bar{c})^2\} \quad (5.25)$$

for a triple $(p_0, \bar{a}, \bar{b}, \bar{c})$ appearing in Table 2.2. As $\mathrm{Frob}_p\,|_K = \left(\frac{K/\mathbb{Q}}{p}\right)$, it follows from Proposition 5.2.1 that

$$\mathrm{tr}\,\rho_{E,2}(\mathrm{Frob}_p) \neq \mathrm{tr}\,\rho_{E',2}(\mathrm{Frob}_p),$$

and consequently $a_p(E) \neq a_p(E')$.

Let us consider more closely the Chebotarev bound of the form $(\bar{a}\log|d_{\tilde{K}}|+2|\delta(G)|\bar{b}+\bar{c})^2$. By Theorem 3.0.1, A has good reduction at some prime ℓ if and only if both E and E' have good reduction at ℓ. Thus, $K/\mathbb{Q}$ is unramified outside of the prime divisors of $m = 2\mathfrak{N}$. As $\tilde{K}$ is the compositum of K and $\mathbb{Q}(\sqrt{m})$, the primes that ramify in $\tilde{K}$ are precisely those that ramify in K or in $\mathbb{Q}(\sqrt{m})$. Thus, since $\mathrm{rad}(d_{\mathbb{Q}(\sqrt{m})}) = \mathrm{rad}(2m) = \mathrm{rad}(2\mathfrak{N})$, and $\mathrm{rad}(d_K) \mid \mathrm{rad}(2\mathfrak{N})$,

$$\mathrm{rad}(d_{\tilde{K}}) = \mathrm{rad}(d_K d_{\mathbb{Q}(\sqrt{m})}) = \mathrm{rad}(2\mathfrak{N}). \quad (5.26)$$

Now, applying Lemma 4.1.6 to (5.25) gives us

$$p \leq (\bar{a}((2|\delta(G)|-1)\log\mathrm{rad}(2\mathfrak{N}) + 2|\delta(G)|\log(2|\delta(G)|)) + 2|\delta(G)|\bar{b} + \bar{c})^2$$

which matches (5.2).

To prove the final statement, note that if the mod 2 representations of E and E', $\bar{\rho}_{E,2}$ and $\bar{\rho}_{E',2}$, are isomorphic and absolutely irreducible, then we instead apply Proposition 5.2.3 over Proposition 5.2.1, in which case $\delta(G)$ is replaced with $\varphi(G)$; in particular, $|\delta(G)|$ can be replaced with $|\varphi(G)|$ in (5.2). □

Now we give a proof of Theorem 5.0.3.

Proof of Theorem 5.0.3. We split our analysis into two cases. First, assume that the mod 2 representations $\rho_1 = \bar{\rho}_{E,2}$ and $\rho_2 = \bar{\rho}_{E',2}$ are isomorphic and irreducible. Apply Proposition 5.0.2, and replace $\delta(G)$ with $\varphi(G)$ (since the mod 2 representations are isomorphic and absolutely irreducible) to get a prime p such that $a_p(E) \neq a_p(E')$ and satisfying

$$p \leq (\bar{a}((2|\varphi(G)|-1)\log\mathrm{rad}(2\mathfrak{N}) + 2|\varphi(G)|\log(2|\varphi(G)|)) + 2|\varphi(G)|\bar{b} + \bar{c})^2.$$

From (5.11) and Lemma 5.1.8, we have for any $g \in G$,

$$\begin{aligned} \det(\rho_2(g)) &= \det((I_2 + 2^\beta \theta(g))\rho_1(g)) \\ &= (1 + 2^\beta \operatorname{tr}(\theta(g)) + O(2^{2\beta})) \det(\rho_1(g)). \end{aligned} \tag{5.27}$$

Note that by (2.10), we have $\det \rho_1 = \det \rho_2$, so the above can be rewritten as $0 = 2^\beta \operatorname{tr}(\theta(g)) + O(2^{2\beta})$, which, after multiplying through by $2^{-\beta}$ implies

$$\operatorname{tr}(\theta) \equiv 0 \pmod 2.$$

In particular, the map φ from Proposition 5.1.6 takes values in $M_2^0(\mathbb{F}_2) \rtimes \mathrm{GL}_2(\mathbb{F}_2)$, where $M_2^0(\mathbb{F}_2)$ denotes the matrices with trace 0 with entries in $\mathbb{F}_2$. Therefore, we have $|\varphi(G)| \leq |M_2^0(\mathbb{F}_2) \rtimes \mathrm{GL}_2(\mathbb{F}_2)| = 8 \cdot 6 = 48$. Choosing $\bar{a} = 1.257$ and $\bar{c} = 7.3$ from Table 2.2, we find

$$\begin{aligned} p &\leq (1.257(95 \log \operatorname{rad}(2\mathfrak{N}) + 96 \log(96)) + 96 \cdot 0 + 7.3)^2 \\ &\leq (119.415 \log \operatorname{rad}(2\mathfrak{N}) + 550.79 + 7.3)^2 \\ &\leq (120 \log \operatorname{rad}(2\mathfrak{N}) + 559)^2. \end{aligned} \tag{5.28}$$

We consider the above bound with the maximum of 3100065 so that the inequality remains true regardless of the value of the discriminant.

If the mod 2 representations are not isomorphic, then mod 2 already distinguishes the traces. Apply Lemma 4.1.4 to the field $L = \mathbb{Q}(E[2], E'[2])$, the conjugacy class C_2 given in (4.14), and the set $S = P(L) \cup \{2\}$, so that we get a prime p unramified in L such that $a_p(E) \not\equiv a_p(E') \pmod 2$ (implying $a_p(E) \neq a_p(E')$) and satisfying

$$p \leq (2\bar{a} \log |d_L| + \bar{c})^2.$$

Applying Lemma 4.1.6 to the above yields

$$p \leq (2\bar{a} \left((n_L - 1) \log \operatorname{rad}(d_L) + n_L \log(n_L)\right) + \bar{c}).$$

Taking $[L : \mathbb{Q}] \leq |\mathrm{GL}_2(\mathbb{F}_2)|^2 = 6^2 = 36$ and the 4-tuple $(3072167, 1.212, 0, 20.56)$ (again absorbing $\bar{b}$ into $\bar{c}$) gives us

$$\begin{aligned} p &\leq (2 \cdot 1.212 \left((35) \log \operatorname{rad}(d_L) + 36 \log(36)\right) + 20.56)^2 \\ &\leq (84.84 \log \operatorname{rad}(2\mathfrak{N}) + 333.28)^2 \\ &\leq (85 \log \operatorname{rad}(2\mathfrak{N}) + 334)^2. \end{aligned} \tag{5.29}$$

Considering the maximum of the above with 3072167 completes this case.

To complete the result, we consider the maximum of all possible cases together (that is, the maximum of (5.28) and (5.29)) which is exactly (5.3). □

Remark 5.3.1. We note that the choice to absorb $\bar{b}$ into $\bar{c}$ is deliberate, as per Remark 4.1.5. For example, when looking at the final case in the proof of Theorem 5.0.3, specifically (5.29), if we had instead taken the 4-tuple $(3072167, 1.212, 0.240, 8.80)$ and absorbed $\bar{b} = 0.240$ into $\bar{a} = 1.212$, we get a 4-tuple $(3072167, 1.649, 0, 8.80)$, which gives us

$$\begin{aligned} p &\leq \left(2 \cdot 1.649 \left((35) \log \operatorname{rad}(d_L) + 36 \log(36)\right) + 8.80\right)^2 \\ &\leq \left(115.43 \log \operatorname{rad}(2\mathfrak{N}) + 434.26\right)^2 \\ &\leq \left(116 \log \operatorname{rad}(2\mathfrak{N}) + 435\right)^2. \end{aligned}$$

which we see, in comparison to (5.29), is worse. This provides further justification, as in Remark 2.4.21, for why we wish $\bar{a}$ to be smaller. For a bound of the form $(A \log(d) + B)^2$, after applying Lemma 4.1.6, we see larger $\bar{a}$ increases both A and B, whereas larger $\bar{c}$ increases B only.

Remark 5.3.2. Let us quickly compare the result of Mayle-Wang in (5.1) and our result in (5.3). In the best case, the two elliptic curves would have bad reduction at only 2, so $\operatorname{rad}(2\mathfrak{N}) = 2$. In this case, (5.1) gives

$$p \leq (482 \log(2) + 2880)^2 \approx 10330419.15.$$

In the case of (5.3), we get

$$p \leq \max\{3100065, (120 \log(2) + 599)^2\} = 3100065,$$

which beats (5.1). In fact, $(120 \log(2\mathfrak{N}) + 599)^2 \leq 3100065$ if and only if

$$\operatorname{rad}(2\mathfrak{N}) \leq \exp\left(\frac{\sqrt{3100065} - 559}{120}\right) \approx 22340.755.$$

Given our result is constant for $\operatorname{rad}(2\mathfrak{N}) \leq 22340$, whereas (5.1) is logarithmic in the size of $2\mathfrak{N}$, we beat Mayle-Wang for small $\operatorname{rad}(2\mathfrak{N})$. When the inequality flips, and $3100065 \leq (120 \log(2) + 599)^2$, then we notice $(120 \log \operatorname{rad}(2\mathfrak{N}) + 599)^2 \leq (482 \log \operatorname{rad}(2\mathfrak{N}) + 2880)^2$ by inspection. Thus, we see (5.3) beats (5.1) in all cases.

Remark 5.3.3. More precise bounds than that which appears in (5.3) are possible. If one were to work with explicit elliptic curves, one could imagine computing the discriminant d_K of the field $K = \mathbb{Q}(E[\ell], E'[\ell])$ for some specially chosen prime ℓ using, say, Kraus and the results discussed in Chapter 3. Then, when $\log |d_K|$ is known, one can take a triple (a, b, c) from Table 2.1 rather than Table 2.2, which would eliminate the need to rely on a coarser estimate for p_0. For elliptic curves whose torsion field has a small discriminant, this can further improve the constants appearing in the final result of Theorem 5.0.3, as well as removing the maximum between p_0 and the bound $(\bar{a} \log \operatorname{rad}(2\mathfrak{N}) + \bar{c})^2$. Additionally, as

we see in (5.29), if we know the mod 2 representations are not isomorphic, we can further reduce the size of the constants appearing in (5.3). In this thesis, we have chosen to give an improvement to Mayle-Wang using a new method that works well for large values of $\mathrm{rad}(d_K)$.

Bibliography

[1] Tom M. Apostol. *Introduction to Analytic Number Theory.* Springer, 1976.

[2] Eric Bach and Jonathan Sorenson. Explicit bounds for primes in residue classes. *Mathematics of Computation*, 65(216):1717–1735, 1996.

[3] Kevin Broughan. *Equivalents of the Riemann Hypothesis*, volume 1 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 2017.

[4] Kevin Broughan. *Equivalents of the Riemann Hypothesis*, volume 2 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, 2017.

[5] Alina Bucur and Kiran S. Kedlaya. An application of the effective sato-tate conjecture. *Contemporary Mathematics*, 663:45–56, 2016.

[6] D. Bump, J. W. Cogdell, E. de Shalit, D. Gaitsgory, E. Kowalski, and S. S. Kudla. *An introduction to the Langlands program.* Birkhäuser Boston, Inc., Boston, MA, 2003. Lectures presented at the Hebrew University of Jerusalem, Jerusalem, March 12–16, 2001.

[7] Élie Cali and Alain Kraus. Sur la p-différente du corps des points de ℓ-torsion des courbes elliptiques, $\ell \neq p$. *Acta Arithmetica*, 104(1):1–21, 2002.

[8] Henri Carayol. Formes modulaires et représentations galoisiennes à valeurs dans un anneau local complet. *Contemprary Mathematics*, 165:213–237, 1994.

[9] Gabriel Chênevert. *Exponential sums, hypersurfaces with many symmetries and Galois representations.* PhD thesis, McGill University, 2008.

[10] Gary Cornell and Joseph H. Silverman. *Arithmetic Geometry.* Springer-Verlag, 1986.

[11] Henri Darmon, Fred Diamond, and Richard Taylor. Fermat's last theorem. *Current Developments in Mathematics*, 1:1–157, 1995.

[12] Gerd Faltings. Endlichkeitssätze für abelsche varietäten über zahlkörpern. *Inventiones Mathematicae*, 73:349–366, 1983.

[13] Aleksander Ivić. *The Riemann Zeta Function: Theory and Applications.* Dover, 2003.

[14] Gordon James and Martin Liebeck. *Representations and Characters of Groups.* Cambridge University Press, 2001.

[15] Alain Kraus. Sur la p-différente du corps des points de p-torsion des courbes elliptiques. *Bulletin of the Australian Mathematical Society*, 60(3):407–428, 1999.

[16] J. C. Lagarias and A. M. Odlyzko. Effective versions of the Chebotarev density theorem. *Algebraic Number Fields*, 54:271–296, 1979.

[17] Robert P. Langlands. *Problems in the theory of automorphic forms.* Springer, 1970.

[18] Jacob Mayle and Tian Wang. On the effective version of Serre's open image theorem. *Bulletin of the London Mathematical Society*, page to appear, 2024.

[19] James S. Milne. Algebraic number theory (v3.08), 2020. Available at www.jmilne.org/math/.

[20] J.S. Milne. Class field theory (v4.03), 2020. Available at www.jmilne.org/math/.

[21] Władysław Narkiewicz. *Elementary and Analytic Theory of Algebraic Numbers.* Springer Berlin, Heidelberg, 2004.

[22] Jürgen Neukirch. *Algebraic Number Theory.* Springer, 1999.

[23] D. J. Newman. Simple analytic proof of the prime number theorem. *The American Mathematical Monthly*, 87(9):693–696, 1980.

[24] J. Oesterlé. Versions effectives du théorème de Chebotarev sous l'hypothèse de Riemann généralisée. *Astérisque*, 61:165–167, 1979.

[25] Paulo Ribenboim. *Classical Theory of Algebraic Numbers.* Springer, 2001.

[26] Ignasi Sánchez Rodríguez. Comparing galois representations and the Falting-Serre-Livné method. Master's thesis, Universitat de Barcelona, 2020.

[27] Jean-Pierre Serre. *Abelian ℓ-Adic Representations and Elliptic Curves.* W. A. Benjamin, Inc., 1968.

[28] Jean-Pierre Serre. Propriétés galoisiennes des points d'ordre fini des courbes elliptiques. *Invent. Math.*, 15(4):259–331, 1972.

[29] Jean-Pierre Serre. *Linear representations of finite groups*, volume Vol. 42 of *Graduate Texts in Mathematics.* Springer-Verlag, New York-Heidelberg, french edition, 1977.

[30] Jean-Pierre Serre. Quelques applications du théorème de densité de Chebotarev. *Publications Mathématiques de l'IHÉS*, 54:123–201, 1981.

[31] Jean-Pierre Serre. Résumé des cours de 1984-1985. In *Annuaire du Collège de France*, pages 85–90. 1985.

[32] Jean-Pierre Serre and John Tate. Good reduction of abelian varieties. *Annals of Mathematics*, 88(3):492–517, 1968.

[33] Joseph H. Silverman. *The Arithmetic of Elliptic Curves.* Springer, second edition, 2009.

[34] John Tate. Endomorphisms of abelian varieties over finite fields. *Inventiones Mathematicae*, 2:134–144, 1966.

[35] Richard Taylor and Andrew Wiles. Ring-theoretic properties of certain hecke algebras. *Annals of Mathematics*, 141(3):553–572, 1995.

[36] Andrew Wiles. Modular elliptic curves and Fermat's last theorem. *Annals of Mathematics*, 141(3):443–551, 1995.

Appendix A

Code

The following is supplemental code which was used to prove Corollary 2.4.19. We pick a pivot tuple (a, b, c) from Table 2.1, then accomplish two tasks: first we compute the number p_0 (which is computed by examining the tuples which appear before our pivot) and then we verify our pivot is larger than all other triples which appear afterwards for that range of $n_K = [K : \mathbb{Q}]$. By larger, we shall mean if we have a second tuple (a', b', c'), we say the bound appearing from the tuple (a, b, c) is larger than the bound appearing from the tuple (a', b', c') if

$$(a' \log d_K + b' n_K + c')^2 \leq (a \log d_K + b n_K + c)^2 \tag{A.1}$$

for all appropriate values of $\log d_K$ and n_K. The language used here is `pari/gp`.

```
1  /* arrays which correspond to columns in Table 2.1. Each entry is given in the
       form [a,b,c,dLower,dUpper], where the range for a,b,c is valid for
      discriminants between dLower and dUpper. Each case also is valid only for
      a certain range of n; for example, case1 holds only for n=2, case2 for n
      between 3 and 4, etc. A value of dUpper = -1 is used to denote unbounded
      disciminant, or "infinity" */
2  case1 = [[3.29,1.48,4.9,1,5], [2.662,0.75,4.8,5,10], [2.301,0.52,5,10,25],
      [1.881,0.34,5.5,25,100], [1.446,0.23,6.8,100,1000],
      [1.125,0.63,10.9,1000,10000], [1.032,0.44,20.2,10000,100000],
      [1.008,-0.06,47.7,100000,-1]];
3  case2 = [[2.808,0.58,4.7,5,10], [2.524,0.45,4.9,10,25],
      [2.035,0.27,5.3,25,100], [1.527,0.17,6.4,100,1000],
      [1.148,0.5,10.2,1000,10000], [1.038,0.5,18.7,10000,100000],
      [1.01,-0.03,41.9,100000,-1]];
4  case3 = [[2.736,0.35,4.7,10,25], [2.231,0.21,5.1,25,100],
      [1.629,0.11,6.1,100,1000], [1.178,0.37,9.5,1000,10000],
      [1.046,0.56,17.3,10000,100000], [1.012,0,37.8,100000,-1]];
5  case4 = [[2.303,0.19,4.8,10,25], [2.297,0.19,5,25,100],
      [1.667,0.09,6,100,1000], [1.189,0.32,9.2,1000,10000],
      [1.049,0.59,16.8,10000,100000], [1.012,0,37.8,100000,-1]];
6  case5 = [[2.228,0.1,4.9,25,100], [1.745,0.04,5.8,100,1000],
      [1.212,0.24,8.8,1000,10000], [1.054,0.63,16,10000,100000],
      [1.014,0.02,35.9,100000,-1]];
7  case6 = [[1.755,0,5.7,100,1000], [1.257,0,7.3,1000,10000],
      [1.095,0,8.2,10000,100000], [1.017,0.07,31.8,100000,-1]];
8
9  /*
```

```
10 Function which condenses a given column of Bach-Sorenson into a single entry
       corresponding to the size of n.
11
12 nStart and nEnd --- determine the range for n_K, the degree of our field
       extension. We use this to determine which column we are in, so there are
       only 6 possible values nStart and nEnd should take:
13   nStart = 2, nEnd = 2 (corresponds to the first column)
14   nStart = 3, nEnd = 4 (corresponds to the second column)
15   nStart = 5, nEnd = 9 (corresponds to the third column)
16   nStart = 10, nEnd = 14 (corresponds to the fourth column)
17   nStart = 15, nEnd = 49 (corresponds to the fifth column)
18   nStart = 50, nEnd = 73 (corresponds to a smaller range of the sixth column)
19
20 pivotNum --- determines which triple we take to be our pivot. We note 1 <
       pivotNum < length(case) (where case is determined by the range of n above)
21 */
22 verifyCase(nStart, nEnd, pivotNum) = {
23   my(case, p_0=-1,a,b,c,isValid = true);
24
25   /* check range for n, and assign case the appropriate array above (
         corresponds to the proper column of triples (a,b,c) for this range of n)
          */
26   if(nStart == 2 && nEnd == 2, case = case1);
27   if(nStart == 3 && nEnd == 4, case = case2);
28   if(nStart == 5 && nEnd == 9, case = case3);
29   if(nStart == 10 && nEnd == 14, case = case4);
30   if(nStart == 15 && nEnd == 49, case = case5);
31   if(nStart == 50 && nEnd == 73, case = case6);
32
33   for(n = nStart, nEnd,
34     printf("Case n = %d \n", n);
35
36     /* initialize p_0 using the generic constants a = 4, b = 2.5, c = 5. This
           takes care of any possible gap in Table 2.1 (see Remark 13 in Mayle-
           Wang) */
37     p_0 = (4*case[1][4] + n*2.5 + 5)^2;
38
39     /* loop that computes p_0. We simply take the equation (a*log(d_K) + b*n +
           c)^2 and replace log(d_K) with its lower bound and upper bound, then
           take the maximum of all possible cases */
40     for(i = 1, pivotNum-1,
41       p_0 = max(p_0, max((case[i][1]*case[i][4] + case[i][2]*n + case[i][3])
             ^2, (case[i][1]*case[i][5] + case[i][2]*n + case[i][3])^2));
42       printf("%0.3f <= log d\_K <= %0.3f: %0.6f <= p_0 <= %0.6f \n", case[i
           ][4], case[i][5], (case[i][1]*case[i][4] + case[i][2]*n + case[i
           ][3])^2, (case[i][1]*case[i][5] + case[i][2]*n + case[i][3])^2);
43     );
44
45     printf("---------- \n verify inequality after log d\_K > %d: \n", case[
         pivotNum][4]);
46     /* get a,b,c from the pivot tuple */
47     a = case[pivotNum][1];
48     b = case[pivotNum][2];
49     c = case[pivotNum][3];
50
51     /* verify that the tuples appearing after the pivot are smaller than the
         pivot */
52     for(i = pivotNum+1, length(case),
```

```
53        aTilde = case[i][1];
54        bTilde = case[i][2];
55        cTilde = case[i][3];
56        r = ((cTilde+bTilde*n)-(c+b*n))/(a-aTilde); /* r is the value such that
              log(d_K) >= r for the given inequality to hold */
57
58        /* check if r is too large for the given range; if it is, then the
              result is not valid */
59        if(case[i][5] != -1 && r > case[i][4] && r > case[i][5], isValid = False
              );
60
61        printf("%d < log d\_K <= %d: For the bound from tuple (%0.4g, %0.4g,
              %0.4g) to be larger than the bound from tuple (%0.4g, %0.4g, %0.4g),
               we require log d\_K >= %0.6f \n", case[i][4], case[i][5], a, b, c,
              aTilde, bTilde, cTilde, r);
62      );
63      printf("---------- \n\n");
64    );
65
66    printf("RESULT \n");
67    if(!isValid, printf("invalid result \n"));
68    printf("If log d\_K <= %d, then p <= %0.6f \n", case[pivotNum][4], p_0);
69    printf("If log d\_K > %d, then p <= (%0.6f \* log d\_K + %0.6f n + %0.6f)^2
          \n", case[pivotNum][4], a,b,c);
70  }
```

The calls to the code that produce Table 2.2 are as follows:

```
1       verifyCase(2, 2, 6)
2       verifyCase(3, 4, 5)
3       verifyCase(5, 9, 4)
4       verifyCase(10, 14, 4)
5       verifyCase(15, 49, 3)
6       verifyCase(50, 74, 2)
```

An example of the output of the code (for the second call, to demonstrate what happens when the range for n does not contain only a single value) looks like:

```
1  Case n = 3
2  5.000 <= log d_K <= 10.000: 419.430400 <= p_0 <= 1191.630400
3  10.000 <= log d_K <= 25.000: 991.620100 <= p_0 <= 4809.422500
4  25.000 <= log d_K <= 100.000: 3247.290225 <= p_0 <= 43936.352100
5  100.000 <= log d_K <= 1000.000: 25475.352100 <= p_0 <= 2352879.888100
6  ----------
7   verify inequality after log d_K > 1000:
8  10000 < log d_K <= 100000: For the bound from tuple (1.148, 0.5000, 10.20) to
       be larger than the bound from tuple (1.038, 0.5000, 18.70), we require log
        d_K >= 77.272727
9  100000 < log d_K <= -1: For the bound from tuple (1.148, 0.5000, 10.20) to be
       larger than the bound from tuple (1.010, -0.03000, 41.90), we require log
       d_K >= 218.188406
10 ----------
11
12 Case n = 4
13 5.000 <= log d_K <= 10.000: 443.523600 <= p_0 <= 1232.010000
14 10.000 <= log d_K <= 25.000: 1020.163600 <= p_0 <= 4872.040000
15 25.000 <= log d_K <= 100.000: 3278.135025 <= p_0 <= 44049.614400
16 100.000 <= log d_K <= 1000.000: 25529.648400 <= p_0 <= 2353401.446400
```

```
17 ----------
18  verify inequality after log d_K > 1000:
19 10000 < log d_K <= 100000: For the bound from tuple (1.148, 0.5000, 10.20) to
       be larger than the bound from tuple (1.038, 0.5000, 18.70), we require log
        d_K >= 77.272727
20 100000 < log d_K <= -1: For the bound from tuple (1.148, 0.5000, 10.20) to be
       larger than the bound from tuple (1.010, -0.03000, 41.90), we require log
       d_K >= 214.347826
21 ----------
22
23 RESULT
24 If log d_K <= 1000, then p <= 2353401.446400
25 If log d_K > 1000, then p <= (1.148000 * log d_K + 0.500000 n + 10.200000)^2
```

Let us examine this output. Lines 1 through 20 print the result for the case when $n = 3$ and $n = 4$. Let us examine closer the output for the case $n = 3$. First, there are four lines describing the bound on p_0 for each given range of $\log d_K$. It is worth noting that these ranges are all the ranges which appear before our chosen pivot. Then, we verify that all tuples (a', b', c') appearing *after* our pivot are in fact smaller than our pivot. For example, here we have chosen our pivot to be the tuple $(1.148, 0.5, 10.2)$. Appearing after line 7 in the output is a check that for the other tuples appearing, $(1.148, 0.5, 10.2)$ provides a larger bound. On line 8, for example, we compare the tuple $(1.148, 0.5, 10.2)$ to the tuple $(1.038, 0.5, 18.7)$, in particular line 8 is verifying the inequality $(1.527 \log d_K + 0.17 \cdot 3 + 6.4)^2 \leq (2.035 \log d_K + 0.27 \cdot 3 + 5.3)^2$, which holds for $\log d_K \geq 77.273$. Though we have a check in the code to determine if there are any problems with the required range for $\log d_K$ being outside the assumed range (see line 56 in the code), we also manually verify this fact for each case (`pari/gp` does not include an `assert` function, and so we must instead use a boolean variable `isValid`). This same analysis can be done for the case $n = 4$. Then, lines 23 to 25 display the result. Line 24 prints the maximum value of p_0, whereas line 25 simply prints our pivot, provided no errors were detected.