Computing elliptic curves over $\mathbb{Q}$ via Thue-Mahler equations and related problems

by

Adela Gherga

B.Sc. Mathematics, McMaster University, 2006

M.Sc. Mathematics, McMaster University, 2010

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Mathematics)

The University of British Columbia

(Vancouver)

July 2019

Abstract

This document provides brief instructions for using the `ubcdiss` class to write a UBC-conformant dissertation in LaTeX. This document is itself written using the `ubcdiss` class and is intended to serve as an example of writing a dissertation in LaTeX. This document has embedded Unique Resource Locators (URLS) and is intended to be viewed using a computer-based Portable Document Format (PDF) reader.

Note: Abstracts should generally try to avoid using acronyms.

Note: at University of British Columbia (UBC), both the Graduate and Postdoctoral Studies (GPS) Ph.D. defence programme and the Library's online submission system restricts abstracts to 350 words.

Lay Summary

The lay or public summary explains the key goals and contributions of the research/ scholarly work in terms that can be understood by the general public. It must not exceed 150 words in length.

Preface

At UBC, a preface may be required. Be sure to check the GPS guidelines as they may have specific content to be included.

Contents

List of Tables

List of Figures

Glossary

This glossary uses the handy `acroynym` package to automatically maintain the glossary. It uses the package's `printonlyused` option to include only those acronyms explicitly referenced in the LaTeX source.

GPS Graduate and Postdoctoral Studies

PDF Portable Document Format

URL Unique Resource Locator, used to describe a means for obtaining some resource on the world wide web

Acknowledgments

Thank those people who helped you.

Don’t forget your parents or loved ones.

You may wish to acknowledge your funding sources.

Chapter 1

Introduction

i mean, the beginning is the part you're not comfortable writing, right? the longer it went on, the better it flowed. at that point you're quoting and weaving results you know well, referencing the little mental web you have woven. it seems cohesive, but also i don't understand it. the beginning bit seems thrown together like Mike told you to include bits about DEs and so you begrudgingly injected something ?? like the very very beginning bit anyway, i'll e-mail you back the tex file and the pdf. i know you're not asking for this advice but it's coming from ozgur and yaniv and they are very smart and i trust them lots:

1. be very careful about whether you're using colloquial language, and how it might be interpreted. e.g. be careful not to insult people's work, and try to not to flip flop on how hand wavey you are being. I think I have a couple of notes in the file pertaining to each of these points
2. when citing work, either use the author names every time or don't. don't mix and match unless appropriate. why would you deny some the respect of appearing in your work, but not others?
3. if you're going to write notes to yourself in your thesis/papers, you must have a way of ensuring that you'll see them later before you send it off. caps lock is not sufficient and yaniv and ozgur can provide examples if you need. I included a little command for you so that you can just Cmd+F (or C-s ??) for all appearances of in the .tex file if you use it. Has the added advantage of making PDF text blue so that everyone reading too knows that it doesn't belong.

also my disclaimer for edits: 1. for some reason my brain is tired today; 2. I don't know the culture of your field nor some of the very elementary things you're presenting 3. Because of 1, I tried to communicate what I wanted to say using the best language I could, but may not have always succeeded at clarity/intent/approachability ?? So basically, remember that it's possible that my edits deserve to be treated with a grain of salt. ??

This start feels outside the realm of where you're going — it seems at once abrupt and off-topic. It would be nice to have an introductory sentence or two to get the reader on track before discussing the "required background" material. A Diophantine equation is a polynomial equation in several variables defined over the integers. The term *Diophantine* refers to the Greek mathematician Diophantus of Alexandria, who studied such equations in the 3rd century A.D. why the history lesson? maybe you could use this as one way of motivating/introducing DEs: "look at these things. look how long they've been studied. here's why, and here are the ways people study them. . . . or something. . .

remove separate paragraph if it's the same thought — f is a DE right? If so, then these next lines are providing additional information to what was given above, not starting a new thread. Let $f(x_1, \ldots, x_n)$ be a polynomial with integer coefficients. We wish to study the set of solutions $(x_1, \ldots, x_n) \in \mathbb{Z}^n$ to the equation

$$f(x_1, \ldots, x_n) = 0. \tag{1.1}$$

There are several different approaches for doing so, arising from three basic problems concerning Diophantine equations. The first such problem is to determine whether ~~or not~~ (1.1) has any solutions ~~at all~~ (too colloquial imo). Indeed, one of the most famous theorems in mathematics, Fermat's Last Theorem, proven by Wiles in 1995, states that for $f(x, y, z) = x^n + y^n - z^n$, where $n \geq 3$, there are no solutions in the positive integers x, y, z (there are so many commas in this sentence. You can remove at least 2 of 7 by splicing and/or rearranging). Qualitative questions of this type are often studied using algebraic methods.

Suppose now that (1.1) is solvable, that is, has at least one solution. The second basic problem is to determine whether the number of solutions is finite or infinite.

For example, consider the *Thue equation*,

$$f(x, y) = a, \tag{1.2}$$

where $f(x, y)$ is an integral binary form of degree $n \geq 3$ (feels like you really jump into the language here. you spelled out what a DE was, but now assume the reader knows the definition of an integral binary form. Personally, I knew the former but the latter reads like domain-specific jargon to me) and a is a fixed nonzero rational integer. In 1909, Thue [REF] proved that this equation has only finitely many solutions. This result followed from a sharpening of Liouville's inequality, an observation that algebraic numbers do not admit very strong approximation by rational numbers. That is, if α is a real algebraic number of degree $n \geq 2$ and p, q are integers, Liouville's ([REF]) observation states that

$$\left|\alpha - \frac{p}{q}\right| > \frac{c_1}{q^n}, \tag{1.3}$$

where $c_1 > 0$ is a value depending explicitly on α. The finitude of the number of solutions to (1.2) follows directly from a sharpening of (1.3) of the type

$$\left|\alpha - \frac{p}{q}\right| > \frac{\lambda(q)}{q^n}, \quad \lambda(q) \to \infty. \tag{1.4}$$

what is the limit $\lambda \to \infty$ with respect to? Indeed, if α is a real root of $f(x, 1)$ and $\alpha^{(i)}$, $i = 1, \ldots, n$ are its conjugates, it follows from (1.2) that

$$\prod_{i=1}^{n} \left|\alpha^{(i)} - \frac{x}{y}\right| = \frac{a}{|a_0||y|^n}$$

where a_0 is the leading coefficient of the polynomial $f(x, 1)$. If the Thue equation has integer solutions with arbitrarily large $|y|$, the product $\prod_{i=1}^{n} |\alpha^{(i)} - x/y|$ must take arbitrarily small values for solutions x, y of (1.2). As all the $\alpha^{(i)}$ are different, x/y must be correspondingly close to one of the real numbers $\alpha^{(i)}$, say α. Thus we obtain

$$\left|\alpha - \frac{x}{y}\right| < \frac{c_2}{|y|^n}$$

where c_2 depends only on a_0, n, and the conjugates $\alpha^{(i)}$. Comparison of this

inequality with (1.4) shows that $|y|$ cannot be arbitrarily large, and so the number of solutions of the Thue equation is finite. Using this argument, an explicit bound can be constructed on the solutions of (1.2) provided that an effective (descriptive? explicit? tight? tractable?) inequality (1.4) is known. The sharpening of the Liouville inequality however, especially in effective form, proved to be very difficult. REF? also "very difficult" seems a subjective qualification; is that okay for your audience?

In [REF:THUE], Thue published a proof that

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\frac{n}{2}+1+\varepsilon}}$$

has only finitely many solutions in integers $p, q > 0$ for all algebraic numbers α of degree $n \geq 3$ and any $\varepsilon > 0$. In essence, he obtained the inequality (1.4) with $\lambda(q) = c_3 q^{\frac{1}{2}n-1-\varepsilon}$ this function does not match the one appearing above in displaymath. is that supposed to be the case? might have something to do with the $<$ not matching the $>$ in (1.4)? it is not clear to me, but hopefully it will be to typical reader, where $c_3 > 0$ depends on α and ε, thereby confirming that all Thue equations have only finitely many solutions. Unfortunately, Thue's arguments do not allow one to find the explicit dependence of c_3 on α and ε, and so the bound for the number of solutions of the Thue equation cannot be given in explicit form either. That is, Thue's proof is ineffective, meaning that it provides no means to ~~actually~~ find the solutions to (1.2). I feel like I would dance more carefully around calling someone's proof ineffective.

Nonetheless, the investigation of Thue's equation and its generalizations was central to the development of the theory of Diophantine equations in the early 20th century when it was discovered that many Diophantine equations in two unknowns could be reduced to it. In particular, the thorough development and enrichment of Thue's method led Siegel to his theorem on the finitude of the number of integral points on an algebraic curve of genus greater than zero [REF?]. However, as Siegel's result relies on Thue's rational approximation to algebraic numbers, it too is ineffective in the above sense.

Shortly following Thue's result, Goormaghtigh conjectured that the only non-trivial

integer solutions of the exponential Diophantine equation

$$\frac{x^m - 1}{x - 1} = \frac{y^n - 1}{y - 1} \tag{1.5}$$

satisfying $x > y > 1$ and $n, m > 2$ are

$$31 = \frac{2^5 - 1}{2 - 1} = \frac{5^3 - 1}{5 - 1} \quad \text{and} \quad 8191 = \frac{2^{13} - 1}{2 - 1} = \frac{90^3 - 1}{90 - 1}.$$

These correspond to the known solutions $(x, y, m, n) = (2, 5, 5, 3)$ and $(2, 90, 13, 3)$ to what is nowadays termed *Goormaghtigh's equation*. The Diophantine equation (1.5) asks for integers having all digits equal to one with respect to two distinct bases, yet whether it has finitely many solutions is still unknown. By fixing the exponents m and n however, Davenport, Lewis, and Schinzel ([REF]) were able to prove that (1.5) has only finitely many solutions. Unfortunately, this result rests on Siegel's aforementioned finiteness theorem, and is therefore ineffective.

In 1933, Mahler [REF] published a paper on the investigation of the Diophantine equation

$$f(x, y) = p_1^{z_1} \cdots p_v^{z_v}, \quad (x, y) = 1,$$

in which $S = \{p_1, \ldots, p_v\}$ denotes a fixed set of prime numbers, $x, y, z_i \geq 0$, $i = 1, \ldots, v$ are unknown integers, and $f(x, y)$ is an integral irreducible binary form of degree $n \geq 3$. Generalizing the classical result of Thue, Mahler proved that this equation has only finitely many solutions. Unfortunately, like Thue, Mahler's argument is also ineffective each time I read this, I believe more strongly that a different word should be used to describe their work. ineffective seems like an attack, and a broad stroke that misses the precise critique you're looking to discuss.

This leads us to the third basic problem regarding Diophantine equations and the main focus of this thesis: given a solvable Diophantine equation, determine all of its solutions. Until long after Thue's work, no method was known for the construction of bounds for the number of solutions of a Thue equation in terms of the parameters of the equation. Only in 1968 was such a method introduced by Baker [REF], based on his theory of bounds for linear forms in the logarithms of alge-

braic numbers. Generalizing Baker's ground-breaking result to the p-adic case, Sprindžuk and Vinogradov [CITE] and Coates [CITE] proved that the solutions of any *Thue-Mahler equation*,

$$f(x, y) = ap_1^{z_1} \cdots p_v^{z_v}, \quad (x, y) = 1, \tag{1.6}$$

where a is a fixed integer, could, at least in principal, be effectively determined. The first practical method for solving the general Thue-Mahler equation (1.6) over $\mathbb{Z}$ is attributed to Tzanakis and de Weger [CITE], whose ideas were inspired in part by the method of Agrawal, Coates, Hunt, and van der Poorten [CITE] in their work to solve the specific Thue-Mahler equation

$$x^3 - x^2y + xy^2 + y^3 = \pm 11^{z_1}.$$

Using optimized bounds arising from the theory of linear forms in logarithms, a refined, automated version of this explicit method has since been implemented by Hambrook as a MAGMA package [REF?].

As for Goormaghtigh's equation, when m and n are fixed and

$$\gcd(m - 1, n - 1) > 1, \tag{1.7}$$

Davenport, Lewis, and Schinzel ([REF]) were able to replace Siegel's result by an effective argument due to Runge. This result was improved by Nesterenko and Shorey ([REF]) and Bugeaud and Shorey ([REF]) using Baker's theory of linear forms in logarithms. In either case, in order to deduce effectively computable bounds (I like this use of effectively) upon the polynomial variables x and y, one must impose the constraints upon m and n that either $m = n + 1$, or that the assumption (1.7) holds. In the extensive literature on this problem, there are a number of striking results that go well beyond what we have mentioned here. By way of example, work of Balasubramanian and Shorey ([REF]) shows that equation (1.5) has at most finitely many solutions if we fix only the set of prime divisors of x and y, while Bugeaud and Shorey ([REF]) prove an analogous finiteness result, under the additional assumption of (1.7), provided the quotient $(m - 1)/(n - 1)$ is

bounded above. Additional results on special cases of equation (1.5) are available in, for example, [**?**], [**?**], [**?**] and [**?**]. An excellent overview of results on this problem can be found in the survey of Shorey [**?**].

1.1 Statement of the results

The novel contributions of this thesis concern the development and implementation of efficient algorithms to determine all solutions of certain Goormaghtigh equations and Thue-Mahler equations. In particular, we follow [REF: BeGhKr] to prove that, in fact, under assumption (1.7), equation (1.5) has at most finitely many solutions which may be found effectively, even if we fix only a single exponent.

Theorem 1.1.1 (BeGhKr)**.** *If there is a solution in integers x, y, n and m to equation* (1.5)*, satisfying* (1.7)*, then*

$$x < (3d)^{4n/d} \leq 36^n. \tag{1.8}$$

In particular, if n is fixed, there is an effectively computable constant $c = c(n)$ such that $\max\{x, y, m\} < c$.

We note that the latter conclusion here follows immediately from (1.8), in conjunction with, for example, work of Baker ([REF]). The constants present in our upper bound (1.8) may be sharpened somewhat at the cost of increasing the complexity of our argument. By refining our approach, in conjunction with some new results from computational Diophantine approximation, we are able to achieve the complete solution of equation (1.5), subject to condition (1.7), for small fixed values of n.

Theorem 1.1.2 (BeGhKr)**.** *If there is a solution in integers x, y and m to equation* (1.5)*, with $n \in \{3, 4, 5\}$ and satisfying* (1.7)*, then*

$$(x, y, m, n) = (2, 5, 5, 3) \ \textit{and} \ (2, 90, 13, 3).$$

In the case $n = 5$ of Theorem (1.1.2) "off-the-shelf" techniques for finding integral points on models of elliptic curves or for solving *Ramanujan-Nagell* equations of the shape $F(x) = z^n$ (where F is a polynomial and z a fixed integer) do not apparently permit the full resolution of this problem in a reasonable amount of time. Instead, we sharpen the existing techniques of [TdW] and [Hambrook] for solving Thue-Mahler equations and specialize them to this problem.

A direct consequence and primary motivation for developing an efficient Thue-Mahler algorithm is the computation of elliptic curves over $\mathbb{Q}$. Let S be a finite set of rational primes. In 1963, Shafarevich [CITE] proved that there are at most finitely many $\mathbb{Q}$-isomorphism classes of elliptic curves defined over $\mathbb{Q}$ having good reduction outside S. The first effective proof of this statement was provided by Coates [CITE] in 1970 for the case $K = \mathbb{Q}$ and $S = \{2, 3\}$ using bounds for linear forms in p-adic and complex logarithms. Early attempts to make these results explicit for fixed sets of small primes overlap with the arguments of [COATES], in that they reduce the problem to that of solving a number of degree 3 Thue-Mahler equations of the form

$$F(x, y) = au,$$

where u is an integer whose prime factors all lie in S.

In the 1950's and 1960's, Taniyama and Weil asked whether all elliptic curves over $\mathbb{Q}$ of a given conductor N are related to modular functions. While this conjecture is now known as the Modularity Theorem, until its proof in 2001 [**?**], attempts to verify it sparked a large effort to tabulate all elliptic curves over $\mathbb{Q}$ of given conductor N. In 1966, Ogg ([**?**], [**?**]) determined all elliptic curves defined over $\mathbb{Q}$ with conductor of the form 2^a. Coghlan, in his dissertation [**?**], studied the curves of conductor $2^a 3^b$ independently of Ogg, while Setzer [**?**] computed all $\mathbb{Q}$-isomorphism classes of elliptic curves of conductor p for certain small primes p. Each of these examples corresponds, via the [BR] approach, to cases with reducible forms. The first analysis on irreducible forms in (**??**) was carried out by Agrawal, Coates, Hunt and van der Poorten [**?**], who determined all elliptic curves of conductor 11 defined over $\mathbb{Q}$ to verify the (then) conjecture of Taniyama-Weil.

There are very few, if any, subsequent attempts in the literature to find elliptic

curves of given conductor via Thue-Mahler equations. Instead, many of the approaches involve a completely different method to the problem, using modular forms. This method relies upon the Modularity Theorem of Breuil, Conrad, Diamond and Taylor [**?**], which was still a conjecture (under various guises) when these ideas were first implemented. Much of the success of this approach can be attributed to Cremona (see e.g. [**?**], [**?**]) and his collaborators, who have devoted decades of work to it. In fact, using this method, all elliptic curves over $\mathbb{Q}$ of conductor N have been determined for values of N as follows

- Antwerp IV (1972): $N \leq 200$
- Tingley (1975): $N \leq 320$
- Cremona (1988): $N \leq 600$
- Cremona (1990): $N \leq 1000$
- Cremona (1997): $N \leq 5077$
- Cremona (2001): $N \leq 10000$
- Cremona (2005): $N \leq 130000$
- Cremona (2014): $N \leq 350000$
- Cremona (2015): $N \leq 364000$
- Cremona (2016): $N \leq 390000$.

In this thesis, we follow [BeGhRe] wherein we return to techniques based upon solving Thue-Mahler equations, using a number of results from classical invariant theory. In particular, we illustrate the connection between elliptic curves over $\mathbb{Q}$ and cubic forms and subsequently describe an effective algorithm for determining all elliptic curves over $\mathbb{Q}$ having good reduction outside S. This result can be summarized as follows. If we wish to find an elliptic curves E of conductor $N = p_1^{a_1} \cdots p_v^{a_v}$ for some $a_i \in \mathbb{N}$, by Theorem 1 of [BeGhRe], there exists an integral binary cubic form F of discriminant $N_0 \mid 12N$ and relatively prime integers u, v

satisfying

$$F(u,v) = w_0u^3 + w_1u^2v + w_2uv^2 + w_3v^3 = 2^{\alpha_1}3^{\beta_1}\prod_{p|N_0} p^{\kappa_p}$$

for some $\alpha_1, \beta_1, \kappa_p$. Then E is isomorphic over $\mathbb{Q}$ to the elliptic curve $E_{\mathcal{D}}$, where $E_{\mathcal{D}}$ is determined by the form F and (u,v). It is worth noting that Theorem 1 of [BeGhRe] very explicitly describes how to generate $E_{\mathcal{D}}$; once a solution (u,v) to the Thue-Mahler equation F is known, a quick computation of the Hessian and Jacobian discriminant of F evaluated at (u,v) yields the coefficients of $E_{\mathcal{D}}$. Using this theorem, all $E/\mathbb{Q}$ of conductor N may be computed by generating all of the relevant binary cubic forms, solving the corresponding Thue-Mahler equations, and outputting the elliptic curves that arise. The first and last steps of this process are straightforward. Indeed, Bennett and Rechnitzer describe an efficient algorithm for carrying out the first step REF. In fact, they having carried out a one-time computation of all irreducible forms that can arise in Theorem 1 of absolute discriminant bounded by 10^{10}. The bulk of the work is therefore concentrated in step 2, solving a large number of degree 3 Thue-Mahler equations.

Unfortunately, despite many refinements, [Hambrook's] MAGMA implementation of a Thue-Mahler solver encounters a multitude of bottlenecks which often yield unavoidable timing and memory problems, even when parallelization is considered. As our aim is to use the results of [BeGhRe] to generate all elliptic curves over $\mathbb{Q}$ of conductor $N < 10^6$, in its current state, the Hambrook algorithm is inefficient for this task, and in many cases, simply unusable due to its memory requirements. The main novel contribution of this thesis is therefore the efficient resolution of an arbitrary degree 3 Thue-Mahler equation and the implementation of this algorithm as a MAGMA package. This work is based on ideas of Matshke, von Kanel [CITE], and Siksek and is summarized in the following steps.

Step 1. Following [TdW] and [Hambrook], we reduce the problem of solving the given Thue-Mahler equation to the problem of solving a collection of finitely many S-unit equations in a certain algebraic number field K. These are equations of the

form

$$\mu_0 y - \lambda_0 x = 1 \tag{1.9}$$

for some $\mu_0, \lambda_0 \in K$ and unknowns x, y. The collection of forms is such that if we know the solutions of each equation in the collection, then we can easily derive all of the solutions of the Thue-Mahler equation. This reduction is performed in two steps. First, (1.6) is reduced to a finite number of ideal equations over K. Here, we employ new results by Siksek [Cite?] to significantly reduce the number of ideal equations to consider. Next, we reduce each ideal equation to a number of certain S-unit equations (1.9) via a finite number of principalization tests. The method of [TdW] reduces (1.6) to $(m/2)h^v$ S-unit equations, where m is the number of roots of unity of K, h is the class number, and v is the number of rational primes $p_1, \ldots, p_v$. The method of Siksek that we employ gives only $m/2$ S-unit equations. The principle computational work here consists of computing an integral basis, a system of fundamental units, and a splitting field of K, as well as computing the class group of K and the factorizations of the primes $p_1, \ldots, p_v$ into prime ideals in the ring of integers of K.

The remaining steps are performed for each of the S-unit equations in our collection.

Step 2. In place of the logarithmic sieves used in [TdW] to derive a large upper bound, we work with the global logarithmic Weil height

$$h : \mathbb{G}_m(\overline{\mathbb{Q}}) \to \mathbb{R}_{\geq 0}.$$

For a given (1.9), we show that the height $h(1/x)$ admits a decomposition into local heights at each place of K appearing in the S-unit equation. Using [CITE : Matshke, von Kanel], we generate a very large upper bound on the height $h(1/x)$, and subsequently, on the local heights. This step is a straightforward computation, whereas the analogous step in Hambrook and TdW is a complex and lengthy derivation which involves factoring rational primes into prime ideals in a splitting field of K and computing heights of certain elements of the splitting field.

Step 3. For each place of K appearing in (1.9), we drastically reduce the upper

bounds derived in Step 2 by using computational Diophantine approximation techniques applied to the intersection of a certain ellipsoid and translated lattice. This technique involves using the Finke-Pohst algorithm to enumerate all short vectors in the intersection. Here, working with the Weil height $h(1/x)$ has the advantage that it leads to ellipsoids whose volumes are smaller than the ellipsoids implicitly used in [TdW] by a factor of $\sim r^{r/2}$ for r the number of places of K appearing in our S-unit equation. In this way, we reduce the number of short vectors appearing from the Fincke-Pohst algorithm, and consequently reduce our running time and memory requirements.

Step 4. Samir's sieve - this may not be done in time as we only just received Samir's writeup and explanation as pertaining to Thue-Mahler equations.

Step 5. Finally, we use a sieving procedure to find all the solutions of the Diophantine equation that live in the box defined by the bounds derived in the previous steps. To carry out this step, we run through all the possible solutions in the box and sieve out the vast majority of non-solutions. This is done via certain low-cost congruence tests. The candidate solutions passing this test are then verified directly against (1.9). Though we expect the bounds defining the box to be small, there can still be a very large number of possible solutions to check, especially if the number of rational primes involved in the Thue-Mahler equation is large. The computations performed on each individual candidate solution are relatively simple, but the sheer number of candidates often makes this step the computational bottleneck of the entire algorithm.

Step 6. Having performed Steps 2-5 for each S-unit equation in our collection, we now have all the solutions of each such equation, and we use this knowledge to determine all the solutions of the Thue-Mahler equation.

The reader will notice several parallels between this refined algorithm and the aforementioned Goormaghtigh equation solver in the case $n = 5$. In particular, both algorithms share the same setup and refinements of the [TdW] and [Hambrook] solver. For (1.5), however, we are left to solve

$$f(y) = x^m,$$

a Thue-Mahler-like equation of degree 4 in explicit values of x and unknown integers y and m. In this case, we are permitted simplifications which allow us to omit the Fincke-Pohst algorithm and final congruence sieves. Instead, for each x, we rely on only a few iterations of the LLL algorithm to reduce our initial bound on the exponents before entering a naive search to complete our computation. Of course, this algorithm can be refined further for efficiency, however, in the context of [BeGhKr], such improvements are not needed.

The outline of this thesis is as follows. ADD

Chapter 2

Preliminaries

1. algebraic number theory background [edited once] Section 2.1.
2. p-adics [DONE - rough] Section 2.2
3. Setup with Lemmata from Samir
4. LLL [DONE - roughly]
5. Fincke-Pohst with changes from Benjamin [DONE- roughly]
6. linear forms in logs
7. Elliptic curves [DONE - rough]

2.1 Algebraic number theory

Add some better intro: maybe see masters thesis
In this section we recall some basic results from algebraic number theory that we use throughout the remaining chapters. We refer to Marcus and Neukirch for full details. Establish notation. The background for the material presented in this chapter is taken primarily from Marcus and Neukirch, and the material presented in Section 2.2 can be found in [5]

Let K be a finite algebraic extension of $\mathbb{Q}$ of degree $n = [K : \mathbb{Q}]$. There are n embeddings $\sigma : K \to \mathbb{C}$. These embeddings can be described by writing $K = \mathbb{Q}(\theta)$ for some $\theta \in \mathbb{C}$ and observing that θ can be sent to any one of its conjugates. Let s denote the number of real embeddings of K and let t denote the number of conjugate pairs of complex embeddings of K, where $n = s + 2t$. By Dirichlet's Unit Theorem, the group of units of K is the direct product of a finite cyclic group consisting of the roots of unity in K and a free abelian group of rank $r = s + t - 1$. Equivalently, there exists a system of r independent units, $\varepsilon_1, \ldots, \varepsilon_r$ such that the group of units of K is given by

$$\left\{\zeta \cdot \varepsilon_1^{a_1} \cdots \varepsilon_r^{a_r} \mid \zeta \text{ a root of unity }, a_i \in \mathbb{Z} \text{ for } i = 1, \ldots, r\right\}.$$

Any set of independent units that generate the torsion-free part of the unit group is called a system of *fundamental units*.

An element $\alpha \in K$ is called an *algebraic integer* if its minimal polynomial over $\mathbb{Z}$ is monic. The set of algebraic integers in K forms a ring, denoted $\mathcal{O}_K$. We refer to this ring as the *ring of integers* or *number ring* corresponding to the number field K. For any $\alpha \in K$, we define the *norm* of α as

$$N_{K/\mathbb{Q}}(\alpha) = \prod_{\sigma : K \to \mathbb{C}} \sigma(\alpha)$$

where the product is taken over all embeddings σ of K. For algebraic integers, $N_{K/\mathbb{Q}}(\alpha) \in \mathbb{Z}$. The units are precisely the elements of norm ± 1. Two elements α, β of K are called *associates* if there exists a unit ε such that $\alpha = \varepsilon\beta$. Let $(\alpha)\mathcal{O}_K$ denote the ideal generated by α. Associated elements generate the same ideal, and distinct generators of an ideal are associated. There exist only finitely many non-associated algebraic integers in K with given norm.

Any element of the ring of integers can be written as a product of *irreducible* elements. These are non-zero non-unit elements of $\mathcal{O}_K$ which have no integral divisors but their own associates. Unfortunately, number rings are not alway unique factorization domains: this decomposition into irreducible elements may not be unique. However, every number ring is a Dedekind domain. This means that every

ideal can be decomposed into a product of prime ideals and this decomposition is unique. A *principal* ideal is an ideal generated by a single element α. Two fractional ideals are called equivalent if their quotient is principal. It is well known that there are only finitely many equivalence classes of fractional ideals. The number of classes is called the *class number* of $\mathcal{O}_K$ and is denoted by h_K. For an ideal $\mathfrak{a}$, it is always true that $\mathfrak{a}^{h_K}$ is principal. The norm of the (integral) ideal $\mathfrak{a}$ is defined by $N_{K/\mathbb{Q}}(\mathfrak{a}) = \#(\mathcal{O}_K/\mathfrak{a})$. If $\mathfrak{a} = (\alpha)\mathcal{O}_K$ is a principal ideal, then $N_{K/\mathbb{Q}}(\mathfrak{a}) = \left|N_{K/\mathbb{Q}}(\alpha)\right|$.

Let L be a finite field extension of K with ring of integers $\mathcal{O}_L$. Every prime ideal $\mathfrak{P}$ of $\mathcal{O}_L$ *lies over* a unique prime ideal $\mathfrak{p}$ in $\mathcal{O}_K$. That is, $\mathfrak{P}$ divides $\mathfrak{p}$. The *ramification index* $e(\mathfrak{P}|\mathfrak{p})$ is the largest power to which $\mathfrak{P}$ divides $\mathfrak{p}$. The field $\mathcal{O}_L/\mathfrak{P}$ is an extension of finite degree $f(\mathfrak{P}|\mathfrak{p})$ over $\mathcal{O}_K/\mathfrak{p}$. We call $f(\mathfrak{P}|\mathfrak{p})$ the *inertial degree* of $\mathfrak{P}$ over $\mathfrak{p}$. For $\mathfrak{p}$ lying over the rational prime p, this is the integer such that

$$N_{K/\mathbb{Q}}(\mathfrak{p}) = p^{f(\mathfrak{p}|p)}.$$

The ramification index and inertial degree are multiplicative in a tower of fields. In particular, if $\mathfrak{P}$ lies over $\mathfrak{p}$ which lies over the rational prime p, then

$$e(\mathfrak{P}|p) = e(\mathfrak{P}|\mathfrak{p})e(\mathfrak{p}|p) \quad \text{and} \quad f(\mathfrak{P}|p) = f(\mathfrak{P}|\mathfrak{p})f(\mathfrak{p}|p).$$

Let $\mathfrak{P}_1, \ldots, \mathfrak{P}_m$ be the primes of $\mathcal{O}_L$ lying over a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$. Denote by $e(\mathfrak{P}_1|\mathfrak{p}), \ldots, e(\mathfrak{P}_m|\mathfrak{p})$ and $f(\mathfrak{P}_1|\mathfrak{p}), \ldots, f(\mathfrak{P}_m|\mathfrak{p})$ the corresponding ramification indices and inertial degrees. Then

$$\sum_{i=1}^{r} e(\mathfrak{P}_i|\mathfrak{p})f(\mathfrak{P}_i|\mathfrak{p}) = [L : K].$$

If L is normal over K and $\mathfrak{P}_i$ and $\mathfrak{P}_j$ are two prime ideals lying over $\mathfrak{p}$, then $e(\mathfrak{P}_i|\mathfrak{p}) = e(\mathfrak{P}_j|\mathfrak{p})$ and $f(\mathfrak{P}_i|\mathfrak{p}) = f(\mathfrak{P}_j|\mathfrak{p})$. In this case, $\mathfrak{p}$ factors into

$$\mathfrak{p}\mathcal{O}_L = (\mathfrak{P}_1 \cdots \mathfrak{P}_m)^e,$$

in $\mathcal{O}_L$, where the $\mathfrak{P}_i$ are distinct prime ideals all having the same ramification

degree e and inertial degree f over $\mathfrak{p}$. Moreover, it follows that

$$mef = [L : K].$$

2.2 p-adic valuations

fix intro: In this section we give a concise exposition of p-adic valuations. Everything in this document is based off of "Elementary and analytic theory of algebraic numbers" by W. Narkiewicz. Throughout this thesis, we denote the algebraic closure of $\mathbb{Q}_p$ by $\overline{\mathbb{Q}}_p$. The completion of $\overline{\mathbb{Q}}_p$ with respect to the absolute value of $\overline{\mathbb{Q}}_p$ is denoted by $\mathbb{C}_p$.

Let $g(t)$ be an irreducible polynomial in $\mathbb{Q}[t]$ of degree n and let $K = \mathbb{Q}(\theta)$, where $g(\theta) = 0$. A homomorphism $v : K^* \to \mathbb{R}_{\geq 0}$ of the multiplicative group of K into the group of positive real numbers is called a *valuation* if it satisfies the condition

$$v(x + y) \leq v(x) + v(y).$$

By setting $v(0) = 0$, we extend v to the whole field K. If

$$v(x + y) \leq \max(v(x), v(y))$$

holds for all $x, y \in K$, then v is called a *non-Archimedean valuation*. All remaining valuations on K are called *Archimedean*.

Every valuation $v(x)$ induces a metric in K via $d(x, y) = v(x - y)$. Under this identification, the additive and multiplicative groups of K become topological groups, and moreover, K becomes a topological field. We say that two valuations are *equivalent* if they define the same topology. A *place* of a number field K is an equivalence class of absolute values on K.

For any non-zero prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$, let $\mathrm{ord}_{\mathfrak{p}}(\mathfrak{a})$ denote the exact power to which $\mathfrak{p}$ divides the ideal $\mathfrak{a}$. For fractional ideals $\mathfrak{a}$ this number can of course be negative. For $\alpha \in K$, we write $\mathrm{ord}_{\mathfrak{p}}(\alpha)$ for $\mathrm{ord}_{\mathfrak{p}}((\alpha)\mathcal{O}_K)$. Every prime ideal defines a

discrete non-Archimedean valuation on K via

$$v(x) := \left(\frac{1}{N_{K/\mathbb{Q}}(\mathfrak{p})}\right)^{\mathrm{ord}_{\mathfrak{p}}(x)}.$$

Moreover, every embedding of K into the complex field defines an Archimedean valuation. Conversely, every discrete valuation on K arises in this way by a prime ideal of $\mathcal{O}_K$, while every Archimedean valuation of K is equivalent to $|\sigma(x)|$, where σ is an embedding of K into $\mathbb{C}$. Valuations defined by different prime ideals are non-equivalent, and two valuations defined by different embeddings of K into $\mathbb{C}$ are equivalent if and only if those embeddings are complex conjugated.

The topology induced in K by a prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ is called the $\mathfrak{p}$-*adic topology*. In the ring $\mathbb{Q}$, the prime ideals are generated by the rational primes p, and the resulting topology in the field $\mathbb{Q}$ of rational numbers is called the p-*adic topology*. In particular, if $v(x)$ is a non-trivial valuation of $\mathbb{Q}$, then either $v(x)$ is equivalent to the ordinary absolute value $|x|$, or it is equivalent to one of the p-adic valuations induced by rational primes.

Let V be the set of all valuations of an algebraic number field K. Then for every non-zero element $a \in K$ we have

$$\prod_{v \in V} v(a) = 1.$$

There are one-to-one correspondences between each of the following four sets of objects:

1. the prime ideals in (the ring of integers of) K that divide p.
2. The irreducible polynomial factors of $g(t)$ in $\mathbb{Q}_p[t]$.
3. The classes of conjugate embeddings of K into the algebraic closure $\overline{\mathbb{Q}_p}$ of $\mathbb{Q}_p$.
4. The extensions of the p-adic valuation ord_p on $\mathbb{Q}$ to K.

In what follows, we will describe the important features of these correspondences.

Note that two embeddings of K into $\overline{\mathbb{Q}_p}$ are called *conjugate* if they map θ to the roots of hte same irreducible polynomial in $\mathbb{Q}_p[t]$. Note also that what we call a p-adic valuation is sometimes called a p-adic order.

Let

$$(p)\mathcal{O}_K = \mathfrak{p}_1^{e_1} \cdots \mathfrak{p}_m^{e_m}$$

change notation of the e_i be the decomposition of $(p)\mathcal{O}_K$ into prime ideals of $\mathcal{O}_K$, with inertial degree f_i for $\mathfrak{p}_i$ over p. Let $K_{\mathfrak{p}_i}$ denote the completion of K with respect to $\mathrm{ord}_{\mathfrak{p}_i}$. Let

$$g(t) = g_1(t) \cdots g_m(t)$$

be the decomposition of $g(t)$ into irreducible polynomials in $\mathbb{Q}_p[t]$. For each $i \in \{1, \ldots, m\}$, let $n_i = \deg g_i(t)$. The correspondence between $\mathfrak{p}_i$ and $g_i(t)$ is such that $n_i = e_i f_i$ and $K_{\mathfrak{p}_i} \simeq \mathbb{Q}_p(\theta_i)$, where $g_i(\theta_i) = 0$.

There are n embeddings of K into $\overline{\mathbb{Q}_p}$, and each one fixes $\mathbb{Q}$ and maps θ to a root of g in $\overline{\mathbb{Q}_p}$. Let $\theta_i^{(1)}, \ldots, \theta_i^{(n_i)}$ denote the roots of $g_i(t)$ in $\overline{\mathbb{Q}_p}$. For $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$, let σ_{ij} be the embedding of K into $\mathbb{Q}_p(\theta_i^{(j)})$ defined by $\theta \mapsto \theta_i^{(j)}$. The m classes of conjugate embeddings are $\{\sigma_{i1}, \ldots, \sigma_{in_i}\}$ for $i = 1, \ldots, m$. Note that σ_{ij} coincides with the embedding $K \hookrightarrow K_{\mathfrak{p}_i} \simeq \mathbb{Q}(\theta_i) \simeq \mathbb{Q}_p\left(\theta_i^{(j)}\right)$.

For any finite extension L of $\mathbb{Q}_p$, the p-adic valuation of $\mathbb{Q}_p$ extends uniquely to L as

$$\mathrm{ord}_p(x) = \frac{1}{[L : \mathbb{Q}_p]} \mathrm{ord}_p(N_{L/\mathbb{Q}_p}(x)).$$

This definition is independent of the field L containing x. So, since each element of $\overline{\mathbb{Q}_p}$ is by definition contained in some finite extension of $\mathbb{Q}_p$, this definition can be used to define the p-adic valuation of any $x \in \overline{\mathbb{Q}_p}$. Every finite extension of $\mathbb{Q}_p$ is complete with respect to ord_p, but $\overline{\mathbb{Q}_p}$ is not. The completion of $\overline{\mathbb{Q}_p}$ with respect to ord_p is denoted by $\mathbb{C}_p$. Note that the formulas

$$\mathrm{ord}_p(xy) = \mathrm{ord}_p(x) + \mathrm{ord}_p(y), \quad \mathrm{ord}_p(x+y) \geq \min\{\mathrm{ord}_p(x), \mathrm{ord}_p(y)\}$$

still hold when $x, y \in \mathbb{C}_p$. It is convenient to record here that an element $x \in \mathbb{C}_p$ having $\mathrm{ord}_p(x) = 0$ is called a p-adic *unit*.

The m extensions of the p-adic valuation on $\mathbb{Q}$ to K are just multiples of the $\mathfrak{p}_i$-adic valuation on K:

$$\mathrm{ord}_p(x) = \frac{1}{e_i}\,\mathrm{ord}_{\mathfrak{p}_i}(x) \quad \text{for } i = 1, \ldots, m.$$

We also view these extensions as arising from various embeddings of K into $\overline{\mathbb{Q}_p}$. Indeed, the extension to $\mathbb{Q}_p(\theta_i^{(j)})$ of the p-adic valuation on $\mathbb{Q}_p$ induces a p-adic valuation on K via the embedding σ_{ij} as

$$\mathrm{ord}_p(x) = \mathrm{ord}_p(\sigma_{ij}(x)),$$

and we have

$$\mathrm{ord}_p(\sigma_{ij}(x)) = \frac{1}{e_i}\,\mathrm{ord}_{\mathfrak{p}_i}(x) \quad \text{for } i = 1, \ldots, m, j = 1, \ldots, n_i.$$

Weil height

Let k be a number field and at each place v of k, let k_v denote the completion of k at v. Then

$$\sum_{v|p} [k_v : \mathbb{Q}_v] = [k : \mathbb{Q}]$$

for all places p of $\mathbb{Q}$. We will use two normalized absolute values $|\ |_v$ and $\|\ \|_v$ on k which we now define. If $v|\infty$, then $\|\ \|_v$ restricted to $\mathbb{Q}$ is the usual Archimedean absolute value; if $v|p$ for a rational prime p, then $\|\ \|_v$ restricted to $\mathbb{Q}$ is the usual p-adic absolute value. We then set

$$|\ |_v = \|\ \|_v^{[k_v:\mathbb{Q}_v]/[k:\mathbb{Q}]}.$$

The *logarithmic Weil height* $h : \overline{\mathbb{Q}} \to [0, \infty)$ is now defined as follows. Given $\alpha \in \overline{\mathbb{Q}}$, select any number field k containing α, and let

$$h(\alpha) = \frac{1}{[k : \mathbb{Q}]} \sum_v \log^+ |\alpha|_v,$$

the sum being taken over all places v of k. The height does not depend on the

choice of k containing α. We note that if α is an algebraic unit, then $|\alpha|_v = 1$ for all finite places v, and therefore $h(\alpha)$ can be taken over the infinite places only. In particular, if $\alpha \in \mathbb{Q}$, then with $\alpha = p/q$ for $p, q \in \mathbb{Z}$ with $\gcd(p, q) = 1$, we have $h(\alpha) = \log \max\{|p|, |q|\}$ and if $\alpha \in \mathbb{Z}$ then $h(\alpha) = \log |\alpha|$. check this.

The p-adic logarithm Every non-zero $\alpha \in \mathbb{Q}_p$ has a p-adic expansion

$$\alpha = \sum_{i=k}^{\infty} u_i p^i$$

where $k = \mathrm{ord}_p(\alpha)$ and the p-adic digits u_i are in $\{0, \ldots, p-1\}$ with $u_k \neq 0$. If $\mathrm{ord}_p(\alpha) \geq 0$ then α is called a p-adic *integer*. The set of p-adic integers is denoted $\mathbb{Z}_p$. A p-adic *unit* is an $\alpha \in \mathbb{Q}_p$ with $\mathrm{ord}_p(\alpha) = 0$. For any p-adic integer α and $\mu \in \mathbb{N}_0$ there exists a unique rational integer $\alpha^{(\mu)} = \sum_{i=0}^{\mu-1} u_i p^i$ such that

$$\mathrm{ord}_p(\alpha - \alpha^{(\mu)}) \geq \mu, \quad \text{and} \quad 0 \leq \alpha^{(\mu)} \leq p^\mu - 1.$$

For $\mathrm{ord}_p(\alpha) \geq k$ we also write $\alpha \equiv 0 \mod p^k$. The p-adic norm is defined by

$$|\alpha|_p = p^{-\mathrm{ord}_p(\alpha)}.$$

We have seen how to define $\mathrm{ord}_\mathfrak{p}$ and ord_p on algebraic extensions of $\mathbb{Q}$. For any $z \in \mathbb{C}_p$ with $\mathrm{ord}_p(z-1) > 0$, we can also define the p-adic logarithm of z by

$$\log_p(z) = -\sum_{i=1}^{\infty} \frac{(1-z)^i}{i}.$$

By the n^{th} term test, this series converges precisely in the region where $\mathrm{ord}_p(z-1) > 0$. Three important properties of the p-adic logarithm are

1. $\log_p(xy) = \log_p(x) + \log_p(y)$ whenever $\mathrm{ord}_p(x-1) > 0$ and $\mathrm{ord}_p(y-1) > 0$.
2. $\log_p(z^k) = k \log(p)$ whenever $\mathrm{ord}_p(z-1) > 0$ and $k \in \mathbb{Z}$.
3. $\mathrm{ord}_p(\log_p(z)) = \mathrm{ord}_p(z-1)$ whenever $\mathrm{ord}_p(z-1) > 1/(p-1)$.

where to find proofs?

We shall use the following lemma to extend the definition of the p-adic logarithm to all p-adic units in $\overline{\mathbb{Q}_p}$.

Lemma 2.2.1. *Let z be a p-adic unit belonging to a finite extensions L of $\mathbb{Q}_p$. Let e and f be the ramification index and inertial degree of L.*

1. *There is a positive integer r such that $\mathrm{ord}_p(z^r - 1) > 0$.*
2. *If r is the smallest positive integer having $\mathrm{ord}_p(z^r - 1) > 0$, then r divides $p^f - 1$, and an integer q satisfies $\mathrm{ord}_p(z^q - 1) > 0$ if and only if it is a multiple of r.*
3. *If r is a nonzero integer with $\mathrm{ord}_p(z^r - 1) > 0$, and if k is an integer with $p^k(p-1) > e$, then*
$$\mathrm{ord}_p(z^{rp^k} - 1) > \frac{1}{p-1}.$$

proofs?

For z a p-adic unit in $\overline{\mathbb{Q}_p}$ we define

$$\log_p z = \frac{1}{q} \log_p z^q,$$

where q is an arbitrary non-zero integer such that $\mathrm{ord}_p(z^q - 1) > 0$. To see that this definition is independent of q, let r be the smallest positive integer with $\mathrm{ord}_p(z^r - 1) > 0$, and note that q/r is an integer, and use the second property of p-adic logarithms above to write

$$\frac{1}{q} \log_p z^q = \frac{1}{r(q/r)} \log_p z^{r(q/r)} = \frac{1}{r} \log_p z^r.$$

Choosing q such that $\mathrm{ord}_p(z^q - 1) > 1/(p-1)$ helps to speed up and control the convergence of the series defining $\log_p$ refs.

It is straightforward to see that Properties 1 and 2 above extend to the case where x, y, z are p-adic units. Combining this with Property 3, we obtain

Lemma 2.2.2. *Let $z_1, \ldots, z_m \in \overline{\mathbb{Q}_p}$ be p-adic units and let $b_1, \ldots, b_m \in \mathbb{Z}$. If*

$$\mathrm{ord}_p(z_1^{b_1} \cdots z_m^{b_m} - 1) > \frac{1}{p-1}$$

then

$$\mathrm{ord}_p(b_1 \log_p z_1 + \cdots + b_m \log_p z_m) = \mathrm{ord}_p(z_1^{b_1} \cdots z_m^{b_m} - 1).$$

2.2.1 Lattices: LLL and the Fincke-Pohst algorithm

references for this are the two books living in the $x + y = z$ folder on the desktop

Lattices An n-dimensional lattice is a discrete subgroup of $\mathbb{R}^n$ of the form

$$\Gamma = \left\{ \sum_{i=1}^{n} x_i \mathbf{c}_i \ : \ x_i \in \mathbb{Z} \right\},$$

where $\mathbf{c}_1, \ldots, \mathbf{c_n}$ are vectors forming a basis for $\mathbb{R}^n$. We say that the vectors $\mathbf{c}_1, \ldots, \mathbf{c_n}$ form a *basis* for Γ, or that they generate Γ. Let B denote the matrix whose columns are the vectors $\mathbf{c}_1, \ldots, \mathbf{c_n}$. Any lattice element $\mathbf{v}$ may be expressed as $\mathbf{v} = B\mathbf{x}$ for some $\mathbf{x} \in \mathbb{Z}^n$.

A *bilinear form* on a lattice Γ is a function $\Phi : \Gamma \times \Gamma \to \mathbb{Z}$ satisying

1. $\Phi(\mathbf{u}, \mathbf{v} + \mathbf{w}) = \Phi(\mathbf{u}, \mathbf{v}) + \Phi(\mathbf{u}, \mathbf{w})$
2. $\Phi(\mathbf{u} + \mathbf{v}, \mathbf{w}) = \Phi(\mathbf{u}, \mathbf{w}) + \Phi(\mathbf{v}, \mathbf{w})$
3. $\Phi(a\mathbf{u}, \mathbf{w}) = a\Phi(\mathbf{u}, \mathbf{w})$
4. $\Phi(\mathbf{u}, a\mathbf{w}) = a\Phi(\mathbf{u}, \mathbf{w})$

for all $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ in Γ and any $a \in \mathbb{R}$.

In particular, given a basis we can define a specific bilinear form on our lattice Γ as part of its structure. In the case of integral lattices, we have $\Phi : \Gamma \times \Gamma \to \mathbb{Z}$. This form describes a kind of distance between elements $\mathbf{x}$ and $\mathbf{y}$ of the lattice defined

by $\Phi(\mathbf{x}, \mathbf{y})$.

A *quadratic form* is a homogeneous polynomial of degree 2. A form Q is called positive definite if $Q(\mathbf{x})$ is strictly positive for any nonzero $\mathbf{x}$. A lattice is called *positive definite* if its quadratic form is positive definite.

A bilinear form has an associated quadratic form $Q : \Gamma \to \mathbb{Z}$ which is simply defined by $Q(\mathbf{x}) = \Phi(\mathbf{x}, \mathbf{x})$. The bilinear forms (and their associated quadratic forms) that we will be using come from the usual inner product on vectors in $\mathbb{R}^n$, also known as the dot product $\mathbf{u} \cdot \mathbf{v}$ for $\mathbf{u}, \mathbf{v} \in \Gamma$, and multiplication with the basis matrix for coordinate vectors. That is, if $\mathbf{u} = B\mathbf{x}$ and $\mathbf{v} = B\mathbf{y}$ for a basis B, we have $\Phi(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T B^T B\mathbf{y}$.

If $\mathbf{v} = B\mathbf{x}$, the *norm* of the vector $\mathbf{v} \in \Gamma$ is defined by the quadratic form. We will be using the inner product $\mathbf{v} \cdot \mathbf{v}$. The norm of the coordinate vector $\mathbf{x}$ is then

$$\mathbf{v}^T\mathbf{v} = (B\mathbf{x})^T(B\mathbf{x}) = \mathbf{x}^T B^T B\mathbf{x}.$$

Notice that this is also $\mathbf{x}^T A\mathbf{x}$ where $B^T B = A$. Here, A is an example of the Gram matrix of Γ. The *Gram matrix* of a lattice with basis B with respect to a bilinear form Φ is defined to be the matrix A with entries $a_{ij} = \Phi(\mathbf{b}_i, \mathbf{b}_j)$.

The bilinear form on L can be written with respect to either embedded or coordinate vectors. Using another basis to express the lattice elements is possible, and sometimes preferable. But the Gram matrix is specific to the bilinear form on the lattice, and should not change when operating on embedded vectors. If it is operating on coordinate vectors, the change of basis must be accounted for.

If A and B are invertible $n \times n$ real matrices, then the lattice generated by the columns of A is equal to the lattice generated by the columns of B if and only if there is a unimodular matrix U such that $AU = B$.

LLL Intro-ish: taken from Cohen p103 Among al1 the Z bases of a lattice L, some are better than others. The ones whose elements are the shortest (for the corresponding norm associated to the quadratic form q) are called reduced. Since the bases al1 have the same determinant, to be reduced implies also that a basis is

not too far from being orthogonal. The notion of reduced basis is quite old, and in fact in some sense one can even define an optimal notion of reduced basis. The problem with this is that no really satisfactory algorithm is known to find such a basis in a reasonable time, except in dimension 2 (Algorithm 1.3.14), and quite recently in dimension 3 from the work of B. Valle [Val]. A real breakthrough came in 1982when A. K. Lenstra, H. W. Lenstra and L. Lovkz succeeded in giving a new notion of reduction (what is now called 2.6 Lattice Reduction Algorithms 85 LLL-reduction) and simultaneously a reduction algorithm which is determin- istic and polynomial time (see [LLL]).This has proved invaluable.

Let Γ be a lattice with $\mathbf{c}_1, \ldots, \mathbf{c}_n$. Define the vectors $\mathbf{c}_i^*$ for $i = 1, \ldots, n$ and real numbers μ_{ij} $(1 \leq j < i \leq n)$ inductively by

$$\mathbf{c}_i^* = \mathbf{c}_i - \sum_{j=1}^{i-1} \mu_{ij} \mathbf{c}_j^*, \quad \mu_{ij} = \frac{\langle \mathbf{c}_i, \mathbf{c}_j^* \rangle}{\langle \mathbf{c}_j, \mathbf{c}_j^* \rangle}$$

(This is just the Gram-Scmidt process). The basis $_1, \ldots, _n$ is called *LLL-reduced* it

$$|\mu_{ij}| \leq \frac{1}{2} \quad \text{for } 1 \leq j < i \leq n,$$

$$\frac{3}{4}|^*_{i-1}|^2 \leq |^*_i + \mu_{ii-1}{}^*_{i-1}|^2 \quad \text{for } 1 < i \leq n.$$

These properties implies that an LLL-reduced basis is approximately orthogonal, and that, generically, its constituent vectors are roughly of the same length. Every n-dimensional lattice has an LLL-reduced basis and such a basis can be computed very quickly using the so-called LLL algorithm (ref). The LLL algorithm takes as input an arbitrary basis for a lattice and outputs an LLL-reduced basis for the lattice. The algorithm is typically modified to additionally output a unimodular matrix U such that $B = AU$, where A is the matrix whose column-vectors are the input basis and B is the matrix whose column-vectors are the LLL-reduced output basis. Several versions of this algorirhm are implemented in MAGMA, including de Weger's exact integer version. (ref).

For Γ an n-dimensional lattice and $\mathbf{y}$ a vector in $\mathbb{R}^n$, we define

$$l(\Gamma, \mathbf{y}) = \min_{\mathbf{x} \in \Gamma \setminus \{\mathbf{y}\}} |\mathbf{x} - \mathbf{y}|.$$

The most important property of an LLL-reduced basis for us is the following lemma.

Lemma 2.2.3. *lemma 18.1*

refs of where this lemma can be found - Cohen, for 1 Note that the assumption in lemma cite is equivalent to $\mathbf{y} \notin \Gamma$.

Cohen: We see that the vector bl in a reduced basis is, in a very precise sense, not too far from being the shortest non-zero vector of L. In fact, it often is the shortest, and when it is not, one can, most of the time, work with bl instead of the actual shortest vector. As has already been mentioned, what makes al1these notions and theorems so valuable is that there is a very simple and efficient algorithm to find a reduced basis in a lattice. We now describe this algorithm in its simplest form.

Fincke-Pohst

We show how to modify the Fincke-Pohst algorithm to output short vectors in a translated lattice. That is, we compute the set of vectors x such that

$$(x - c)^t B^t B (x - c) \leq C$$

where c is some vector over $\mathbb{Q}$ which represents the translation of our lattice.

We begin with the usual Fincke-Pohst method for

$$x^t B^t B x \leq C.$$

We call a vector $\mathbf{v}$ *small* if its norm $\Phi(\mathbf{v}, \mathbf{v})$ is less than a constant C. This clearly depends on the basis which is given, and can vary depending on the choice of basis. If a particular basis is not specified, it is assumed to be the matrix B which defines the Gram matrix $A = B^t B$. This is equivalent to solving the inequality

$\Phi(\mathbf{y}, \mathbf{y}) \leq C$ where $\Phi(\mathbf{y}, \mathbf{y}) = \mathbf{y}^t\mathbf{y}$ denotes the norm of the vector computed with respect to the lattice. Let B denote the matrix whose columns are the basis vectors of the lattice $\mathcal{L}$. As an element of the lattice, $\mathbf{y} = B\mathbf{x}$ for some coordinate vector $\mathbf{x} \in \mathbb{Z}^n$. So our inequality becomes

$$\Phi(\mathbf{y}, \mathbf{y}) = \mathbf{y}^t\mathbf{y} = \mathbf{x}^t B^t B\mathbf{x} \leq C.$$

We consider the quadratic form $Q(\mathbf{x}) = \mathbf{x}^t B^t B\mathbf{x}$ and solve $Q(\mathbf{x}) \leq C$.

Quadratic Completion

To solve our inequality, it helps to first rearrange the terms of our quadratic form. This reformulation is called the quadratic completion or quadratic complementation. Here we assume the lattice is positive definite. That is, every nonzero element has a positive norm. With this, we can find the Cholesky decomposition $A = LL^t$, where L is a lower triangular matrix. Equivalently, we can express this as $A = R^t R$, where R is an upper triangular matrix. Since Fincke-Pohst uses upper triangular matrices, this is what we will use. The formulas below will reflect this. We now express Q as:

$$Q(\mathbf{x}) = \sum_{i=1}^{m} q_{ii} \left(x_i + \sum_{j=i+1}^{m} q_{ij} x_j \right)^2.$$

Our coefficients q_{ij} are defined from R and stored in a matrix for convenience.

$$q_{ij} = \begin{cases} \frac{r_{ij}}{r_{ii}} & \text{if } i < j \\ r_{ii}^2 & \text{if } i = j \end{cases}.$$

Since R is upper triangular, the matrix $Q = [q_{ij}]$ will be as well.

To obtain the upper triangular matrix R from our matrix A, we compute the diag-

onal and non-diagonal entries as follows:

$$r_{ii} = \sqrt{a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2}$$

$$r_{ij} = \frac{1}{r_{ii}} \left(a_{ij} - \sum_{k=1}^{j-1} r_{ki} r_{kj} \right).$$

Using these, we can reformulate the construction of the coefficients of Q to use values from A. We will soon see how it is possible to do away with using the Cholesky decomposition entirely.

$$q_{ii} = a_{ii} - \sum_{k=1}^{i-1} r_{ki}^2$$

$$q_{ij} = \frac{1}{r_{ii}^2} \left(a_{ij} - \sum_{k=1}^{j-1} r_{ki} r_{kj} \right).$$

By putting this construction in terms of the coefficients of Q only, we arrive at the following

$$q_{ii} = a_{ii} - \sum_{k=1}^{i-1} q_{ki}^2 q_{kk}$$

$$q_{ij} = \frac{1}{q_{ii}} \left(a_{ij} - \sum_{k=1}^{j-1} q_{ki} q_{kj} q_{kk} \right).$$

We can then calculate these coefficients, starting with q_{11} and calculating q_{1j} for $1 \leq j \leq m$. Then we continue by calculating q_{22} and q_{2j} for $2 \leq j \leq m$. We proceed by first always calculating the diagonal entry q_{ii} and then q_{ij} for $i \leq j \leq m$ until we reach q_{mm}. In practice, this is how we compute the coefficients for our form. However, it is equally possible to first compute the Cholesky Decomposition using available methods, and then computing the entries of Q from this. In fact, we do exactly this, by first computing the Cholesky decomposition.

The usual Fincke-Pohst way to bound x_i

Since the sum $Q(x)$ is less than C, the individual term $q_{mm}x_m^2$ must also be less than C.

$$\begin{aligned}\sum_{i=1}^{m} q_{ii}\left(x_i + \sum_{j=i+1}^{m} q_{ij}x_j\right)^2 &\leq C \\ q_{mm}x_m^2 &\leq C \\ x_m^2 &\leq \frac{C}{q_{mm}}.\end{aligned}$$

In fact, x_m is bounded above by $\sqrt{C/q_{mm}}$ and below by $-\sqrt{C/q_{mm}}$.

This illustrates the first step in establishing bounds on a specific entry x_i. Adding more terms from the outer sum to this sequence, a pattern emerges.

$$\begin{aligned}q_{mm}x_m^2 &\leq C \\ q_{m-1,m-1}\left(x_{m-1} + q_{m-1,m}x_m\right)^2 &\leq C - q_{mm}x_m^2 \\ q_{m-2,m-2}\left(x_{m-2} + \sum_{j=m-1}^{m} q_{m-2,j}x_j\right)^2 &\leq C - q_{mm}x_m^2 - q_{m-1,m-1}\left(x_{m-1} + q_{m-1,m}x_m\right)^2\end{aligned}$$

Let

$$U_k = \sum_{j=k+1}^{m} q_{kj}x_j$$

so that we can rewrite $Q(\mathbf{x})$ as

$$Q(\mathbf{x}) = \sum_{i=1}^{m} q_{ii}\left(x_i + \sum_{j=i+1}^{m} q_{ij}x_j\right)^2 = \sum_{i=1}^{m} q_{ii}\left(x_i + U_i\right)^2$$

In general,

$$q_{kk}(x_k + U_k)^2 \leq C - \sum_{i=k+1}^{m} q_{ii}(x_i + U_i)^2.$$

Let T_k denote the bound on the right-hand side. That is

$$T_k = C - \sum_{i=k+1}^{m} q_{ii}(x_i + U_i)^2,$$

so that $T_m = C$, $T_{m-1} = C - q_{mm}x_m^2$ and

$$T_{m-2} = C - q_{mm}x_m^2 - q_{m-1,m-1}\left(x_{m-1} + q_{m-1,m}x_m\right)^2.$$

We set T_m as C and find each subsequent T_k by subtracting the next term from the outer summand:

$$T_k = C - \sum_{i=k+1}^{m} q_{ii}(x_i + U_i)^2,$$

$$T_k = T_{k+1} - q_{k+1,k+1}(x_{k+1} + U_{k+1})^2.$$

Now, we have an upper bound for each summand.

$$q_{kk}(x_k + U_k)^2 \leq T_k.$$

Using this, we can estimate upper and lower bounds for each x_k in the coordinate vector $\mathbf{x}$. We start by computing the last entries of $\mathbf{x}$ and their bounds first. Assuming that the last several entries of $\mathbf{x}$ have been assigned, upper and lower bounds on x_k can be determined. Now that we have established a bound on a term in the outer sum, we can determine bounds on the specific entry x_k. Take the above equation, and solve for x_k. Take the above equation and solve for x_k:

$$\begin{aligned}(x_k + U_k)^2 &\leq T_k/q_{kk} \\ x_k + U_k &\leq \sqrt{T_k/q_{kk}} \\ x_k &\leq \sqrt{T_k/q_{kk}} - U_k.\end{aligned}$$

Similarly, we have a lower bound:

$$x_k \geq -\sqrt{T_k/q_{kk}} - U_k.$$

Since x_k must be an integer, we can restrict our bounds further. Let $t_k = \sqrt{T_k/q_{kk}}$.

$$UB_k = \lfloor t_k - U_k \rfloor$$

$$LB_k = \lceil -t_k - U_k \rceil$$

Here UB_k is the upper bound on x_k and LB_k is the lower bound on x_k.

$$LB_x \leq x_k \leq UB_k.$$

To enumerate all of the vectors $\mathbf{x}$ such that $Q(\mathbf{x}) \leq C$, begin with the last entry x_m (letting all other $x_j = 0$). Determine the upper and lower bounds UB_m and LB_m by first calculating $t_m = \sqrt{T_m/q_{mm}}$. We define $U_m = 0$, and by definition remember that $T_m = C$.

For each entry x_i, starting with x_m and going down to x_1, we initialize the value to be $x_i = LB_i$. After the value is initialized, we begin to increment the values of all the entries, adding 1 to each entry until we either reach the last index (in which case we have found a solution) or we exceed the upper bound on a particular entry (we will need to readjust the previously assigned entries). If at any time the lower bound exceeds the upper bound for a given entry, it will become immediately apparent when the value for that entry is initialized. We must then backtrack to our previous entries (that is, entries with a higher index). If we reach x1 without exceeding the upper bounds for any entry, then we have found a complete vector $\mathbf{x}$ which satisfies $Q(\mathbf{x}) \leq C$.

We will know we have found all the short vectors when we reach the zero vector. This is because we start by assigning each value x_i its lower bound, which is calculated with respect to the values $x_{i+1}, \ldots, x_n$. We increase x_i incrementally, until it exceeds the corresponding calculated upper bound. When this happens we revisit x_{i+1}, increasing its value. Since x_{i+1} was originally assigned its own lower bound,

it starts off as a negative integer and increases steadily until it reaches 0. Likewise, the other values will start off negative at each iteration and slowly increase in value. It is only when all entries are 0 that the algorithm terminates. When we add each vector, we also add the vector with entries $-x_i$ for each i. In this we capture all the small vectors without having to check positive values for x_n.

Before beginning the search, first find the coefficients of the quadratic form expressed as above. Initialize T_k, U_k, UB_k and x_k to be 0 for all k. Begin with $i = m$ and $T_i = C$ as the value bounding our vectors.

It is noted in the Fincke-Pohst paper that if we label the columns of R by $\mathbf{r}_i$ (from the Cholesky decomposition $\mathbf{x}^t R^t R \mathbf{x}$) and the rows of R^{-1} by $\mathbf{r}'_i$, then we see that

$$x_i^2 = \left(\mathbf{r}_i'^t \left(\sum_{k=1}^{m} x_k \mathbf{r}_k \right) \right)^2 \leq \mathbf{r}_i'^t \mathbf{r}_i (\mathbf{x}^t R^t R \mathbf{x}) \leq \|\mathbf{r}'_i\|^2 C.$$

So it may behoove us to reduce the rows of R^{-1} in order to reduce our search space. Furthermore, it helps to put the smallest basis vectors first, so reordering the columns may also be beneficial.

Express this reduction with a unimodular matrix V^{-1} so that $R_1^{-1} = V^{-1}R^{-1}$. Then reorder the columns of R_1 with a permutation matrix P. Since $R_1 = RV$, we then have that $R_2 = (RV)P$.

Then $R_2^{-1} = P^{-1}V^{-1}R^{-1}$. If we find a solution to the inequality $\mathbf{y}^t R_2^t R_2 \mathbf{y} \leq C$, we can recover a solution to our original inequality by $\mathbf{x} = VP\mathbf{y}$. Since $R_2^{-1} = P^{-1}V^{-1}R^{-1}$, we know that $R_2 = RVP$.

$$\begin{aligned}
\mathbf{y}^t R_2^t R_2 \mathbf{y} &\leq C \\
\mathbf{y}^t (P^t V^t R^t)(RVP)\mathbf{y} &\leq C \\
(\mathbf{y}^t P^t V^t) R^t R (VP\mathbf{y}) &\leq C \\
(VP\mathbf{y})^t R^t R (VP\mathbf{y}) &\leq C \\
\mathbf{x}^t R^t R \mathbf{x} &\leq C.
\end{aligned}$$

This improves the search time by giving us a nicer quadratic form to work with. Once we find solutions to the inequality given by $Q_2(\mathbf{y}) = \mathbf{y}^t R_2^t R_2 \mathbf{y} \leq C$, it is a simple matter of translating them into solutions of our original inequality.

2.2.2 Translated Lattices

We now explain how to apply Fincke-Pohst to the case

$$(x-c)^t B^t B (x-c) \leq C.$$

In place of the usual reduction listed above, we use MAGMA's built-in LLLGram function on the symmetric positive-definite matrix $A = B^t B$. Here, since A is symmetric and positive-definite, it can be written as $A = R^t R$ for some upper triangular matrix R (via Cholesky Decomposition).The function LLLGram, with input A, computes a matrix G which is the Gram matrix corresponding to a LLL-reduced form of the matrix R. This function returns three values:

- A LLL-reduced Gram matrix G of the Gram matrix A;
- A unimodular matrix U in the matrix ring over $\mathbb{Z}$ whose degree is the number of rows of A such that $G = U^t A U$ (technically it returns $G = U A U^t$, but we change this here to simplify our computations later);
- The rank of A (which equals the dimension of the lattice generated by R).

Thus

$$(U^{-1})^t G U^{-1} = A$$

and we have

$$\begin{aligned}
(x-c)^t B^t B (x-c) &\leq C \\
(x-c)^t A (x-c) &\leq C \\
(x-c)^t (U^{-1})^t G U^{-1} (x-c) &\leq C \\
\left(U^{-1}(x-c)\right)^t G \left(U^{-1}(x-c)\right) &\leq C \\
(y-d)^t G (y-d) &\leq C
\end{aligned}$$

where

$$y = U^{-1}x \quad \text{and} \quad d = U^{-1}c.$$

Now, we are in position to enumerate the short vectors y satisfying

$$(y-d)^t G (y-d) \leq C.$$

We retrieve our solutions x via $x = Uy$.

As before, we generate the matrix Q such that

$$Q(\mathbf{x}) = \sum_{i=1}^{m} q_{ii} \left(y_i - d_i + \sum_{j=i+1}^{m} q_{ij}(y_j - d_j) \right)^2 .$$

Since the sum $Q(x)$ is less than C, the individual term $q_{mm}(y_m - d_m)^2$ must also be less than C.

$$\begin{aligned}
\sum_{i=1}^{m} q_{ii} \left(y_i - d_i + \sum_{j=i+1}^{m} q_{ij}(y_j - d_j) \right)^2 &\leq C \\
q_{mm}(y_m - d_m)^2 &\leq C.
\end{aligned}$$

Here, in place of the usual method of bounding $y_m - d_m$ by $\sqrt{C/q_{mm}}$ and $-\sqrt{C/q_{mm}}$, we instead let y_m vary between $-\lfloor(-d_m)\rfloor$ and $-\lceil(-d_m)\rceil$. In this way, we simply

need to verify that, for these choices of y_m, the equivalence

$$q_{mm}(y_m - d_m)^2 \leq C$$

is satisfied. If it is, we store this value of y_m, otherwise we let $y_m = y_m + 1$. This illustrates the first step in establishing bounds on a specific entry y_i. Adding more terms from the outer sum to this sequence, a pattern emerges.

Let

$$U_i = -d_i + \sum_{j=i+1}^{m} q_{ij}(y_j - d_j)$$

so that we can rewrite $Q(\mathbf{x})$ as

$$Q(\mathbf{x}) = \sum_{i=1}^{m} q_{ii} \left(y_i - d_i + \sum_{j=i+1}^{m} q_{ij}(y_j - d_j) \right)^2 = \sum_{i=1}^{m} q_{ii} \left(y_i + U_i \right)^2$$

In general,

$$q_{kk}(y_k + U_k)^2 \leq C - \sum_{i=k+1}^{m} q_{ii}(y_i + U_i)^2.$$

Let T_k denote the bound on the right-hand side. That is

$$T_k = C - \sum_{i=k+1}^{m} q_{ii}(y_i + U_i)^2,$$

so that $T_m = C$, $T_{m-1} = C - q_{mm}(y_m - d_m)^2$ and

$$T_{m-2} = C - q_{mm}(y_m - d_m)^2 - q_{m-1,m-1} \left(y_{m-1} - d_{m-1} + q_{m-1,m}(y_m - d_m) \right)^2 .$$

We set T_m as C and find each subsequent T_k by subtracting the next term from the outer summand:

$$T_k = C - \sum_{i=k+1}^{m} q_{ii}(y_i + U_i)^2,$$

$$T_k = T_{k+1} - q_{k+1,k+1}(y_{k+1} + U_{k+1})^2.$$

Now, we have an upper bound for each summand.

$$q_{kk}(y_k + U_k)^2 \leq T_k.$$

Using this, we can estimate upper and lower bounds for each y_k in the coordinate vector $\mathbf{y}$. We start by computing the last entries of $\mathbf{y}$ and their bounds first. Assuming that the last several entries of $\mathbf{y}$ have been assigned, upper and lower bounds on y_k can be determined. Now that we have established a bound on a term in the outer sum, we can determine bounds on the specific entry y_k. The following diagram illustrates the scenario. In the usual Fincke-Pohst algorithm, we take the above equation and solve for y_k:

$$\begin{aligned}(y_k + U_k)^2 &\leq T_k/q_{kk} \\ y_k + U_k &\leq \sqrt{T_k/q_{kk}} \\ y_k &\leq \sqrt{T_k/q_{kk}} - U_k.\end{aligned}$$

Similarly, we have a lower bound:

$$y_k \geq -\sqrt{T_k/q_{kk}} - U_k.$$

Since x_k must be an integer, we can restrict our bounds further. Let $t_k = \sqrt{T_k/q_{kk}}$.

$$UB_k = \lfloor t_k - U_k \rfloor$$

$$LB_k = \lceil -t_k - U_k \rceil$$

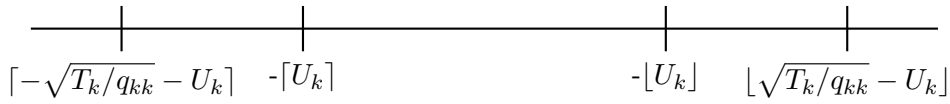
Here UB_k is the upper bound on y_k and LB_k is the lower bound on y_k.

$$LB_k \leq y_k \leq UB_k.$$

2.2.3 Refinements

We note here that computing LB_k and UB_k is highly inefficient as it often requires high precision to accurately compute $\sqrt{T_k/q_{kk}}$. Instead, we adopt the following

bounds, as per Matshke's algorithm. To help justify this process, we refer to the following diagram



As stated above,

$$\lceil -\sqrt{T_k/q_{kk}} - U_k \rceil = LB_k \leq y_k \leq UB_k = \lfloor \sqrt{T_k/q_{kk}} - U_k \rfloor.$$

In our old implementation for non-translated lattices, we set each $y_k = LB_k$ and increased each term until we reached the zero (centre) vector. Here since the centre vector is non-zero, we instead set each $y_k = -\lceil U_k \rceil$ and increase each y_k successively until $y_k > \lfloor \sqrt{T_k/q_{kk}} - U_k \rfloor$. This is equivalent to the above computation and generates only half of the vectors, assuming symmetry. This symmetry can only be applied if the centre vector is defined over $\mathbb{Z}$, otherwise we must compute all vectors. To do (we can also break symmetry and compute all vectors in the $\mathbb{Z}$ case), we also set $y_k = \lceil U_k \rceil - 1$ and successively decrease this term until $y_k < \lceil -\sqrt{T_k/q_{kk}} - U_k \rceil$.

Of course, in this refinement, we want to avoid computing $\sqrt{T_k/q_{kk}}$, and so instead of verifying whether $y_k > \lfloor \sqrt{T_k/q_{kk}} - U_k \rfloor$ or $y_k < \lceil -\sqrt{T_k/q_{kk}} - U_k \rceil$, we compute $q_{kk}(y_k + U_k)^2$ in each case and verify whether

$$q_{kk}(y_k + U_k)^2 \leq C - \sum_{i=k+1}^{m} q_{ii}(y_i + U_i)^2$$

holds. In the first round, if this does not hold and if $y_k < -\lfloor U_k \rfloor$, we continue to iterate $y_k = y_k + 1$, otherwise we simply iterate $y_k = y_k + 1$. Once this equivalence does not hold and $y_k \geq -\lfloor U_k \rfloor$, we stop this loop. We then reset $y_k = \lceil U_k \rceil - 1$ and search in the other direction, by successively subtracting 1 if

$$q_{kk}(y_k + U_k)^2 \leq C - \sum_{i=k+1}^{m} q_{ii}(y_i + U_i)^2$$

holds. We stop searching in this direction only once this equivalence does not hold.

2.2.4 Preliminaries: Elliptic Curves

Let K be a field. An *elliptic curve* E over K is a nonsingular curve of the form

$$E : y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6 \tag{2.1}$$

with $a_i \in K$, having a specified base point, $\mathcal{O} \in E$. An equation of the form (2.1) is called a *Weierstrass equation*. For an elliptic curve E over K, this equation is unique up to a coordinate transformation of the form

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

with $r, s, t, u \in K, u \neq 0$. Writing

$$b_2 = a_1^2 + 4a_2, \quad b_4 = a_1a_3 + 2a_4, \quad b_6 = a_3^2 + 4a_6,$$

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

$$c_4 = b_2^2 - 24b_4, \quad \text{and} \quad c_6 = -b_2^3 + 36b_2b_4 + 9b_2b_4b_6,$$

if char$(K) \neq 2, 3$, we can make several linear changes of variables so that, using these values, our elliptic curve has equation

$$E : y^2 = x^3 - 27c_4x - 54c_6. \tag{2.2}$$

Associated to this curve are the quantities

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6 \quad \text{and} \quad j = c_4^3/\Delta,$$

where Δ is called the *discriminant* of the Weierstrass equation, and the quantity j is called the *j-invariant* of the elliptic curve. The condition of being nonsingular is equivalent to Δ being non-zero. Additionally, one may show that two elliptic curves are isomorphic over $\bar{K}$, the algebraic closure of K, if and only if they both

have the same j-invariant.

When $K = \mathbb{Q}$, we can choose the Weierstrass model (2.1) with the $a_i \in \mathbb{Z}$ and the p-order of Δ minimal for each prime p. Supposing (2.1) is such a global minimal model for an elliptic curve E over $\mathbb{Q}$, reducing the coefficients modulo a prime p, we obtain a (possibly singular) curve over $\mathbb{F}_p$, namely

$$\tilde{E} : y^2 + \tilde{a_1}xy + \tilde{a_3}y = x^3 + \tilde{a_2}x^2 + \tilde{a_4}x + \tilde{a_6}, \tag{2.3}$$

with $\tilde{a_i} \in \mathbb{F}_p$. This "reduced" curve $\tilde{E}/\mathbb{F}_p$ is called the *reduction of E modulo p*. It is nonsingular provided that $\Delta \not\equiv 0 \mod p$, in which case it is an elliptic curve defined over $\mathbb{F}_p$. The curve E is said to have *good reduction* modulo p if $\tilde{E}/\mathbb{F}_p$ is nonsingular, otherwise, we say E has *bad reduction* modulo p.

The bad reduction of E is measured by the *conductor* of E,

$$N = \prod_{p \text{ prime}} p^{f_p},$$

where $f_p \neq 0$ if $p \nmid \Delta$ (so $f_p = 0$ for all but finitely many primes p), while $f_p = 1$ if the singularity is a node, and $f_p \geq 2$ if the singularity is a cusp. The f_p, hence the conductor, are invariant under isogeny. Hence, roughly speaking, the conductor N is the product of primes at which E has bad reduction raised to small powers, while the discriminant Δ is a product of the same primes, but they may sometimes appear to large powers.

2.2.5 Preliminaries: Cubic Forms

Let a, b, c and d be integers, and consider the binary cubic form

$$F(x, y) = ax^3 + bx^2y + cxy^2 + dy^3.$$

Two such forms F_1 and F_2 are called *equivalent* if they are equivalent under the $GL_2(\mathbb{Z})$-action. That is, if there exist integers a_1, a_2, a_3, and a_4 such that

$$F_1(a_1x + a_2y, a_3x + a_4y) = F_2(x, y)$$

for all x, y, where $a_1a_4 - a_2a_3 = \pm 1$. In this case, we write $F_1 \sim F_2$. The *discriminant* D_F of such a form is given by

$$D_F = -27a^2d^2 + b^2c^2 + 18abcd - 4ac^3 - 4b^3d = a^4 \prod_{i<j} (\alpha_i - \alpha_j)^2$$

where α_1, α_2 and α_3 are the roots of the polynomial $F(x, 1)$. We observe that if $F_1 \sim F_2$, then $D_{F_1} = D_{F_2}$.

Associated to F is the Hessian $H_F(x, y)$, given by

$$\begin{aligned} H_F(x, y) &= -\frac{1}{4}\left(\frac{\partial^2 F}{\partial x^2}\frac{\partial^2 F}{\partial y^2} - \left(\frac{\partial^2 F}{\partial x \partial y}\right)^2\right) \\ &= (b^2 - 3ac)x^2 + (bc - 9ad)xy + (c^2 - 3bd)y^2, \end{aligned}$$

and the Jacobian determinant of F and H, a cubic form $G_F(x, y)$ defined via

$$\begin{aligned} G_F(x, y) &= \frac{\partial F}{\partial x}\frac{\partial H}{\partial y} - \frac{\partial F}{\partial y}\frac{\partial H}{\partial x} \\ &= (-27a^2d + 9abc - 2b^3)x^3 + (-3b^2c - 27abd + 18ac^2)x^2y + \\ &\quad + (3bc^2 - 18b^2d + 27acd)xy^2 + (-9bcd + 2c^3 + 27ad^2)y^3. \end{aligned}$$

Appendix A

Supporting Materials

This would be any supporting material not central to the dissertation. For example:

- additional details of methodology and/or data;
- diagrams of specialized equipment developed.;
- copies of questionnaires and survey instruments.