TM Algorithm Background

August 26, 2020

Contents

Chapter 1

To do list

All to do items for the Thue-Mahler code background, Magma code, and subsequent paper that we will hopefully, eventually publish.

1.1 Theory

1. initial precision heuristic (p-adic, real, complex)
2. complex case
3. update non-Archimedean ellipsoid and sieve
4. update Archimedean ellipsoid and sieve
5. heuristic for choosing vectors $\mathbf{l}$ in reduction
6. refined reduction process?
7. Samir has updated the PIRL; should update corresponding writeup
8. probably don't need to use Ficnke-Pohst on translated lattices

1.2 Code

1. back up data onto dropbox (fix error on github repo)
2. Thue equations - regenerate Thue equations as in GenerateSUnitEquations Alpha.m

3. Thue equations from NoSUnitEqNeeded.csv
4. Python code to rerun missing TM equations arising from Magma inernal errors
5. Samir has updated the PIRL; should update corresponding code
6. reducible forms
7. change bound Ω to be $[K : \mathbb{Q}]\Omega$
8. ensure ad-bc in matR is not 0

1.3 Paper

1. probably don't need subsection 3.6.3
2. chapter/section/subsection labels

Chapter 2

Preliminaries

In this chapter, we give some of the primary algorithms needed to solve an arbitrary Thue-Mahler equation. The methods presented here follow somewhat [**?**] and [**?**], with new results and modifications from [**?**].

2.1 p-adic valuations

In this section we give a concise exposition of p-adic valuations. As references for this material we give [**?**] (especially Theorem 3 in Chapter 4, Section 2), [**?**] (especially Lemma 2.1 in Chapter 9), [**?**] (especially Chapter 18), [**?**] (especially Chapter 3, Section 2), and [**?**] (especially Theorem 6.1).

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 K be an arbitrary number field. 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

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

This definition may be extended to all of K by setting $v(0) = 0$. 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 valua-

tions on K are called *Archimedean*.

Every valuation v induces on K the structure of a metric topological space which may or may not be complete. We say that two valuations are *equivalent* if they define the same topology and we call an equivalence class of absolute values a *place* of K. It is an elementary result of topology that every metric space may be embedded in a complete metric space, and this can be done in an essentially unique way. For the field K, the resulting complete metric space may be given a field structure. Equivalently, there exists a field L with a valuation w such that L is complete in the topology induced by w. The field K is contained in L and the valuations v and w coincide in K. Moreover, the completion L of K is unique up to topological isomorphism.

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 may 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

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

Furthermore, 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*. The completion of K under this valuation is denoted by $K_{\mathfrak{p}}$ or K_v and called the $\mathfrak{p}$-*adic field*. Let V be the set of all valuations of an algebraic number field K. Then for every non-zero element $\alpha \in K$ we have

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

In the ring of integers of $\mathbb{Q}$, the prime ideals are generated by the rational primes p, and the resulting topology in the field $\mathbb{Q}$ is called the p-*adic topology*. The completion of $\mathbb{Q}$ under this valuation is denoted by $\mathbb{Q}_p$. 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. Analogous to $\mathrm{ord}_{\mathfrak{p}}$, for any prime p we define the p-adic order of $x \in \mathbb{Q}$ as the largest exponent of p dividing x. Then, the p-adic valuation

v is defined as

$$v(x) = p^{-\operatorname{ord}_p(x)}.$$

If $K_{\mathfrak{p}}$ is a $\mathfrak{p}$-adic field, it is necessarily a finite extension of a certain $\mathbb{Q}_p$.

Consider now $K/\mathbb{Q}$ where $n = [K : \mathbb{Q}]$ and let $g(t)$ denote the minimal polynomial of K over $\mathbb{Q}$. Suppose p is a rational prime and let $g(t) = g_1(t) \cdots g_m(t)$ be the decomposition of $g(t)$ into irreducible polynomials $g_i(t) \in \mathbb{Q}_p[t]$ of degree $n_i = \deg g_i(t)$. The prime ideals in K dividing p are in one-to-one correspondence with $g_1(t), \ldots, g_m(t)$. More precisely, we have in K the following decomposition of $(p)\mathcal{O}_K$

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

with $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ distinct prime ideals and ramification indices $e(\mathfrak{p}_1|p), \ldots, e(\mathfrak{p}_m|p) \in \mathbb{N}$. For $i = 1, \ldots, m$ the inertial degree of $\mathfrak{p}_i$ is denoted by $f(\mathfrak{p}_i|p)$. Then $n_i = e(\mathfrak{p}_i|p)f(\mathfrak{p}_i|p)$ and $K_{\mathfrak{p}_i} \simeq \mathbb{Q}_p(\theta_i)$, where $g(\theta_i) = 0$.

By $\overline{\mathbb{Q}_p}$ we denote the algebraic closure of $\mathbb{Q}_p$. 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 $K \hookrightarrow K_{\mathfrak{p}_i} \simeq \mathbb{Q}_p(\theta_i) \simeq \mathbb{Q}_p(\theta_i^{(j)})$.

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

$$v(x) = |N_{L/\mathbb{Q}_p}(x)|^{1/[L:\mathbb{Q}_p]}.$$

Here, we define the p-adic order of $x \in L$ by

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

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 v of any $x \in \overline{\mathbb{Q}_p}$. Every finite extension of $\mathbb{Q}_p$ is complete with respect to v, but $\overline{\mathbb{Q}_p}$ is not. The completion of $\overline{\mathbb{Q}_p}$ with respect to v is denoted by $\mathbb{C}_p$.

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

$$v(x) = |N_{K_{\mathfrak{p}_i}/\mathbb{Q}_p}(\sigma_{ij}(x))|^{1/n_i}.$$

Here, as before, $n_i = \deg g_i(t) = [K_{\mathfrak{p}_i} : \mathbb{Q}_p]$. Furthermore,

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

and we have

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

Of course, in the special case $x \in \mathbb{Q}_p$, we can write

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

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

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

For $\mathrm{ord}_p(x) \geq k$ we also write $x \equiv 0 \mod p^k$.

2.2 p-adic logarithms

We have seen how to extend p-adic valuations to 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 $\text{ord}_p(z-1) > 0$. Three important properties of the p-adic logarithm are

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

Proofs of the first and last property can be found in [**?**] (pp. 264-265). The second property follows from the first.

We will 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.*

(a) There is a positive integer r such that $\text{ord}_p(z^r - 1) > 0$.

(b) If r is the smallest positive integer having $\text{ord}_p(z^r - 1) > 0$, then r divides $p^f - 1$, and an integer q satisfies $\text{ord}_p(z^q - 1) > 0$ if and only if it is a multiple of r.

(c) If r is a nonzero integer with $\text{ord}_p(z^r - 1) > 0$, and if k is an integer with $p^k(p-1) > e$, then

$$\text{ord}_p(z^{rp^k} - 1) > \frac{1}{p-1}.$$

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 $\text{ord}_p(z^q - 1) > 0$. To see that this definition is independent of q, let r be the smallest positive integer with $\text{ord}_p(z^r - 1) > 0$, 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 $\text{ord}_p(z^q - 1) > 1/(p-1)$ helps to speed up and control the convergence of the series defining $\log_p$ (cf. [**?**] (pp. 28-30) and [**?**] (pp. 263-265)).

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

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

2.3 The 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 absolute values $|\cdot|_v$ and $\|\cdot\|_v$ on K which we now define. If $v|\infty$, then $\|\cdot\|_v$ restricted to $\mathbb{Q}$ is the usual Archimedean absolute value; if $v|p$ for a rational prime p, then $\|\cdot\|_v$ restricted to $\mathbb{Q}$ is the usual p-adic valuation. We then set

$$|\cdot|_v = \|\cdot\|_v^{[K_v:\mathbb{Q}_v]/[K:\mathbb{Q}]}.$$

Let $x \in K$ and let $\log^+(\cdot)$ denote the real-valued function $\max\{\log(\cdot), 0\}$ on $\mathbb{R}_{\geq 0}$. We define the *logarithmic Weil height* $h(x)$ by

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

where the sum is take over all places v of K. If x is an algebraic unit, then $|x|_v = 1$ for all non-Archimedean places v, and therefore $h(x)$ can be taken over the Archimedean places only. In particular, if $x \in \mathbb{Q}$, then with $x = p/q$ for $p, q \in \mathbb{Z}$ with $\gcd(p, q) = 1$, we have $h(x) = \log \max\{|p|, |q|\}$, and if $x \in \mathbb{Z}$ then $h(x) = \log |x|$.

2.4 Elliptic curves

Let K be a field of characteristic $\mathrm{char}(K) \neq 2, 3$. 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*. This equation is unique up to a coordinate transformation of the form

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

with $r, s, t, u \in K, u \neq 0$. Applying several linear changes of variables and writing

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

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

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

E can be written as

$$E : y^2 = x^3 - 27c_4x - 54c_6.$$

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. 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 $K = \mathbb{Q}$, the Weierstrass model (2.1) can be chosen so that Δ has minimal p-adic order for each rational prime p and $a_i \in \mathbb{Z}$. Suppose (2.1) is such a global minimal model for an elliptic curve E over $\mathbb{Q}$. Reducing the coefficients modulo a rational prime p yields a (possibly singular) curve over $\mathbb{F}_p$

$$\tilde{E} : y^2 + \tilde{a_1}xy + \tilde{a_3}y = x^3 + \tilde{a_2}x^2 + \tilde{a_4}x + \tilde{a_6}, \tag{2.2}$$

where $\tilde{a_i} \in \mathbb{F}_p$. 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 reduction type of E at a rational prime p is measured by the *conductor*,

$$N = \prod_p p^{f_p}$$

where the product runs over all primes p and $f_p = 0$ for all but finitely many primes. In particular, $f_p \neq 0$ if p does not divide Δ. Equivalently, E has bad reduction at p if and only if $p \mid N$. Suppose E has bad reduction at p so that $f_p \neq 0$. The reduction type of E at p is said to be *multiplicative* (E has a node over $\mathbb{F}_p$) or *additive* (E has a cusp over $\mathbb{R}_p$) depending on whether $f_p = 1$ or $f_p \geq 2$, respectively. The f_p, hence the conductor, are invariant under isogeny.

2.5 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_F, a cubic form $G_F(x, y)$ defined by

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

2.6 Lattices

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

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

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

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

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 $a \in \mathbb{Z}$.

Given a basis, we can define a specific bilinear form on our lattice Γ as part of its structure. This form describes a kind of distance between elements $\mathbf{u}$ and $\mathbf{v}$ and we say the lattice is *defined* by Φ. Associated to this bilinear form is a quadratic form $Q : \Gamma \to \mathbb{Z}$ defined by $Q(\mathbf{v}) = \Phi(\mathbf{v}, \mathbf{v})$. A lattice is called *positive definite* if its quadratic form is positive definite.

The bilinear forms (and their associated quadratic forms) that we will be using come from the usual inner product on vectors in $\mathbb{R}^n$. This is simply the dot product $\Phi(\mathbf{u}, \mathbf{v}) = \mathbf{u} \cdot \mathbf{v}$ for embedded vectors, $\mathbf{u}, \mathbf{v}$. For the coordinate vectors $\mathbf{x}, \mathbf{y}$ associated to these vectors,

this translates to multiplication with the basis matrix. Precisely, if $\mathbf{u} = B\mathbf{x}$ and $\mathbf{v} = B\mathbf{y}$, we have $\Phi(\mathbf{u}, \mathbf{v}) = \mathbf{x}^T B^T B\mathbf{y}$.

If $\mathbf{v} = B\mathbf{x}$, the *norm* of the vector $\mathbf{v} \in \Gamma$ is defined to be the inner product $\Phi(\mathbf{v}, \mathbf{v})$. In terms of the corresponding coordinate vector $\mathbf{x}$, this is

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

Equivalently, we write $\mathbf{x}^T A\mathbf{x}$ where $A = B^T B$ is the Gram matrix of Γ with basis B and bilinear form Φ. The entries of the matrix A are $a_{ij} = \Phi(\mathbf{b}_i, \mathbf{b}_j)$.

Two basis matrices B_1 and B_2 define the same lattice Γ if and only if there is a unimodular matrix U such that $B_1 U = B_2$. The bilinear form on Γ can be written with respect to either embedded or coordinate vectors. Using another basis to express the lattice elements is possible, and sometimes preferable. However, 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.

2.7 Continued fractions

In this section, we give a brief introduction to continued fractions. The background developed here will be important to us later when we will need to generate rational lattices. In particular, we will have to generate rational approximations to $\log(p)$ for various primes p, and the theory discussed here will enable us to generate such an approximation with relatively high precision. This is all taken from http://pi.math.cornell.edu/ gautam/ContinuedFractions.pdf (basics); the more advanced concepts are from https://www.math.ru.nl/ bosma/Students/CF.pdf and I can probably edit this section to sound nicer and/or remove anything extra.

Definition 2.7.1. A *simple continued fraction* is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}}$$

where a_i are non-negative integers, for $i > 0$, and $a_0 \in \mathbb{Z}$.

For ease of notation, we let

$$[a_0, a_1, a_2, a_3, \dots]$$

denote the continued fraction above.

Definition 2.7.2. For $0 \leq m \leq n$, we call $[a_0, \dots, a_m]$ the m^{th} *convergent* to $[a_0, \dots, a_n]$.

Of course, every rational number has a simple continued fraction expansion which is finite, and every finite simple continued fraction expansion is a rational number. Moreover, if a rational number x is representable by a simple continued fraction with an odd (respectively, even) number of convergents, it is also representable by one with an even (respectively, odd) number. In fact, there are exactly two ways to represent a rational number as a finite simple continued fraction: if $a_n \geq 2$,

$$[a_0, a_1, \dots, a_n] = [a_0, a_1, \dots, a_n - 1, 1].$$

If $a_n = 1$,

$$[a_0, a_1, \dots, a_{n-1}, 1] = [a_0, a_1, \dots, a_{n-1} + 1, 1].$$

If p_n and q_n are defined by

$$p_0 = a_0, \quad p_1 = a_1 a_0 + 1, \quad p_n = a_n p_{n-1} + p_{n-2} \quad \text{for } n \geq 2$$

$$q_0 = 1, \quad q_1 = a_1, \quad q_n = a_n q_{n-1} + q_{n-2} \quad \text{for } n \geq 2,$$

then

$$[a_0, \dots, a_n] = \frac{p_n}{q_n}.$$

It follows that the n^{th} convergent is

$$\frac{p_n}{q_n} = \frac{a_n p_{n-1} + p_{n-2}}{a_n q_{n-1} + q_{n-2}}.$$

To compute the continued fraction expansion of a number x, rational or not, one uses the following algorithm, based on the Euclidean algorithm: Let $x \in \mathbb{R}$ and set $x_0 = x$.

1. Set a_m to be the integral part of x_m.
2. Set $\psi_m := x_m - a_m$.
3. If $\psi_m \neq 0$, set $\frac{q}{\psi_m}$ as x_{m+1} and return to step 1 to compute a_{m+1}.

4. If $\psi = 0$, terminate this algorithm.

Using this algorithm, we may see rounding errors in computing $\frac{1}{\psi_m}$. Once these intermediate fractions become close to 0, we stop the calculations to obtain a good approximation of x.

Let x be given by $[a_0, \ldots, a_n]$ and let x_m denote the m^{th} convergent p_m/q_m. The following theorems hold.

Theorem 2.7.3. *The even convergents x_{2m} increase strictly with m, while the odd convergents x_{2m+1} decrease strictly.*

We thus have

$$x_0 < x_2 < x_4 < x_6 < \ldots \quad \text{and} \quad x_1 > x_3 > x_5 > x_7 > \ldots.$$

Theorem 2.7.4. *Every odd convergent is greater than any even convergent.*

Theorem 2.7.5. *The value of the continued fraction is greater than and of its even convergents and less than any of its odd convergents (except where it is equal to the last convergent).*

Consider now the best approximation to a given number with small denominators.

Definition 2.7.6. The rational number p/q is the *best approximation* to a real number x if $|p/q - x| \leq |P/Q - x|$, where P/Q is any other rational number such that $Q \leq q$.

Theorem 2.7.7. *The convergents to a simple continued fraction are in their lowest terms.*

Theorem 2.7.8. *The denominators of the convergents satisfy the following inequalities*

$$q_n \geq n, \text{ with strict inequality when } n > 3.$$

Theorem 2.7.9. *For any number x with convergents $\frac{p_m}{q_m}$,*

$$\frac{1}{2q_m q_{m+1}} < \frac{1}{q_m(q_m + q_{m+1})} < \left| x - \frac{p_m}{q_m} \right| < \frac{1}{q_m q_{m+1}} < \frac{1}{q_m^2}$$

for $m \geq 1$.

Theorem 2.7.10. *$\frac{p_m}{q_m}$ is the best approximation to x with denominator $\leq q_m$.*

All convergents to x are best approximations to x, but these are not all the best approximations! Consider the following example:

Example 2.7.11.

$$x = 0.18421052631 = \frac{7}{38} = [0, 5, 2, 3], \qquad \text{with convergents } \frac{0}{1}, \frac{1}{5}, \frac{2}{11}, \frac{7}{38}.$$

If we look at all fractions with denominators ≤ 28, then $\frac{5}{27}$ is the best approximation to x, with an error of 0.00097465886, but it is not one of the convergents. Indeed,

$$\frac{5}{27} = 0.18518518518 = [0, 5, 2, 2].$$

This leads us to the following two questions:

1. What are all the best approximations?
2. How are they related to the convergents?

Definition 2.7.12. A *semi-convergent* or *secondary convergent* to x is a number of the form

$$\frac{p_k + rp_{k+1}}{q_k + rq_{k+1}},$$

where $\frac{p_k}{q_k}$ and $\frac{p_{k+1}}{q_{k+1}}$ are two consecutive convergents to $x = [a_0, a_1, a_2, \dots]$ and r is an integer such that $0 \leq r \leq a_k$.

Note in particular that the convergents to x are also semi-convergents.

Theorem 2.7.13. *If x is any real number and a/b is a not a semi-convergent to x, then a/b is not the best approximation to x with denominator $\leq b$.*

The previous theorem tells us that the only candidates for the best approximations are the semi-convergents. In fact, there is a precise statement which says which semi-convergents are the best approximations which we will not state here but we actually should include this becasue it's important. Roughtly speaking, about half the semi-convergents between p_k/q_k and p_{k+2}/q_{k+2} which are closest to p_{k+2}/q_{k+2} are the best approximations to x.

In our example above, $5/27$ is a semi-convergent to $7/38$ but not a convergent. All the best approximations to $7/38 = [0, 5, 2, 3]$ are $[0], [0, 3], [0, 4], [0, 5], [0, 5, 2], [0, 5, 2, 2], [0, 5, 2, 3]$, where the convergents are $[0], [0, 5], [0, 5, 2], [0, 5, 2, 3]$.

When considering best approximations, we consider the distance between the fraction p/q and x as a measure of how well it approximates x. However, if the denominator increases, we should expect better approximations to have smaller distance from x. To take this into account, one could consider instead the distance $|qx - p|$.

Definition 2.7.14. A fraction p/q is a *best approximation of the second kind* to a real number x if for every fraction a/b with $b \leq q$, we have $|qx - p| < |bx - a|$.

Theorem 2.7.15. *The convergents to a real number x are precisely all the best approximations of the second kind to x.*

Consider now the coordinate plane where all points with integer coordinates are marked (ie. a lattice on the plane). Fix some α and draw the line $y = \alpha x$. If α is rational, then it would intersect the lattice in inifintely many points. If α is irrational, then it will not intersect the lattice at any point other than the origin.

Theorem 2.7.16. *The closest points in the lattice to the line $y = \alpha x$ are in one-to-one correspondence with the convergents to the continued fraction of α. Let these points be labelled as $A_m = (q_m, p_m)$. The point is above the line if m is odd and below otherwise.*

Given just the points on the lattice closest to the line, it is also possible to recover the continued fraction.

Theorem 2.7.17. *Let $[a_0, \ldots, a_n]$ be the continued fraction for α and let A_m be the marked points in the lattice as before. Then a_m is the integral distance between A_m and A_{m+2}.*

Here the integral distance between two points refers to the number of points on the line segment joining two points minus 1. These two theorems give us a geometric way of going back and forth between continued fractions and approximations.

One of the main interests of continued fractions is in its application to the representation of irrationals. Suppose $a_0, a_1, a_2, \ldots$ is a sequence of positive integers such that

$$x_n = [a_0, a_1, \ldots, a_n]$$

is a simple continued fraction of a rational number x_n for every n. If these $x_n \to x$ when $n \to \infty$, then we say that $[a_0, a_1, \ldots]$ is x and we write

$$x = [a_0, a_1, \ldots].$$

Theorem 2.7.18. *If $a_0, a_1, a_2, \ldots$ is a sequence of positive integers, then $x = [a_0, a_1, \ldots] \to x$ as $n \to \infty$.*

In particular, this means that we can always talk about $[a_0, a_1, a_2, \ldots]$ as it is a well-defined real number.

Theorem 2.7.19. *An infinite simple continued fraction is less than any of its odd convergents and greater than any of its even convergents.*

Theorem 2.7.20. *Every irrational number can be expressed in just one way as an infinite simple continued fraction.*

2.7.1 General continued fractions

Definition 2.7.21. The *general continued fraction* is a simple continued fraction in which the numerators and denominators can assume arbitrary complex values. A generalized continued fraction is an expression of the form

$$x = a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \dots}}}.$$

Working with general continued fractions instead of simple continued fractions does not provide any new information, but in some cases, makes it easier to explicitly compute the former rather than the latter.

If $\{c_1, c_2, c_3, \dots\}$ is any infinite sequence of non-zero complex numbers, the following equivalence holds

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \cfrac{b_4}{a_4 + \dots}}}} = a_0 + \cfrac{c_1 b_1}{c_1 a_1 + \cfrac{c_1 c_2 b_2}{c_2 a_2 + \cfrac{c_2 c_3 b_3}{c_3 a_3 + \cfrac{c_3 c_4 b_4}{c_4 a_4 + \dots}}}}.$$

That is, the successive convergents of the continued fraction on the left are exactly the same as the convergents of the fraction on the right. The equivalence transformation is perfectly general, however, if none of the b_i are zero, one can choose the sequence $\{c_i\}$ so

as to obtain a simple continued fraction

$$a_0 + \cfrac{b_1}{a_1 + \cfrac{b_2}{a_2 + \cfrac{b_3}{a_3 + \cfrac{b_4}{a_4 + \ldots}}}} = a_0 + \cfrac{1}{c_1 a_1 + \cfrac{1}{c_2 a_2 + \cfrac{1}{c_3 a_3 + \cfrac{1}{c_4 a_4 + \ldots}}}}$$

via

$$c_1 = \frac{1}{b_1}, \quad c_{n+1} = \frac{1}{b_{n+1} c_n}.$$

For our purposes, we will need the continued fration expansion of $\log(p)$. Indeed, for

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

$$= \cfrac{2z}{1 - \cfrac{z^2}{3 - \cfrac{4z^2}{5 - \cfrac{9z^2}{7 - \cfrac{16z^2}{9 - \cfrac{25z^2}{11 - \cfrac{36z^2}{13 - \ldots}}}}}}}$$

for all $z \in \mathbb{C}/((-\infty,-1] \cup [1,\infty))$. Writing $z = \frac{p-1}{p+1} < 1$, we obtain

$$\begin{aligned}\log(p) &= \log\left(\frac{1+z}{1-z}\right) \\ &= \cfrac{2z}{1-\cfrac{z^2}{3-\cfrac{4z^2}{5-\cfrac{9z^2}{7-\cfrac{16z^2}{9-\cfrac{25z^2}{11-\cfrac{36z^2}{13-\ldots}}}}}}}\end{aligned}$$

The above does not work in Magma. Using this approach does not give us fast enough convergence, and in fact, the bound in Theorem 2.7.9 does not even hold! It's far better to compute the convergents as

```
Convergents(ContinuedFraction(RealField(prec)!Log(RealField(prec)!p)));
```

This generates a 2×2 matrix with integer coefficients

$$\begin{bmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{bmatrix}.$$

The quotients p_{n-1}/q_{n-1} and p_n/q_n form the last two convergents for $\log(p)$ as provided by

```
ContinuedFraction(RealField(prec)!Log(RealField(prec)!p));
```

Meanwhile, the above function takes $\log(p)$ as input, with given precision, and returns a sequence of integers that form the partial quotient for the (regular/simple) continued fraction expansion for $\log(p)$. The length of the sequence is determined in such a way that the last significant partial quotient is obtained (determined by the precision with which $\log(p)$ is known). Note finally that if n is even, the resulting continued fraction p_n/q_n will have

$$\frac{p_n}{q_n} < \log(p).$$

Taking n odd will yield the desired result, $\log(p) < \frac{p_n}{q_n}$, however.

Chapter 3

Algorithms for Thue-Mahler Equations

In this chapter, we give some of the primary algorithms needed to solve an arbitrary Thue-Mahler equation. The methods presented here follow somewhat [**?**] and [**?**], with new results and modifications from [**?**].

3.1 First steps

Fix a nonzero integer c and let $S = \{p_1, \ldots, p_v\}$ be a set of rational primes. Let

$$F(X, Y) = c_0 X^n + c_1 X^{n-1} Y + \cdots + c_{n-1} X Y^{n-1} + c_n Y^n$$

be an irreducible binary form over $\mathbb{Z}$ of degree $n \geq 3$. We want to solve the Thue–Mahler equation

$$F(X, Y) = c p_1^{Z_1} \cdots p_v^{Z_v} \tag{3.1}$$

for unknowns $X, Y, Z_1, \ldots, Z_v$ with $\gcd(X, Y) = 1$ and $Z_i \geq 0$ for $i = 1, \ldots, v$. To do so, we first reduce (3.1) to the special case where $c_0 = 1$ and $\gcd(c, p_i) = 1$ for $i = 1, \ldots, v$, loosely following [**?**].

As F is irreducible by assumption, at least one of the coefficients c_0 and c_n is nonzero. Hence, we may transform the given Thue–Mahler equation to one with $c_0 \neq 0$ by interchanging X and Y and by renaming the coefficients c_i appropriately. In particular, solving

(3.1) is equivalent to solving

$$c_0'\overline{X}^n + c_1'\overline{X}^{n-1}\overline{Y} + \cdots + c_{n-1}'\overline{XY}^{n-1} + c_n'\overline{Y}^n = cp_1^{Z_1}\cdots p_v^{Z_v},$$

where $c_i' = c_{n-1}$ for $i = 0, \ldots, n$, $\overline{X} = Y$, and $\overline{Y} = X$.

Let $\mathcal{D}$ be the set of all positive integers m dividing c_0 such that $\mathrm{ord}_p(m) \leq \mathrm{ord}_p(c)$ for each rational prime $p \notin S$. Equivalently, $\mathcal{D}$ is precisely the set of all possible integers d such that $d = \gcd(c_0, Y)$. To see this, let $q_1, \ldots, q_w$ denote the distinct prime divisors of a not contained in S. Then

$$c = \prod_{i=1}^{w} q_i^{b_i} \cdot \prod_{i=1}^{v} p_i^{\mathrm{ord}_{p_i}(c)}$$

for some integers $b_i > 0$. If $(X, Y, Z_1, \ldots, Z_v)$ is a solution of the Thue-Mahler equation in question, it follows that

$$F(X,Y) = cp_1^{Z_1}\ldots p_v^{Z_v} = \prod_{i=1}^{w} q_i^{b_i} \cdot \prod_{i=1}^{v} p_i^{\mathrm{ord}_{p_i}(c)+Z_i}.$$

Suppose $\gcd(c_0, Y) = d$. Since d divides $F(X,Y)$, it necessarily divides

$$\prod_{i=1}^{w} q_i^{b_i} \cdot \prod_{i=1}^{v} p_i^{\mathrm{ord}_{p_i}(c)+Z_i}.$$

In particular,

$$d = \prod_{i=1}^{w} q_i^{s_i} \cdot \prod_{i=1}^{v} p_i^{t_i}$$

for some non-negative integers $s_1, \ldots, s_w, t_1, \ldots, t_v$ such that

$$s_i \leq \min\{\mathrm{ord}_{q_i}(c), \mathrm{ord}_{q_i}(c_0)\} \quad \text{and} \quad t_i \leq \min\{\mathrm{ord}_{p_i}(c) + Z_i, \mathrm{ord}_{p_i}(c_0)\}.$$

From here, it is easy to see that $\mathrm{ord}_p(d) \leq \mathrm{ord}_p(c)$ for each rational prime $p \notin S$ so that $d \in \mathcal{D}$.

Conversely, suppose $d \in \mathcal{D}$ so that $\mathrm{ord}_p(d) \leq \mathrm{ord}_p(c)$ for all $p \notin S$. That is, the right-hand side of

$$\mathrm{ord}_p(d) \leq \mathrm{ord}_p(c) = \mathrm{ord}_p\left(\prod_{i=1}^{w} q_i^{b_i} \cdot \prod_{i=1}^{v} p_i^{\mathrm{ord}_{p_i}(c)}\right)$$

is non-trivial only at the primes $\{q_1, \ldots, q_w\}$. In particular,

$$d = \prod_{i=1}^{w} q_i^{s_i} \cdot \prod_{i=1}^{v} p_i^{t_i}$$

for non-negative integers $s_1, \ldots, s_w, t_1, \ldots, t_v$ such that

$$s_i \leq \min\{\mathrm{ord}_{q_i}(c), \mathrm{ord}_{q_i}(c_0)\} \quad \text{and} \quad t_i \leq \mathrm{ord}_{p_i}(c_0).$$

It follows that $d = \gcd(c_0, Y)$ for some solution $(X, Y, Z_1, \ldots, Z_v)$ of equation (3.1).

For any $d \in \mathcal{D}$, we define the rational numbers

$$u_d = c_0^{n-1}/d^n \quad \text{and} \quad c_d = \mathrm{sgn}(u_d c) \prod_{p \notin S} p^{\mathrm{ord}_p(u_d c)}.$$

On using that $d \in \mathcal{D}$, we see that the rational number c_d is in fact an integer coprime to S.

Suppose $(X, Y, Z_1, \ldots, Z_v)$ is a solution of (3.1) with $\gcd(X, Y) = 1$ and $d = \gcd(c_0, Y)$. Define the homogeneous polynomial $f(x, y) \in \mathbb{Z}[x, y]$ of degree n by

$$f(x, y) = x^n + C_1 x^{n-1} y + \cdots + C_{n-1} x y^{n-1} + C_n y^n,$$

where

$$x = \tfrac{c_0 X}{d}, \quad y = \tfrac{Y}{d} \quad \text{and} \quad C_i = c_i c_0^{i-1} \quad \text{for } i = 1, \ldots, n.$$

Since $\gcd(X, Y) = 1$, the numbers x and y are also coprime integers by definition of d. We observe that

$$f(x, y) = u_d F(X, Y) = u_d c \prod_{i=1}^{v} p_i^{Z_i} = c_d \prod_{p \in S} p^{Z_i + \mathrm{ord}_p(u_d c)}.$$

Setting $z_i = Z_i + \mathrm{ord}_p(u_d c)$ for all $i \in \{1, \ldots, v\}$, we obtain

$$f(x, y) = x^n + C_1 x^{n-1} y + \cdots + C_{n-1} x y^{n-1} + C_n y^n = c_d p_1^{z_1} \cdots p_v^{z_v}, \tag{3.2}$$

where $\gcd(x, y) = 1$ and $\gcd(c_d, p_i) = 1$ for all $i = 1, \ldots, v$.

Since there are only finitely many choices for $d = \gcd(c_0, Y)$, there are only finitely many choices for $\{c_d, u_d, d\}$. Then, solving (3.1) is equivalent to solving the finitely many Thue-

Mahler equations (3.2) for each choice of $\{c_d, u_d, d\}$. For each such choice, the solution $\{x, y, z_1, \ldots, z_v\}$ is related to $\{X, Y, Z_1, \ldots, Z_v\}$ via

$$X = \frac{dx}{c_0}, \quad Y = dy \quad \text{and} \quad Z_i = z_i - \text{ord}_p(u_d c).$$

Lastly, we observe that the polynomial $f(x, y)$ of (3.2) remains the same for any choice of $\{c_d, u_d, d\}$. Thus, to solve the family of equations (3.2), we need only to enumerate over every possible c_d. Now, if $\mathcal{C}$ denotes the set of all $\{c_d, u_d, d\}$ and $d_1, d_2 \in \mathcal{D}$, we may have $\{c_{d_1}, u_{d_1}, d_1\}, \{c_{d_2}, u_{d_2}, d_2\} \in \mathcal{C}$ where $c_{d_1} = c_{d_2}$. That is, d_1, d_2 may yield the same value of c_d, reiterating that we need only solve (3.2) for each distinct c_d.

3.2 The relevant algebraic number field

For the remainder of this chapter, we consider the Thue-Mahler equation

$$f(x, y) = x^n + C_1 x^{n-1} y + \cdots + C_{n-1} x y^{n-1} + C_n y^n = c p_1^{z_1} \cdots p_v^{z_v} \tag{3.3}$$

where $\gcd(x, y) = 1$ and $\gcd(c, p_i) = 1$ for $i = 1, \ldots, p_v$.

Following [**?**], put

$$g(t) = f(t, 1) = t^n + C_1 t^{n-1} + \cdots + C_{n-1} t + C_n$$

and note that $g(t)$ is irreducible in $\mathbb{Z}[t]$. Let $K = \mathbb{Q}(\theta)$ with $g(\theta) = 0$. Now (3.3) is equivalent to the norm equation

$$N_{K/\mathbb{Q}}(x - y\theta) = c p_1^{z_1} \ldots p_v^{z_v}. \tag{3.4}$$

Let p_i be any rational prime and let

$$(p_i)\mathcal{O}_K = \prod_{j=1}^{m_i} \mathfrak{p}_{ij}^{e(\mathfrak{p}_{ij}|p_i)}$$

be the factorization of p_i into prime ideals in the ring of integers $\mathcal{O}_K$ of K. Let $f(\mathfrak{p}_{ij}|p_i)$ be the inertial degree of $\mathfrak{p}_{ij}$ over p_i. Since $N(\mathfrak{p}_{ij}) = p_i^{f_{ij}}$, (3.4) leads to finitely many ideal

equations of the form

$$(x - y\theta)\mathcal{O}_K = \mathfrak{a}\prod_{j=1}^{m_1}\mathfrak{p}_{1j}^{z_{1j}}\cdots\prod_{j=1}^{m_v}\mathfrak{p}_{vj}^{z_{vj}} \tag{3.5}$$

where $\mathfrak{a}$ is an ideal of norm $|c|$ and the z_{ij} are unknown integers related to z_i by

$$\sum_{j=1}^{m_i} f(\mathfrak{p}_{ij}|p_i)z_{ij} = z_i$$

for $i \in \{1, \ldots, v\}$.

Our first task is to cut down the number of variables appearing in (3.5). We will do this by showing that only a few prime ideals can divide $(x - y\theta)\mathcal{O}_K$ to a large power.

3.3 The prime ideal removing lemma

In this section, we establish some key results that will allow us to cut down the number of prime ideals that can appear to a large power in the factorization of $(x - y\theta)\mathcal{O}_K$. It is of particular importance to note that we do not appeal to the Prime Ideal Removing Lemma of Tzanakis and de Weger ([**?**]) here and instead apply the following results of [**?**].

Let $p \in \{p_1, \ldots, p_v\}$. We will produce the following two finite lists L_p and M_p. The list L_p will consist of certain ideals $\mathfrak{b}$ of $\mathcal{O}_K$ supported at the prime ideals above p. The list M_p will consist of certain pairs $(\mathfrak{b}, \mathfrak{p})$ where $\mathfrak{b}$ is supported at the prime ideals above p and $\mathfrak{p}$ is a prime ideal lying over p satisfying $e(\mathfrak{p}|p) = f(\mathfrak{p}|p) = 1$. These lists will satisfy the following property: if $(x, y, z_1, \ldots, z_v)$ is a solution to the Thue-Mahler equation (3.3) then

(i) either there is some $\mathfrak{b} \in L_p$ such that

$$\mathfrak{b} \mid (x - y\theta)\mathcal{O}_K, \qquad (x - y\theta)\mathcal{O}_K/\mathfrak{b} \text{ is coprime to } (p)\mathcal{O}_K; \tag{3.6}$$

(ii) or there is a pair $(\mathfrak{b}, \mathfrak{p}) \in M_p$ and a non-negative integer v_p such that

$$(\mathfrak{b}\mathfrak{p}^{v_p}) \mid (x - y\theta)\mathcal{O}_K, \qquad (x - y\theta)\mathcal{O}_K/(\mathfrak{b}\mathfrak{p}^{v_p}) \text{ is coprime to } (p)\mathcal{O}_K. \tag{3.7}$$

To generate the lists M_p, L_p we consider two affine patches, $p \nmid y$ and $p \mid y$. We begin with

the following lemmata.

Lemma 3.3.1. *Let $(x, y, z_1, \ldots, z_v)$ be a solution of (3.3) with $p \nmid y$, let t be a positive integer, and suppose $x/y \equiv u \pmod{p^t}$, where $u \in \{0, 1, 2, \ldots, p^t - 1\}$. If $\mathfrak{q}$ is a prime ideal of $\mathcal{O}_K$ lying over p, then*

$$\mathrm{ord}_{\mathfrak{q}}(x - y\theta) \geq \min\{\mathrm{ord}_{\mathfrak{q}}(u - \theta), t \cdot e(\mathfrak{q}|p)\}.$$

Moreover, if $\mathrm{ord}_{\mathfrak{q}}(u - \theta) < t \cdot e(\mathfrak{q}|p)$, then

$$\mathrm{ord}_{\mathfrak{q}}(x - y\theta) = \mathrm{ord}_{\mathfrak{q}}(u - \theta).$$

Lemma 3.3.2. *Let $(x, y, z_1, \ldots, z_v)$ be a solution of (3.3) with $p \mid y$ (and thus $p \nmid x$), let t be a positive integer, and suppose $y/x \equiv u \pmod{p^t}$, where $u \in \{0, 1, 2, \ldots, p^t - 1\}$. If $\mathfrak{q}$ is a prime ideal of $\mathcal{O}_K$ lying over p, then*

$$\mathrm{ord}_{\mathfrak{q}}(x - y\theta) \geq \min\{\mathrm{ord}_{\mathfrak{q}}(1 - \theta u), t \cdot e(\mathfrak{q}|p)\}.$$

Moreover, if $\mathrm{ord}_{\mathfrak{q}}(1 - \theta u) < t \cdot e(\mathfrak{q}|p)$, then

$$\mathrm{ord}_{\mathfrak{q}}(x - y\theta) = \mathrm{ord}_{\mathfrak{q}}(1 - \theta u).$$

Proof of Lemmas 3.3.1 and 3.3.2. Suppose $p \nmid y$. Thus $\mathrm{ord}_{\mathfrak{q}}(y) = 0$ and hence

$$\mathrm{ord}_{\mathfrak{q}}(x - y\theta) = \mathrm{ord}_{\mathfrak{q}}(x/y - \theta).$$

Since $x/y - \theta = u - \theta + x/y - u$, we have

$$\begin{aligned} \mathrm{ord}_{\mathfrak{q}}(x/y - \theta) &= \mathrm{ord}_{\mathfrak{q}}(u - \theta + x/y - u) \\ &\geq \min\{\mathrm{ord}_{\mathfrak{q}}(u - \theta), \mathrm{ord}_{\mathfrak{q}}(x/y - u)\}. \end{aligned}$$

By assumption,

$$\mathrm{ord}_{\mathfrak{q}}(x/y - u) \geq \mathrm{ord}_{\mathfrak{q}}(p^t) = t \cdot e(\mathfrak{q}|p),$$

completing the proof of Lemma 3.3.1. The proof of Lemma 3.3.2 is similar. □

The following algorithm computes the lists L_p and M_p that come from the first patch $p \nmid y$. We denote these respectively by $\mathcal{L}_p$ and $\mathcal{M}_p$.

Lemma 3.3.3. *Algorithm 1 terminates.*

To compute $\mathcal{L}_p$ and $\mathcal{M}_p$:

Step (1) Let

$$\mathcal{L}_p \leftarrow \emptyset, \qquad \mathcal{M}_p \leftarrow \emptyset,$$
$$t \leftarrow 1, \quad \mathcal{U} \leftarrow \{w : w \in \{0, 1, \ldots, p-1\}\}.$$

Step (2) Let

$$\mathcal{U}' \leftarrow \emptyset.$$

Loop through the elements $u \in \mathcal{U}$. Let

$$\mathcal{P}_u = \{\mathfrak{q} \text{ lying above } p \ : \ \mathrm{ord}_{\mathfrak{q}}(u-\theta) \geq t \cdot e(\mathfrak{q}|p)\}$$

and

$$\mathfrak{b}_u = \prod_{\mathfrak{q}|p} \mathfrak{q}^{\min\{\mathrm{ord}_{\mathfrak{q}}(u-\theta), t\cdot e(\mathfrak{q}|p)\}} = (u-\theta)\mathcal{O}_K + p^t\mathcal{O}_K.$$

(i) If $\mathcal{P}_u = \emptyset$ then

$$\mathcal{L}_p \leftarrow \mathcal{L}_p \cup \{\mathfrak{b}_u\}.$$

(ii) Else if $\mathcal{P}_u = \{\mathfrak{p}\}$ with $e(\mathfrak{p}|p) = f(\mathfrak{p}|p) = 1$ and there is at least one $\mathbb{Z}_p$-root α of $g(t)$ satisfying $\alpha \equiv u \pmod{p^t}$, then

$$\mathcal{M}_p \leftarrow \mathcal{M}_p \cup \{(\mathfrak{b}_u, \mathfrak{p})\}.$$

(iii) Else

$$\mathcal{U}' \leftarrow \mathcal{U} \cup \{u + p^t w : w \in \{0, \ldots, p-1\}\}.$$

Step (3) If $\mathcal{U}' \neq \emptyset$ then let

$$t \leftarrow t+1, \qquad \mathcal{U} \leftarrow \mathcal{U}',$$

and return to Step (2). Else output $\mathcal{L}_p$, $\mathcal{M}_p$.

Proof. Suppose otherwise. Write $t_0 = 1$ and $t_i = t_0 + i$ for $i = 1, 2, 3, \ldots$. Then there is an infinite sequence of congruence classes $u_i \mod p^{t_i}$ such that $u_{i+1} \equiv u_i \mod p^{t_i}$, and such that the u_i fail the hypotheses of both (i) and (ii). This means that $\mathcal{P}_{u_i}$ is non-empty for every $i \in \mathbb{N}_{>0}$. By the pigeon-hole principle, some prime ideal $\mathfrak{p}$ of $\mathcal{O}_K$ appears in infinitely many of the $\mathcal{P}_{u_i}$. Thus $\mathrm{ord}_{\mathfrak{p}}(u_i - \theta) \geq t_i \cdot e(\mathfrak{p}|p)$ infinitely often. However, the sequence $\{u_i\}_{i=1}^{\infty}$ converges to some $\alpha \in \mathbb{Z}_p$ so that $\alpha = \theta$ in $K_{\mathfrak{p}}$. This forces $e(\mathfrak{p}|p) = f(\mathfrak{p}|p) = 1$ and α to be a $\mathbb{Z}_p$-root of $g(t)$. In this case, $\mathfrak{p}$ corresponds to the factor $(t - \alpha)$ in the p-adic factorisation of $g(t)$. There can be at most one such $\mathfrak{p}$, forcing $\mathcal{P}_{u_i} = \{\mathfrak{p}\}$ for all i. In particular, the hypothesis of (ii) are satisfied and we reach a contradiction. □

Lemma 3.3.4. *Let $p \in \{p_1, \ldots, p_v\}$ and let $\mathcal{L}_p$, $\mathcal{M}_p$ be as given by Algorithm 1. Let $(x, y, z_1, \ldots, z_v)$ be a solution to* (3.3)*. Then*

- *either there is some $\mathfrak{b} \in \mathcal{L}_p$ such that* (3.6) *is satisfied;*
- *or there is some $(\mathfrak{b}, \mathfrak{p}) \in \mathcal{M}_p$ with $e(\mathfrak{p}|p) = f(\mathfrak{p}|p) = 1$ and integer $v_p \geq 0$ such that* (3.7) *is satisfied.*

Proof. Let

$$t_0 = 1 \quad \text{and} \quad \mathcal{U}_0 = \{w \ : \ w \in \{0, 1, \ldots, p-1\}\}$$

be the initial values for t and $\mathcal{U}$ in the algorithm. Then $x/y \equiv u_0 \pmod{p^{t_0}}$ for some $u_0 \in \mathcal{U}_0$. Write $\mathcal{U}_i$ for the value of $\mathcal{U}$ after i iterations of the algorithm and let $t_i = t_0 + i$. As the algorithm terminates, $\mathcal{U}_i = \emptyset$ for some sufficiently large i. Hence there is some i such that $x/y \equiv u_i \mod p^{t_i}$ where $u_i \in \mathcal{U}_i$, but there is no element in $\mathcal{U}_{i+1}$ congruent to x/y modulo $p^{t_{i+1}}$. In other words, u_i must satisfy the hypotheses of either step (i) or (ii) of algorithm 1. Write $u = u_i$ and $t = t_i$ for $x/y \equiv u \mod p^t$ and consider the ideal $\mathfrak{b}_u$ generated in this step. By Lemma 3.3.1, $\mathfrak{b}_u$ divides $(x - y\theta)\mathcal{O}_K$. Furthermore, by definition of $\mathcal{P}_u$, if $\mathfrak{q}$ is a prime ideal of $\mathcal{O}_K$ not contained in $\mathcal{P}_u$, then $(x - y\theta)\mathcal{O}_K/\mathfrak{b}_u$ is not divisible by $\mathfrak{q}$.

Suppose first that the hypothesis of (i) is satisfied: $\mathcal{P}_u = \emptyset$. The algorithm adds $\mathfrak{b}_u$ to the set $\mathcal{L}_p$, with the above remarks ensuring that (3.6) is satisfied.

Suppose next that the hypothesis of (ii) is satisfied: $\mathcal{P}_u = \{\mathfrak{p}\}$ where $e(\mathfrak{p}|p) = f(\mathfrak{p}|p) = 1$ and there is a unique $\mathbb{Z}_p$ root α of $g(t)$ such that $\alpha \equiv u \mod p^t$. The algorithm adds $(\mathfrak{b}_u, \mathfrak{p})$ to the set $\mathcal{M}_p$. By the above, $(x - y\theta)\mathcal{O}_K/\mathfrak{b}_u$ is an integral ideal, not divisible by any prime ideal $\mathfrak{q} \neq \mathfrak{p}$ lying over p. Thus there is some positive integer $v_p \geq 0$ such that (3.7) is satisfied, concluding the proof. □

Having computed the lists arising from the affine patch $p \nmid y$, we initialize L_p and M_p as $\mathcal{L}_p$ and $\mathcal{M}_p$, respectively, and append to these lists the elements from the second patch, $p \mid y$, using the following algorithm.

To compute L_p and M_p.

Step (1) Let

$$L_p \leftarrow \mathcal{L}_p, \qquad M_p \leftarrow \mathcal{M}_p,$$

where $\mathcal{L}_p$, $\mathcal{M}_p$ are computed by Algorithm 1.

Step (2) Let

$$t \leftarrow 2, \qquad \mathcal{U} \leftarrow \{pw \ : \ w \in \{0, 1, \dots, p-1\}\}.$$

Step (3) Let

$$\mathcal{U}' \leftarrow \emptyset.$$

Loop through the elements $u \in \mathcal{U}$. Let

$$\mathcal{P}_u = \{\mathfrak{q} \text{ lying above } p \ : \ \mathrm{ord}_{\mathfrak{q}}(1 - u\theta) \geq t \cdot e(\mathfrak{q}|p)\},$$

and

$$\mathfrak{b}_u = \prod_{\mathfrak{q}|p} \mathfrak{q}^{\min\{\mathrm{ord}_{\mathfrak{q}}(1-u\theta), t \cdot e(\mathfrak{q}|p)\}} = (1 - u\theta)\mathcal{O}_K + p^t \mathcal{O}_K.$$

(i) If $\mathcal{P}_u = \emptyset$ then

$$L_p \leftarrow L_p \cup \{\mathfrak{b}_u\}.$$

(ii) Else

$$\mathcal{U}' \leftarrow \mathcal{U}' \cup \{u + p^t w : w \in \{0, \dots, p-1\}\}.$$

Step (4) If $\mathcal{U}' \neq \emptyset$ then let

$$t \leftarrow t + 1, \qquad \mathcal{U} \leftarrow \mathcal{U}',$$

and return to Step (3). Else output L_p, M_p.

Lemma 3.3.5. *Algorithm 2 terminates.*

Proof. Suppose that the algorithm does not terminate. Let $t_0 = 2$ and $t_i = t_0 + i$ for $i \in \mathbb{N}$. Then there is an infinite sequence of congruence classes $\{u_i\}_{i=0}^{\infty}$ and corresponding sets $\{\mathcal{P}_{u_i}\}_{i=0}^{\infty}$ such that $u_{i+1} \equiv u_i \mod t_i$ and $\mathcal{P}_{u_i} \neq \emptyset$ for all i. Moreover, $p \mid u_0$. Let α be the limit of $\{u_i\}_{i=0}^{\infty}$ in $\mathbb{Z}_p$. By the pigeon-hole principle, there is some ideal $\mathfrak{q}$ in $\mathcal{O}_K$ above p which appears in infinitely many sets $\mathcal{P}_{u_i}$. It follows that $\mathrm{ord}_{\mathfrak{q}}(1 - u_i\theta) \geq t_i \cdot e(\mathfrak{q}|p)$ and thus $1 - \alpha\theta = 0$ in $K_{\mathfrak{q}}$. But as $p \mid u_0$, we have $\mathrm{ord}_p(\alpha) \geq 1$, and so $\mathrm{ord}_{\mathfrak{q}}(\theta) < 0$. This contradicts the fact that θ is an algebraic integer. Therefore the algorithm must terminate. □

Lemma 3.3.6. *Let $p \in \{p_1, \dots, p_v\}$ and let L_p, M_p be as given by Algorithm 2. Let*

$(x, y, z_1, \ldots, z_v)$ *be a solution to* (3.3)*. Then*

- *either there is some* $\mathfrak{b} \in L_p$ *such that* (3.6) *is satisfied;*
- *or there is some* $(\mathfrak{b}, \mathfrak{p}) \in M_p$ *with* $e(\mathfrak{p}|p) = f(\mathfrak{p}|p) = 1$ *and integer* $v_p \geq 0$ *such that* (3.7) *is satisfied.*

Proof. Let $(x, y, z_1, \ldots, z_v)$ be a solution to (3.3). In view of Lemma 3.3.4 we may suppose $p \mid y$. Then $\mathrm{ord}_{\mathfrak{q}}(x) = 0$ and $\mathrm{ord}_{\mathfrak{q}}(x - y\theta) = \mathrm{ord}_{\mathfrak{q}}(1 - (y/x)\theta)$ for any prime ideal $\mathfrak{q}$ lying over p. The remainder of the proof is analogous to the proof of Lemma 3.3.4. □

3.3.1 Computational remarks and refinements

In implementing Algorithms 1 and 2, we reduce the number of prime ideals appearing to a large power in the factorization of $(x - y\theta)\mathcal{O}_K$. The Prime Ideal Removing Lemma, as originally stated in Tzanakis - de Weger outlines a similar process by comparing the valuations of $(x-y\theta)\mathcal{O}_K$ at two prime ideals $\mathfrak{p}_1$ and $\mathfrak{p}_2$ above p. Of course if $\mathfrak{p}_1 \mid (x-y\theta)\mathcal{O}_K$, we restrict the possible values for x and y modulo p. However any choice of x and y modulo p affects the valuations of $(x-y\theta)\mathcal{O}_K$ at all prime ideals above p. In the present refinement outlined by Lemma 3.3.1 and Lemma 3.3.2, we instead study the valuations of $(x-y\theta)\mathcal{O}_K$ at all prime ideals above p simultaneously. This presents us with considerably less ideal equations of the form (3.5) to resolve.

Moreover, this variant of the Prime Ideal Removing Lemma permits the following additional refinements:

- Let $\mathfrak{b} \in L_p$. If there exists a pair $(\mathfrak{b}', \mathfrak{p}) \in M_p$ with $\mathfrak{b}' \mid \mathfrak{b}$ and $\mathfrak{b}/\mathfrak{b}' = \mathfrak{p}^w$ for some $w \geq 0$, then we may delete $\mathfrak{b}$ from L_p. In doing so, the conclusion to Lemma 3.3.6 continues to hold.
- Suppose $(\mathfrak{b}, \mathfrak{p})$, $(\mathfrak{b}', \mathfrak{p}) \in M_p$ with $\mathfrak{b}' \mid \mathfrak{b}$, and $\mathfrak{b}/\mathfrak{b}' = \mathfrak{p}^w$ for some $w \geq 0$. Then, we may delete $(\mathfrak{b}, \mathfrak{p})$ from M_p without affecting the conclusion to Lemma 3.3.6.

3.4 Factorization of the Thue-Mahler equation

After applying Algorithm 1 and Algorithm 2, we are reduced to solving finitely many ideal equations of the form

$$(x - y\theta)\mathcal{O}_K = \mathfrak{a}\mathfrak{p}_1^{u_1} \cdots \mathfrak{p}_\nu^{u_\nu} \tag{3.8}$$

in integer variables $x, y, u_1, \ldots, u_\nu$ with $u_i \geq 0$ for $i = 1, \ldots, \nu$, where $0 \leq \nu \leq v$. Here

- for $i \in \{1, \ldots, \nu\}$, $\mathfrak{p}_i$ is a prime ideal of $\mathcal{O}_K$ arising from Algorithm 1 and Algorithm 2 applied to $p \in \{p_1, \ldots, p_v\}$, such that $(\mathfrak{b}, \mathfrak{p}_i) \in M_p$ for some ideal $\mathfrak{b}$;
- for $i \in \{\nu + 1, \ldots, v\}$, the corresponding rational prime $p_i \in S$ yields $M_{p_i} = \emptyset$, in which case we set $u_i = 0$;
- $\mathfrak{a}$ is an ideal of $\mathcal{O}_K$ of norm $|c| \cdot p_1^{t_1} \cdots p_v^{t_v}$ such that $u_i + t_i = z_i$.

For each choice of $\mathfrak{a}$ and prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_\nu$, we reduce equation (3.8) to a number of so-called "S-unit equations". We present two different algorithms for doing so and outline the advantages and disadvantages of each. In practicality, we do not know a priori which of these two options is more efficient. Instead, we implement and use both algorithms simultaneously and selecting the most computationally efficient option as it appear.

3.4.1 Avoiding the class group $\mathbf{Cl}(K)$

For $i = 1, \ldots, \nu$ let h_i be the smallest positive integer for which $\mathfrak{p}_i^{h_i}$ is principal and let r_i be a positive integer satisfying $0 \leq r_i < h_i$. Let

$$\mathbf{a}_i = (a_{1i}, \ldots, a_{\nu i}).$$

where $a_{ii} = h_i$ and $a_{ji} = 0$ for $j \neq i$. We let A be the matrix with columns $\mathbf{a}_1, \ldots, \mathbf{a}_\nu$. Hence A is a $\nu \times \nu$ diagonal matrix over $\mathbb{Z}$ with diagonal entries h_i. Now, if (3.8) has a solution $\mathbf{u} = (u_1, \ldots, u_\nu)$, it necessarily must be of the form $\mathbf{u} = A\mathbf{n} + \mathbf{r}$, where $\mathbf{n} = (n_1, \ldots, n_\nu)$ and $\mathbf{r} = (r_1, \ldots, r_\nu)$. The vector $\mathbf{n}$ is comprised of integers n_i which we solve for. The vector $\mathbf{r}$ is comprised of the values r_i satisfying $0 \leq r_i < h_i$ for $i = 1, \ldots, \nu$.

Using the above notation, we let

$$\mathfrak{c}_i = \tilde{\mathfrak{p}}^{\mathbf{a}_i} = \mathfrak{p}_1^{a_{1i}} \cdot \mathfrak{p}_2^{a_{2i}} \cdots \mathfrak{p}_\nu^{a_{\nu i}} = \mathfrak{p}_i^{h_i}$$

for all $i \in \{1, \ldots, \nu\}$.

Thus, we can write (3.8) as

$$(x - y\theta)\mathcal{O}_K = \mathfrak{a}\tilde{\mathfrak{p}}^{\mathbf{u}} = (\mathfrak{a} \cdot \tilde{\mathfrak{p}}^{\mathbf{r}}) \cdot \mathfrak{c}_1^{n_1} \cdots \mathfrak{c}_\nu^{n_\nu}.$$

By definition of h_i, each $i \in \{1, \ldots, \nu\}$ yields an element $\gamma_i \in K^*$ such that

$$\mathfrak{c}_i = (\gamma_i)\mathcal{O}_K.$$

Furthermore, if $\mathbf{u}$ is a solution of (3.8) with corresponding vectors $\mathbf{n}, \mathbf{r}$, there exists some $\alpha \in K^*$ such that

$$\mathfrak{a} \cdot \tilde{\mathfrak{p}}^{\mathbf{r}} = (\alpha)\mathcal{O}_K.$$

3.4.2 Using the class group $\mathbf{Cl}(K)$

Let $\mathbf{u} = (u_1, \ldots, u_\nu)$ be a solution of (3.8) and consider the map

$$\phi : \mathbb{Z}^\nu \to \mathrm{Cl}(K), \qquad (x_1, \ldots, x_\nu) \mapsto [\mathfrak{p}_1]^{x_1} \cdots [\mathfrak{p}_\nu]^{x_\nu},$$

where $[\mathfrak{q}]$ denotes the equivalence class of the fractional ideal $\mathfrak{q}$. Since the product of $\mathfrak{a}$ and $\mathfrak{p}_1^{u_1} \cdots \mathfrak{p}_\nu^{u_\nu}$ defines a principal ideal, the map ϕ implies

$$\phi(\mathbf{u}) = [\mathfrak{a}]^{-1}.$$

In particular, if $[\mathfrak{a}]^{-1}$ does not belong to the image of ϕ then (3.8) has no solutions. We therefore suppose that $[\mathfrak{a}]^{-1}$ belongs to the image. Let $\mathbf{r} = (r_1, \ldots, r_\nu)$ denote a preimage of $[\mathfrak{a}]^{-1}$ and observe that $\mathbf{u} - \mathbf{r}$ belongs to the kernel of ϕ. The kernel is a subgroup of $\mathbb{Z}^v$ of rank ν. Let $\mathbf{a}_1, \ldots, \mathbf{a}_\nu$ be a basis for the kernel, where

$$\mathbf{a}_i = (a_{1i}, \ldots, a_{\nu i}) \quad \text{for } i = 1, \ldots, \nu.$$

Let

$$\mathbf{u} - \mathbf{r} = n_1\mathbf{a}_1 + \cdots + n_\nu\mathbf{a}_\nu$$

for some integers $n_i \in \mathbb{Z}$ and let A denote the $\nu \times \nu$ matrix over $\mathbb{Z}$ with columns $\mathbf{a}_1, \ldots, \mathbf{a}_\nu$. It follows that $\mathbf{u} = A\mathbf{n} + \mathbf{r}$ where $\mathbf{n} = (n_1, \ldots, n_\nu)$.

For $\mathbf{a}_i = (a_{1i}, \ldots, a_{\nu i}) \in \mathbb{Z}^\nu$, we adopt the notation

$$\tilde{\mathfrak{p}}^{\mathbf{a}} := \mathfrak{p}_1^{a_{1i}} \cdot \mathfrak{p}_2^{a_{2i}} \cdots \mathfrak{p}_\nu^{a_{\nu i}}.$$

Let

$$\mathfrak{c}_1 = \tilde{\mathfrak{p}}^{\mathbf{a}_1}, \ldots, \mathfrak{c}_\nu = \tilde{\mathfrak{p}}^{\mathbf{a}_\nu}.$$

Thus, we can rewrite (3.8) as

$$(x - y\theta)\mathcal{O}_K = \mathfrak{a}\tilde{\mathfrak{p}}^{\mathbf{u}} = (\mathfrak{a} \cdot \tilde{\mathfrak{p}}^{\mathbf{r}}) \cdot \mathfrak{c}_1^{n_1} \cdots \mathfrak{c}_\nu^{n_\nu}.$$

Consider the ideal equivalence class of $(\mathfrak{a} \cdot \tilde{\mathfrak{p}}^{\mathbf{r}})$ in $\mathrm{Cl}(K)$ and note that

$$[\mathfrak{a} \cdot \tilde{\mathfrak{p}}^{\mathbf{r}}] = [\mathfrak{a}] \cdot [\mathfrak{p}_1]^{r_1} \cdots [\mathfrak{p}_\nu]^{r_\nu} = [\mathfrak{a}] \cdot \phi(r_1, \ldots, r_\nu) = [1]$$

as $\phi(r_1, \ldots, r_\nu) = [\mathfrak{a}]^{-1}$ by construction. This means

$$\mathfrak{a} \cdot \tilde{\mathfrak{p}}^{\mathbf{r}} = (\alpha)\mathcal{O}_K$$

for some $\alpha \in K^*$. Furthermore,

$$[\mathfrak{c}_i] = [\tilde{\mathfrak{p}}^{\mathbf{a}_i}] = \phi(\mathbf{a}_i) = [1] \quad \text{for } i = 1, \ldots, \nu,$$

as the $\mathbf{a}_i$ are a basis for the kernel of ϕ. For all $i \in \{1, \ldots, \nu\}$, we therefore have

$$\mathfrak{c}_i = (\gamma_i)\mathcal{O}_K$$

for some $\gamma_i \in K^*$.

3.4.3 The S-unit equation

subsection 3.4.1 and subsection 3.4.2 outline two different algorithms to reduce the ideal equation (3.8) to a number of certain "S-unit equations", which we define shortly. Regardless of which method we use, under both algorithms outlined above, equation (3.8) becomes

$$(x - y\theta)\mathcal{O}_K = (\alpha \cdot \gamma_1^{n_1} \cdots \gamma_\nu^{n_\nu})\mathcal{O}_K \tag{3.9}$$

for some vector $\mathbf{n} = (n_1, \ldots, n_\nu) \in \mathbb{Z}^\nu$. The ideal generated by α in K has norm

$$|c| \cdot p_1^{t_1 + r_1} \cdots p_\nu^{t_\nu + r_\nu} p_{\nu+1}^{t_{\nu+1}} \cdots p_v^{t_v}$$

and the n_i are related to the z_i via

$$z_i = u_i + t_i = \sum_{j=1}^{\nu} n_j a_{ij} + r_i + t_i \quad \text{for } i = 1, \ldots, v.$$

where $u_i = r_i = 0$ for all $i \in \{\nu+1, \ldots, v\}$.

Fix a complete set of fundamental units $\{\varepsilon_1, \ldots, \varepsilon_r\}$ of $\mathcal{O}_K$. Here $r = s + t - 1$, where s denotes the number of real embeddings of K into $\mathbb{C}$ and t denotes the number of complex conjugate pairs of non-real embeddings of K into $\mathbb{C}$. Then, under either method, equation (3.8) reduces to a finite number of equations in K of the form

$$x - y\theta = \alpha\zeta\varepsilon_1^{a_1} \cdots \varepsilon_r^{a_r}\gamma_1^{n_1} \cdots \gamma_\nu^{n_\nu} \tag{3.10}$$

with unknowns $a_i \in \mathbb{Z}$, $n_i \in \mathbb{Z}$, and ζ in the set T of roots of unity in $\mathcal{O}_K$. Since T is finite, we treat ζ as another parameter.

Let $p \in \{p_1, \ldots, p_v, \infty\}$. Recall that $g(t)$ is an irreducible polynomial in $\mathbb{Z}[t]$ arising from (3.3) such that

$$g(t) = f(t, 1) = t^n + C_1 t^{n-1} + \cdots + C_{n-1}t + C_n.$$

Denote the roots of $g(t)$ in $\overline{\mathbb{Q}_p}$ (where $\overline{\mathbb{Q}_\infty} = \overline{\mathbb{R}} = \mathbb{C}$) by $\theta^{(1)}, \ldots, \theta^{(n)}$. Let $i_0, j, k \in \{1, \ldots, n\}$ be distinct indices and consider the three embeddings of K into $\overline{\mathbb{Q}_p}$ defined by $\theta \mapsto \theta^{(i_0)}, \theta^{(j)}, \theta^{(k)}$. We use $z^{(i)}$ to denote the image of z under the embedding $\theta \mapsto \theta^{(i)}$. From the Siegel identity

$$(\theta^{(i_0)} - \theta^{(j)})(x - y\theta^{(k)}) + (\theta^{(j)} - \theta^{(k)})(x - y\theta^{(i_0)}) + (\theta^{(k)} - \theta^{(i_0)})(x - y\theta^{(j)}) = 0,$$

applying the embeddings to $\beta = x - y\theta$ yields the so-called "S-unit equation"

$$\delta_1 \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)^{a_i} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}}\right)^{n_i} - 1 = \delta_2 \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(i_0)}}{\varepsilon_i^{(j)}}\right)^{a_i} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}}\right)^{n_i}, \tag{3.11}$$

where

$$\delta_1 = \frac{\theta^{(i_0)} - \theta^{(j)}}{\theta^{(i_0)} - \theta^{(k)}} \cdot \frac{\alpha^{(k)}\zeta^{(k)}}{\alpha^{(j)}\zeta^{(j)}}, \quad \delta_2 = \frac{\theta^{(j)} - \theta^{(k)}}{\theta^{(k)} - \theta^{(i_0)}} \cdot \frac{\alpha^{(i_0)}\zeta^{(i_0)}}{\alpha^{(j)}\zeta^{(j)}}$$

are constants.

To summarize, our original problem of solving (3.3) for $(x, y, z_1, \ldots, z_v)$ has been reduced to solving finitely many equations of the form (3.11) for the variables $(n_1, \ldots, n_\nu, a_1, \ldots, a_r)$.

3.4.4 Computational remarks and comparisons

In subsection 3.4.1, we follow closely the method of [**?**] to reduce the ideal equation (3.8) to the S-unit equation (3.11). To implement this reduction, we begin by computing all h_i for which $\mathfrak{p}_i^{h_i}$ is principal for $i = 1, \ldots, \nu$. In doing so, we generate all possible values for r_i, the non-negative integer satisfying $0 \leq r_i < h_i$. We then generate every possible vector $\mathbf{r} = (r_1, \ldots, r_\nu)$ and test the corresponding ideal product $\mathfrak{a} \cdot \tilde{\mathfrak{p}}^{\mathbf{r}}$ for principality. Those vectors which pass this test yield an S-unit equation (3.11). In the worst case scenario, this method reduces to h_K^ν such equations, where h_K is the class number of K. Moreover, this process needs to be applied to every ideal equation (3.8), yielding what may be a very large number of principalization tests and subsequent large number of S-unit equations to solve.

In contrast, the method in subsection 3.4.2 reduces (3.8) to only $\#T/2$ S-unit equations to solve, where T is the set of roots of unity in K. In particular, the sum total of S-unit equations does not drastically increase. If $[\mathbf{a}]^{-1}$ is not in the image of ϕ, we reach a contradiction. If $[\mathbf{a}]^{-1}$ is in the image of ϕ then we obtain only $\#T/2$ corresponding equations (3.11). In particular, the number of principalization tests in this method is limited by the number of ideal equations (3.8), where each such equation yields only $(1 + \nu)$ tests.

However, when generating the vectors $\mathbf{r} = (r_1, \ldots, r_\nu)$ using the class group, we observe that some of the integers r_i may be negative, so we do not expect α to be an algebraic integer in general. This can be problematic later in the algorithm when we compute the embedding of K into our p-adic fields. In those instances, the precision on our p-adic fields may not be high enough, and as a result, some non-zero elements of K may be erroneously mapped to 0. To avoid this, we force the r_i to be positive by adding sufficiently many multiples of the class number.

In most cases, the method described in subsection 3.4.2 is far more efficient than that of subsection 3.4.1. However, computing the class group may be a very costly computation. Indeed, for some Thue-Mahler equations, this may be the bottle-neck of the algorithm. In this case, it may happen that computing the class group will take longer than directly checking each potential S-unit equation arising from the alternative method. Unfortunately, we cannot know a priori how long computing $\mathrm{Cl}(K)$ will take in so much that we cannot know a priori how long solving all S-unit equations from the other algorithm will take. In practicality, generating the class group in Magma is a process which cannot

be terminated without exiting the program. For this reason, we cannot simply apply a timeout in Magma if computing $\mathrm{Cl}(K)$ is exceeding what we deem a reasonable amount of time. Adding to this, Magma does not support parallelization, so we cannot implement both algorithms simultaneously. Our compromise to solve a single Thue-Mahler equation is to run two separate instances of Magma in parallel, each generating the S-unit equations using the two aforementioned algorithms. When one of these instances finishes, the other is forced to terminate. In this way, though far from ideal, we are able to select the most computationally efficient option.

3.5 A small upper bound for u_l in a special case

We now restrict our attention to those $p \in \{p_1, \ldots, p_\nu\}$ and study the p-adic valuations of the numbers appearing in (3.11). In particular, for $l \in \{1, \ldots, \nu\}$, we identify conditions in which $\sum_{j=1}^{\nu} n_j a_{lj}$ can be bounded by a small explicit constant, where a_{lj} is the $(l,j)^{\text{th}}$ entry of the matrix A derived in either subsection 3.4.1 or subsection 3.4.2. Recall that $u_l + r_l = \sum_{j=1}^{\nu} n_j a_{lj}$, where r_l is known, so that a bound on $\sum_{j=1}^{\nu} n_j a_{lj}$ yields a bound on the exponent u_l in (3.8).

Fix a rational prime $p_l \in \{p_1, \ldots, p_\nu\}$ and recall that $z \in \mathbb{C}_{p_l}$ having $\mathrm{ord}_{p_l}(z) = 0$ is called a p_l-adic unit. Part (i) of the Corollary of Lemma 7.2 of [**?**] tells us that $\frac{\varepsilon_1^{(i_0)}}{\varepsilon_1^{(j)}}, \ldots, \frac{\varepsilon_r^{(i_0)}}{\varepsilon_r^{(j)}}$ and $\frac{\varepsilon_1^{(k)}}{\varepsilon_1^{(j)}}, \ldots, \frac{\varepsilon_r^{(k)}}{\varepsilon_r^{(j)}}$ are p_l-adic units.

Let $g_l(t)$ be the irreducible factor of $g(t)$ in $\mathbb{Q}_{p_l}[t]$ corresponding to the prime ideal $\mathfrak{p}_l$. Since $\mathfrak{p}_l$ has ramification index and residue degree equal to 1, $\deg(g_l(t)) = 1$. We now choose $i_0 \in \{1, \ldots, 4\}$ so that $\theta^{(i_0)}$ is the root of $g_l(t)$. We fix this choice of index i_0 for the remainder of this chapter. The indices of j, k are fixed, but arbitrary.

Lemma 3.5.1.

(i) Let $i \in \{1, \ldots, \nu\}$. Then $\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}}$ are p_l-adic units.

(ii) Let $i \in \{1, \ldots, \nu\}$. Then $\mathrm{ord}_{p_l}\left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}}\right) = a_{li}$, where $\mathbf{a_i} = (a_{1i}, \ldots, a_{\nu i})$ is the i^{th} column of the matrix A of either subsection 3.4.1 or subsection 3.4.2.

Proof. Consider the factorization $g(t) = g_1(t) \cdots g_m(t)$ of $g(t)$ in $\mathbb{Q}_{p_l}[t]$. Note that $\theta^{(j)}$ is a root of some $g_h(t) \neq g_l(t)$. Let $\mathfrak{p}_h$ be the corresponding prime ideal above p_l and $e(\mathfrak{p}_h | p_l)$

be its ramification index. Then $\mathfrak{p} \neq \mathfrak{p}_l$ and since

$$(\gamma_i)\mathcal{O}_K = \mathfrak{p}_1^{a_{1i}} \cdots \mathfrak{p}_v^{a_{vi}},$$

we have

$$\operatorname{ord}_{p_l}(\gamma_i^{(j)}) = \frac{1}{e(\mathfrak{p}_h|p_l)} \operatorname{ord}_{\mathfrak{p}_h}(\gamma_i) = 0.$$

An analogous argument gives $\operatorname{ord}_{p_l}(\gamma_i^{(k)}) = 0$. On the other hand,

$$\operatorname{ord}_{p_l}(\gamma_i^{(i_0)}) = \frac{1}{e(\mathfrak{p}_l|p_l)} \operatorname{ord}_{\mathfrak{p}_l}(\gamma_i) = \operatorname{ord}_{\mathfrak{p}_l}(\mathfrak{p}_1^{a_{1i}} \cdots \mathfrak{p}_v^{a_{vi}}) = a_{li}.$$

□

The next lemma deals with a special case in which the sum $\sum_{j=1}^{\nu} n_j a_{lj}$ can be computed directly. This lemma is analogous to Lemma 7.3 of [**?**].

Recall the constants

$$\delta_1 = \frac{\theta^{(i_0)} - \theta^{(j)}}{\theta^{(i_0)} - \theta^{(k)}} \cdot \frac{\alpha^{(k)}\zeta^{(k)}}{\alpha^{(j)}\zeta^{(j)}}, \quad \delta_2 = \frac{\theta^{(j)} - \theta^{(k)}}{\theta^{(k)} - \theta^{(i_0)}} \cdot \frac{\alpha^{(i_0)}\zeta^{(i_0)}}{\alpha^{(j)}\zeta^{(j)}}$$

of (3.11).

Lemma 3.5.2. *Let $l \in \{1, \ldots, v\}$. If $\operatorname{ord}_{p_l}(\delta_1) \neq 0$, then*

$$\sum_{i=1}^{\nu} n_i a_{li} = \min\{\operatorname{ord}_{p_l}(\delta_1), 0\} - \operatorname{ord}_{p_l}(\delta_2).$$

Proof. Apply the Corollary of Lemma 7.2 of [**?**] and Lemma 3.5.1 to both expressions of λ in (3.11). On the one hand, we obtain that

$$\operatorname{ord}_{p_l}(\lambda) = \min\{\operatorname{ord}_{p_l}(\delta_1), 0\},$$

and on the other hand,

$$\begin{aligned} \operatorname{ord}_{p_l}(\lambda) &= \operatorname{ord}_{p_l}(\delta_2) + \sum_{i=1}^{\nu} \operatorname{ord}_{p_l}\left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}}\right)^{n_i} \\ &= \operatorname{ord}_{p_l}(\delta_2) + \sum_{i=1}^{\nu} n_i a_{li}. \end{aligned}$$

□

For the remainder of this section, we assume $\mathrm{ord}_{p_l}(\delta_1) = 0$. Here, it is convenient to use the notation

$$b_1 = 1, \quad b_{1+i} = n_i \ \text{ for } i \in \{1, \dots, \nu\},$$

and

$$b_{1+\nu+i} = a_i \ \text{ for } i \in \{1, \dots, r\}.$$

Put

$$\alpha_1 = \log_{p_l} \delta_1, \quad \alpha_{1+i} = \log_{p_l} \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}} \right) \ \text{ for } i \in \{1, \dots, \nu\},$$

and

$$\alpha_{1+\nu+i} = \log_{p_l} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right) \ \text{ for } i \in \{1, \dots, r\}.$$

Define

$$\Lambda_l = \sum_{i=1}^{1+\nu+r} b_i \alpha_i.$$

Let L be a finite extension of $\mathbb{Q}_{p_l}$ containing δ_1, $\frac{\gamma_1^{(k)}}{\gamma_1^{(j)}}, \dots, \frac{\gamma_\nu^{(k)}}{\gamma_\nu^{(j)}}$, and $\frac{\varepsilon_1^{(k)}}{\varepsilon_1^{(j)}}, \dots, \frac{\varepsilon_r^{(k)}}{\varepsilon_r^{(j)}}$. Since finite p_l-adic fields are complete, $\alpha_i \in L$ for $i = 1, \dots, 1 + \nu + r$ as well. Choose $\phi \in \overline{\mathbb{Q}_{p_l}}$ such that $L = \mathbb{Q}_{p_l}(\phi)$ and $\mathrm{ord}_{p_l}(\phi) > 0$. Let $G(t)$ be the minimal polynomial of ϕ over $\mathbb{Q}_{p_l}$ and let s be its degree. For $i = 1, \dots, 1 + \nu + r$ write

$$\alpha_i = \sum_{h=1}^{s} \alpha_{ih} \phi^{h-1}, \quad \alpha_{ih} \in \mathbb{Q}_{p_l}.$$

Then

$$\Lambda_l = \sum_{h=1}^{s} \Lambda_{lh} \phi^{h-1}, \tag{3.12}$$

with

$$\Lambda_{lh} = \sum_{i=1}^{1+\nu+r} b_i \alpha_{ih}$$

for $h = 1, \dots, s$.

Lemma 3.5.3. *For every $h \in \{1, \dots, s\}$, we have*

$$\mathrm{ord}_{p_l}(\Lambda_{lh}) > \mathrm{ord}_{p_l}(\Lambda_l) - \frac{1}{2} \mathrm{ord}_{p_l}(Disc(G(t))).$$

Proof. For $h = 1, \dots, s$, taking the images of (3.12) under conjugation $\phi \mapsto \phi^{(h)}$ yields

$$\begin{bmatrix} \Lambda_l^{(1)} \\ \vdots \\ \Lambda_l^{(s)} \end{bmatrix} = \begin{bmatrix} 1 & \phi^{(1)} & \dots & \phi^{(1)s-1} \\ \vdots & \vdots & & \vdots \\ 1 & \phi^{(s)} & \dots & \phi^{(s)s-1} \end{bmatrix} \begin{bmatrix} \Lambda_{l1} \\ \vdots \\ \Lambda_{ls} \end{bmatrix}.$$

The $s \times s$ matrix $(\phi^{(h)i-1})$ above is invertible, with inverse

$$\frac{1}{\prod\limits_{1 \le j < k \le s} (\phi^{(k)} - \phi^{(j)})} \begin{bmatrix} \gamma_{11} & \dots & \gamma_{1s} \\ \vdots & & \vdots \\ \gamma_{s1} & \dots & \gamma_{ss} \end{bmatrix},$$

where γ_{jk} is an integral polynomial in the entries of $(\phi^{(h)i-1})$. As $\operatorname{ord}_{p_l}(\phi) > 0$ and $\operatorname{ord}_{p_l}(\phi^{(h)}) = \operatorname{ord}_{p_l}(\phi)$ for all $h = 1, \dots, s$, it follows that $\operatorname{ord}_{p_l}(\gamma_{jk}) > 0$ for every γ_{jk}. Therefore, as

$$\Lambda_{lh} = \frac{1}{\prod\limits_{1 \le j < k \le s} (\phi^{(k)} - \phi^{(j)})} \sum_{i=1}^{s} \gamma_{hi} \Lambda_l^{(i)},$$

we have

$$\begin{aligned} \operatorname{ord}_{p_l}(\Lambda_{lh}) &= \min_{1 \le i \le s} \left\{ \operatorname{ord}_{p_l}(\gamma_{hi}) + \operatorname{ord}_{p_l}(\Lambda_l^{(i)}) \right\} - \frac{1}{2} \operatorname{ord}_{p_l}(\operatorname{Disc}(G(t))) \\ &\ge \min_{1 \le i \le s} \operatorname{ord}_{p_l}(\Lambda_l^{(i)}) + \min_{1 \le i \le s} \operatorname{ord}_{p_l}(\gamma_{hi}) - \frac{1}{2} \operatorname{ord}_{p_l}(\operatorname{Disc}(G(t))) \\ &= \operatorname{ord}_{p_l} \Lambda_l + \min_{1 \le i \le s} \operatorname{ord}_{p_l}(\gamma_{hi}) - \frac{1}{2} \operatorname{ord}_{p_l}(\operatorname{Disc}(G(t))) \end{aligned}$$

for every $h \in \{1, \dots, s\}$. □

Lemma 3.5.4. *If*

$$\sum_{i=1}^{\nu} n_i a_{li} > \frac{1}{p_l - 1} - \operatorname{ord}_{p_l}(\delta_2),$$

then

$$\operatorname{ord}_{p_l}(\Lambda_l) = \sum_{i=1}^{\nu} n_i a_{li} + \operatorname{ord}_{p_l}(\delta_2).$$

Proof. Immediate from Lemma 2.2.2. □

Lemma 3.5.5. *Let*

$$w_l = \left\lfloor \frac{1}{p_l - 1} - \mathrm{ord}_{p_l}(\delta_2) \right\rfloor.$$

(i) If $\mathrm{ord}_{p_l}(\alpha_1) < \min_{2 \le i \le 1+\nu+r} \mathrm{ord}_{p_l}(\alpha_i)$, *then*

$$\sum_{i=1}^{\nu} n_i a_{li} \le \max\left\{ w_l, \left\lceil \min_{2 \le i \le 1+\nu+r} \mathrm{ord}_{p_l}(\alpha_i) - \mathrm{ord}_{p_l}(\delta_2) \right\rceil - 1 \right\}$$

(ii) For all $h \in \{1, \ldots, s\}$, *if* $\mathrm{ord}_{p_l}(\alpha_{1h}) < \min_{2 \le i \le 1+\nu+r} \mathrm{ord}_{p_l}(\alpha_{ih})$, *then*

$$\sum_{i=1}^{\nu} n_i a_{li} \le \max\left\{ w_l, \left\lceil \min_{2 \le i \le 1+\nu+r} \mathrm{ord}_{p_l}(\alpha_{ih}) - \mathrm{ord}_{p_l}(\delta_2) + d_l \right\rceil - 1 \right\},$$

where

$$d_l = \frac{1}{2} \mathrm{ord}_{p_l}(Disc(G(t))).$$

Proof.

(i) We prove the contrapositive. Suppose

$$\sum_{i=1}^{\nu} n_i a_{li} > \frac{1}{p_l - 1} - \mathrm{ord}_{p_l}(\delta_2),$$

and

$$\sum_{i=1}^{\nu} n_i a_{li} \ge \min_{2 \le i \le 1+\nu+r} \mathrm{ord}_{p_l}(\alpha_i) - \mathrm{ord}_{p_l}(\delta_2).$$

Observe that

$$\begin{aligned} \mathrm{ord}_{p_l}(\alpha_1) &= \mathrm{ord}_{p_l}\left(\Lambda_l - \sum_{i=2}^{1+\nu+r} b_i \alpha_i \right) \\ &\ge \min\left\{ \mathrm{ord}_{p_l}(\Lambda_l), \min_{2 \le i \le 1+\nu+r} \mathrm{ord}_{p_l}(b_i \alpha_i) \right\}. \end{aligned}$$

Therefore, it suffices to show that

$$\mathrm{ord}_{p_l}(\Lambda_l) \ge \min_{2 \le i \le 1+\nu+r} \mathrm{ord}_{p_l}(b_i \alpha_i).$$

By Lemma 2.2.2, the first inequality implies

$$\mathrm{ord}_{p_l}(\Lambda_l) = \sum_{i=1}^{\nu} n_i a_{li} + \mathrm{ord}_{p_l}(\delta_2),$$

from which the result follows.

(ii) Similar to the proof of (i).

□

3.6 Lattice-Based Reduction

At this point in solving the Thue-Mahler equation, we proceed to solve each S-unit equation (3.11) for the exponents $(n_1, \ldots, n_\nu, a_1, \ldots, a_r)$. To do so, we generate a very large upper bound on the exponents and reduce this bound via Diophantine approximation computations. The specific details of this process are described in chapter 5 and **??**. In general, from each S-unit equation, we generate several linear forms in logarithms to which we associate an integral lattice Γ. It will be important in this reduction process to enumerate all short vectors in these lattices. In this section, we describe two algorithms used in the short vector enumeration process.

3.6.1 The L^3-lattice basis reduction algorithm

Let Γ be an n-dimensional lattice with basis vectors $\mathbf{b}_1, \ldots, \mathbf{b}_n$ equipped with a bilinear form $\Phi : \Gamma \times \Gamma \to \mathbb{Z}$. Recall that Φ defines a norm on Γ via the usual inner product on $\mathbb{R}^n$. For $i = 1, \ldots, n$, define the vectors $\mathbf{b}_i^*$ inductively by

$$\mathbf{b}_i^* = \mathbf{b}_i - \sum_{j=1}^{i-1} \mu_{ij} \mathbf{b}_j^*, \quad \mu_{ij} = \frac{\Phi(\mathbf{b}_i, \mathbf{b}_j^*)}{\Phi(\mathbf{b}_j^*, \mathbf{b}_j)},$$

where $\mu_{ij} \in \mathbb{R}$ for $1 \leq j < i \leq n$. This is the usual Gram-Schmidt process. The basis $\mathbf{b}_1, \ldots, \mathbf{b}_n$ is called *LLL-reduced* if

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

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

where $|\cdot|$ is the usual Euclidean norm in $\mathbb{R}^n$,

$$|\mathbf{v}| = \Phi(\mathbf{v}, \mathbf{v}) = \mathbf{v}^T\mathbf{v}.$$

These properties imply 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 ([**?**]). This algorithm takes as input an arbitrary basis for a lattice and outputs an LLL-reduced basis. The algorithm is typically modified to additionally output a unimodular matrix U such that $A = BU$, where B is the matrix whose column-vectors are the input basis and A is the matrix whose column-vectors are the LLL-reduced output basis. Several versions of this algorithm are implemented in Magma, including de Weger's exact integer version. ([**?**]).

We remark that a lattice may have more than one reduced basis, and that the ordering of the basis vectors is not arbitrary. The properties of reduced bases that are of most interest to us are the following. Let $\mathbf{v}$ a vector in $\mathbb{R}^n$ and denote by $l(\Gamma, \mathbf{v})$ the distance from $\mathbf{v}$ to the nearest point in the lattice Γ, viz.

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

From an LLL-reduced basis for Γ, we can compute lower bounds for $l(\Gamma, \mathbf{v})$, according to the following results.

Lemma 3.6.1. *Let Γ be a lattice with LLL-reduced basis $\mathbf{c}_1, \ldots, \mathbf{c}_n$ and let $\mathbf{v}$ be a vector in $\mathbb{R}^n$.*

(a) If $\mathbf{v} = \mathbf{0}$, then $l(\Gamma, \mathbf{v}) \geq 2^{-(n-1)/2}|\mathbf{c}_1|$.

(b) Assume $\mathbf{v} = s_1\mathbf{c}_1 + \cdots + s_n\mathbf{c}_n$, where $s_1, \ldots, s_n \in \mathbb{R}$ with not all $s_i \in \mathbb{Z}$. Put

$$J = \{j \in \{1, \ldots, n\} \ : \ s_j \notin \mathbb{Z}\}.$$

For $j \in J$, set

$$\delta(j) = \begin{cases} \max_{i>j} \|s_i\||\mathbf{c}_i| & \text{if } j < n \\ 0 & \text{if } j = n, \end{cases}$$

where $\|\cdot\|$ denotes the distance to the nearest integer. We have

$$l(\Gamma, \mathbf{v}) \geq \max_{j \in J} \left(2^{-(n-1)/2} \|s_j\| |\mathbf{c}_1| - (n-j)\delta(j) \right).$$

Lemma 3.6.1 (a) is Proposition 1.11 in [**?**]; proofs can be found in [**?**], [**?**] (Section 3.4), or [**?**] (Section V.3). Lemma 3.6.1 (b) is a combination of Lemmas 3.5 and 3.6 in [**?**]. Note that the assumption in Lemma 3.6.1 (b) is equivalent to $\mathbf{v} \notin \Gamma$.

We see that the vector $\mathbf{c}_1$ in a reduced basis is, in a very precise sense, not too far from being the shortest non-zero vector of Γ. As has already been mentioned, what makes this result so valuable is that there is a very simple and efficient algorithm to find a reduced basis in a lattice, namely the LLL algorithm.

3.6.2 The Fincke-Pohst algorithm

Sometimes it is not sufficient to have a lower bound for $l(\Gamma, \mathbf{v})$ only. It may be useful to know exactly all vectors $\mathbf{u} \in \Gamma$ such that $|\mathbf{u}| = \Phi(\mathbf{u}, \mathbf{u}) \leq C$ for a given constant C. This can be done efficiently using an algorithm of Fincke-Pohst (cf. [**?**], [**?**]). A version of this algorithm with some improvements due to Stehlé is implemented in Magma. As input this algorithm takes a matrix B, whose columns span the lattice Γ, and a constant $C > 0$. The output is a list of all lattice points $\mathbf{u} \in \Gamma$ with $|\mathbf{u}| \leq C$, apart from $\mathbf{u} = \mathbf{0}$. In this section, we outline the main steps in this algorithm.

We begin by letting B denote the basis matrix associated to the lattice Γ, with corresponding bilinear form Φ. We call a vector $\mathbf{u} \in \Gamma$ *small* if its norm $\Phi(\mathbf{u}, \mathbf{u})$ is less than a constant C. As an element of the lattice, $\mathbf{u} = B\mathbf{x}$ for some coordinate vector $\mathbf{x} \in \mathbb{Z}^n$. Let Q be the quadratic form associated to Φ and let $A = B^T B$. Now finding the short vectors $\mathbf{u} \in \Gamma$ is equivalent to solving

$$Q(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} \leq C. \tag{3.13}$$

Let $\mathbf{x} = (x_1, \ldots, x_n)$. To solve this inequality, we first rearrange the terms of the quadratic form via quadratic completion. Here we assume that Γ is positive definite so that every nonzero element of the lattice has a positive norm. With this, we find the Cholesky

decomposition $A = R^T R$, where R is an upper triangular matrix, and express Q as

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

The coefficients q_{ij} are defined from R and stored in a matrix $\tilde{Q}$ for convenience. In particular,

$$q_{ij} = \begin{cases} \frac{r_{ij}}{r_{ii}} & \text{if } i < j \\ r_{ii}^2 & \text{if } i = j. \end{cases} \tag{3.14}$$

Since R is upper triangular, the matrix $\tilde{Q}$ is as well. This yields the following reformulation of (3.13)

$$\sum_{i=1}^{n} q_{ii} \left(x_i + \sum_{j=i+1}^{n} q_{ij} x_j \right)^2 \leq C.$$

From here we observe that the individual term $q_{nn} x_n^2$ must also be less than C. Specifically,

$$x_n^2 \leq \frac{C}{q_{nn}}$$

so that x_n is bounded above by $\sqrt{C/q_{nn}}$ and below by $-\sqrt{C/q_{nn}}$. 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. Let

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

where $U_n = 0$, and rewrite $Q(\mathbf{x})$ as

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

In general,

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

Let T_k denote the bound on the right-hand side,

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

We set $T_n = C$ and find each subsequent T_k by subtracting the next term from the outer summand,

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

This yields the upper bound

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

so that x_k is bounded above by $\sqrt{T_k/q_{kk}} - U_k$ and below by $-\sqrt{T_k/q_{kk}} - U_k$. In this way, we iteratively enumerate all vectors $\mathbf{x}$ satisfying $Q(\mathbf{x}) \leq C$, beginning with the entry x_n of $\mathbf{x}$ and working down towards x_1.

3.6.3 Computational remarks and translated lattices

Recall that the Cholesky decomposition of $A = B^T B$ yields the upper triangular matrix R where $A = R^T R$. It is noted in the [**?**] that if we label the columns of R by $\mathbf{r}_i$ and the rows of R^{-1} by $\mathbf{r}'_i$, then

$$x_k^2 = \left({\mathbf{r}'_k}^T \cdot \sum_{i=1}^{n} x_i \mathbf{r}_i \right)^2 \leq {\mathbf{r}'_k}^T \mathbf{r}_k (\mathbf{x}^T R^T R \mathbf{x}) \leq |\mathbf{r}'_k|^2 C.$$

To reduce the search space, it is thus beneficial to reduce the rows of R^{-1}. Furthermore, rearranging the columns of R so that the shortest column vector is first helps reduce the total running time of the Fincke-Pohst algorithm. In particular, doing so leads to progressively smaller intervals in which x_k may exist.

We express this reduction with a unimodular matrix V^{-1} so that $R_1^{-1} = V^{-1} R^{-1}$. Applying an appropriate permutation matrix P, we then reorder the columns of R_1. Since $R_1 = RV$, this yields $R_2 = (RV)P$. Finally, we compute the solutions $\mathbf{y}$ to $\mathbf{y}^T R_2^T R_2 \mathbf{y} \leq C$ and recover the short vectors $\mathbf{x}$ satisfying the original inequality (3.13) via $\mathbf{x} = VP\mathbf{y}$.

As before, let Γ be an n-dimensional lattice with basis matrix B, quadratic form Φ, and associated bilinear form Q. In subsection 3.6.2, it is noted that an implementation of the Fincke-Pohst algorithm is available in Magma. Unfortunately, this implementation does

not support *translated* lattices, a variant of the Fincke-Pohst algorithm which we will need in chapter 5. By a translated lattice, we mean the discrete subgroup of $\mathbb{R}^n$ of the form

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

where $\mathbf{b}_1, \ldots, \mathbf{b}_n$ form the columns of B and $\mathbf{w} \in \mathbb{R}^n$. In the remainder of this section, we describe how to modify the Fincke-Pohst algorithm and its refinements to support translated lattices.

Analogous to the non-translated case, any embedded vector $\mathbf{u}$ of $\Gamma + \mathbf{w}$ may be expressed as $\mathbf{u} = B\mathbf{x} + \mathbf{w}$ for a corresponding coordinate vector $\mathbf{x}$. In this case, we call the vector $\mathbf{u} \in \Gamma + \mathbf{w}$ *small* if

$$(\mathbf{x} - \mathbf{c})^T B^T B (\mathbf{x} - \mathbf{c}) \leq C \tag{3.15}$$

for some $C \geq 0$, where $\mathbf{c} = -\mathbf{w}$.

As in the usual short vectors process, we begin by applying Cholesky decomposition to the positive definite matrix $A = B^T B$ to obtain an upper triangular matrix R satisfying $A = R^T R$. We then generate the matrices R_1, R_2, V, and P described earlier in this section. This allows us to write $A = U^T G U$ for a unimodular matrix U and Gram matrix G given by

$$U = P^{-1} V^{-1} \quad \text{and} \quad G = R_2^T R_2.$$

Thus the inequality (3.15) becomes

$$(\mathbf{y} - \mathbf{d})^T G (\mathbf{y} - \mathbf{d}) \leq C \tag{3.16}$$

where

$$\mathbf{y} = U\mathbf{x} \quad \text{and} \quad \mathbf{d} = U\mathbf{c}.$$

To enumerate the vectors $\mathbf{y}$ which satisfy this inequality, we consider the bilinear form Q associated to the lattice Γ. We express this form as

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

As in the usual Fincke-Pohst algorithm, the coefficients q_{ij} are defined from the matrix R

via equation (3.14). Let

$$U_k = -d_k + \sum_{j=k+1}^{n} q_{kj}(y_j - d_j),$$

where $U_n = -d_n$, and rewrite $Q(\mathbf{y} - \mathbf{d})$ as

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

From here, we proceed as in the usual Fincke-Pohst algorithm described in subsection 3.6.2. Once we compute all vectors $\mathbf{y}$ which satisfy (3.16), we recover $\mathbf{x}$ using $\mathbf{x} = U^{-1}\mathbf{y}$.

As a final remark about Fincke-Pohst for translated lattices, it is worth noting that one could use the variant implemented in Magma simply by increasing the dimension of the lattice Γ and appropriately redefining the basis vectors $\mathbf{b}_i$. This is highly ill-advised as it increases the search space and subsequent running time of the algorithm.

Generally speaking, the use of Fincke-Pohst in our applications poses one of the main bottlenecks in solving Thue-Mahler and Thue-Mahler-like equations. Specifically, this algorithm often yields upwards of hundreds of millions of short vectors, each one needing to be stored and, in our case, appropriately manipulated. This creates both timing and memory problems, often leading to gigabytes of data usage. Deleting these vectors does not release the memory and, as with the class group function, Magma's built-in Fincke-Pohst process cannot be terminated without exiting the program. The primary advantage of implementing and using our own version of Fincke-Pohst, as described in this section, is therefore the ability to add a fail-stop should the number of vectors found become too large.

Chapter 4

Computing Elliptic Curves over $\mathbb{Q}$

In the chapter at hand, we outline an algorithm to compute elliptic curves over $\mathbb{Q}$, based upon techniques of solving Thue-Mahler equations. Our aim is to give a straightforward demonstration of the link between the conductors of the elliptic curves in question and the corresponding equations, and to make the Diophantine approximation problem that follows as easy to tackle as possible. It is worth noting here that these connections are quite straightforward for primes $p > 3$, but require careful analysis at the primes 2 and 3.

4.1 Elliptic curves

Our basic problem is to find a model for each isomorphism class of elliptic curves over $\mathbb{Q}$ with a given conductor. Let $S = \{p_1, p_2, \ldots, p_k\}$ where the p_i are distinct primes, and fix a conductor $N = p_1^{\eta_1} \cdots p_k^{\eta_k}$ for $\eta_i \in \mathbb{N}$. Any curve of conductor N has a minimal model

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

with the a_i integral and discriminant

$$\Delta_E = (-1)^{\delta} p_1^{\gamma_1} \cdots p_k^{\gamma_k},$$

where the γ_i are positive integers satisfying $\gamma_i \geq \eta_i$, for each $i = 1, 2, \ldots, k$, and $\delta \in \{0, 1\}$.

Writing

$$b_2 = a_1^2 + 4a_2, \;\; b_4 = a_1a_3 + 2a_4, \;\; b_6 = a_3^2 + 4a_6, \;\; c_4 = b_2^2 - 24b_4$$

and

$$c_6 = -b_2^3 + 36b_2b_4 - 216b_6,$$

we have $1728\Delta_E = c_4^3 - c_6^2$ and $j_E = c_4^3/\Delta_E$. It follows that

$$c_6^2 = c_4^3 + (-1)^{\delta+1}2^6 \cdot 3^3 \cdot p_1^{\gamma_1} \cdots p_k^{\gamma_k}. \tag{4.1}$$

In fact, it is equation (4.1) that lies at the heart of our method (see also Cremona and Lingham [**?**] for an approach to the problem that takes as its starting point equation (4.1), but subsequently heads in a rather different direction).

Let $\nu_p(x)$ be the largest power of a prime p dividing a nonzero integer x. Since our model is minimal, we may suppose (via Tate's algorithm; see, for example, Papadopoulos [**?**]) that

$$\min\{3\nu_p(c_4), 2\nu_p(c_6)\} < 12 + 12\nu_p(2) + 6\nu_p(3),$$

for each prime p, while

$$\nu_p(N_E) \leq 2 + \nu_p(1728).$$

For future use, it will be helpful to have a somewhat more precise determination of the possible values of $\nu_p(c_4)$ and $\nu_p(c_6)$ we encounter. We compile this data from Papadopoulos [**?**] and summarize it in Tables 4.1, 4.2 and 4.3.

4.2 Cubic forms: the main theorem and algorithm

Having introduced the notation we require for elliptic curves, we now turn our attention to cubic forms and our main result. Fix integers a, b, c and d, and consider the binary cubic form

$$F(x, y) = ax^3 + bx^2y + cxy^2 + dy^3, \tag{4.2}$$

$\nu_2(c_4)$	$\nu_2(c_6)$	$\nu_2(\Delta_E)$	$\nu_2(N)$
0	0	≥ 0	$\min\{1, \nu_2(\Delta_E)\}$
≥ 4	3	0	0
≥ 4	5	4	2, 3 or 4
≥ 4	≥ 6	6	5 or 6
4	6	7	7
4	6	8	2, 3 or 4
4	6	9	5
4	6	10 or 11	3 or 4
4	6	≥ 12	4
5	7	8	7
≥ 6	7	8	2, 3 or 4

$\nu_2(c_4)$	$\nu_2(c_6)$	$\nu_2(\Delta_E)$	$\nu_2(N)$
5	≥ 8	9	8
≥ 6	8	10	6
6	≥ 9	12	5 or 6
6	9	≥ 14	6
7	9	12	5
≥ 8	9	12	4
6	9	13	7
7	10	14	7
7	≥ 11	15	8
≥ 8	10	14	6

Table 4.1: The possible values of $\nu_2(c_4), \nu_2(c_6), \nu_2(\Delta_E)$ and $\nu_2(N)$.

$\nu_3(c_4)$	$\nu_3(c_6)$	$\nu_3(\Delta_E)$	$\nu_3(N)$
0	0	≥ 0	$\min\{1, \nu_3(\Delta_E)\}$
1	≥ 3	0	0
≥ 2	3	3	2 or 3
2	4	3	3
2	≥ 5	3	2
2	3	4	4
2	3	5	3
2	3	≥ 6	2
≥ 3	4	5	5
3	5	6	4

$\nu_3(c_4)$	$\nu_3(c_6)$	$\nu_3(\Delta_E)$	$\nu_3(N)$
3	≥ 6	6	2
≥ 4	5	7	5
≥ 4	6	9	2 or 3
4	7	9	3
4	≥ 8	9	2
4	6	10	4
4	6	11	3
≥ 5	7	11	5
5	8	12	4
≥ 6	8	13	5

Table 4.2: The possible values of $\nu_3(c_4), \nu_3(c_6), \nu_3(\Delta_E)$ and $\nu_3(N)$.

with discriminant

$$D_F = -27a^2d^2 + b^2c^2 + 18abcd - 4ac^3 - 4b^3d. \tag{4.3}$$

To any such form, we can associate a pair of covariants, the Hessian $H = H_F$:

$$H = H_F(x, y) = -\frac{1}{4}\left(\frac{\partial^2 F}{\partial x^2}\frac{\partial^2 F}{\partial y^2} - \left(\frac{\partial^2 F}{\partial x \partial y}\right)^2\right)$$

$\nu_p(c_4)$	$\nu_p(c_6)$	$\nu_p(\Delta_E)$	$\nu_p(N)$
0	0	≥ 1	1
≥ 1	1	2	2
1	≥ 2	3	2
≥ 2	2	4	2
≥ 2	≥ 3	6	2

$\nu_p(c_4)$	$\nu_p(c_6)$	$\nu_p(\Delta_E)$	$\nu_p(N)$
2	3	≥ 7	2
≥ 3	4	8	2
3	≥ 5	9	2
≥ 4	5	10	2

Table 4.3: The possible values of $\nu_p(c_4), \nu_p(c_6), \nu_p(\Delta_E)$ and $\nu_p(N)$ when $p > 3$ is prime and $p \mid \Delta_E$.

and the Jacobian determinant of F and H, a cubic form $G = G_F$ defined by

$$G = G_F(x, y) = \frac{\partial F}{\partial x}\frac{\partial H}{\partial y} - \frac{\partial F}{\partial y}\frac{\partial H}{\partial x}.$$

A quick computation reveals that, explicitly,

$$H = (b^2 - 3ac)x^2 + (bc - 9ad)xy + (c^2 - 3bd)y^2$$

and

$$\begin{aligned} G = (-27a^2d + 9abc - 2b^3)x^3 + (-3b^2c - 27abd + 18ac^2)x^2y \\ +(3bc^2 - 18b^2d + 27acd)xy^2 + (-9bcd + 2c^3 + 27ad^2)y^3. \end{aligned}$$

These satisfy the syzygy

$$4H(x, y)^3 = G(x, y)^2 + 27D_F F(x, y)^2 \tag{4.4}$$

as well as the resultant identities:

$$\mathrm{Res}(F, G) = -8D_F^3 \quad \text{and} \quad \mathrm{Res}(F, H) = D_F^2. \tag{4.5}$$

Note here that we could just as readily work with $-G$ instead of G here (corresponding to taking the Jacobian determinant of H and F, rather than of F and H). Indeed, as we shall observe in Section 4.4.4, for our applications we will, in some sense, need to consider both possibilities.

Notice that if we set $(x, y) = (1, 0)$ and multiply through by $\mathcal{D}^6/4$ (for any rational $\mathcal{D}$),

then this syzygy can be rewritten as

$$(\mathcal{D}^2 H(1,0))^3 - \left(\frac{\mathcal{D}^3}{2} G(1,0)\right)^2 = 1728 \cdot \frac{\mathcal{D}^6 D_F}{256} F(1,0)^2.$$

Given an elliptic curve with corresponding invariants c_4, c_6 and Δ_E, we will show that it is always possible to construct a binary cubic form F, with corresponding $\mathcal{D}$ for which

$$\mathcal{D}^2 H(1,0) = c_4, \;\; -\frac{1}{2}\mathcal{D}^3 G(1,0) = c_6 \;\text{ and }\; \Delta_E = \frac{\mathcal{D}^6 D_F F(1,0)^2}{256}$$

(and hence equation (4.1) is satisfied). This is the basis of the proof of our main result, which provides an algorithm for computing all isomorphism classes of elliptic curves $E/\mathbb{Q}$ with conductor a fixed positive integer N. Though we state our result for curves with $j_E \neq 0$, the case $j_E = 0$ is easy to treat separately (see Section 4.2.1).

Theorem 4.2.1. *Let $E/\mathbb{Q}$ be an elliptic curve of conductor $N = 2^\alpha 3^\beta N_0$, where N_0 is coprime to 6 and $0 \leq \alpha \leq 8$, $0 \leq \beta \leq 5$. Suppose further that $j_E \neq 0$. Then there exists an integral binary cubic form F of discriminant*

$$D_F = sign(\Delta_E) 2^{\alpha_0} 3^{\beta_0} N_1,$$

and relatively prime integers u and v with

$$F(u,v) = \omega_0 u^3 + \omega_1 u^2 v + \omega_2 u v^2 + \omega_3 v^3 = 2^{\alpha_1} \cdot 3^{\beta_1} \cdot \prod_{p | N_0} p^{\kappa_p}, \tag{4.6}$$

such that E is isomorphic over $\mathbb{Q}$ to $E_{\mathcal{D}}$, where

$$E_{\mathcal{D}} \;\; : \;\; 3^{[\beta_0/3]} y^2 = x^3 - 27\mathcal{D}^2 H_F(u,v) x + 27\mathcal{D}^3 G_F(u,v) \tag{4.7}$$

and, for $[r]$ the greatest integer not exceeding a real number r,

$$\mathcal{D} = \prod_{p | \gcd(c_4(E), c_6(E))} p^{\min\{[\nu_p(c_4(E))/2], [\nu_p(c_6(E))/3]\}}. \tag{4.8}$$

The α_0, α_1, β_0, β_1 and N_1 are nonnegative integers satisfying $N_1 \mid N_0$,

$$(\alpha_0, \alpha_1) = \begin{cases} (2,0) \text{ or } (2,3) & \text{if } \alpha = 0, \\ (3, \geq 3) \text{ or } (2, \geq 4) & \text{if } \alpha = 1, \\ (2,1), (4,0) \text{ or } (4,1) & \text{if } \alpha = 2, \\ (2,1), (2,2), (3,2), (4,0) \text{ or } (4,1) & \text{if } \alpha = 3, \\ (2, \geq 0), (3, \geq 2), (4,0) \text{ or } (4,1) & \text{if } \alpha = 4, \\ (2,0) \text{ or } (3,1) & \text{if } \alpha = 5, \\ (2, \geq 0), (3, \geq 1), (4,0) \text{ or } (4,1) & \text{if } \alpha = 6, \\ (3,0) \text{ or } (4,0) & \text{if } \alpha = 7, \\ (3,1) & \text{if } \alpha = 8 \end{cases}$$

and

$$(\beta_0, \beta_1) = \begin{cases} (0,0) & \text{if } \beta = 0, \\ (0, \geq 1) \text{ or } (1, \geq 0) & \text{if } \beta = 1, \\ (3,0), (0, \geq 0) \text{ or } (1, \geq 0) & \text{if } \beta = 2, \\ (\beta, 0) \text{ or } (\beta, 1) & \text{if } \beta \geq 3. \end{cases}$$

The κ_p are nonnegative integers with

$$\nu_p(\Delta_E) = \begin{cases} \nu_p(D_F) + 2\kappa_p & \text{if } p \nmid \mathcal{D}, \\ \nu_p(D_F) + 2\kappa_p + 6 & \text{if } p \mid \mathcal{D} \end{cases} \tag{4.9}$$

and

$$\kappa_p \in \{0,1\} \quad \text{whenever} \quad p^2 \mid N_1. \tag{4.10}$$

Further, we have

$$\text{if} \quad \beta_0 \geq 3, \quad \text{then} \quad 3 \mid \omega_1 \text{ and } 3 \mid \omega_2, \tag{4.11}$$

and

$$\text{if } \nu_p(N) = 1, \text{ for } p \geq 3, \text{ then } p \mid D_F F(u,v). \tag{4.12}$$

Here, as we shall make explicit in the next subsection, the form F corresponding to the curve E in Theorem 4.2.1 determines the 2-division field of E. This connection was noted by Rubin and Silverberg [**?**] in a somewhat different context – they proved that if K is a

field of characteristic $\neq 2, 3$, $F(u, v)$ is a binary cubic form defined over K, E is an elliptic curve defined by $y^2 = F(x, 1)$, and E_0 is another elliptic curve over K with the property that $E[2] \cong E_0[2]$ (as Galois modules), then E_0 is isomorphic to the curve

$$y^2 = x^3 - 3H_F(u, v)x + G_F(u, v),$$

for some $u, v \in K$.

4.2.1 Remarks

Before we proceed, there are a number of observations we should make regarding Theorem 4.2.1.

Historical comments

Theorem 4.2.1 is based upon a generalization of classical work of Mordell [**?**] (see also Theorem 3 of Chapter 24 of Mordell [**?**]), in which the Diophantine equation

$$X^2 + kY^2 = Z^3$$

is treated through reduction to binary cubic forms and their covariants, under the assumption that X and Z are coprime. That this last restriction can, with some care, be eliminated, was noted by Sprindzuk (see Chapter VI of [**?**]). A similar approach to this problem can be made through the invariant theory of binary quartic forms, where one is led to solve, instead, equations of the shape

$$X^2 + kY^3 = Z^3.$$

We will not carry out the analogous analysis here.

2-division fields and reducible forms

It might happen that the form F whose existence is guaranteed by Theorem 4.2.1 is reducible over $\mathbb{Z}[x, y]$. This occurs precisely when the elliptic curve E has a nontrivial rational 2-torsion point. This follows from the more general fact that the cubic form $F(u, v) = \omega_0 u^3 + \omega_1 u^2 v + \omega_2 u v^2 + \omega_3 v^3$ corresponding to an elliptic curve E has the

property that the splitting field of $F(u,1)$ is isomorphic to the 2-division field of E. This is almost immediate from the identity

$$\begin{aligned} 3^3\,\omega_0^2\,F\left(\tfrac{x-\omega_1}{3\omega_0},1\right) &= x^3+(9\omega_0\omega_2-3\omega_1^2)x+27\omega_0^2\omega_3-9\omega_0\omega_1\omega_2+2\omega_1^3 \\ &= x^3-3H_F(1,0)x+G_F(1,0). \end{aligned}$$

Indeed, from (4.7), the elliptic curve defined by the equation $y^2=x^3-3H_F(1,0)x+G_F(1,0)$ is a quadratic twist of that given by the model $y^2=x^3-27c_4(E)x-54c_6(E)$, and hence also of E (whereby they have the same 2-division field).

Imprimitive forms

It is also the case that the cubic forms arising need not be primitive (in the sense that $\gcd(\omega_0,\omega_1,\omega_2,\omega_3)=1$). This situation can occur if each of the coefficients of F is divisible by some integer $g\in\{2,3,6\}$. Since the discriminant is a quartic form in the coefficients of F, for this to take place one requires that

$$D_F\equiv 0 \mod g^4.$$

This is a necessary but not sufficient condition for the form F to be imprimitive. It follows, if we wish to restrict attention to primitive forms in Theorem 4.2.1, that the possible values for $\nu_p(D_F)$ that can arise are

$$\nu_2(D_F)\in\{0,2,3,4\},\ \ \nu_3(D_F)\in\{0,1,3,4,5\} \tag{4.13}$$

$$\text{and } \nu_p(D_F)\in\{0,1,2\},\ \ \text{for } p>3. \tag{4.14}$$

Possible twists

We note that necessarily

$$\mathcal{D}\mid 2^3\cdot 3^2\cdot\prod_{p\mid N_0}p, \tag{4.15}$$

so that, given N, there is a finite set of $E_{\mathcal{D}}$ to consider (we can restrict our attention to quadratic twists of the curve defined via $y^2=x^3-3H_F(1,0)x+G_F(1,0)$, by squarefree divisors of $6N$). In case we are dealing with squarefree conductor N (i.e. for semistable

curves E), then, from Tables 4.1, 4.2 and 4.3, it follows that $\mathcal{D} \in \{1, 2\}$.

Necessity, but not sufficiency

If we search for elliptic curves of conductor N, say, there may exist a cubic form F for which the corresponding Thue-Mahler equation (4.6) has a solution, where all of the conditions of Theorem 4.2.1 are satisfied, but for which the corresponding $E_{\mathcal{D}}$ has conductor $N_{E_{\mathcal{D}}} \neq N$ for all possible $\mathcal{D}$. This can happen when certain local conditions at primes dividing $6N$ are not met; these local conditions are, in practice, easy to check and only a minor issue when performing computations. Indeed, when producing tables of elliptic curves of conductor up to some given bound, we will, in many cases, apply Theorem 4.2.1 to find all curves with good reduction outside a fixed set of primes – in effect, working with multiple conductors simultaneously. For such a computation, the conductor of every twist $E_{\mathcal{D}}$ we encounter will be of interest to us.

Special binary cubic forms

If, for a given binary form $F(x, y) = ax^3+bx^2y+cxy^2+dy^3$, 3 divides both the coefficients b and c (say $b = 3b_0$ and $c = 3c_0$), then $27 \mid D_F$ and, consequently, we can write $D_F = 27\widetilde{D}_F$, where

$$\widetilde{D}_F = -a^2d^2 + 6ab_0c_0d + 3b_0^2c_0^2 - 4ac_0^3 - 4b_0^3d.$$

One can show that the set of binary cubic forms with $b \equiv c \equiv 0 \mod 3$ is closed within the larger set of all binary cubic forms in $\mathbb{Z}[x, y]$, under the action of either $\mathrm{SL}_2(\mathbb{Z})$ or $\mathrm{GL}_2(\mathbb{Z})$. Also note that for such forms we have

$$\widetilde{H}_F(x, y) = \frac{H_F(x, y)}{9} = (b_0^2 - ac_0)x^2 + (b_0c_0 - ad)xy + (c_0^2 - b_0d)y^2$$

and $\widetilde{G}_F(x, y) = G_F(x, y)/27$, so that

$$\begin{aligned}\widetilde{G}_F(x, y) = (-a^2d + 3ab_0c_0 - 2b_0^3)x^3 + 3(-b_0^2c_0 - ab_0d + 2ac_0^2)x^2y \\ +3(b_0c_0^2 - 2b_0^2d + ac_0d)xy^2 + (-3b_0c_0d + 2c_0^3 + ad^2)y^3.\end{aligned}$$

The syzygy now becomes

$$4\widetilde{H}_F(x,y)^3 = \widetilde{G}_F(x,y)^2 + \widetilde{D}_F F(x,y)^2. \tag{4.16}$$

We note, from Theorem 4.2.1, that we will be working exclusively with forms of this shape whenever we wish to treat elliptic curves of conductor $N \equiv 0 \mod 3^3$.

The case $j_E = 0$

This case is treated over a general number field in Proposition 4.1 of Cremona and Lingham [**?**]. The elliptic curves $E/\mathbb{Q}$ with $j_E = 0$ and a given conductor N are particularly easy to determine, since a curve with this property is necessarily isomorphic over $\mathbb{Q}$ to a *Mordell* curve with a model of the shape $Y^2 = X^3 - 54c_6$ where $c_6 = c_6(E)$. Such a model is minimal except possibly at 2 and 3 and has discriminant $-2^6 \cdot 3^9 \cdot c_6^2$ (whereby any primes $p > 2$ which divide c_6 necessarily also divide N). Here, without loss of generality, we may suppose that c_6 is sixth-power-free. Further, from Tables 4.1, 4.2, and 4.3, we have that $\nu_2(N) \in \{0, 2, 3, 4, 6\}$, that $\nu_3(N) \in \{2, 3, 5\}$, and that $\nu_p(N) = 2$ whenever $p \mid N$ for $p > 3$. Given a positive integer N satisfying these constraints, it is therefore a simple matter to check to see if there are elliptic curves $E/\mathbb{Q}$ with conductor N and j-invariant 0. One needs only to compute the conductors of the curves given by $Y^2 = X^3 - 54c_6$ for each sixth-power-free integer (positive or negative) c_6 dividing $64N^3$.

4.2.2 The algorithm

It is straightforward to convert Theorem 4.2.1 into an algorithm for finding all $E/\mathbb{Q}$ of conductor N. We can proceed as follows.

1. Begin by finding all $E/\mathbb{Q}$ of conductor N with $j_E = 0$, as outlined in Section 4.2.1.
2. Next, compute $\mathrm{GL}_2(\mathbb{Z})$-representatives for every binary form F with discriminant

   $$\Delta_F = \pm 2^{\alpha_0} 3^{\beta_0} N_1$$

   for each divisor N_1 of N_0, and each possible pair (α_0, β_0) given in the statement of Theorem 4.2.1 (see (4.13) for specifics). We describe an algorithm for listing these forms in Section 4.4.

3. Solve the corresponding Thue-Mahler equations, finding pairs of integers (u, v) such that $F(u, v)$ is an S-unit, where $S = \{p \text{ prime } : p \mid N\} \cup \{2\}$ and $F(u, v)$ satisfies the additional conditions given in the statement of Theorem 4.2.1.

4. For each cubic form F and pair of integers (u, v), consider the elliptic curve

$$E_1 \quad : \quad y^2 = x^3 - 27H_F(u, v)x + 27G_F(u, v)$$

and all its quadratic twists by squarefree divisors of $6N$. Output those curves with conductor N (if any).

The first, second and fourth steps here are straightforward; the first and second can be done efficiently, while the fourth is essentially trivial. The main bottleneck is step (3). While there is a deterministic procedure for carrying this out (see Tzanakis and de Weger [**?**], [**?**]), it is both involved and, often, computationally taxing.

4.3 Proof of Theorem Theorem 4.2.1

Proof. Given an elliptic curve $E/\mathbb{Q}$ of conductor $N = 2^\alpha 3^\beta N_0$ and invariants $c_4 = c_4(E) \neq 0$ and $c_6 = c_6(E)$, we will construct a corresponding cubic form F explicitly. In fact, our form F will have the property that its leading coefficient will be supported on the primes dividing $6N$, i.e. that

$$F(1, 0) = 2^{\alpha_1} \cdot 3^{\beta_1} \cdot \prod_{p \mid N_0} p^{\kappa_p}.$$

Define $\mathcal{D}$ as in (4.8), i.e. take $\mathcal{D}$ to be the largest integer whose square divides c_4 and whose cube divides c_6. We then set

$$X = c_4/\mathcal{D}^2 \quad \text{and} \quad Y = c_6/\mathcal{D}^3,$$

whereby, from (4.1),

$$Y^2 = X^3 + (-1)^{\delta+1}M, \tag{4.17}$$

for

$$M = \mathcal{D}^{-6} \cdot 2^6 \cdot 3^3 \cdot |\Delta_E|.$$

Note that the assumption that $c_4(E) \neq 0$ ensures that both the j-invariant $j_E \neq 0$ and that $X \neq 0$.

It will prove useful to us later to understand precisely the possible common factors among $X, Y, \mathcal{D}$ and M. For any $p > 3$, we have $\nu_p(N) \leq 2$. When $\nu_p(N) = 1$, from Table 4.3 we find that

$$(\nu_p(\mathcal{D}), \nu_p(X), \nu_p(Y), \nu_p(M)) = (0, 0, 0, \geq 1), \tag{4.18}$$

while, if $\nu_p(N) = 2$, then either

$$\nu_p(\mathcal{D}) = 1 \text{ and } \min\{\nu_p(X), \nu_p(Y)\} = 0,\ \nu_p(M) = 0, \tag{4.19}$$

or

$$\nu_p(\mathcal{D}) \leq 1,\ (\nu_p(X), \nu_p(Y), \nu_p(M)) = (0, 0, \geq 1), (\geq 1, 1, 2), (1, \geq 2, 3) \tag{4.20}$$

$$\text{or } (\geq 2, 2, 4). \tag{4.21}$$

Things are rather more complicated for the primes 2 and 3; we summarize this in Tables 4.4 and 4.5 (which are, in turn, compiled from the data in Tables 4.1 and 4.2).

$\nu_2(N)$	$(\nu_2(X), \nu_2(Y), \nu_2(M), \nu_2(\mathcal{D}))$
0	$(\geq 2, 0, 0, 1)$ or $(0, 0, 6, 0)$
1	$(0, 0, \geq 7, 0)$
2	$(\geq 2, 2, 4, 1), (\geq 2, 1, 2, 2)$ or $(0, 0, 2, 2)$
3	$(\geq 2, 2, 4, 1), (\geq 2, 1, 2, 2)$ or $(0, 0, t, 2), t = 2, 4$ or 5
4	$(\geq 2, 2, 4, 1), (\geq 2, 1, 2, 2), (\geq 2, 0, 0, 3)$ or $(0, 0, t, 2), t = 2$ or $t \geq 4$
5	$(\geq 0, \geq 0, 0, 2), (0, \geq 0, 0, 3), (0, 0, 3, 2)$ or $(1, 0, 0, 3)$
6	$(\geq 0, \geq 0, 0, 2), (0, \geq 0, 0, 3), (\geq 2, 2, 4, 2), (\geq 2, 1, 2, 3)$ or $(0, 0, \geq 2, 3)$
7	$(0, 0, 1, 2), (0, 0, 1, 3), (1, 1, 2, 2)$ or $(1, 1, 2, 3)$
8	$(1, \geq 2, 3, 2)$ or $(1, \geq 2, 3, 3)$.

Table 4.4: The possible values of $\nu_2(N), \nu_2(X), \nu_2(Y), \nu_2(M)$ and $\nu_2(D)$

We will construct a cubic form

$$F_1(x, y) = ax^3 + 3b_0x^2y + 3c_0xy^2 + dy^3,$$

one coefficient at a time; our main challenge will be to ensure that the a, b_0, c_0 and d we produce are actually integral rather than just rational. The form F whose existence is asserted in the statement of Theorem 4.2.1 will turn out to be either F_1 or $F_1/3$.

$\nu_3(N)$	$(\nu_3(X), \nu_3(Y), \nu_3(M), \nu_3(\mathcal{D}))$
0	$(1, \geq 3, 3, 0)$ or $(0, 0, 3, 0)$
1	$(0, 0, \geq 4, 0)$
2	$(\geq 0, 0, 0, 1), (0, \geq 2, 0, 1), (0, 0, \geq 3, 1), (1, \geq 3, 3, 1), (\geq 0, 0, 0, 2)$ or $(0, \geq 2, 0, 2)$
3	$(\geq 0, 0, 0, 1), (\geq 0, 0, 0, 2), (0, 1, 0, 1), (0, 1, 0, 2), (0, 0, 2, 1)$ or $(0, 0, 2, 2)$
4	$(0, 0, 1, 1), (0, 0, 1, 2), (1, 2, 3, 1)$ or $(1, 2, 3, 2)$
5	$(\geq 1, 1, 2, 1), (\geq 1, 1, 2, 2), (\geq 2, 2, 4, 1)$ or $(\geq 2, 2, 4, 2)$.

Table 4.5: The possible values of $\nu_3(N), \nu_3(X), \nu_3(Y), \nu_3(M)$ and $\nu_3(D)$

Let us write

$$M = M_1 \cdot M_2$$

where M_2 is the largest integer divisor of M that is coprime to X, so that

$$M_1 = \prod_{p \,|\, X} p^{\nu_p(M)} \quad \text{and} \quad M_2 = \prod_{p \nmid X} p^{\nu_p(M)}.$$

We define

$$a_1 = \prod_{p | M_1} p^{\left\lceil \frac{\nu_p(M)-1}{2} \right\rceil} \tag{4.22}$$

and set

$$a_2 = \begin{cases} 3^{-1} \prod_{p|M_2} p^{\left\lceil \frac{\nu_p(M)}{2} \right\rceil} & \text{if } \nu_3(X) = 0,\ \nu_3(M) = 2t,\ t \in \mathbb{Z}, t \geq 2, \\ \prod_{p|M_2} p^{\left\lceil \frac{\nu_p(M)}{2} \right\rceil} & \text{otherwise.} \end{cases} \tag{4.23}$$

Define $a = a_1 \cdot a_2$. It follows that $a_1^2 \mid M_1$ and, from (4.18), (4.19), (4.20), and Tables 4.4 and 4.5, that both

$$a_1 \mid X \quad \text{and} \quad a_1^2 \mid Y.$$

We write $X = a_1 \cdot X_1$ and observe that $a_2^2 \mid M_2$. Note that a_2 is coprime to X and hence to a_1. Since $a^2 \mid M$, we may thus define a positive integer K via $K = M/a^2$, so that (4.17) becomes

$$Y^2 - X^3 = (-1)^{\delta+1} K a^2.$$

From the fact that $\gcd(a_2, X) = 1$ and $X \neq 0$, we may choose B so that

$$a_2 B \equiv -Y/a_1 \mod X^3,$$

whereby

$$aB + Y \equiv 0 \mod a_1 X^3. \tag{4.24}$$

Note that, since $a_1^2 \mid Y$ and $a_1 \mid X$, it follows that $a_1 \mid B$. Let us define

$$b_0 = \frac{aB + Y}{X}, \quad c_0 = \frac{b_0^2 - X}{a} \quad \text{and} \quad d = \frac{b_0 c_0 - 2B}{a}. \tag{4.25}$$

We now demonstrate that these are all integers. That $b_0 \in \mathbb{Z}$ is immediate from (4.24). Since $b_0 X - Y = aB$, we know that $b_0 X \equiv Y \mod a$. Squaring both sides thus gives

$$b_0^2 X^2 \equiv Y^2 \equiv X^3 + (-1)^{\delta+1} K a^2 \equiv X^3 \mod a_1 \cdot a_2,$$

and, since $\gcd(a_2, X) = 1$,

$$b_0^2 \equiv X \mod a_2.$$

From (4.24), we have $b_0 \equiv 0 \mod a_1 X^2$, whereby, since $a_1 \mid X$,

$$b_0^2 \equiv X \equiv 0 \mod a_1.$$

The fact that $\gcd(a_1, a_2) = 1$ thus allows us to conclude that $b_0^2 \equiv X \mod a$ and hence that $c_0 \in \mathbb{Z}$.

It remains to show that d is an integer. Let us rewrite ad as

$$ad = b_0 c_0 - 2B = \left(\frac{aB + Y}{aX}\right)\left(\left(\frac{aB + Y}{X}\right)^2 - X\right) - 2B,$$

so that

$$ad = \left(\frac{aB + Y}{aX}\right)\left(\frac{(-1)^{\delta+1} K a^2 + 2aBY + a^2 B^2}{X^2}\right) - 2B.$$

Expanding, we find that

$$X^3 d = (-1)^{\delta+1} KY + 3YB^2 + aB^3 + (-1)^{\delta+1} 3KaB. \tag{4.26}$$

We wish to show that

$$(-1)^{\delta+1}KY + 3YB^2 + aB^3 + (-1)^{\delta+1}3KaB \equiv 0 \mod X^3.$$

From (4.24), we have that

$$(-1)^{\delta+1}KY + 3YB^2 + aB^3 + (-1)^{\delta+1}3KaB \equiv 2Y\left(B^2 + (-1)^{\delta}K\right) \mod a_1X^3.$$

Multiplying congruence (4.24) by $aB - Y$ (which, from our prior discussion, is divisible by a_1^2), we find that

$$a^2B^2 \equiv Y^2 \equiv X^3 + (-1)^{\delta+1}Ka^2 \mod a_1^3X^3$$

and hence, dividing through by a_1^2,

$$a_2^2B^2 \equiv a_1X_1^3 + (-1)^{\delta+1}Ka_2^2 \mod a_1X^3.$$

It follows that

$$B^2 + (-1)^{\delta}K \equiv a_2^{-2}a_1X_1^3 \mod a_1X^3, \tag{4.27}$$

and so, since $a_1^2 \mid Y$,

$$Y\left(B^2 + (-1)^{\delta}K\right) \equiv 0 \mod X^3,$$

whence we conclude that d is an integer, as desired.

With these values of a, b_0, c_0 and d, we can then confirm (with a quick computation) that the cubic form

$$F_1(x, y) = ax^3 + 3b_0x^2y + 3c_0xy^2 + dy^3$$

has discriminant

$$D_{F_1} = \frac{108}{a^2}(X^3 - Y^2) = (-1)^{\delta} \cdot 2^2 \cdot 3^3 \cdot K$$

We also note that

$$F_1(1, 0) = a, \;\; \widetilde{H}_{F_1}(1, 0) = b_0^2 - ac_0 = X$$

and

$$-\frac{1}{2}\widetilde{G}_{F_1}(1, 0) = \frac{1}{2}(a^2d - 3ab_0c_0 + 2b_0^3) = Y,$$

where $\widetilde{G}_F$ and $\widetilde{H}_F$ are as in Section 4.2.1.

Summarizing Table 4.5, we find that we are in one of the following four cases :

(i) $\nu_3(X) = 1,\ \nu_3(Y) = 2,\ \nu_3(M) = 3$ and $\nu_3(N) = 4$,

(ii) $\nu_3(X) \geq 2,\ \nu_3(Y) = 2,\ \nu_3(M) = 4,\ \nu_3(N) = 5$,

(iii) $\nu_3(M) \leq 2$ and $\nu_3(N) \geq 2$, or

(iv) $\nu_3(M) \geq 3$ and either $\nu_3(XY) = 0$ or $\nu_3(X) = 1$, $\nu_3(Y) \geq 3$.

In cases (i), (ii), and (iii), we choose $F = F_1$, i.e.

$$(\omega_0, \omega_1, \omega_2, \omega_3) = (a, 3b_0, 3c_0, d),$$

so that

$$F(1,0) = a, \quad D_F = (-1)^\delta 2^2 \cdot 3^3 \cdot K, \quad c_4 = \mathcal{D}^2 \widetilde{H}_F(1,0)$$

and

$$c_6 = -\frac{1}{2}\mathcal{D}^3 \widetilde{G}_F(1,0).$$

It follows that E is isomorphic over $\mathbb{Q}$ to the curve

$$y^2 = x^3 - 27c_4 x - 54c_6 = x^3 - 3\mathcal{D}^2 H_F(1,0)x + \mathcal{D}^3 G_F(1,0).$$

In case (iv), observe that, from definitions (4.22) and (4.23),

$$\nu_3(a) = \left[\frac{\nu_3(M) - 1}{2}\right] \quad \text{and} \quad \nu_3(K) = \nu_3(M) - 2\nu_3(a), \tag{4.28}$$

so that $3 \mid a$ and $3 \mid K$. From equation (4.26), $3 \mid X^3 d$. If $\nu_3(X) = 0$ this implies that $3 \mid d$. On the other hand, if $\nu_3(X) = 1$, then, from (4.27), we may conclude that $3 \mid B$. Since each of a, B and K is divisible by 3, while $\nu_3(X) = 1$ and $\nu_3(Y) \geq 3$, equation (4.26) once again implies that $3 \mid d$. In this case, we can therefore write $a = 3a_0$ and $d = 3d_0$, for integers a_0 and d_0 and set $F = F_1/3$, i.e. take

$$(\omega_0, \omega_1, \omega_2, \omega_3) = (a_0, b_0, c_0, d_0).$$

We have

$$F(1,0) = a/3, \quad D_F = (-1)^\delta 2^2 \cdot K/3, \quad c_4 = \mathcal{D}^2 H_F(1,0)$$

and

$$c_6 = -\frac{1}{2}\mathcal{D}^3 G_F(1,0).$$

The curve E is now isomorphic over $\mathbb{Q}$ to the model

$$y^2 = x^3 - 27c_4x - 54c_6 = x^3 - 27\mathcal{D}^2 H_F(1,0)x + 27\mathcal{D}^3 G_F(1,0).$$

Since $|D_F|/D_F = (-1)^\delta$ and $a^2K \mid 1728\Delta_E$, we may write

$$F(1,0) = 2^{\alpha_1} \cdot 3^{\beta_1} \cdot \prod_{p \mid N_0} p^{\kappa_p} \quad \text{and} \quad D_F = (|\Delta_E|/\Delta_E) 2^{\alpha_0} 3^{\beta_0} N_1,$$

for nonnegative integers $\alpha_0, \alpha_1, \beta_0, \beta_1, \kappa_p$ and a positive integer N_1, divisible only by primes dividing N_0. More explicitly, we have

$$\alpha_0 = \nu_2(K) + 2 \quad \text{and} \quad \beta_0 = \nu_3(K) + \begin{cases} 3 & \text{in case (i), (ii) or (iii), or} \\ -1 & \text{in case (iv),} \end{cases}$$

and

$$\alpha_1 = \nu_2(a) \quad \text{and} \quad \beta_1 = \nu_3(a) + \begin{cases} 0 & \text{in case (i), (ii) or (iii), or} \\ -1 & \text{in case (iv).} \end{cases}$$

It remains for us to prove that these integers satisfy the conditions listed in the statement of the theorem. It is straightforward to check this, considering in turn each possible triple (X, Y, M) from (4.18), (4.19), (4.20), and Tables 4.4 and 4.5, and using the fact that $K = M/a^2$.

In particular, if $p > 3$, we have $\nu_p(\Delta_E) = 6\nu_p(\mathcal{D}) + \nu_p(D_F) + 2\kappa_p$. From Table 4.3 and (4.8), we have $\nu_p(\mathcal{D}) \leq 1$, whereby (4.9) follows. Further,

$$\nu_p(a) = \begin{cases} \left[\frac{\nu_p(M)-1}{2}\right] & \text{if } p \mid X, \\ \left[\frac{\nu_p(M)}{2}\right] & \text{if } p \nmid X, \end{cases} \tag{4.29}$$

and so, if $p \nmid X$,

$$\nu_p(M) - 2\nu_p(a) \leq 1.$$

Since $a^2K = M$, if $p^2 \mid D_F$, then $\nu_p(N) = 2$ and it follows that we are in case (4.20), with $p \mid X$. We may thus conclude that $\nu_p(M) \in \{2, 3, 4\}$ and hence, from (4.29), that $\nu_p(a) \leq 1$. This proves (4.10).

For (4.11), note that, in cases (i), (ii) and (iii), we clearly have that $3 \mid \omega_1$ and $3 \mid \omega_2$. In case (iv), from (4.28),

$$\beta_0 = \nu_3(D_F) = \nu_3(K) - 1 = \nu_3(M) - 2\left[\frac{\nu_3(M)-1}{2}\right] - 1 \in \{0, 1\}.$$

Finally, to see (4.12), note that if $\nu_p(N) = 1$, for $p > 3$, then we have (4.18) and hence

$$\nu_p(D_F) + 2\nu_p(F(u, v)) = \nu_p(M) \geq 1,$$

whereby $p \mid D_F$ or $p \mid F(u, v)$. We may also readily check that the same conclusion obtains for $p = 3$ (since, equivalently, $\beta_0 + \beta_1 \geq 1$). This completes the proof of Theorem 4.2.1.

□

To illustrate this argument, suppose we consider the elliptic curve (denoted 109a1 in Cremona's database) defined via

$$E \ : \ y^2 + xy = x^3 - x^2 - 8x - 7,$$

with $\Delta_E = -109$. We have

$$c_4(E) = 393 \quad \text{and} \quad c_6(E) = 7803,$$

so that $\gcd(c_4(E), c_6(E)) = 3$. It follows that

$$\mathcal{D} = 1, \ X = 393, \ Y = 7803, \ \delta = 1, \ M = 2^6 \cdot 3^3 \cdot 109,$$

and hence we have

$$M_1 = 3^3, \ M_2 = 2^6 \cdot 109, \ a_1 = 3, \ a_2 = 2^3, \ a = 2^3 \cdot 3 \ \text{ and } \ K = 3 \cdot 109.$$

We solve the congruence $8B \equiv -2601 \mod 393^3$ to find that we may choose $B = 7586982$, so that

$$b_0 = 463347, \ \ c_0 = 8945435084 \quad \text{and} \quad d = 172701687278841.$$

We are in case (iv) and thus set

$$F(x, y) = 8x^3 + 463347x^2y + 8945435084xy^2 + 57567229092947y^3,$$

with discriminant $D_F = -4 \cdot 109$,

$$G_F(1,0) = -15606 = -2c_6(E) \quad \text{and} \quad H_F(1,0) = 393 = c_4(E).$$

The curve E is thus isomorphic to the model

$$E_{\mathcal{D}} \ : \ y^2 = x^3 - 27\mathcal{D}^2 H_F(1,0)x + 27\mathcal{D}^3 G_F(1,0) = x^3 - 10611x - 421362. \tag{4.30}$$

We observe that the form F is $\mathrm{GL}_2(\mathbb{Z})$-equivalent to a "reduced" form (see Section 4.4 for details), given by

$$\tilde{F}(x,y) = x^3 + 3x^2y + 4xy^2 + 6y^3.$$

In fact, this is the only form (up to $\mathrm{GL}_2(\mathbb{Z})$-equivalence) of discriminant $\pm 4 \cdot 109$. We can check that the solutions to the Thue equation $\tilde{F}(u,v) = 8$ are given by $(u,v) = (2,0)$ and $(u,v) = (-7,3)$. The minimal quadratic twist of

$$y^2 = x^3 - 27H_{\tilde{F}}(2,0)x + 27G_{\tilde{F}}(2,0)$$

has conductor $2^5 \cdot 109$ and hence cannot correspond to E. For the solution $(u,v) = (-7,3)$, we find that the curve given by the model

$$y^2 = x^3 - 27H_{\tilde{F}}(-7,3)x + 27G_{\tilde{F}}(-7,3) = x^3 - 10611x + 421362,$$

is the quadratic twist by -1 of the curve (4.30). This situation arises from the fact that G_F is an $\mathrm{SL}_2(\mathbb{Z})$-covariant, but not a $\mathrm{GL}_2(\mathbb{Z})$-covariant of F (we will discuss this more in the next section).

4.4 Finding representative forms

As Theorem 4.2.1 illustrates, we are able to tabulate elliptic curves over $\mathbb{Q}$ with good reduction outside a given set of primes, by finding a set of representatives for $\mathrm{GL}_2(\mathbb{Z})$-equivalence classes of binary cubic forms with certain discriminants, and then solving a number of Thue-Mahler equations. In this section, we will provide a brief description of techniques to find distinguished *reduced* representatives for equivalence classes of cubic forms over a given range of discriminants. For both positive and negative discriminants, the notion of *reduction* arises from associating a particular definite quadratic form to a

given cubic form.

4.4.1 Irreducible Forms

For forms of positive discriminant, there is a well developed classical theory of reduction dating back to work of Hermite [**?**], [**?**] and, later, Davenport (see e.g. [**?**], [**?**] and [**?**]). We can actually apply this method to both reducible and irreducible forms. Initially, though, we will assume the forms are irreducible, since we will treat the elliptic curves corresponding to reducible forms by a somewhat different approach (see Section 4.4.2). Note that when one speaks of "irreducible, reduced forms", as Davenport observes, "the terminology is unfortunate, but can hardly be avoided" ([**?**], page 184).

In each of Belabas [**?**], Belabas and Cohen [**?**] and Cremona [**?**], we find very efficient algorithms for computing cubic forms of both positive and negative discriminant, refining classical work of Hermite, Berwick and Mathews [**?**], and Julia [**?**]. These are readily translated into computer code to loop over valid (a, b, c, d)-values (with corresponding forms $ax^3+bx^2y+cxy^2+dy^3$). The running time in each case is linear in the upper bound X. Realistically, this step (finding representatives for our cubic forms) is highly unlikely to be the bottleneck in our computations.

4.4.2 Reducible forms

One can make similar definitions of reduction for reducible forms (see [**?**] for example). However, for our purposes, it is sufficient to note that a reducible form is equivalent to

$$F(x, y) = bx^2y + cxy^2 + dy^3 \quad \text{with} \quad 0 \leq d \leq c,$$

which has discriminant

$$\Delta_F = b^2(c^2 - 4bd).$$

To find all elliptic curves with good reduction outside $S = \{p_1, p_2, \ldots, p_k\}$, corresponding to reducible cubics in Theorem 4.2.1 (i.e. those E with at least one rational 2-torsion point), it is enough to find all such triples (b, c, d) for which there exist integers x and y so that both

$$b^2(c^2 - 4bd) \quad \text{and} \quad bx^2y + cxy^2 + dy^3$$

are S^*-units (with $S^* = S \cup \{2\}$). For this to be true, it is necessary that each of the integers

$$b, \;\; c^2 - 4bd, \;\; y \quad \text{and} \quad \mu = bx^2 + cxy + dy^2$$

is an S^*-unit. Taking the discriminant of μ as a function of x, we thus require that

$$(c^2 - 4bd)y^2 + 4b\mu = Z^2, \tag{4.31}$$

for some integer Z. This is an equation of the shape

$$X + Y = Z^2 \tag{4.32}$$

in S^*-units X and Y.

While *a priori* equation (4.32) arises as only a necessary condition for the existence of an elliptic curve of the desired form, given any solution to (4.32) in S^*-units X and Y and integer Z, the curves

$$E_1(X,Y) \;\; : \;\; y^2 = x^3 + Zx^2 + \frac{X}{4}x$$

and

$$E_2(X,Y) \;\; : \;\; y^2 = x^3 + Zx^2 + \frac{Y}{4}x$$

have nontrivial rational 2-torsion (i.e. the point corresponding to $(x,y) = (0,0)$) and discriminant X^2Y and XY^2, respectively (and hence good reduction at all primes outside S^*).

4.4.3 Computing forms of fixed discriminant

For our purposes, we will typically compute and tabulate a large list of irreducible forms of absolute discriminant bounded by a given positive number X (of size up to 10^{12} of so, beyond which storage becomes problematical). In certain situations, however, we will want to compute all forms of a given fixed, larger discriminant (perhaps up to size 10^{15}). To carry this out and find desired forms of the shape $ax^3 + bx^2y + cxy^2 + dy^3$, we can argue as in, for example, Cremona [**?**], to restrict our attention to $O(X^{3/4})$ triples (a,b,c). From (4.3), the definition of D_F, we have that

$$d = \frac{9abc - 2b^3 \pm \sqrt{4(b^2 - 3ac)^3 - 27a^2D_F}}{27a^2}$$

and hence it remains to check that the quantity $4(b^2 - 3ac)^3 - 27a^2 D_F$ is an integer square, that the relevant conditions modulo $27a^2$ are satisfied, and that a variety of further inequalities from [**?**] are satisfied. The running time for finding forms with discriminants of absolute value of size X via this approach is of order $X^{3/4}$.

4.4.4 $\mathbf{GL}_2(\mathbb{Z})$ vs $\mathbf{SL}_2(\mathbb{Z})$

One last observation which is very important to make before we proceed, is that while G_F^2 is $\mathrm{GL}_2(\mathbb{Z})$-covariant, the same is not actually true for G_F (it is, however, an $\mathrm{SL}_2(\mathbb{Z})$-covariant). This may seem like a subtle point, but what it means for us in practice is that, having found our $\mathrm{GL}_2(\mathbb{Z})$-representative forms F and corresponding curves of the shape $E_{\mathcal{D}}$ from Theorem 4.2.1, we need, in every case, to also check to see if

$$\tilde{E}_{\mathcal{D}} \quad : \quad 3^{[\beta_0/3]} y^2 = x^3 - 27\mathcal{D}^2 H_F(u,v)x - 27\mathcal{D}^3 G_F(u,v),$$

the quadratic twist of $E_{\mathcal{D}}$ by -1, yields a curve of the desired conductor.

Chapter 5

Towards Efficient Resolution of Thue-Mahler Equations

Let c denote a nonzero integer and let $S = \{p_1, \ldots, p_v\}$ be a set of rational primes. In this section, we specialize the results of chapter 3 to the degree 3 Thue–Mahler equation

$$F(X,Y) = c_0X^3 + c_1X^2Y + c_2XY^2 + c_3Y^3 = cp_1^{Z_1} \cdots p_v^{Z_v}, \tag{5.1}$$

where $(X,Y) \in \mathbb{Z}^2$, $\gcd(X,Y) = 1$, and $Z_i \geq 0$ for $i = 1, \ldots, v$. In particular, to enumerate the set of solutions $\{X, Y, Z_1, \ldots, Z_v\}$ to this equation, we follow section 3.4 to reduce the problem of solving (5.1) to solving finitely many so-called "S-unit" equations

$$\lambda = \delta_1 \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}}\right)^{n_i} \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)^{a_i} - 1 = \delta_2 \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}}\right)^{n_i} \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(i_0)}}{\varepsilon_i^{(j)}}\right)^{a_i}, \tag{5.2}$$

where

$$\delta_1 = \frac{\theta^{(i_0)} - \theta^{(j)}}{\theta^{(i_0)} - \theta^{(k)}} \cdot \frac{\alpha^{(k)}\zeta^{(k)}}{\alpha^{(j)}\zeta^{(j)}}, \quad \delta_2 = \frac{\theta^{(j)} - \theta^{(k)}}{\theta^{(k)} - \theta^{(i_0)}} \cdot \frac{\alpha^{(i_0)}\zeta^{(i_0)}}{\alpha^{(j)}\zeta^{(j)}}$$

are constants. Here, we adopt the notation of chapter 3 and recall that we reduce (5.1) to a homogenous equation of the form

$$f(x,y) = x^3 + C_1x^2y + C_2xy^2 + C_3y^3 = cp_1^{z_1} \cdots p_v^{z_v}, \tag{5.3}$$

where $\gcd(x, y) = 1$ and $\gcd(c, p_i) = 1$ for $i = 1, \ldots, p_v$. Moreover, we set

$$g(t) = f(t, 1) = t^3 + C_1 t^2 + C_2 t + C_3 \tag{5.4}$$

so that $K = \mathbb{Q}(\theta)$ with $g(\theta) = 0$. Recall that ζ in (5.2) denotes a root of unity in K, while $\{\varepsilon_1, \ldots, \varepsilon_r\}$ is a set of fundamental units of $\mathcal{O}_K$. In this case, as K is a degree 3 extension of $\mathbb{Q}$, we either have 3 real embeddings of K into $\mathbb{C}$, or one real embedding of K into $\mathbb{C}$ and a pair of complex conjugate embeddings of K into $\mathbb{C}$. Thus either $r = 1$ or $r = 2$.

In this chapter, we describe new techniques to solve equation (5.2) via a global Weil height.

5.1 Decomposition of the Weil height

The sieves of [**?**] involve logarithms which are of local nature. To obtain a global sieve, we work instead with the global logarithmic Weil height. This height is invariant under conjugation and admits a decomposition into local heights which can be related to complex and p-adic logarithms.

Let $n_1, \ldots, n_\nu, a_1, \ldots, a_r$ be a solution to (5.2) and set $z = \frac{\delta_2}{\lambda}$, where

$$z = \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(j)}}{\gamma_i^{(i_0)}} \right)^{n_i} \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(j)}}{\varepsilon_i^{(i_0)}} \right)^{a_i}.$$

Given the global Weil height of z, or all the local heights of z, we will construct several ellipsoids 'containing' $n_1, \ldots, n_\nu, a_1, \ldots, a_r$ such that the volume of the ellipsoids are as small as possible. We begin by computing the height of z.

Let L be the splitting field of K. Recall that for cubic extensions K, the Galois group $\mathrm{Gal}(L/\mathbb{Q})$ is isomorphic to either the alternating group A_3 or the symmetric group S_3.

Lemma 5.1.1. *Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$ and let $\mathfrak{P}$ denote an ideal of $\mathcal{O}_L$ lying above it. Suppose $\sigma_{i_0} : L \to L$, $\theta \mapsto \theta^{(i_0)}$ and $\sigma_j : L \to L$, $\theta \mapsto \theta^{(j)}$ are two automorphisms of L such that (i_0, j, k) forms a subgroup of $\mathrm{Gal}(L/\mathbb{Q})$ of order 3. Let $\mathfrak{P}^{(i_0)} = \sigma_{i_0}(\mathfrak{P})$ and*

$\mathfrak{P}^{(j)} = \sigma_j(\mathfrak{P})$ *be the prime ideals lying over* $\mathfrak{p}^{(i_0)}$, $\mathfrak{p}^{(j)}$ *respectively. For* $i = 1, \ldots, \nu$,

$$\left(\frac{\gamma_i^{(j)}}{\gamma_i^{(i_0)}}\right)\mathcal{O}_L = \left(\prod_{\mathfrak{P}|\mathfrak{p}_1} \frac{\mathfrak{P}^{(j)\ e(\mathfrak{P}^{(j)}|\mathfrak{p}_1^{(j)})}}{\mathfrak{P}^{(i_0)\ e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_1^{(i_0)})}}\right)^{a_{1i}} \cdots \left(\prod_{\mathfrak{P}|\mathfrak{p}_\nu} \frac{\mathfrak{P}^{(j)\ e(\mathfrak{P}^{(j)}|\mathfrak{p}_\nu^{(j)})}}{\mathfrak{P}^{(i_0)\ e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_\nu^{(i_0)})}}\right)^{a_{\nu i}}$$

where $\mathfrak{P}^{(j)} \neq \mathfrak{P}^{(i_0)}$ *for all* $\mathfrak{P}$ *lying above* $\mathfrak{p}$ *in* K.

Proof. Since

$$(\gamma_i)\mathcal{O}_K = \mathfrak{p}_1^{a_{1i}} \cdots \mathfrak{p}_\nu^{a_{\nu i}},$$

for $i = 1, \ldots, \nu$, where

$$\mathfrak{p}_i\mathcal{O}_L = \prod_{\mathfrak{P}|\mathfrak{p}_i} \mathfrak{P}^{e(\mathfrak{P}|\mathfrak{p}_i)},$$

it holds that

$$(\gamma_i)\mathcal{O}_L = \left(\prod_{\mathfrak{P}|\mathfrak{p}_1} \mathfrak{P}^{e(\mathfrak{P}|\mathfrak{p}_1)}\right)^{a_{1i}} \cdots \left(\prod_{\mathfrak{P}|\mathfrak{p}_\nu} \mathfrak{P}^{e(\mathfrak{P}|\mathfrak{p}_\nu)}\right)^{a_{\nu i}}.$$

Let $\mathfrak{P}^{(i_0)}, \mathfrak{P}^{(j)}$ denote the ideal $\mathfrak{P}$ under the automorphisms of L

$$\sigma_{i_0} : L \to L, \quad \theta \mapsto \theta^{(i_0)} \quad \text{and} \quad \sigma_j : L \to L, \quad \theta \mapsto \theta^{(j)},$$

respectively. That is, $\mathfrak{P}^{(i_0)} = \sigma_{i_0}(\mathfrak{P})$ and $\mathfrak{P}^{(j)} = \sigma_j(\mathfrak{P})$. Then

$$\left(\frac{\gamma_i^{(j)}}{\gamma_i^{(i_0)}}\right)\mathcal{O}_L = \left(\prod_{\mathfrak{P}|\mathfrak{p}_1} \frac{\mathfrak{P}^{(j)\ e(\mathfrak{P}^{(j)}|\mathfrak{p}_1^{(j)})}}{\mathfrak{P}^{(i_0)\ e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_1^{(i_0)})}}\right)^{a_{1i}} \cdots \left(\prod_{\mathfrak{P}|\mathfrak{p}_\nu} \frac{\mathfrak{P}^{(j)\ e(\mathfrak{P}^{(j)}|\mathfrak{p}_\nu^{(j)})}}{\mathfrak{P}^{(i_0)\ e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_\nu^{(i_0)})}}\right)^{a_{\nu i}}.$$

To show that $\mathfrak{P}^{(j)} \neq \mathfrak{P}^{(i_0)}$ for all $\mathfrak{P}$ lying above $\mathfrak{p}$ in K, we consider the decomposition group of $\mathfrak{P}$,

$$D(\mathfrak{P}|p) = \{\sigma \in G \ : \ \sigma(\mathfrak{P}) = \mathfrak{P}\}.$$

Iterating through all possible decompositions of $\mathfrak{p}$ in L, we observe that $\mathfrak{P}^{(i_0)} \neq \mathfrak{P}^{(j)}$ whenever $D(\mathfrak{P}_i|p)$ does not have cardinality 2. Since (i_0, j, k) forms an order 3 subgroup of $\mathrm{Gal}(L/\mathbb{Q})$, it cannot coincide with $D(\mathfrak{P}|p)$ and therefore cannot lead to $\mathfrak{P}^{(i_0)} = \mathfrak{P}^{(j)}$. □

For the remainder of this paper, we assume that (i_0, j, k) are automorphisms of L selected as in Lemma 5.1.1.

Lemma 5.1.2. *Let $\mathfrak{p}$ be a prime ideal of $\mathcal{O}_K$ and let $\mathfrak{P}$ denote an ideal of $\mathcal{O}_L$ lying above it. Let $\mathfrak{P}^{(i_0)} = \sigma_{i_0}(\mathfrak{P})$ and $\mathfrak{P}^{(j)} = \sigma_j(\mathfrak{P})$ be the prime ideals lying over $\mathfrak{p}^{(i_0)}$, $\mathfrak{p}^{(j)}$ respectively. We have*

$$\mathrm{ord}_{\mathfrak{P}}\left(\frac{\delta_2}{\lambda}\right) = \begin{cases} (u_l - r_l)e(\mathfrak{P}^{(j)}|\mathfrak{p}_l^{(j)}) & \text{if } \mathfrak{P}^{(j)} \mid p_l,\ p_l \in \{p_1, \ldots, p_\nu\} \\ (r_l - u_l)e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_l^{(i_0)}) & \text{if } \mathfrak{P}^{(i_0)} \mid p_l,\ p_l \in \{p_1, \ldots, p_\nu\} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. By Lemma 5.1.1, we have

$$\begin{aligned} \left(\frac{\delta_2}{\lambda}\right)\mathcal{O}_L &= \left(\frac{\gamma_1^{(j)}}{\gamma_1^{(i_0)}}\right)^{n_1} \cdots \left(\frac{\gamma_\nu^{(j)}}{\gamma_\nu^{(i_0)}}\right)^{n_\nu} \mathcal{O}_L \\ &= \left(\prod_{\mathfrak{P}|\mathfrak{p}_1} \frac{\mathfrak{P}^{(j)\ e(\mathfrak{P}^{(j)}|\mathfrak{p}_1^{(j)})}}{\mathfrak{P}^{(i_0)\ e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_1^{(i_0)})}}\right)^{u_1 - r_1} \cdots \left(\prod_{\mathfrak{P}|\mathfrak{p}_\nu} \frac{\mathfrak{P}^{(j)\ e(\mathfrak{P}^{(j)}|\mathfrak{p}_\nu^{(j)})}}{\mathfrak{P}^{(i_0)\ e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_\nu^{(i_0)})}}\right)^{u_\nu - r_\nu}. \end{aligned}$$

It follows that

$$\mathrm{ord}_{\mathfrak{P}}\left(\frac{\delta_2}{\lambda}\right) = \begin{cases} (u_l - r_l)e(\mathfrak{P}^{(j)}|\mathfrak{p}_l^{(j)}) & \text{if } \mathfrak{P}^{(j)} \mid p_l,\ p_l \in \{p_1, \ldots, p_\nu\} \\ (r_l - u_l)e(\mathfrak{P}^{(i_0)}|\mathfrak{p}_l^{(i_0)}) & \text{if } \mathfrak{P}^{(i_0)} \mid p_l,\ p_l \in \{p_1, \ldots, p_\nu\} \\ 0 & \text{otherwise.} \end{cases}$$

□

Let $\log^+(\cdot)$ denote the real valued function $\max(\log(\cdot), 0)$ on $\mathbb{R}_{\geq 0}$.

Proposition 5.1.3. *The height $h(z)$ admits a decomposition*

$$h(z) = \frac{1}{[K:\mathbb{Q}]} \sum_{l=1}^{\nu} \log(p_l)|u_l - r_l| + \frac{1}{[L:\mathbb{Q}]} \sum_{w: L \to \mathbb{C}} \log^+ |w(z)|. \tag{5.5}$$

In particular, when $\deg(g(t)) = 3$,

$$\sum_{w: L \to \mathbb{C}} \log^+ |w(z)| \leq 2 \max_{w: L \to \mathbb{C}} \log^+ |w(z)|$$

where this bound becomes an equality if $\mathrm{Gal}(L/\mathbb{Q}) \cong S_3$.

For ease of notation, let $S^* = S \cup \{w : L \to \mathbb{C}\}$ and write

$$\begin{aligned} h(z) &= \frac{1}{[K:\mathbb{Q}]} \sum_{l=1}^{\nu} \log(p_l)|u_l - r_l| + \frac{1}{[L:\mathbb{Q}]} \sum_{w:L\to\mathbb{C}} \log^+ |w(z)| \\ &= \frac{1}{[K:\mathbb{Q}]} \sum_{l=1}^{\nu} h_{p_l}(z) + \frac{1}{[K:\mathbb{Q}]} \sum_{w:L\to\mathbb{C}} h_w(z) \end{aligned}$$

so that

$$h(z) = \frac{1}{[K:\mathbb{Q}]} \sum_{v \in S^*} h_v(z).$$

By Proposition 5.1.3, when $v = p_l$ is a finite place,

$$h_v(z) = \log(p_l)|u_l - r_l|,$$

whereas we write

$$h_v(z) = \frac{1}{[L:K]} \log^+ |w(z)|$$

for all infinite places $v = w : L \to \mathbb{C}$. When $\deg(g(t)) = 3$, we may thus write

$$\begin{aligned} h(z) &= \frac{1}{[K:\mathbb{Q}]} \sum_{l=1}^{\nu} h_{p_l}(z) + \frac{1}{[K:\mathbb{Q}]} \sum_{w:L\to\mathbb{C}} h_w(z) \\ &\leq \frac{1}{[K:\mathbb{Q}]} \left(\sum_{l=1}^{\nu} h_{p_l}(z) + 2 \max_{w:L\to\mathbb{C}} h_w(z) \right). \end{aligned}$$

Proof of Proposition 5.1.3. Since $z \in L$, the definition of the absolute logarithmic Weil height gives

$$h(z) = \frac{1}{[L:\mathbb{Q}]} \sum_{w \in M_L} \log^+ \|z\|_w$$

where $||z||_w$ and M_L are the usual norms and set of inequivalent absolute values on L, respectively. In particular, if $w : L \to \mathbb{C}$ is an infinite place,

$$\log^+ \|z\|_w = \log^+ |w(z)|.$$

If $w = \mathfrak{P}$ is a finite place, we have

$$\log^+ \|z\|_w = \log^+ \left(\frac{1}{N(\mathfrak{P})^{\mathrm{ord}_{\mathfrak{P}}(z)}} \right).$$

Let $p_l \in S$ and $\mathfrak{P}^{(j)} \mid p_l$. Applying Lemma 5.1.2, we obtain

$$\begin{aligned}\log^+ \|z\|_w &= \log^+ \left(\frac{1}{N(\mathfrak{P})^{(u_l - r_l)e(\mathfrak{P}^{(j)} | \mathfrak{p}_l^{(j)})}} \right) \\ &= \max \left\{ -(u_l - r_l) f(\mathfrak{P}^{(j)} \mid p_l) e(\mathfrak{P}^{(j)} \mid \mathfrak{p}_l^{(j)}) \log(p_l), 0 \right\}.\end{aligned}$$

For each $p_l \in S$, there is only one unique prime ideal $\mathfrak{p}_l \in \mathcal{O}_K$ in the ideal equation (3.8) lying above p_l. Hence, each $\mathfrak{P}$ lying over p_l must also lie over $\mathfrak{p}_l$. Now, for $w = \mathfrak{P}^{(j)}$ lying over p_l,

$$\begin{aligned}\sum_{w | \mathfrak{p}_l^{(j)}} \log^+ \|z\|_w &= \max \left\{ (r_l - u_l) \log(p_l), 0 \right\} f(\mathfrak{p}_l^{(j)} \mid p_l)[L : \mathbb{Q}(\theta^{(j)})] \\ &= \max \left\{ (r_l - u_l) \log(p_l), 0 \right\} f(\mathfrak{p}_l^{(j)} \mid p_l)[L : K],\end{aligned}$$

where the last inequality follows from $K = \mathbb{Q}(\theta) \cong \mathbb{Q}(\theta^{(j)})$.

Similarly, applying Lemma 5.1.2 to all $w = \mathfrak{P}^{(i_0)}$ lying over $p_l \in S$, we obtain

$$\sum_{w | \mathfrak{p}_l^{(i_0)}} \log^+ \|z\|_w = \max \left\{ (u_l - r_l) \log(p_l), 0 \right\} f(\mathfrak{p}_l^{(i_0)} \mid p_l)[L : K].$$

Lastly, if $w = \mathfrak{P}$ such that $\mathfrak{P} \neq \mathfrak{P}^{(i_0)}, \mathfrak{P}^{(j)}$, we have $\log^+ \|z\|_w = 0$. Putting this all together yields the first result (5.5).

To prove the second statement, write z as the quotient $z = d^{(j)}/d^{(i_0)} \in L$. The orbit of z is

$$\begin{cases} \left\{ \frac{d^{(j)}}{d^{(i_0)}}, \frac{d^{(k)}}{d^{(j)}}, \frac{d^{(i_0)}}{d^{(k)}}, \frac{d^{(j)}}{d^{(k)}}, \frac{d^{(k)}}{d^{(i_0)}}, \frac{d^{(i_0)}}{d^{(j)}} \right\} & \text{if } \mathrm{Gal}(L/\mathbb{Q}) \cong S_3 \\ \left\{ \frac{d^{(j)}}{d^{(i_0)}}, \frac{d^{(k)}}{d^{(j)}}, \frac{d^{(i_0)}}{d^{(k)}} \right\} & \text{if } \mathrm{Gal}(L/\mathbb{Q}) \cong A_3. \end{cases}$$

Choose $a, b, c \in \{i_0, j, k\}$ such that

$$|d^{(a)}| \geq |d^{(b)}| \geq |d^{(c)}|.$$

If $\mathrm{Gal}(L/\mathbb{Q}) \cong S_3$,

$$\begin{aligned}\sum_{w:L\to\mathbb{C}} \log^+|w(z)| &= \log^+\left|\frac{d^{(a)}}{d^{(b)}}\right| + \log^+\left|\frac{d^{(b)}}{d^{(c)}}\right| + \log^+\left|\frac{d^{(c)}}{d^{(a)}}\right| \\ &\quad + \log^+\left|\frac{d^{(a)}}{d^{(c)}}\right| + \log^+\left|\frac{d^{(a)}}{d^{(c)}}\right| + \log^+\left|\frac{d^{(c)}}{d^{(b)}}\right| \\ &= \log\left|\frac{d^{(a)}}{d^{(b)}}\right| + \log\left|\frac{d^{(b)}}{d^{(c)}}\right| + \log\left|\frac{d^{(a)}}{d^{(c)}}\right| \\ &= 2\log\left|\frac{d^{(a)}}{d^{(c)}}\right| \\ &= 2\max_{w:L\to\mathbb{C}} \log^+|w(z)|.\end{aligned}$$

Alternatively, if $\mathrm{Gal}(L/\mathbb{Q}) \cong A_3$,

$$\begin{aligned}\sum_{w:L\to\mathbb{C}} \log^+|w(z)| &= \log^+\left|\frac{d^{(a)}}{d^{(b)}}\right| + \log^+\left|\frac{d^{(b)}}{d^{(c)}}\right| + \log^+\left|\frac{d^{(c)}}{d^{(a)}}\right| \\ &= \log\left|\frac{d^{(a)}}{d^{(b)}}\right| + \log\left|\frac{d^{(b)}}{d^{(c)}}\right| \\ &\leq 2\max_{w:L\to\mathbb{C}} \log^+|w(z)|.\end{aligned}$$

□

5.2 Initial height bounds

We seek solutions to equaition (5.2). We recall this equation presently,

$$\lambda = \delta_1 \prod_{i=1}^{\nu}\left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}}\right)^{n_i}\prod_{i=1}^{r}\left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)^{a_i} - 1 = \delta_2 \prod_{i=1}^{\nu}\left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}}\right)^{n_i}\prod_{i=1}^{r}\left(\frac{\varepsilon_i^{(i_0)}}{\varepsilon_i^{(j)}}\right)^{a_i}.$$

To simplify notation, we write

$$\tilde{y} = \prod_{i=1}^{\nu}\left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}}\right)^{n_i}\prod_{i=1}^{r}\left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)^{a_i}, \quad \tilde{x} = \prod_{i=1}^{\nu}\left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}}\right)^{n_i}\prod_{i=1}^{r}\left(\frac{\varepsilon_i^{(i_0)}}{\varepsilon_i^{(j)}}\right)^{a_i}$$

so that equation (5.2) becomes

$$\delta_1 \tilde{y} - \delta_2 \tilde{x} = 1. \tag{5.6}$$

Let $z = \frac{1}{\tilde{x}} = \frac{\delta_2}{\lambda}$ and denote by Σ the set of pairs $(\tilde{x}, \tilde{y})$ satisfying (5.6). That is, Σ denotes the set of tuples $(n_1, \ldots, n_\nu, a_1, \ldots, a_r)$ corresponding to $(\tilde{x}, \tilde{y})$ which satisfy (5.6).

Let $\mathbf{l}, \mathbf{h} \in \mathbb{R}^{\nu+m}$ with $\mathbf{0} \leq \mathbf{l} \leq \mathbf{h}$. Then we define $\Sigma(\mathbf{l}, \mathbf{h})$ as the set of all $(\tilde{x}, \tilde{y}) \in \Sigma$ such that $(h_v(z)) \leq \mathbf{h}$ and such that $(h_v(z)) \not\leq \mathbf{l}$, and write $\Sigma(\mathbf{h}) = \Sigma(\mathbf{l}, \mathbf{h})$ if $\mathbf{l} = \mathbf{0}$. Additionally, for each place w, we denote by $\Sigma_w(\mathbf{l}, \mathbf{h})$ the set of all $(\tilde{x}, \tilde{y}) \in \Sigma(\mathbf{h})$ such that $h_w(z) > l_w$.

Recall the minimal polynomial $g(t)$ of K, (5.4), derived from

$$f(x, y) = x^3 + C_1 x^2 y + C_2 x y^2 + C_3 y^3 = c p_1^{z_1} \cdots p_v^{z_v}.$$

For $S = \{p_1, \ldots, p_v\}$, let $N_S = \prod_{p \in S} p$ and set

$$b_S = 1728 N_S^2 \prod_{p \notin S} p^{\min(2, \mathrm{ord}_p(b))}$$

for any integer b. In particular, we take $b = 432\Delta c^2$ with Δ the discriminant of f. Denote by $h(f - c)$ the maximum logarithmic Weil heights of the coefficients of the polynomial $f - c$,

$$h(f - c) = \max(\log |C_1|, \log |C_2|, \log |C_3|, \log |c|).$$

Now, setting

$$\Omega' = 2 b_S \log(b_S) + 172 h(f - c),$$

we obtain, by Corollary J (ii) of [**?**], the following height bound on any solution (x, y) of (5.3)

$$\max(h(x), h(y)) \leq \Omega'.$$

To translate this result for use with our logarithmic Weil height (5.5), we have the following lemma.

Lemma 5.2.1. *Let* $\mathbf{m} = (n_1, \ldots, n_\nu, a_1, \ldots, a_r) \in \mathbb{R}^{r+\nu}$ *be any solution of* (5.2) *and let*

$$\Omega = [K : \mathbb{Q}](2h(\alpha) + 4\Omega' + 2h(\theta) + 2\log(2)). \tag{5.7}$$

If $\mathbf{h} \in \mathbb{R}^{\nu+m}$ *with* $\mathbf{h} = (\Omega)$, *then* $\mathbf{m} \in \Sigma(h)$.

Proof. Let $(\tilde{x}, \tilde{y}) \in \Sigma$. We show that the corresponding value $z = \frac{1}{\tilde{x}} = \frac{\delta_2}{\lambda}$ arising from this choice of $\tilde{x}, \tilde{y}$ satisfies

$$\mathbf{0} < (h_v(z)) \leq \mathbf{h}.$$

As stated earlier, any solution x, y of $f(x, y) = cp_1^{z_1} \cdots p_v^{z_v}$ satisfies

$$\max(h(x), h(y)) \leq \Omega'.$$

Taking the height of

$$\beta = x - y\theta = \alpha\zeta\varepsilon_1^{a_1} \cdots \varepsilon_r^{a_r} \cdot \gamma_1^{n_1} \cdots \gamma_\nu^{n_\nu},$$

we obtain

$$h(\beta) = h(x) + h(\theta) + h(y) + \log 2 \leq 2\Omega' + h(\theta) + \log 2.$$

In particular, as $h(\beta) = h(\beta^{(i)})$,

$$h(\beta^{(i)}) \leq 2\Omega' + h(\theta) + \log 2.$$

Now,

$$\delta_2\tilde{x} = \frac{\theta^{(j)} - \theta^{(k)}}{\theta^{(k)} - \theta^{(i_0)}} \cdot \frac{\alpha^{(i_0)}\zeta^{(i_0)}}{\alpha^{(j)}\zeta^{(j)}} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}} \right)^{n_i} \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(i_0)}}{\varepsilon_i^{(j)}} \right)^{a_i} = \frac{\theta^{(j)} - \theta^{(k)}}{\theta^{(k)} - \theta^{(i_0)}} \cdot \frac{\beta^{(i_0)}}{\beta^{(j)}},$$

meaning that $\tilde{x}$ may be written as

$$\tilde{x} = \frac{\beta^{(i_0)}}{\beta^{(j)}} \cdot \frac{\alpha^{(j)}\zeta^{(j)}}{\alpha^{(i_0)}\zeta^{(i_0)}}.$$

Hence,

$$h(\tilde{x}) = 2h(\beta) + 2h(\alpha) \leq 4\Omega' + 2h(\theta) + 2\log 2 + 2h(\alpha).$$

Finally, we observer that

$$h(z) = h(1/\tilde{x}) \leq 4\Omega' + 2h(\theta) + 2\log 2 + 2h(\alpha).$$

Together with $\dfrac{1}{[K : \mathbb{Q}]} h_v(z) \leq h(z)$, this implies

$$h_v(z) \leq [K : \mathbb{Q}](4\Omega' + 2h(\theta) + 2\log 2 + 2h(\alpha)) = \Omega.$$

Of course, by definition, we have $h_v(z) \geq 0$, so that $(\tilde{x}, \tilde{y}) \in \Sigma(h)$ as required. □

5.3 Coverings of Σ

From section 5.2, we now know that all solutions $(\tilde{x}, \tilde{y}) \in \Sigma$ satisfy $\mathbf{m} \in \Sigma(h)$ if $\mathbf{h} = (\Omega')$. In the notation of section 5.2, we have the following result.

Lemma 5.3.1. *Let* $\mathbf{l}, \mathbf{h} \in \mathbb{R}^{\nu+m}$ *with* $\mathbf{0} \leq \mathbf{l} \leq \mathbf{h}$. *It holds that* $\Sigma(\mathbf{h}) = \Sigma(\mathbf{l}, \mathbf{h}) \cup \Sigma(\mathbf{l})$ *and* $\Sigma(\mathbf{l}, \mathbf{h}) = \cup_{v \in S^*} \Sigma_v(\mathbf{l}, \mathbf{h})$.

Proof. Suppose $(\tilde{x}, \tilde{y}) \in \Sigma(\mathbf{h})$. By definition this means, $(h_v(z)) \leq \mathbf{h}$ and that $h_v(z) > 0$ for at least one coordinate v. Since $\mathbf{0} \leq \mathbf{l} \leq \mathbf{h}$, it follows that either $(h_v(z)) \leq \mathbf{l}$ or $(h_v(z)) \not\leq \mathbf{l}$. That is, either all coordinates satisfy $h_v(z) \leq l_v$, or there is at least one coordinate for which $h_v(z) > l_v$. This means that either $(\tilde{x}, \tilde{y}) \in \Sigma(\mathbf{l})$ or $(\tilde{x}, \tilde{y}) \in \Sigma(\mathbf{l}, \mathbf{h})$, and so $\Sigma(\mathbf{h}) \subseteq \Sigma(\mathbf{l}, \mathbf{h}) \cup \Sigma(\mathbf{l})$.

Conversely, suppose $(\tilde{x}, \tilde{y}) \in \Sigma(\mathbf{l}, \mathbf{h}) \cup \Sigma(\mathbf{l})$. It follows that either $(h_v(z)) \leq \mathbf{h}$ and $(h_v(z)) \not\leq \mathbf{l}$ or $(h_v(z)) \leq \mathbf{l}$ and $(h_v(z)) \not\leq \mathbf{0}$. In either case, this means that $(h_v(z)) \leq \mathbf{h}$ and $(h_v(z)) \not\leq \mathbf{0}$. Hence $(\tilde{x}, \tilde{y}) \in \Sigma(\mathbf{h})$ and $\Sigma(\mathbf{h}) \supseteq \Sigma(\mathbf{l}, \mathbf{h}) \cup \Sigma(\mathbf{l})$.

To prove the second equality, let $(\tilde{x}, \tilde{y}) \in \Sigma(\mathbf{l}, \mathbf{h})$. Then there exists $w \in S^*$ with $h_w(z) > l_w$ so that $(\tilde{x}, \tilde{y})$ lies in $\Sigma_w(\mathbf{l}, \mathbf{h})$. Hence $\Sigma(\mathbf{l}, \mathbf{h}) \subseteq \cup_{v \in S^*} \Sigma_v(\mathbf{l}, \mathbf{h})$. Lastly, since each set $\Sigma_v(\mathbf{l}, \mathbf{h})$ is contained in $\Sigma(\mathbf{l}, \mathbf{h})$ it follows that $\Sigma(\mathbf{l}, \mathbf{h}) = \cup_{v \in S^*} \Sigma_v(\mathbf{l}, \mathbf{h})$ as required. □

Let $\mathbf{h}_0 = (\Omega', \ldots, \Omega')$ denote the vector consisting of the initial bound Ω'. By Proposition 5.2.1, every solution of (5.2) is contained in $\mathbf{h}_0$. Therefore, we write $\Sigma = \Sigma(\mathbf{h}_0)$. Consider the pairs $(\mathbf{l}_n, \mathbf{h}_n) \in \mathbb{R}^{\nu+m} \times \mathbb{R}^{\nu+m}$ with $\mathbf{0} \leq \mathbf{l}_n \leq \mathbf{h}_n$ and $\mathbf{h}_{n+1} = \mathbf{l}_n$ for $n = 0, \ldots, N$. Then we can cover Σ:

$$\Sigma = \Sigma(\mathbf{l}_N) \cup \left(\cup_{n=0}^{N} \cup_{v \in S^*} \Sigma_v(\mathbf{l}_n, \mathbf{h}_n) \right).$$

Indeed this follows directly by applying Lemma 5.3.1 N times. In particular, Lemma 5.3.1 gives

$$\Sigma = \Sigma(\mathbf{h}_0), \quad \Sigma(\mathbf{h}) = \Sigma(\mathbf{l}, \mathbf{h}) \cup \Sigma(\mathbf{l}) \quad \text{and} \quad \Sigma(\mathbf{l}, \mathbf{h}) = \cup_{v \in S^*} \Sigma_v(\mathbf{l}, \mathbf{h}).$$

After choosing a good sequence of lower and upper bounds $\mathbf{l}_n, \mathbf{h}_n$ covering the whole space Σ, we are reduced to computing $\Sigma_v(\mathbf{l}, \mathbf{h})$ for each $v \in S^*$. In the following section, we construct the ellipsoids associated to each $\Sigma_v(\mathbf{l}, \mathbf{h})$, after which we describe the sieve allowing us to compute the solutions of each $\Sigma_v(\mathbf{l}, \mathbf{h})$.

5.4 Construction of the ellipsoids

In section 5.3, we establish that for a suitable pair of vectors $\mathbf{l}, \mathbf{h}$, solving (5.2) reduces to computing $\Sigma_v(\mathbf{l}, \mathbf{h})$ for each $v \in S^*$. In this section, we construct the ellipsoids associated to each $\Sigma_v(\mathbf{l}, \mathbf{h})$, which will subsequently allow us to compute all solutions of $\Sigma_v(\mathbf{l}, \mathbf{h})$.

We begin with the quadratic form $q_f = A^T D^2 A$ on $\mathbb{Z}^\nu$, where D is a $\nu \times \nu$ diagonal matrix with diagonal entries $\log^*(p_i)$ for $p_i \in S$. Here, by $\log^*(p_i)$ we mean the best continued fraction approximation P_n/Q_n to $\log(p_i)$ as determined by the precision on $\log(p_i)$, and such that $P_n/Q_n \leq \log(p_i)$. That is, the convergent P_n/Q_n has n odd.

Recall that A is the matrix generated in either subsection 3.4.1 or subsection 3.4.1. As A is invertible, our choice of entries in D guarantees that this quadratic form is positive definite. This will become very important later in the sieve when we will need to apply many instances of the Fincke-Pohst algorithm.

Lemma 5.4.1. *For any solution* $(n_1, \ldots, n_\nu, a_1, \ldots, a_r)$ *of* (5.2) *with* $\mathbf{n} = (n_1, \ldots, n_\nu)$, *we have*

$$q_f(\mathbf{n}) \leq \sum_{l=1}^{\nu} \log(p_l)^2 |u_l - r_l|^2.$$

Proof. Recall from subsection 3.4.1 and subsection 3.4.2 that

$$A\mathbf{n} = \mathbf{u} - \mathbf{r}.$$

$$q_f(\mathbf{n}) = (A\mathbf{n})^{\mathrm{T}} D^2 A\mathbf{n} = (\mathbf{u} - \mathbf{r})^{\mathrm{T}} D^2 (\mathbf{u} - \mathbf{r}) = \sum_{l=1}^{\nu} \log^*(p_l)^2 |u_l - r_l|^2 \leq \sum_{l=1}^{\nu} \log(p_l)^2 |u_l - r_l|^2.$$

□

We now briefly re-examine the decomposition of $h(z)$ into local heights,

$$\begin{aligned} h(z) &= \frac{1}{[K:\mathbb{Q}]} \sum_{l=1}^{\nu} \log(p_l)|u_l - r_l| + \frac{1}{[L:\mathbb{Q}]} \sum_{w:L\to\mathbb{C}} \log^+ |w(z)| \\ &= \frac{1}{[K:\mathbb{Q}]} \left(\sum_{l=1}^{\nu} \log(p_l)|u_l - r_l| + \frac{1}{[L:K]} \sum_{w:L\to\mathbb{C}} \log^+ |w(z)| \right) \\ &\leq \frac{1}{[K:\mathbb{Q}]} \left(\sum_{l=1}^{\nu} \log(p_l)|u_l - r_l| + 2 \max_{w:L\to\mathbb{C}} \frac{\log^+ |w(z)|}{[L:K]} \right) \end{aligned}$$

For every finite place v, Lemma 5.4.1 tells us that any set of bounds $\{h_v\}_{v\in S}$ on the set $\{h_v(z)\}_{v\in S}$ yields a bound on $q_f(\mathbf{n})$. In particular, we have

$$q_f(\mathbf{n}) \leq \sum_{l=1}^{\nu} \log(p_l)^2 |u_l - r_l|^2 \leq \sum_{l=1}^{\nu} h_l^2.$$

In the remainder of this section, we build analogous bounds on the exponents $a_1, \ldots, a_r$ of the fundamental units.

Recall $r = 1$ or $r = 2$ for the degree 3 Thue-Mahler equation (5.3) in question. Choose a set I of embeddings $L \to \mathbb{C}$ of cardinality r. For $r = 1$, consider the matrix

$$R = \left(\log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1} \right| \right),$$

where $I = \{\iota_1\}$. Clearly, as long as we choose ι_1 such that $\log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1} \right| \neq 0$, this matrix is invertible.

When $r = 2$, we let I be the set of embeddings $L \to \mathbb{C}$ of cardinality 2 such that for any $\alpha \in K$, it holds that $I\alpha^{(i_0)} \cup I\alpha^{(j)} = \mathrm{Gal}(L/\mathbb{Q})\alpha$. For $I = \{\iota_1, \iota_2\}$, let R be the 2×2 matrix

$$R = \begin{pmatrix} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1} \right| & \log \left| \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_1} \right| \\ \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_2} \right| & \log \left| \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_2} \right| \end{pmatrix}.$$

Lemma 5.4.2. *When $r = 2$, the matrix R has an inverse,*

$$R^{-1} = \begin{pmatrix} \overline{r}_{11} & \overline{r}_{12} \\ \overline{r}_{21} & \overline{r}_{22} \end{pmatrix}.$$

Proof. Suppose that $\mathbf{m} \in \mathbb{Z}^2$ satisfies $R\mathbf{m} = \mathbf{0}$. Then for each $\iota \in I$ it holds that

$$m_1 \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota} \right| + m_2 \log \left| \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota} \right| = 0,$$

and hence

$$\left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota} \right|^{m_1} \cdot \left| \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota} \right|^{m_2} = 1.$$

This together with $I(i) \cup I(j) = Gal(L/\mathbb{Q})$ implies that all conjugates of $\alpha = \varepsilon_1^{m_1} \varepsilon_2^{m_2}$ have the same absolute value. Since all ε_i are units of $\mathcal{O}_K$, it follows that $|\alpha|^{[L:\mathbb{Q}]} = N(\alpha) = 1$ and hence α is a root of unity in K. On using that the elements ε_i are multiplicatively independent, we obtain that $\mathbf{m} = \mathbf{0}$. Then linear algebra gives $R^{-1} \in \mathbb{R}^{2\times 2}$, completing the proof. □

For the remainder of this chapter, we specialize to the real case, $r = 2$. The setup for $r = 1$ follows closely the work described here, yet poses other difficulties when defining the corresponding sieves.

Now, for any solution $(n_1, \ldots, n_\nu, a_1, a_2)$ of (5.2), set

$$\varepsilon = \begin{pmatrix} a_1 & a_2 \end{pmatrix}^{\mathrm{T}}.$$

We have

$$R\varepsilon = \begin{pmatrix} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1 \ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_1 \ a_2} \right| \\ \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_2 \ a_2} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_2 \ a_2} \right| \end{pmatrix}.$$

Since R is invertible with $R^{-1} = (\overline{r}_{nm})$, we find

$$\varepsilon = \begin{pmatrix} \overline{r}_{11} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1 \ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_1 \ a_2} \right| + \overline{r}_{12} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_2 \ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_2 \ a_2} \right| \\ \overline{r}_{21} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1 \ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_1 \ a_2} \right| + \overline{r}_{22} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_2 \ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_2 \ a_2} \right| \end{pmatrix},$$

giving

$$a_l = \overline{r}_{l1} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1\ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_1\ a_2} \right| + \overline{r}_{l2} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_2\ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_2\ a_2} \right|$$

for $l = 1, 2$.

To estimate $|a_l|$, we begin to estimate the sum on the right hand side. For this, we consider

$$z = \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(j)}}{\gamma_i^{(i_0)}} \right)^{n_i} \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{a_1} \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{a_2}.$$

For any embedding $\iota : L \to \mathbb{C}$, we have

$$(z)^{\iota} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}} \right)^{\iota\ n_i} = \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota\ a_1} \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota\ a_2}.$$

In particular

$$\left| (z)^{\iota} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(i_0)}}{\gamma_i^{(j)}} \right)^{\iota\ n_i} \right| = \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota\ a_1} \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota\ a_2} \right|,$$

so that

$$\log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota\ a_1} \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota\ a_2} \right| = \log |\iota(z)| - \log \left| \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(j)}}{\gamma_i^{(i_0)}} \right)^{\iota\ n_i} \right|.$$

Hence, for $l = 1, 2$,

$$\begin{aligned} a_l &= \overline{r}_{l1} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_1\ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_1\ a_2} \right| + \overline{r}_{l2} \log \left| \left(\frac{\varepsilon_1^{(j)}}{\varepsilon_1^{(i_0)}} \right)^{\iota_2\ a_1} \cdot \left(\frac{\varepsilon_2^{(j)}}{\varepsilon_2^{(i_0)}} \right)^{\iota_2\ a_2} \right| \\ &= \overline{r}_{l1} \left(\log |\iota_1(z)| - \log \left| \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(j)}}{\gamma_i^{(i_0)}} \right)^{\iota_1\ n_i} \right| \right) + \\ &\quad + \overline{r}_{l2} \left(\log |\iota_2(z)| - \log \left| \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(j)}}{\gamma_i^{(i_0)}} \right)^{\iota_2\ n_i} \right| \right) \\ &= \overline{r}_{l1} \log |\iota_1(z)| + \overline{r}_{l2} \log |\iota_2(z)| - n_1 \beta_{\gamma_1 l} - \cdots - n_\nu \beta_{\gamma_\nu l}, \end{aligned}$$

where

$$\beta_{\gamma_k l} = \left(\overline{r}_{l1} \log \left| \iota_1 \left(\frac{\gamma_k^{(j)}}{\gamma_k^{(i_0)}} \right) \right| + \overline{r}_{l2} \log \left| \iota_2 \left(\frac{\gamma_k^{(j)}}{\gamma_k^{(i_0)}} \right) \right| \right)$$

for $k = 1, \ldots, \nu$. Recall that $\mathbf{n} = A^{-1}(\mathbf{u} - \mathbf{r})$ and suppose $A^{-1} = (\overline{a}_{nm})$. We have

$$\mathbf{n} = A^{-1}(\mathbf{u} - \mathbf{r}) = \begin{pmatrix} \sum_{k=1}^{\nu} \overline{a}_{1k}(u_k - r_k) \\ \vdots \\ \sum_{k=1}^{\nu} \overline{a}_{\nu k}(u_k - r_k) \end{pmatrix},$$

so that we may rewrite each a_l as

$$a_l = \overline{r}_{l1} \log |\iota_1(z)| + \overline{r}_{l2} \log |\iota_2(z)| - \sum_{k=1}^{\nu} (u_k - r_k)\alpha_{\gamma lk},$$

where

$$\alpha_{\gamma lk} = \overline{a}_{1k}\beta_{\gamma_1 l} + \cdots + \overline{a}_{\nu k}\beta_{\gamma_\nu l}.$$

Taking absolute values, we obtain

$$|a_l| \leq |\overline{r}_{l1}||\log |\iota_1(z)|| + |\overline{r}_{l2}||\log |\iota_2(z)|| + \sum_{k=1}^{\nu} |u_k - r_k||\alpha_{\gamma lk}|.$$

Suppose $\log |\iota_1(z)| \geq 0$ and $\log |\iota_2(z)| \geq 0$. Then

$$\begin{aligned} |a_l| &\leq |\overline{r}_{l1}| \log |\iota_1(z)| + |\overline{r}_{l2}| \log |\iota_2(z)| + \sum_{k=1}^{\nu} |u_k - r_k||\alpha_{\gamma lk}| \\ &\leq \max\{|\overline{r}_{l1}|, |\overline{r}_{l2}|\} \sum_{w : L \to \mathbb{C}} \log^+ |w(z)| + \sum_{k=1}^{\nu} |u_k - r_k||\alpha_{\gamma lk}|. \end{aligned}$$

Applying Proposition 5.1.3 yields

$$|a_l| \leq 2 \max\{|\overline{r}_{l1}|, |\overline{r}_{l2}|\} \max_{w : L \to \mathbb{C}} \log^+ |w(z)| + \sum_{k=1}^{\nu} |u_k - r_k||\alpha_{\gamma lk}|.$$

Alternatively, suppose that both $\log |\iota_1(z)| < 0$ and $\log |\iota_2(z)| < 0$. We recall that z is a quotient of elements which are conjugate to one another. By taking the norm of z in L, we obtain $N(z) = 1$. On the other hand, by definition, we have

$$1 = N(z) = \prod_{w : L \to \mathbb{C}} w(z).$$

Taking absolute values and logarithms,

$$0 = \sum_{w:L\to\mathbb{C}} \log|w(z)|$$

so that

$$-\log|\iota(z)| = \sum_{\substack{w\,:\,L\,\to\,\mathbb{C}\\ w\,\neq\,\iota}} \log|w(z)|.$$

In our present case, we use this equivalence to obtain a bound on $|a_l|$ as follows.

$$\begin{aligned}|a_l| &\leq |\overline{r}_{l1}| \sum_{\substack{w\,:\,L\,\to\,\mathbb{C}\\ w\,\neq\,\iota_1}} \log|w(z)| - |\overline{r}_{l2}|\log|\iota_2(z)| + \sum_{k=1}^{\nu}|u_k - r_k||\alpha_{\gamma lk}| \\ &\leq 2\max\{|\overline{r}_{l1}|,|\overline{r}_{l2}|\}\max_{w:L\to\mathbb{C}}\log^+|w(z)| + \sum_{k=1}^{\nu}|u_k - r_k||\alpha_{\gamma lk}|.\end{aligned}$$

This bound is not correct in general, and there is something wrong...

Here, the second inequality follows again by Proposition 5.1.3. Lastly, if, without loss of generality, we have $\log|\iota_1(z)| < 0$ and $\log|\iota_2(z)| \geq 0$, then

$$\begin{aligned}|a_l| &\leq |\overline{r}_{l1}| \sum_{\substack{w\,:\,L\,\to\,\mathbb{C}\\ w\,\neq\,\iota_1}} \log|w(z)| + |\overline{r}_{l2}|\log|\iota_2(z)| + \sum_{k=1}^{\nu}|u_k - r_k||\alpha_{\gamma lk}| \\ &\leq 3\max\{|\overline{r}_{l1}|,|\overline{r}_{l2}|\}\max_{w:L\to\mathbb{C}}\log^+|w(z)| + \sum_{k=1}^{\nu}|u_k - r_k||\alpha_{\gamma lk}|.\end{aligned}$$

Now, let

$$w_{\varepsilon l} = 3\max\{|\overline{r}_{l1}|,|\overline{r}_{l2}|\}, \tag{5.8}$$

and

$$w_{\gamma lk} = \frac{|\alpha_{\gamma lk}|}{\log(p_k)}, \tag{5.9}$$

where

$$\alpha_{\gamma lk} = \overline{a}_{1k}\beta_{\gamma_1 l} + \cdots + \overline{a}_{\nu k}\beta_{\gamma_\nu l},$$

and

$$\beta_{\gamma_k l} = \left(\overline{r}_{l1}\log\left|\left(\frac{\gamma_k^{(j)}}{\gamma_k^{(i_0)}}\right)^{\iota_1}\right| + \overline{r}_{l2}\log\left|\left(\frac{\gamma_k^{(j)}}{\gamma_k^{(i_0)}}\right)^{\iota_2}\right|\right)$$

for $k = 1, \dots, \nu$. We have proven the following lemma.

Lemma 5.4.3. *For any solution $(n_1, \dots, n_\nu, a_1, \dots, a_r)$ of (5.2), for $l = 1, 2$, we have*

$$|a_l| \le w_{\varepsilon l} \max_{w: L \to \mathbb{C}} \log^+ |w(z)| + \sum_{k=1}^{\nu} w_{\gamma lk} \log(p_k) |u_k - r_k|.$$

5.4.1 The Archimedean ellipsoid: the real case

didn't update this section

Let $\tau : L \to \mathbb{R} \subset \mathbb{C}$ be an embedding and let $l_\tau \ge c_\tau$ and $c > 0$ be given real numbers for $c_\tau = \log^+(2|\tau(\delta_2)|)$. We define

$$\alpha_0 = [c \log |\tau(\delta_1)|], \quad \alpha_{\varepsilon 1} = \left[c \log \left| \tau \left(\frac{\varepsilon_1^{(k)}}{\varepsilon_1^{(j)}} \right) \right| \right], \quad \alpha_{\varepsilon 2} = \left[c \log \left| \tau \left(\frac{\varepsilon_2^{(k)}}{\varepsilon_2^{(j)}} \right) \right| \right]. \tag{5.10}$$

For $i = 1, \dots, \nu$, define

$$\alpha_{\gamma i} = \left[c \log \left| \tau \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}} \right) \right| \right]. \tag{5.11}$$

Here, $[\,\cdot\,]$ denotes the nearest integer function.

Let

$$w_\varepsilon = \frac{w_{\varepsilon 1} + w_{\varepsilon 2}}{2}, \qquad w_{\gamma k} = \frac{w_{\gamma 1k} + w_{\gamma 2k}}{2} + \frac{1}{2 \log(p_k)} \sum_{i=1}^{\nu} |\overline{a}_{ik}| \tag{5.12}$$

for $k = 1, \dots, \nu$. Here $w_{\varepsilon 1}, w_{\varepsilon 2}$ and $w_{\gamma 1k}, w_{\gamma 2k}$ are the coefficients (5.8) and (5.9), respectively. Let $\kappa_\tau = 3/2$ and recall that

$$h_\tau(z) = \frac{1}{[L : K]} \log^+ |\tau(z)|$$

denotes the local height of z at τ in the decomposition of $h(z)$.

Lemma 5.4.4. *Let $(n_1, \dots, n_\nu, a_1, \dots, a_r)$ be any solution of (5.2). If $h_\tau(z) > c_\tau$, then*

$$\begin{aligned} &\left| \alpha_0 + \sum_{i=1}^{r} a_i \alpha_{\varepsilon i} + \sum_{i=1}^{\nu} n_i \alpha_{\gamma i} \right| \\ &\qquad \le \frac{1}{2} + w_\varepsilon \max_{w: L \to \mathbb{C}} \log^+ |w(z)| + \sum_{l=1}^{\nu} w_{\gamma l} \log(p_l) |u_l - r_l| + c \kappa_\tau e^{-h_\tau(z)} \end{aligned}$$

Proof. Let

$$\alpha_\tau = \alpha_0 + \sum_{i=1}^{r} a_i \alpha_{\varepsilon i} + \sum_{i=1}^{\nu} n_i \alpha_{\gamma i}$$

and

$$\Lambda_\tau = \log \left| \tau \left(\delta_1 \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right)^{a_i} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}} \right)^{n_i} \right) \right|.$$

We claim that

$$\tau \left(\delta_1 \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right)^{a_i} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}} \right)^{n_i} \right) > 0.$$

Indeed, $h_\tau(z) > c_\tau$ by assumption, hence

$$\max\{|\tau(z)|, 1\} > \max\{2|\tau(\delta_2)|, 1\}.$$

From this inequality, we must have that $\max\{|\tau(z)|, 1\} = |\tau(z)|$ and so

$$2|\tau(\delta_2)| < |\tau(z)| = \frac{|\tau(\delta_2)|}{|\tau(\lambda)|} \implies |\tau(\lambda)| < \frac{1}{2}.$$

Recall that $\delta_1 \tilde{y} - \delta_2 \tilde{x} = 1$. This is the equation (5.6) defined earlier. In particular, observe that $\lambda = \delta_2 \tilde{x}$ so that applying τ gives

$$\tau(\lambda) = \tau(\delta_2 \tilde{x}) = \tau(\delta_1 \tilde{y}) - 1.$$

Thus

$$|\tau(\lambda)| < \frac{1}{2} \implies \tau(\delta_1 \tilde{y}) = \tau(\lambda) + 1 > 0.$$

This proves our claim

$$\tau(\delta_1 \tilde{y}) = \tau \left(\delta_1 \prod_{i=1}^{r} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right)^{a_i} \prod_{i=1}^{\nu} \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}} \right)^{n_i} \right) > 0.$$

Having established this, we may now write

$$\Lambda_\tau = \log\left(\tau\left(\delta_1\right)\right) + \sum_{i=1}^{r} a_i \log \left(\tau \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}} \right) \right) + \sum_{i=1}^{\nu} n_i \log \left(\tau \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}} \right) \right).$$

By the triangle inequality,

$$|\alpha_\tau| \leq |\alpha_\tau - c\Lambda_\tau| + c|\Lambda_\tau|,$$

where

$$|\alpha_\tau - c\Lambda_\tau| \le |[c\log(\tau(\delta_1))] - c\log(\tau(\delta_1))| + \sum_{i=1}^{r}|a_i|\left|\left[c\log\left(\tau\left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)\right)\right] - c\log\left(\tau\left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(j)}}\right)\right)\right| + \sum_{i=1}^{\nu}|n_i|\left|\left[c\log\left(\tau\left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}}\right)\right)\right] - c\log\left(\tau\left(\frac{\gamma_i^{(k)}}{\gamma_i^{(j)}}\right)\right)\right|.$$

Since $[\ \cdot\ ]$ denotes the nearest integer function, it is clear that $|[\ c\ ] - c| \le 1/2$ for any integer c,

$$|\alpha_\tau - c\Lambda_\tau| \le \frac{1}{2} + \frac{1}{2}\sum_{i=1}^{r}|a_i| + \frac{1}{2}\sum_{i=1}^{\nu}|n_i| \le \frac{1}{2}\left(1 + \sum_{i=1}^{r}|a_i| + |u_1 - r_1|\sum_{i=1}^{\nu}|\overline{a}_{i1}| + \cdots + |u_\nu - r_\nu|\sum_{i=1}^{\nu}|\overline{a}_{i\nu}|\right).$$

Applying Lemma 5.4.3, this becomes

$$|\alpha_\tau - c\Lambda_\tau| \le \frac{1}{2} + \frac{(w_{\varepsilon 1} + w_{\varepsilon_2})}{2}\max_{w:L\to\mathbb{C}}\log^+|w(z)| + \log(p_1)|u_1 - r_1|\left(\frac{(w_{\gamma 11} + w_{\gamma 21})}{2} + \frac{1}{2\log(p_1)}\sum_{i=1}^{\nu}|\overline{a}_{i1}|\right) + \cdots + \log(p_\nu)|u_\nu - r_\nu|\left(\frac{(w_{\gamma 1\nu} + w_{\gamma 2\nu})}{2} + \frac{1}{2\log(p_\nu)}\sum_{i=1}^{\nu}|\overline{a}_{i\nu}|\right).$$

In the notation of (5.12), this inequality reduces to

$$|\alpha_\tau - c\Lambda_\tau| \le \frac{1}{2} + w_\varepsilon \max_{w:L\to\mathbb{C}}\log^+|w(z)| + \sum_{l=1}^{\nu} w_{\gamma l}\log(p_l)|u_l - r_l|.$$

Now the following upper bound for $|\Lambda_\tau|$ implies the statement. On using power series definition of exponential function, we obtain

$$\Lambda_\tau(1 + \sum_{n\ge 2}(\Lambda_\tau)^{n-1}/n!) = \Lambda_\tau + \sum_{n\ge 2}(\Lambda_\tau)^n/n! = e^{\Lambda_\tau} - 1 = \tau(\lambda).$$

If $\Lambda_\tau \ge 0$ then $1 + \sum_{n\ge 2}(\Lambda_\tau)^{n-1}/n! > 1$ which implies that $|\Lambda_\tau| \le |\tau(\lambda)|$. Suppose now

that $\Lambda_\tau < 0$. Our assumption $h_\tau(z) \geq \log^+(2|\lambda_0|)$ means that $|\tau(\lambda)| \leq 1/2$ and thus $|\Lambda_\tau| = -\log(\tau(\lambda)+1) \leq -\log(1/2) = \log 2$. Therefore, the absolute value of $\sum_{n\geq 2}(\Lambda_\tau)^{n-1}/n!$ is at most

$$\sum_{n\geq 2}|\Lambda_\tau|^{n-1}/n! = \sum_{n\geq 1}|\Lambda_\tau|^n/(n+1)! \leq \tfrac{1}{2}\sum_{n\geq 1}|\Lambda_\tau|^n/n! \leq \tfrac{1}{2}e^{\log 2} - 1/2 = 1/2.$$

More precisely, for any even $N \geq 2$, we obtain

$$\begin{aligned}
|\sum_{n\geq 2}(\Lambda_\tau)^{n-1}/n!| &= |\sum_{n\geq 1}(\Lambda_\tau)^n/(n+1)!| \\
&\leq |\sum_{N\geq n\geq 1}(\Lambda_\tau)^n/(n+1)!| + \tfrac{1}{N+2}|\sum_{n>N}(\Lambda_\tau)^n/n!| \\
&\leq |\sum_{N\geq n\geq 1}(\Lambda_\tau)^n/(n+1)!| + \tfrac{1}{N+2}e^{|\Lambda_\tau|} \\
&\leq |\sum_{N\geq n\geq 1}(\Lambda_\tau)^n/(n+1)!| + \tfrac{2}{N+2} := k_N.
\end{aligned}$$

We now give an upper bound for k_N. Since $\Lambda_\tau < 0$, we obtain

$$\sum_{n\geq 2}(\Lambda_\tau)^{n-1}/n! = \sum_{N\geq n\geq 1}(\Lambda_\tau)^n/(n+1)! = \sum_{N\geq n\geq 2,\, 2|n}\tfrac{|\Lambda_\tau|^n}{(n+1)!} - \tfrac{|\Lambda_\tau|^{n-1}}{n!}$$

$$= \sum_{N\geq n\geq 2,\, 2|n}\tfrac{|\Lambda_\tau|^{n-1}}{n!}(\tfrac{|\Lambda_\tau|}{n+1} - 1) = \tfrac{|\Lambda_\tau|}{2}(\tfrac{|\Lambda_\tau|}{3} - 1) + \sum_{N\geq n\geq 4,\, 2|n}\tfrac{|\Lambda_\tau|^{n-1}}{n!}(\tfrac{|\Lambda_\tau|}{n+1} - 1)$$

$$\geq \tfrac{\log 2}{2}(\tfrac{\log 2}{4} - 1) + \sum_{N\geq n\geq 4,\, 2|n}\tfrac{(\log 2)^{n-1}}{n!}(\tfrac{3/4(\log 2)}{n+1} - 1) := -k_N.$$

The last inequality follows by distinguishing two cases whether $|\Lambda_\tau| \leq 3/4 \cdot \log 2$ or not; note that $\ln(2)/2 \cdot (\ln(2)/4 - 1)/(-\ln(2) \cdot 3/8) \geq 1$. Now, on using that $-k_N$ is negative, it follows that

$$|1 + \sum_{n\geq 2}(\Lambda_\tau)^{n-1}/n!| \geq 1 - |\sum_{n\geq 2}(\Lambda_\tau)^{n-1}/n!| \geq 1 - k_N$$

and thus

$$|\Lambda_\tau| \leq \kappa_\tau|\tau(x)|, \quad \kappa_\tau = \tfrac{1}{1-k_N}|\tau(\lambda_0)|, \quad c_\tau = \log^+(2|\lambda_0|).$$

The constant κ_τ depends on N which can be taken arbitrarily as long as $N \geq 2$ is even. Further, the value k_N can be slightly improved when one finds the maximum of the

functions $x^{n-1}(\frac{x}{n+1}-1)$ on the interval $[0,\log 2]$ for each even $n \geq 2$. This is our reason for taking $\kappa_\tau = \frac{3}{2}$. Currently this is not the optimal choice of κ_τ, but it suffices for our present case.

Finally, we have

$$|\alpha_\tau| \leq \frac{1}{2} + w_\varepsilon \max_{w:L\to\mathbb{C}} \log^+ |w(z)| + \sum_{l=1}^{\nu} w_{\gamma l} \log(p_l)|u_l - r_l| + c\kappa_\tau e^{-h_\tau(z)}.$$

□

To summarize the results of this section, let $\mathbf{m} = (n_1,\dots,n_\nu,a_1,\dots,a_r) \in \mathbb{R}^{r+\nu}$ be any solution of (5.2) with corresponding vector $\mathbf{n} = (n_1,\dots,n_\nu)$. Take $\mathbf{l},\mathbf{h} \in \mathbb{R}^{\nu+m}$ such that $\mathbf{0} \leq \mathbf{l} \leq \mathbf{h}$ and suppose $h_v(z) \leq h_v$ for all $v \in S^*$. By Lemma 5.4.1, we deduce

$$q_f(\mathbf{n}) \leq \frac{1}{\log(2)^2} \sum_{k=1}^{\nu} \log(p_k)^2 |u_k - r_k|^2 \leq \frac{1}{\log(2)^2} \sum_{k=1}^{\nu} h_k^2 =: b_\gamma. \tag{5.13}$$

For $l = 1, 2$, Lemma 5.4.3 gives us

$$|a_l|^2 \leq \left(w_{\varepsilon l} \max_{w:L\to\mathbb{C}} \log^+ |w(z)| + \sum_{k=1}^{\nu} w_{\gamma lk} \log(p_k)|u_k - r_k| \right)^2 \tag{5.14}$$

$$\leq \left([L:K] w_{\varepsilon l} \max_{w:L\to\mathbb{C}} h_w + \sum_{k=1}^{\nu} w_{\gamma lk} h_k \right)^2 =: b_{\varepsilon_l}. \tag{5.15}$$

Finally, suppose in addition that

$$h_\tau(z) \geq l_\tau > c_\tau.$$

Then by Lemma 5.4.4, we obtain

$$\left|\alpha_0 + \sum_{i=1}^{r} a_i \alpha_{\varepsilon i} + \sum_{i=1}^{\nu} n_i \alpha_{\gamma i}\right|^2 \tag{5.16}$$

$$\leq \left(\frac{1}{2} + w_\varepsilon \max_{w:L\to\mathbb{C}} \log^+ |w(z)| + \sum_{k=1}^{\nu} w_{\gamma k} \log(p_k)|u_k - r_k| + c\kappa_\tau e^{-h_\tau(z)}\right)^2 \tag{5.17}$$

$$\leq \left(\frac{1}{2} + [L:K] w_\varepsilon \max_{w:L\to\mathbb{C}} h_w + \sum_{k=1}^{\nu} w_{\gamma k} h_k + c\kappa_\tau e^{-l_\tau}\right)^2 =: b_{\varepsilon_l^*}. \tag{5.18}$$

It is of particular importance to note that the assumptions $h_\tau(z) \geq l_\tau$ and $h_v(z) \leq h_v$ for all $v \in S^*$ are not arbitrary. Indeed, for the vectors $\mathbf{l}, \mathbf{h}$, these conditions imply precisely that $(\tilde{x}, \tilde{y}) \in \Sigma_\tau(\mathbf{l}, \mathbf{h})$, where $(\tilde{x}, \tilde{y})$ are solutions to (5.2) corresponding to $\mathbf{m}$.

We are finally in position to define the ellipsoid corresponding to $\Sigma_\tau(\mathbf{l}, \mathbf{h})$. Fix any $\varepsilon_l^* \in \{\varepsilon_1, \ldots, \varepsilon_r\}$. For each ε_l in $\{\varepsilon_1, \ldots, \varepsilon_r\}$ such that $\varepsilon_l \neq \varepsilon_l^*$, we associate the bound b_{ε_l}. For ε_l, we associate the value $b_{\varepsilon_l^*}$.

Let

$$\mathbf{x} = (x_1, \ldots, x_\nu, x_{\varepsilon_1}, \ldots, x_{\varepsilon_r}) \in \mathbb{R}^{\nu+r}.$$

Then we define the ellipsoid $\mathcal{E}_\tau \subseteq \mathbb{R}^{r+\nu}$ by

$$\mathcal{E}_\tau = \{q_\tau(\mathbf{x}) \leq (1+r)(b_\gamma b_{\varepsilon_1} \cdots b_{\varepsilon_r});\ \mathbf{x} \in \mathbb{R}^{r+\nu}\} \tag{5.19}$$

where

$$q_\tau(\mathbf{x}) = (b_{\varepsilon_1} \cdots b_{\varepsilon_r}) \left(q_f(x_1, \ldots, x_\nu) + \sum_{i=1}^{r} \frac{b_\gamma}{b_{\varepsilon_i}} x_{\varepsilon_i}^2\right)$$

and

$$q_f(\mathbf{y}) = (A\mathbf{y})^{\mathrm{T}} D^2 A\mathbf{y}.$$

We associate to this ellipsoid a matrix. More precisely, we let $M = M_\tau$ be the matrix

defining the ellipsoid $\mathcal{E}_\tau$. Explicitly, this is the matrix

$$M = \sqrt{b_{\varepsilon_1} \cdots b_{\varepsilon_r}} \begin{pmatrix} DA & 0 & \dots & 0 & 0 \\ 0 & \sqrt{\frac{b_\gamma}{b_{\varepsilon 1}}} & \dots & 0 & 0 \\ 0 & 0 & \sqrt{\frac{b_\gamma}{b_{\varepsilon 2}}} & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \sqrt{\frac{b_\gamma}{b_{\varepsilon *}}} \end{pmatrix}.$$

Note that we never need to compute M, but rather $M^T M$ so that we only ever work with integral matrices. In this case,

$$M^T M = b_{\varepsilon_1} \cdots b_{\varepsilon_r} \begin{pmatrix} A^T D^2 A & 0 & \dots & 0 & 0 \\ 0 & \frac{b_\gamma}{b_{\varepsilon 1}} & \dots & 0 & 0 \\ 0 & 0 & \frac{b_\gamma}{b_{\varepsilon 2}} & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \frac{b_\gamma}{b_{\varepsilon *}} \end{pmatrix}.$$

5.4.2 The non-Archimedean ellipsoid

We now restrict our attention to those $p_v \in \{p_1, \dots, p_\nu\}$ and define the corresponding ellipsoid. As before, let $\mathbf{m} = (n_1, \dots, n_\nu, a_1, \dots, a_r) \in \mathbb{R}^{r+\nu}$ be any solution of (5.2) with corresponding vector $\mathbf{n} = (n_1, \dots, n_\nu)$. Take $\mathbf{l}, \mathbf{h} \in \mathbb{R}^{\nu+m}$ such that $\mathbf{0} \leq \mathbf{l} \leq \mathbf{h}$ and suppose $h_v(z) \leq h_v$ for all $v \in S^*$.

Now, Lemma 5.4.1 and Lemma 5.4.3 still hold here. In particular, we let b, b_{ε_l} be defined as in (5.13) and (5.14), respectively, where $l = 1, \dots, r$:

$$q_f(\mathbf{n}) \leq \sum_{k=1}^{\nu} \log(p_k)^2 |u_k - r_k|^2 \leq \sum_{k=1}^{\nu} h_k^2 =: b. \tag{5.20}$$

For $l = 1, 2$, Lemma 5.4.3 gives us These are off and I need to go back and re-edit the bound on $|a_l|$ in the previous section - but is there a reason that we take the square

value?

$$|a_l|^2 \leq \left(\frac{1}{[L:\mathbb{Q}]} \sum_{\sigma: L \to \mathbb{C}} w_{\varepsilon l \sigma} \log \max \left\{ \left| \sigma \left(\frac{\delta_2}{\lambda} \right) \right|, 1 \right\} \right. \tag{5.21}$$

$$\left. + \frac{1}{[K:\mathbb{Q}]} \sum_{k=1}^{\nu} w_{\gamma l k} \log(p_k) |u_k - r_k| \right)^2 \tag{5.22}$$

$$\leq \left(\frac{1}{[K:\mathbb{Q}]} \sum_{k=1}^{\nu} w_{\gamma l k} h_k + \frac{1}{[L:\mathbb{Q}]} \sum_{\sigma: L \to \mathbb{C}} w_{\varepsilon l \sigma} h_\sigma \right)^2 =: b_{\varepsilon_l}. \tag{5.23}$$

Or is this one?

$$|a_l| \leq w_{\varepsilon l} \max_{w: L \to \mathbb{C}} \log^+ |w(z)| + \sum_{k=1}^{\nu} w_{\gamma l k} \log(p_k) |u_k - r_k|.$$

We do not distinguish any ε_l^*. Instead, we will see later that the condition $h_v(z) \geq l_v$ corresponding to the set $\Sigma_v(\mathbf{l}, \mathbf{h})$ will be used elsewhere.

Let $b := \operatorname{lcm}\left(\sqrt{b_{\varepsilon_1}}, \ldots, \sqrt{b_{\varepsilon_r}}, Q_{n_1}, \ldots, Q_{n_\nu}\right)$, where $Q_{n_1}, \ldots, Q_{n_\nu}$ correspond to the denominators of the convergents of $\log^*(p_1), \ldots, \log^*(p_\nu)$ respectively. Recall that b_{ε_i} is a square so that $\sqrt{b_{\varepsilon_i}}$ is an integer. We define the ellipsoid $\mathcal{E}_v \subseteq \mathbb{R}^{\nu+r}$ by

$$\mathcal{E}_v = \{ q_v(\mathbf{x}) \leq b(b_\gamma + r);\ \mathbf{x} \in \mathbb{R}^{r+\nu} \}, \tag{5.24}$$

where

$$q_v(\mathbf{x}) = b \left(q_f(x_1, \ldots, x_\nu) + \sum_{i=1}^{r} \frac{1}{b_{\varepsilon_i}} x_{\varepsilon_i}^2 \right)$$

and

$$q_f(\mathbf{y}) = (A\mathbf{y})^{\mathrm{T}} D^2 A \mathbf{y}.$$

Similar to the Archimedean case, we let $M = M_v$ be the matrix defining the ellipsoid $\mathcal{E}_v$.

Explicitly, this is the matrix

$$M = b\begin{pmatrix} DA & 0 & \dots & 0 & 0 \\ 0 & \sqrt{\frac{1}{b_{\varepsilon 1}}} & \dots & 0 & 0 \\ 0 & 0 & \sqrt{\frac{1}{b_{\varepsilon 2}}} & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \sqrt{\frac{1}{b_{\varepsilon}}} \end{pmatrix} = \begin{pmatrix} bDA & 0 & \dots & 0 & 0 \\ 0 & \frac{b}{\sqrt{b_{\varepsilon_1}}} & \dots & 0 & 0 \\ 0 & 0 & \frac{b}{\sqrt{b_{\varepsilon_2}}} & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \frac{b}{\sqrt{b_{\varepsilon_r}}} \end{pmatrix}.$$

By definition of b, it follows that M is an integral matrix.

5.5 The Archimedean sieve: the real case

Let $\tau : L \to \mathbb{C}$ be an embedding. We take $\mathbf{l}, \mathbf{h} \in \mathbb{R}^{m+\nu}$ with $\mathbf{0} \leq \mathbf{l} \leq \mathbf{h}$ and $l_\tau \geq \log 2$. Let c be a constant the size of e^{l_τ} and let $\alpha_0, \alpha_{\varepsilon 1}, \dots, \alpha_{\varepsilon r}, \alpha_{\gamma 1}, \dots, \alpha_{\gamma \nu}$ be defined as in (5.10) and (5.11).

Define the $(\nu + r) \times (\nu + r)$-dimensional matrix A_τ as

$$A_\tau = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 & 0 \\ 0 & 1 & \dots & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \dots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & 1 & 0 \\ \alpha_{\gamma 1} & \dots & \alpha_{\gamma \nu} & \alpha_{\varepsilon 1} & \dots & \alpha_{\varepsilon r} \end{pmatrix}$$

and consider the lattice defined by its columns. Let $\mathbf{w} = (0, \dots, 0, \alpha_0)$ be a vector of length $(\nu + r)$. We now consider the translated lattice Γ_τ defined by $A_\tau \mathbf{x} + \mathbf{w}$, where $\mathbf{x}$ is an arbitrary coordinate vector.

Let $\mathcal{E}_\tau = \mathcal{E}_\tau(h, l_\tau)$ be the ellipsoid constructed in (5.19). Let

$$\mathbf{m} = (n_1, \dots, n_\nu, a_1, \dots, a_r) \in \mathbb{R}^{r+\nu}$$

be any solution of (5.2). We say that $\mathbf{m}$ is determined by some $\mathbf{y} \in \Gamma_\tau$ if

$$\mathbf{y} = (y_1, \dots, y_{r+\nu}) = \left(n_1, \dots, n_\nu, a_1, \dots, a_{r-1}, \alpha_0 + \sum_{i=1}^{r} a_i \alpha_{\varepsilon i} + \sum_{i=1}^{\nu} n_i \alpha_{\gamma i} \right)$$

where the missing element a_l corresponds to ε_l^*.

Lemma 5.5.1. *Let* $\mathbf{m} = (n_1, \ldots, n_\nu, a_1, \ldots, a_r) \in \mathbb{R}^{r+\nu}$ *be any solution of* (5.2) *which lies in* $\Sigma_\tau(l, h)$. *Then* $\mathbf{m}$ *is determined by some* $\mathbf{y} \in \Gamma_\tau \cap \mathcal{E}_\tau$.

Proof. Let

$$\mathbf{y} = \left(n_1, \ldots, n_\nu, a_1, \ldots, a_{r-1}, \alpha_0 + \sum_{i=1}^{r} a_i \alpha_{\varepsilon i} + \sum_{i=1}^{\nu} n_i \alpha_{\gamma i} \right).$$

Then $\mathbf{y} \in \Gamma_\tau$ and (5.16) implies that $y_{\varepsilon_l^*}^2 \leq b_{\varepsilon_l^*}$. Further $q_f(y_1, \ldots, y_\nu) \leq b$ by (5.13) and (5.14) provides that $y_{\varepsilon_l}^2 \leq b_{\varepsilon_l}$ for $l = 1, \ldots, r$ with $\varepsilon_l \neq \varepsilon_l^*$. It follows that

$$q_\tau(\mathbf{y}) = (b_{\varepsilon_1} \cdots b_{\varepsilon_r}) \left(q_f(y_1, \ldots, y_\nu) + \sum_{i=1}^{r} \frac{b}{b_{\varepsilon_i}} y_{\varepsilon_i}^2 \right) \leq (1 + r)(b b_{\varepsilon_1} \cdots b_{\varepsilon_r}).$$

This proves that $\mathbf{y} \in \mathcal{E}_\tau$ and hence the statement follows. □

We now explicitly determine $\Gamma_\tau \cap \mathcal{E}_\tau$. Suppose that $\mathbf{y} \in \Gamma_\tau \cap \mathcal{E}_\tau$. Let $M = M_\tau$ be the matrix defining the ellipsoid $\mathcal{E}_\tau$. Since $\mathbf{y} \in \Gamma_\tau \cap \mathcal{E}_\tau$, there exists $\mathbf{x} \in \mathbb{R}^{r+\nu}$ such that $\mathbf{y} = A_\tau \mathbf{x} + \mathbf{w}$ and $\mathbf{y}^t M^t M \mathbf{y} \leq (1 + r)(b b_{\varepsilon_1} \cdots b_{\varepsilon_r})$. We thus have

$$(A_\tau \mathbf{x} + \mathbf{w})^t M^t M (A_\tau \mathbf{x} + \mathbf{w}) \leq (1 + r)(b b_{\varepsilon_1} \cdots b_{\varepsilon_r}).$$

As A_τ is clearly invertible, with matrix inverse

$$A_\tau^{-1} = \begin{pmatrix} 1 & 0 & \ldots & \ldots & 0 & 0 \\ 0 & 1 & \ldots & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \ldots & \vdots & \vdots \\ 0 & 0 & \ldots & \ldots & 1 & 0 \\ -\frac{\alpha_{\gamma 1}}{\alpha_{\varepsilon r}} & \ldots & -\frac{\alpha_{\gamma \nu}}{\alpha_{\varepsilon r}} & -\frac{\alpha_{\varepsilon 1}}{\alpha_{\varepsilon r}} & \ldots & \frac{1}{\alpha_{\varepsilon r}} \end{pmatrix},$$

we can find a vector $\mathbf{c}$ such that $A_\tau \mathbf{c} = -\mathbf{w}$. Indeed, this vector is

$$\mathbf{c} = A_\tau^{-1} \mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ -\frac{\alpha_0}{\alpha_{\varepsilon r}} \end{pmatrix}.$$

Now,

$$\begin{aligned}(1+r)(bb_{\varepsilon_1}\cdots b_{\varepsilon_r}) &\geq (A_\tau\mathbf{x}+\mathbf{w})^t M^t M(A_\tau\mathbf{x}+\mathbf{w}) \\ &= (A_\tau(\mathbf{x}-\mathbf{c}))^T M^T M(A_\tau(\mathbf{x}-\mathbf{c})) \\ &= (\mathbf{x}-\mathbf{c})^T (MA_\tau)^T MA_\tau(\mathbf{x}-\mathbf{c}) \\ &= (\mathbf{x}-\mathbf{c})^T B^T B(\mathbf{x}-\mathbf{c})\end{aligned}$$

where $B = MA_\tau$. That is, we are left to solve

$$(\mathbf{x}-\mathbf{c})^T B^T B(\mathbf{x}-\mathbf{c}) \leq (1+r)(bb_{\varepsilon_1}\cdots b_{\varepsilon_r}).$$

Now, finding all vectors satisfying this inequality amounts to computing all solutions to (5.2) contained in $\Sigma_\tau(\mathbf{l},\mathbf{h})$. The set of vectors $\mathbf{x}$ can be found using the Fincke-Pohst algorithm outlined in subsection 3.6.2.

5.6 The non-Archimedean Sieve

In this section, we describe the p-adic reduction procedure. Throughout this section, we let $v \in \{1,\ldots,\nu\}$, and we denote by $\mathbf{h} = (h_v)$ the vector containing the best current upper bounds on the solution vector $(h_v(z))$.

Recall that $\Sigma(\mathbf{h}) = \Sigma(\mathbf{l},\mathbf{h}) \cup \Sigma(\mathbf{l})$ for $\mathbf{0} \leq \mathbf{l} \leq \mathbf{h}$ a vector in $\mathbb{R}^{\nu+m}$ and moreover, that $\Sigma(\mathbf{l},\mathbf{h}) = \cup_{v\in S^*}\Sigma_v(\mathbf{l},\mathbf{h})$. In this section, we determine all solutions $(\tilde{x},\tilde{y}) \in \Sigma_v(\mathbf{l},\mathbf{h})$ for some appropriately chosen vector $\mathbf{l}$. We should note that in this section, we only choose the v^{th} coordinate of this vector $\mathbf{l}$. However, by applying the non-Archimedean sieve on the remaining rational primes, as well as the Archimedean sieve, we will obtain the vector $\mathbf{l}$, and subsequently obtain all solutions in $\Sigma(\mathbf{l},\mathbf{h})$.

We begin by applying the results of section 3.5. In particular, we consider the form

$$\Lambda_v = \sum_{i=1}^{1+\nu+r} b_i\alpha_i$$

where

$$b_1 = 1, \quad b_{1+i} = n_i \ \text{ for } i \in \{1,\ldots,\nu\},$$

$$b_{1+\nu+i} = a_i \ \text{ for } i \in \{1,\ldots,r\},$$

and

$$\alpha_1 = \log_{p_v} \delta_1, \quad \alpha_{1+i} = \log_{p_v} \left(\frac{\gamma_i^{(k)}}{\gamma_i^{(l)}} \right) \quad \text{for } i \in \{1, \ldots, \nu\},$$

$$\alpha_{1+\nu+i} = \log_{p_v} \left(\frac{\varepsilon_i^{(k)}}{\varepsilon_i^{(l)}} \right) \quad \text{for } i \in \{1, \ldots, r\}.$$

5.6.1 Applying Lemma 3.5.2

We apply Lemma 3.5.2 by which $\sum_{j=1}^{\nu} n_j a_{vj}$ can be computed directly provided $\mathrm{ord}_{p_v}(\delta_1) \neq 0$. If this is the case then what? We need to incorporate this data into the information of the known bounds so that we may reduce the rank and subsequently remove p_v from the set of unbounded rational primes. But how?

However, we speculate that this actually will never happen...

Thus, we assume for the remainder of this chapter that $\mathrm{ord}_{p_v}(\delta_1) = 0$.

5.6.2 Applying Lemma 3.5.5

We now apply Lemma 3.5.5 to obtain a small bound on $\sum_{j=1}^{\nu} n_j a_{vj}$ when $\mathrm{ord}_{p_v}(\alpha_1) < \min_{2 \leq i \leq 1+\nu+r} \mathrm{ord}_{p_v}(\alpha_i)$. If this lemma holds, we may reduce the bound h_v on $h_v(z)$ as follows. Let B denote the bound obtained in Lemma 3.5.5 so that

$$\sum_{j=1}^{\nu} n_j a_{vj} \leq B.$$

Recall that

$$u_v = \sum_{j=1}^{\nu} n_j a_{vj} + r_v,$$

where u_v and r_v are positive integers. It therefore follows that

$$-r_v \leq u_v - r_v = \sum_{j=1}^{\nu} n_j a_{vj} \leq B.$$

Now, if $B < -r_v$ so that $B < 0$, we obtain a contradiction and may thus discard this entire S-unit equation. Conversely, if $B \geq -r_v$ (B may be positive or negative), we obtain

$$|u_v - r_v| = \left| \sum_{j=1}^{\nu} n_j a_{vj} \right| \leq \max\{B, r_v\}.$$

It follows then that

$$h_v(z) = \log(p_v)|u_v - r_v| \leq \log(p_v) \max\{B, r_v\},$$

and we may thus upgrade the bound h_v on $h_v(z)$ to be

$$h_v = \log(p_v) \max\{B, r_v\}.$$

5.6.3 The reduction procedure

Now, if the above lemmatta do not hold, we are in the situation that

$$\mathrm{ord}_{p_v}(\delta_1) = 0 \quad \text{and} \quad \mathrm{ord}_{p_v}(\alpha_1) \geq \min_{2 \leq i \leq 1+\nu+r} \mathrm{ord}_{p_v}(\alpha_i).$$

Indeed, we assume this for the remainder of this chapter.

We now set some notation and give some preliminaries for the p_v-adic reduction procedures. Let I be the set of all indices $i' \in \{2, \ldots, 1 + \nu + r\}$ for which

$$\mathrm{ord}_{p_v}(\alpha_{i'}) = \min_{2 \leq i \leq 1+\nu+r} \mathrm{ord}_{p_v}(\alpha_i).$$

Following [**?**], we are always in the case where there exists an index $i' \in I$ such that $\alpha_i/\alpha_{i'} \in \mathbb{Q}_{p_v}$ for $i = 1, \ldots, 1 + \nu + r$. Thus, let $\hat{i}$ denote this index. We define

$$\beta_i = -\frac{\alpha_i}{\alpha_{\hat{i}}} \quad i = 1, \ldots, 1 + \nu + r,$$

and

$$\Lambda'_v = \frac{1}{\alpha_{\hat{i}}} \Lambda_v = \sum_{i=1}^{1+\nu+r} b_i(-\beta_i).$$

Now, we have $\beta_i \in \mathbb{Z}_{p_v}$ for $i = 1, \ldots, 1 + \nu + r$.

Lemma 5.6.1. *Suppose* $\mathrm{ord}_{p_v}(\delta_1) = 0$ *and*

$$\sum_{i=1}^{\nu} n_i a_{li} > \frac{1}{p_v - 1} - \mathrm{ord}_{p_v}(\delta_2).$$

Then

$$\mathrm{ord}_{p_v}(\Lambda_v') = \sum_{i=1}^{\nu} n_i a_{li} + \mathrm{ord}_{p_l}(\delta_2) - \mathrm{ord}_{p_l}(\alpha_{\hat{i}}).$$

Proof. Immediate from Lemma 3.5.3 and Lemma 3.5.4. □

Fix $l_v \in \mathbb{Z}$ such that

$$\frac{l_v}{\log(p_v)} \geq \max\left(\frac{1}{p_v - 1}, \mathrm{ord}_{p_v}(\delta_1)\right) - \mathrm{ord}_{p_v}(\delta_2).$$

Now, let μ be the largest element of $\mathbb{Z}_{\geq 0}$ at most

$$\mu \leq \frac{l_v}{\log(p_v)} - \mathrm{ord}_{p_v}(\alpha_{\hat{i}}) + \mathrm{ord}_{p_v}(\delta_2).$$

Lemma 5.6.2. *Suppose* $\mathrm{ord}_{p_v}(\delta_1) = 0$ *and* $\sum_{i=1}^{\nu} n_i a_{vi} > \frac{1}{p_v - 1} - \mathrm{ord}_{p_v}(\delta_2)$. *Then*

$$\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_{\hat{i}}) \quad \textit{if and only if} \quad \mathrm{ord}_{p_v}(\Lambda_v') \geq \mu$$

Proof. By Lemma 5.6.1, the assumptions mean that

$$\mathrm{ord}_{p_v}(\Lambda_v') = \sum_{i=1}^{\nu} n_i a_{vi} + \mathrm{ord}_{p_v}(\delta_2) - \mathrm{ord}_{p_v}(\alpha_{\hat{i}}).$$

Now, suppose

$$\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_{\hat{i}}).$$

We thus have

$$\begin{aligned}\mathrm{ord}_{p_v}(\Lambda'_v) &= \sum_{i=1}^{\nu} n_i a_{vi} + \mathrm{ord}_{p_v}(\delta_2) - \mathrm{ord}_{p_v}(\alpha_{\hat{i}}) \\ &\geq \mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_{\hat{i}}) + \mathrm{ord}_{p_v}(\delta_2) - \mathrm{ord}_{p_v}(\alpha_{\hat{i}}) \\ &= \mu.\end{aligned}$$

Conversely, suppose $\mathrm{ord}_{p_v}(\Lambda'_v) \geq \mu$. Then

$$\mu \leq \mathrm{ord}_{p_v}(\Lambda'_v) = \sum_{i=1}^{\nu} n_i a_{vi} + \mathrm{ord}_{p_v}(\delta_2) - \mathrm{ord}_{p_v}(\alpha_{\hat{i}}).$$

That is,

$$\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_{\hat{i}}).$$

Hence, it follows that $\displaystyle\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_{\hat{i}})$ if and only if $\mathrm{ord}_{p_v}(\Lambda'_v) \geq \mu$. $\square$

For each $x \in \mathbb{Z}_{p_v}$, let $x^{\{\mu\}}$ denote the unique rational integer in $[0, p_v^{\mu} - 1]$ such that $\mathrm{ord}_{p_v}(x - x^{\mu}) \geq \mu$ (ie. $x \equiv x^{\{\mu\}} \pmod{p_v^{\mu}}$).

Let Γ_v be the $(\nu + r)$-dimensional lattice determined by column vectors of the matrix MA_v, where M is the matrix defining the ellipsoid $\mathcal{E}_v$ of subsection 5.4.2 and A_v is the diagonal matrix having $\hat{i}^{\text{th}}$ row

$$\left(\beta_2^{\{\mu\}}, \cdots, \beta_{\hat{i}-1}^{\{\mu\}}, p_v^{\mu}, \beta_{\hat{i}+1}^{\{\mu\}}, \cdots, \beta_{1+\nu+r}^{\{\mu\}}\right) \in \mathbb{Z}^{\nu+r}.$$

Here, p_v^{μ} is the $(\hat{i}, \hat{i})$ entry of A_v. That is,

$$A_v = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & 0 & & \\ & & 1 & & & & \\ \beta_2^{\{\mu\}} & \cdots & \beta_{\hat{i}-1}^{\{\mu\}} & p_v^{\mu} & \beta_{\hat{i}+1}^{\{\mu\}} & \cdots & \beta_{1+\nu+r}^{\{\mu\}} \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}.$$

Additionally, let

$$\lambda = \frac{1}{p_v^{\mu}} \sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}}),$$

and set

$$\mathbf{w} = M(0, \ldots 0, -\beta_1^{\{\mu\}}, 0, \ldots, 0)^T \in \mathbb{Z}^{\nu+r}$$

to be the product of M and the vector whose only non-zero entry is the $\hat{i}^{\text{th}}$ element, $-\beta_1^{\{\mu\}}$.

Choose a vector $\mathbf{z} \in \Gamma_v$ that is close to $\mathbf{w}$. An efficient way to do this is as follows. Compute the matrix B_v whose column vectors $\mathbf{c_1}, \ldots, \mathbf{c_{\nu+r}}$ form an LLL-reduced basis for Γ_v and write

$$\mathbf{w} = s_1\mathbf{c_1} + \cdots + s_{\nu+r}\mathbf{c}_{\nu+r}, \quad s_i \in \mathbb{R}.$$

In other words, compute $(s_1, \ldots, s_{\nu+r})^T = B_v^{-1}\mathbf{w}$. Choose $\mathbf{t} \in \mathbb{Z}^{\nu+r}$ such that $|t_i - s_i| \leq 1$ for all i and $|\mathbf{w} - B_v\mathbf{t}|$ is minimal. Then $B_v\mathbf{t}$ is likely the closes lattice vector to $\mathbf{w}$, so we take $\mathbf{z} = B_v\mathbf{t}$.

Of course, we must compute the β_i to p_v-adic precision at least μ in order to avoid errors here.

We say that $\mathbf{m} = (n_1, \ldots, n_\nu, a_1, \ldots, a_r) \in \mathbb{R}^{r+\nu}$ is determined by some $\mathbf{y} \in \Gamma_v$ if the entries of $\mathbf{y}$ are a (fixed) permutation of the entries of $\mathbf{m}$. Let $\mathcal{E}_v$ be the ellipsoid constructed in (5.24).

Lemma 5.6.3. *Any $(\tilde{x}, \tilde{y}) \in \Sigma_v(l, h)$ is determined by some $\mathbf{y} \in \Gamma_v \cap \mathcal{E}_v$.*

In the remainder of this section, we prove this lemma.

Let $\mathbf{y} = (n_1, \ldots, n_\nu, a_1, \ldots, a_r) \in \mathbb{R}^{\nu+r}$ denote a solution to (5.2).

Lemma 5.6.4. *Suppose* $\text{ord}_{p_v}(\delta_1) = 0$ *and*

$$\sum_{i=1}^{\nu} n_i a_{vi} > \frac{1}{p_v - 1} - \text{ord}_{p_v}(\delta_2).$$

Then the following equivalence holds:

$$\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \text{ord}_{p_v}(\delta_2) + \text{ord}_{p_v}(\alpha_{\hat{i}}) \quad \textit{if and only if} \quad \text{ord}_{p_v}(\Lambda_v') \geq \mu$$

$$\textit{if and only if} \quad \mathbf{y} \in \Gamma_v.$$

Proof. By Lemma 5.6.1, the assumption means that

$$\operatorname{ord}_{p_v}(\Lambda'_v) = \sum_{i=1}^{\nu} n_i a_{vi} + \operatorname{ord}_{p_v}(\delta_2) - \operatorname{ord}_{p_v}(\alpha_{\hat{i}}).$$

Now, suppose

$$\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \operatorname{ord}_{p_v}(\delta_2) + \operatorname{ord}_{p_v}(\alpha_{\hat{i}}).$$

We thus have

$$\begin{aligned}\operatorname{ord}_{p_v}(\Lambda'_v) &= \sum_{i=1}^{\nu} n_i a_{vi} + \operatorname{ord}_{p_v}(\delta_2) - \operatorname{ord}_{p_v}(\alpha_{\hat{i}}) \\ &\geq \mu - \operatorname{ord}_{p_v}(\delta_2) + \operatorname{ord}_{p_v}(\alpha_{\hat{i}}) + \operatorname{ord}_{p_v}(\delta_2) - \operatorname{ord}_{p_v}(\alpha_{\hat{i}}) \\ &= \mu.\end{aligned}$$

Conversely, suppose $\operatorname{ord}_{p_v}(\Lambda'_v) \geq \mu$. Then

$$\mu \leq \operatorname{ord}_{p_v}(\Lambda'_v) = \sum_{i=1}^{\nu} n_i a_{vi} + \operatorname{ord}_{p_v}(\delta_2) - \operatorname{ord}_{p_v}(\alpha_{\hat{i}}).$$

That is,

$$\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \operatorname{ord}_{p_v}(\delta_2) + \operatorname{ord}_{p_v}(\alpha_{\hat{i}}).$$

Hence, it follows that $\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \operatorname{ord}_{p_v}(\delta_2) + \operatorname{ord}_{p_v}(\alpha_{\hat{i}})$ if and only if $\operatorname{ord}_{p_v}(\Lambda'_v) \geq \mu$.

Now, suppose $\mathbf{y} = (n_1, \ldots, n_\nu, a_1, \ldots, a_r) \in \mathbb{R}^{\nu+r}$ is a solution to (5.2). Suppose further that $\sum_{i=1}^{\nu} n_i a_{vi} \geq \mu - \operatorname{ord}_{p_v}(\delta_2) + \operatorname{ord}_{p_v}(\alpha_{\hat{i}})$ so that $\operatorname{ord}_{p_v}(\Lambda'_v) \geq \mu$. Let

$$\lambda = \frac{1}{p_v^{\mu}} \sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}})$$

and consider the $(\nu + r)$-dimensional vector

$$\mathbf{x} = (n_1, \ldots, n_{\hat{i-1}}, \lambda, n_{\hat{i+1}}, \ldots, n_\nu, a_1, \ldots, a_r).$$

We claim $\mathbf{x} \in \mathbb{Z}^{\nu+r}$. That is, $\lambda \in \mathbb{Z}$, meaning that $\sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}})$ is divisible by p_v^{μ}, or

equivalently,

$$\operatorname{ord}_{p_v}\left(\sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}})\right) \geq \mu.$$

Indeed, since

$$\operatorname{ord}_{p_v}\left(\beta_i^{\{\mu\}} - \beta_i\right) \geq \mu \quad \text{for } i = 1, \ldots, 1+\nu+r,$$

by definition, it follows that $\beta_i^{\{\mu\}}$ and β_i share the first $\mu - 1$ terms and thus $\operatorname{ord}_{p_v}(\beta_i) = \operatorname{ord}_{p_v}(\beta_i^{\{\mu\}})$. Now, to compute this order, we only need to concern ourselves with the first non-zero term in the series expansion of $\sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}})$. Since $\beta_i^{\{\mu\}}$ and β_i share the first $\mu - 1$ terms, it follows that showing

$$\operatorname{ord}_{p_v}\left(\sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}})\right) \geq \mu$$

is equivalent to showing that

$$\operatorname{ord}_{p_l}(\Lambda_l') \geq \mu.$$

Of course, this latter inequality is true by assumption. Thus $\lambda \in \mathbb{Z}$.

Then, computing $A_v\mathbf{x} + \mathbf{w}$ yields

$$A_v\mathbf{x} + \mathbf{w} = \begin{pmatrix} b_2 \\ \vdots \\ b_{\hat{i}-1} \\ b^* \\ b_{\hat{i}+1} \\ \vdots \\ b_{\nu+r+1} \end{pmatrix},$$

where

$$b^* = b_2\beta_2^{\{\mu\}} + \cdots + b_{\hat{i}-1}\beta_{\hat{i}-1}^{\{\mu\}} + \lambda p_l^{\mu} + b_{\hat{i}+1}\beta_{\hat{i}+1}^{\{\mu\}} + \cdots + b_{\nu+r+1}\beta_{1+\nu+r}^{\{\mu\}} + \beta_1^{\{\mu\}}.$$

Now,

$$\lambda p_v^{\mu} = p_v^{\mu}\frac{1}{p_v^{\mu}}\sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}}) = \sum_{i=1}^{\nu+r+1} b_i(-\beta_i^{\{\mu\}}),$$

hence

$$\begin{aligned}
&b_2\beta_2^{\{\mu\}} + \cdots + b_{\hat{i}-1}\beta_{\hat{i}-1}^{\{\mu\}} + b_{\hat{i}+1}\beta_{\hat{i}+1}^{\{\mu\}} + \cdots + b_{\nu+r+1}\beta_{1+\nu+r}^{\{\mu\}} + \lambda p_l^{\mu} + \beta_1^{\{\mu\}} \\
&= b_{\hat{i}}(-\beta_{\hat{i}}^{\{\mu\}}) \\
&= b_{\hat{i}}
\end{aligned}$$

where the last equality follows from the fact that

$$-\beta_i = \frac{\alpha_{\hat{i}}}{\alpha_{\hat{i}}} = 1.$$

Thus,

$$A_v\mathbf{x} + \mathbf{w} = \begin{pmatrix} b_2 \\ \vdots \\ b_{\hat{i}-1} \\ b_{\hat{i}} \\ b_{\hat{i}+1} \\ \vdots \\ b_{\nu+r+1} \end{pmatrix} = \begin{pmatrix} n_1 \\ \vdots \\ n_\nu \\ a_1 \\ \vdots \\ a_r \end{pmatrix} = \mathbf{y}.$$

and $\mathbf{y} \in \Gamma_v$.

□

Define

$$c_{p_v} = \log p_v \left(\max\left(\tfrac{1}{p_v - 1}, \mathrm{ord}_{p_v}(\delta_1) \right) - \mathrm{ord}_{p_v}(\delta_2) \right).$$

Corollary 5.6.5. *Assume that $h_{p_v}(z) > \max(0, c_{p_v})$. Then the following equivalence holds:*

$$h_{p_v}(z) \geq \log p_v \left(\mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_{\hat{i}}) \right) \quad \textit{if and only if} \quad \mathbf{y} \in \Gamma_v.$$

Proof. Recall from Proposition 5.1.3 that

$$h_{p_v}(z) = \begin{cases} \log(p_v)|u_v - r_v| \\ 0 \end{cases}.$$

Since $h_{p_v}(z) > 0$, it follows that $h_{p_v}(z) = \log(p_v)|u_v - r_v|$. Hence the assumption becomes

$$\log(p_v)|u_v - r_v| = h_{p_v}(z) > \log p_v \left(\max\left(\tfrac{1}{p_v - 1}, \mathrm{ord}_{p_v}(\delta_1) \right) - \mathrm{ord}_{p_v}(\delta_2) \right),$$

or equivalently,

$$\sum_{j=1}^{\nu} n_j a_{vj} > \left(\max\left(\tfrac{1}{p_v-1}, \mathrm{ord}_{p_v}(\delta_1)\right) - \mathrm{ord}_{p_v}(\delta_2)\right).$$

Moreover, the conclusion is equivalent to

$$\log(p_v)|u_v - r_v| \geq \log p_v \left(\mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_i^{\wedge})\right) \quad \text{if and only if} \quad \mathbf{y} \in \Gamma_v,$$

or,

$$\sum_{j=1}^{\nu} n_j a_{vj} \geq \left(\mu - \mathrm{ord}_{p_v}(\delta_2) + \mathrm{ord}_{p_v}(\alpha_i^{\wedge})\right) \quad \text{if and only if} \quad \mathbf{y} \in \Gamma_v,$$

which is the previous lemma.

□

We now prove Lemma 5.6.3.

Proof of Lemma 5.6.3. If $(n_1, \ldots, n_\nu, a_1, \ldots, a_r) \in \mathbb{R}^{r+\nu}$ is a solution of (5.2), then, by definition, it corresponds to a solution $(\tilde{x}, \tilde{y}) \in \Sigma_v(\mathbf{l}, \mathbf{h})$. Hence $h_v(z) > l_v$, where l_v is a constant such that

$$\frac{l_v}{\log(p_v)} \geq \max\left(\frac{1}{p_v - 1}, \mathrm{ord}_{p_v}(\delta_1)\right) - \mathrm{ord}_{p_v}(\delta_2).$$

That is,

$$h_v(z) > l_v \geq \log(p_v)\left(\max\left(\frac{1}{p_v - 1}, \mathrm{ord}_{p_v}(\delta_1)\right) - \mathrm{ord}_{p_v}(\delta_2)\right) = c_p.$$

Now, recall that $\mathbf{l} \geq \mathbf{0}$ so that $l_v \geq 0$. It thus follows that

$$h_v(z) > l_v \geq \begin{cases} 0 \\ c_p \end{cases} \implies h_v(z) > \max(0, c_p).$$

In other words, the conditions of Corollary 5.6.5 are satisfied.

Now, recall that μ is the largest element of $\mathbb{Z}_{\geq 0}$ at most

$$\mu \leq \frac{l_v}{\log(p_v)} - \mathrm{ord}_{p_v}(\alpha_i^{\wedge}) + \mathrm{ord}_{p_v}(\delta_2).$$

That is

$$\frac{l_v}{\log(p_v)} \geq \mu + \mathrm{ord}_{p_v}(\alpha_{\hat{i}}) - \mathrm{ord}_{p_v}(\delta_2)$$

so that

$$h_v(z) > l_v \geq \log(p_v)\left(\mu + \mathrm{ord}_{p_l}(\alpha_{\hat{i}}) - \mathrm{ord}_{p_v}(\delta_2)\right).$$

Now, by Corollary 5.6.5, we must have $\mathbf{y} \in \Gamma_v$. This shows that $(\tilde{x}, \tilde{y})$ is determined by $\mathbf{y} = \mathbf{m}' \in \Gamma_v$, which proves Lemma 5.6.3. □

Finally, suppose that $\mathbf{y} \in \Gamma_v \cap \mathcal{E}_v$. Let $M = M_v$ be the matrix defining the ellipsoid $\mathcal{E}_v$. That is

$$M = \sqrt{b_{\varepsilon_1} \cdots b_{\varepsilon_r}} \begin{pmatrix} DA & 0 & \dots & 0 & 0 \\ 0 & \sqrt{\frac{b}{b_{\varepsilon 1}}} & \dots & 0 & 0 \\ 0 & 0 & \sqrt{\frac{b}{b_{\varepsilon 2}}} & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \dots & \dots & \sqrt{\frac{b}{b_{\varepsilon}}} \end{pmatrix}.$$

Recall that $A_v\mathbf{x} + \mathbf{w}$ defines the lattice Γ_v. In particular, since $\mathbf{y} \in \Gamma_v \cap \mathcal{E}_v$, there exists $\mathbf{x} \in \mathbb{R}^{r+\nu}$ such that $\mathbf{y} = A_v\mathbf{x} + \mathbf{w}$ and $\mathbf{y}^T M^T M \mathbf{y} \leq (1+r)(b b_{\varepsilon_1} \cdots b_{\varepsilon_r})$. We thus have

$$(A_v\mathbf{x} + \mathbf{w})^T M^T M (A_v\mathbf{x} + \mathbf{w}) \leq (1+r)(b b_{\varepsilon_1} \cdots b_{\varepsilon_r}).$$

As A_v is clearly invertible, with matrix inverse

$$A_v^{-1} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & 0 & & \\ & & 1 & & & & \\ -\frac{\beta_2^{\{\mu\}}}{p_l^{\mu}} & \dots & -\frac{\beta_{\hat{i}-1}^{\{\mu\}}}{p_l^{\mu}} & \frac{1}{p_l^{\mu}} & -\frac{\beta_{\hat{i}+1}^{\{\mu\}}}{p_l^{\mu}} & \dots & -\frac{\beta_{1+\nu+r}^{\{\mu\}}}{p_l^{\mu}} \\ & & & & 1 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix},$$

we can find a vector $\mathbf{c}$ such that $A_v\mathbf{c} = -\mathbf{w}$. Indeed, this vector is $\mathbf{c} = A_v^{-1}(-\mathbf{w})$,

where

$$\mathbf{c} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -\frac{\beta_1^{\{\mu\}}}{p^{\{\mu\}}} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Now,

$$\begin{aligned}(1+r)(bb_{\varepsilon_1}\cdots b_{\varepsilon_r}) &\geq (A_v\mathbf{x}+\mathbf{w})^T M^T M(A_v\mathbf{x}+\mathbf{w}) \\ &= (A_v\mathbf{x}-A_v\mathbf{c})^T M^T M(A_v\mathbf{x}-A_v\mathbf{c}) \\ &= (\mathbf{x}-\mathbf{c})^T (MA_v)^T MA_v(\mathbf{x}-\mathbf{c}) \\ &= (\mathbf{x}-\mathbf{c})^T B^T B(\mathbf{x}-\mathbf{c})\end{aligned}$$

where $B = MA_v$. That is, we are left to solve

$$(\mathbf{x}-\mathbf{c})^T B^T B(\mathbf{x}-\mathbf{c}) \leq (1+r)(bb_{\varepsilon_1}\cdots b_{\varepsilon_r}).$$

As in section 5.5 finding all vectors satisfying this inequality amounts to computing all solutions to (5.2) contained in $\Sigma_v(\mathbf{l},\mathbf{h})$. The set of vectors $\mathbf{x}$ can be found using the Fincke-Pohst algorithm outlined in subsection 3.6.2.