Descents on Curves of Genus 1

Samir Siksek

Submitted by Samir Siksek
to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy
in the Faculty of Science
July 1995

This thesis is available for Library use on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

I certify that all the material in this thesis which is not my own work has been clearly identified and that no material is included for which a degree has previously been conferred upon me.

Samir Siksek

Abstract. In this thesis we improve on various methods connected with computing the Mordell-Weil group of an elliptic curve. Our work falls into several parts:

1. We give a new upper bound for the difference of the logarithmic and canonical heights of points on elliptic curves.

2. We give a new method for performing the infinite descent on an elliptic curve. This is essentially a lattice enlargement algorithm.

3. We show how to compute the 2-Selmer group of an elliptic curve defined over the rationals by a method which has complexity

   $$L_D(0.5, c_1) = (e^{(\log D)^{0.5}(\log\log D)^{0.5}})^{c_1+o(1)},$$

   where $D = |\Delta|$ the absolute value of the discriminant of the elliptic curve, and c_1 is a positive constant. This part is based on joint work with N. Smart.

4. We give a recipe for 'higher descents' on homogeneous spaces arising from the 2-descent. This is useful in dealing with homogeneous spaces which are everywhere locally soluble but for which a search for points does not reveal any global points.

5. We give algorithms for checking our homogeneous spaces for solubility over completions of number fields.

Acknowledgements

I am grateful to my supervisor John Cremona for his help and encouragement and for suggesting the topic of the thesis to me. I would also like to thank EPSRC for their financial support, Robin Chapman for putting up with my questions, Ray Miller and Jeremy Bygott for help with LaTeX.

Finally I would like to thank my family for their patience and support.

Contents

Chapter 1

Introduction

In this thesis we improve on various methods connected with computing the Mordell-Weil group of an elliptic curve. This is a deep and non-trivial problem with many interesting applications to diophantine equations. For reasons which will be made clear, it is not within the present “state-of-the-art” to be able to determine the Mordell-Weil group of every elliptic curve, even in theory. However we genuinely believe that the existing methods together with those developed in this thesis will eventually make it practical to compute the Mordell-Weil group of most elliptic curves defined over the rationals with a reasonably small discriminant.

We start by sketching the proof of the Mordell-Weil theorem. We assume that the reader is familiar with the basic theory of elliptic curves. Excellent references on the theory of elliptic curves are [Si2] and [Ca1]. For the basic algorithms concerning computing the Mordell-Weil group over the rationals see [Cre].

1.1 The Mordell-Weil Theorem

Let K be a number field. We shall normally take our elliptic curve defined over K to be in standard Weierstrass form:

$$E \quad : \quad Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6 \tag{1.1}$$

where $a_1, \ldots, a_6$ are in the ring of integers $\mathcal{O}_K$ of K.

Theorem 1.1.1 *$E(K)$ is finitely generated.*

This was proved by Mordell for elliptic curves defined over the rationals, and later extended by Weil to elliptic curves (as well as higher-dimensional abelian varieties) defined over arbitrary number fields. We shall sketch the basic idea of the proof, which falls in to 2 parts: the first is called the weak Mordell-Weil theorem, where one proves that $E(K)/2E(K)$ is finite, and the second is called the infinite descent, where it is shown that this implies that $E(K)$ is finitely generated.

1.1.1 The Weak Mordell-Weil Theorem

By a standard change of variable we may suppose that

$$E \quad : \quad Y^2 = X^3 + AX + B, \tag{1.2}$$

where $A,\ B \in \mathcal{O}_K$. We let $f(X) = X^3 + AX + B$, and let L be the K-algebra defined by

$$L = K[X]/(f(X)).$$

Then L is the sum of as many fields as $f(X)$ has irreducible factors in $K[X]$. We let Θ be the image of X under the natural map

$$K[X] \to L.$$

It turns out that we have a group homomorphism (see [Ca1] page 66 or [Ca6] page 31)

$$\alpha : E(K) \to L^*/L^{*2} \tag{1.3}$$

given explicitly by [1]

$$P = (x, y) \rightarrow (x - \Theta)L^{*2}. \tag{1.4}$$

This homomorphism has kernel $2E(K)$. Moreover its image is some finite subgroup of L^*/L^{*2}. Hence it will follow that the group $E(K)/2E(K)$ is finite. This is the first step in the proof of the Mordell-Weil Theorem.

For the 'generic' case where $f(X)$ is irreducible over K, and hence L is a field, we shall be more explicit about the image of α. It can be shown that the image of α is contained in the group

$$L(R, 2) = \{\beta \in L^*/L^{*2} : \mathrm{Norm}_{L/K}(\beta) \in K^{*2} \text{ and } \mathrm{ord}_{\wp}(\beta) \equiv 0 \pmod 2 \quad \text{if } \ \wp \notin R\}. \tag{1.5}$$

where R is the set of all primes in L which are either infinite or divide the discriminant Δ of the elliptic curve.

It will be seen that to determine $E(K)/2E(K)$ it is sufficient to determine for each $s \in L(R, 2)$ whether or not it is in the image of the map α and if it is to give a $P \in E(K)$ satisfying $\alpha(P) = s$.

Hence given $s \in L(R, 2)$ we must determine if it is possible to have

$$(x - \Theta) = s\epsilon^2 \tag{1.6}$$

for some $x \in K$ and $\epsilon \in L^*$, and if so determine the x (and ϵ) explicitly. Now any such ϵ can be written in the form

$$\epsilon = u_1 + u_2\Theta + u_3\Theta^2$$

where $u_1, \ u_2, \ u_3 \in K$. Substituting in equation (1.6) and comparing coefficients of $1, \ \Theta, \ \Theta^2$ we get

$$Q_1(u_1, u_2, u_3) = x \tag{1.7}$$

$$Q_2(u_1, u_2, u_3) = -1 \tag{1.8}$$

$$Q_3(u_1, u_2, u_3) = 0, \tag{1.9}$$

[1] The definition must be adjusted appropriately to give the correct image of the points of order 2, if there are any. See [Ca1] page 67.

where $Q_1,\ Q_2,\ Q_3$ are ternary homogeneous quadratic forms.

Here we would solve our problem for the particular s if and only if we can find a simultaneous solution to the last two equations above. We will call the simultaneous pair of equations [2]

$$\left.\begin{array}{l} Q_2(u_1,u_2,u_3) = -u_4^2 \\ Q_3(u_1,u_2,u_3) = 0 \end{array}\right\} \tag{1.10}$$

a homogeneous space (see [Si2] page 287). It is convenient to point out here that it is not always possible to determine if our homogeneous space (1.10) has solutions over K (these would be termed global solutions). However, in principle, there is no problem in checking if our homogeneous space has solutions over every local completion of K, and this is plainly a necessary condition for it to have global solutions. When (1.10) has solutions over every local completion of K, we will say that it is everywhere locally soluble. It turns out the set of all $s \in L(R,2)$ for which the corresponding homogeneous space 1.10 is everywhere locally soluble forms a subgroup of $L(R,2)$. This is termed the 2-Selmer group. We note here for later reference that if we determine which of the pairs of equations (1.10) have rational solutions, and for each of these find a point on it, then we will be able to recover a complete set of coset representatives of $E(K)/2E(K)$.

At any rate, for the proof of the Weak Mordell-Weil Theorem it suffices to note that $L(R,2)$ is finite, and hence that $E(K)/2E(K)$ is finite.

1.1.2 The Infinite Descent

We assume that the reader is familiar with the basic theory of heights in projective space and on elliptic curves (see [Si2] page 205-220).

Recall, that if M_K is a complete set of inequivalent valuations on K, then we define the naive height of a point $P = (X,Y) \in E(K)$ by

$$H(P) = \left(\prod_{\upsilon \in M_K} \max\{1, |X|_v\}^{n_v}\right)^{\frac{1}{[K:\mathbb{Q}]}} \tag{1.11}$$

[2] Note that we will prefer to write these in homogeneous form.

where $n_v = [K_v : \mathbb{Q}_v]$.

We define the logarithmic height of the point P by $h(P) = \log H(P)$, and finally the canonical height of the point P by

$$\hat{h}(P) = \lim_{n\to\infty} \left\{4^{-n} h(2^n P)\right\}. \tag{1.12}$$

It turns out that for any constant C, the points for which $H(P) \leq C$ are at most finitely many, and these may be effectively enumerated. To complete the proof of the Mordell-Weil Theorem we use the fact that $E(K)/2E(K)$ is finite to show that there is a C such that the points for which $H(P) \leq C$ generate $E(K)$. The first step here is to use the following theorem of Zagier.

Theorem 1.1.2 *(Zagier) Let $B_1 > 0$ be such that*

$$S = \left\{P \in E(K) : \hat{h}(P) \leq B_1\right\} \tag{1.13}$$

contains a complete set of coset representatives for $mE(K)$ in $E(K)$ ($m \geq 2$). Then the set S generates $E(K)$.

Proof. See [Cre] p61 or [Si1] p740. □

Now it remains to show that if B_1 is given by the above Theorem, then we can obtain a C such that the region $H(P) \leq C$ contains all the points for which $\hat{h}(P) \leq B_1$. This is possible at once since the difference $h(P) - \hat{h}(P)$ is absolutely bounded for any elliptic curve E. For example, Silverman has shown the following.

Theorem 1.1.3 *(Silverman) Let K be a number field and let E/K be given by the Weierstrass equation*

$$E \;:\quad Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6 \tag{1.14}$$

whose coefficients are in the ring of integers of K. Let Δ be the discriminant of the equation (1.14) and let j be the j-invariant of E. Further let

$$b_2 = a_1^2 + 4a_2 \quad \textit{and} \quad 2^* = \begin{cases} 2 & \textit{if } b_2 \neq 0, \\ 1 & \textit{if } b_2 = 0. \end{cases}$$

Define "height of E" (really of the Weierstrass equation (1.14)) by

$$\mu(E) = \frac{1}{12}h(\Delta) + \frac{1}{12}h_\infty(j) + \frac{1}{2}h_\infty(b_2/12) + \frac{1}{2}\log(2^*),$$

where [3], *for $t \in K$,*

$$h_\infty(t) = \frac{1}{[K:\mathbb{Q}]}\sum_{v \in M_K^\infty} n_v \log(\max(1, |t|_v))$$

Then for all $P \in E(\bar{K})$,

$$h(P) - \hat{h}(P) \leq \frac{1}{12}h(j) + 2\mu(E) + 1.946.$$

Proof. See [Si1]. □

If we let B_2 be the bound for $h(P) - \hat{h}(P)$ in Silverman's Theorem above, then we see $E(K)$ is generated by the points satisfying $H(P) \leq C$ where $C = \exp(B_1 + B_2)$. This completes the (sketched) proof of the Mordell-Weil Theorem.

1.2 Outline of the Usual Method of Computing the Mordell-Weil Group

The classical method of computing the Mordell-Weil group of an elliptic curve E over a number field K is via several distinct steps. We outline these below and explain if and why the method involved in each step is in need of improvement. Further we will summarize the contribution we have made towards making each step practical.

1.2.1 Computing the Torsion Subgroup of $E(K)$

This step is completely trivial for elliptic curves over $\mathbb{Q}$ (see [Cre] page 52). We will not consider the problem of computing the torsion subgroup for an elliptic curve defined over a number field.

[3] M_K^∞ is the set of archimedean valuations on K.

1.2.2 Computing the 2-Selmer Group of E

The best method for computing the 2-Selmer group for elliptic curves defined over $\mathbb{Q}$ and of small discriminants is using the algorithm of Birch and Swinnerton-Dyer (see [Bi, SwD]). Indeed Cremona's program `mwrank` (see page 19), which is an implementation of algorithm of Birch and Swinnerton-Dyer, computes 2-Selmer groups of elliptic curves of discriminants of size 10^{15} in a few minutes. In this algorithm the elements of the 2-Selmer group are represented by curves of the form

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

where $g(x)$ is a quartic polynomial with integral coefficients. Birch and Swinnerton-Dyer showed that for each element of 2-Selmer we can choose a representative as above where the coefficients of $g(x)$ lie in a certain region and have given invariants. Hence the algorithm involves searching this region for the polynomials with the given invariants. It turns out that the size of this region is at least $O(|\Delta|^{\frac{1}{2}})$ where Δ is the discriminant of the elliptic curve E (See [Bi, SwD] page 11). This is in fact the obstruction to using the method for elliptic curves of large discriminant.

There is a much older method of determining the 2-Selmer group. Birch and Swinnerton-Dyer comment on this:

> "It is possible to find the elements of G [the 2-Selmer group] by the classical process of descent; and for hand calculation this is probably the easiest way. However for any given curve ... one needs to know the structure of the appropriate algebraic number field, and it is not convenient to investigate this by means of an automatic computer. We have therefore used a different procedure ...". ([Bi, SwD] page 8).

Of course the situation concerning algorithms for computing the structure of algebraic number fields is now very different from that of the time Birch and Swinnerton-Dyer devised their algorithm ([Bi, SwD] appeared in 1963). There are now 'subexponential' algorithms for computing the class groups and fundamental units of algebraic number fields (see for example Chapters 5 and 6 of [Cohen] for a description of some of these). This strongly suggests that the descent via algebraic number fields deserves to be examined again. Indeed we show in Chapter 3, which is based on joint work with N. Smart, that this approach can be refined so that its complexity is

$$L_D(0.5, c_1) = (e^{(\log D)^{0.5}(\log\log D)^{0.5}})^{c_1+o(1)},$$

where $D = |\Delta|$, and c_1 is a positive constant. This is better than the complexity of the method of Birch and Swinnerton-Dyer.

The only implementation of these 'subexponential' algorithms for computing the class groups and fundamental units that is available to us is part of the package `Pari/GP` (see page 19). According to the manual ([Pari] page 46) these programs are "completely experimental", and we have found that they perform rather badly for cubic number fields of large discriminants. It is for this reason that no example is given for computing the 2-Selmer group using the method of Chapter 3. However it is hoped that improved implementations for computing the class group and fundamental units will make this method completely practical in the future.

1.2.3 Computing $E(K)/2E(K)$

Once we have computed the 2-Selmer group we hope to find enough points on $E(K)$ to show that the map from $E(K)/2E(K)$ to the 2-Selmer group is a surjection. Equivalently we wish to show that every homogeneous space corresponding to an element of the 2-Selmer group has point defined over K on it. This is not always possible because of the failure of the local-to-global principle for curves of genus 1. This leads us to the concept of 'higher descent ' discussed in Chapter 4 where we give a method often successful in resolving the problem of which homogeneous spaces have K-rational points. It should be noted that these methods do not always meet with success.

Both the computation of the 2-Selmer group and the method of 'higher descent' require algorithms for testing certain homogeneous spaces for local solubility. We give these in Chapters 5 and 6.

1.2.4 The Infinite Descent: Computing $E(K)$ from $E(K)/2E(K)$

Having obtained a set of generators for $E(K)/mE(K)$ we can compute all the coset representatives for $E(K)/mE(K)$ and hence their canonical heights. If B is an upper bound for these canonical heights then by Zagier's Theorem (1.1.2) we get an upper bound for the canonical heights of all the points of a set S (defined above) which generates $E(K)$. Combining this with Silverman's result (1.1.3) we get an upper bound B' for the logarithmic heights of all the points of S. It follows that the set S can be enumerated, provided of course that this upper bound is not too large.

Unhappily, practical experience suggests that the upper bound B' involved in this method is often too large. This can be for several reasons:

1. It is possible that the Silverman estimate on the difference between the logarithmic and canonical height is very large.
2. It is possible that the canonical heights of the generators of $E(K)/mE(K)$ are large .
3. It is also possible that even though the generators of $E(K)/mE(K)$ have small canonical heights, that some of the coset representatives (particularly if the rank is large) will have large heights.

We stress that the size of the search regions for the points of S increase exponentially with B'. To illustrate, if say $K = \mathbb{Q}$, and if $P = (X, Y) \in S$ then we can write $X = x/z^2$ where x and z are in $\mathbb{Z}$ and satisfy $|x| \leq \exp(B')$ and $|z| \leq \exp(B'/2)$. It follows that the search region here is roughly proportional to $\exp(1.5B')$. For a number field K of degree n over the rationals, the search region is, very roughly, between $\exp(1.5nB')$ and $\exp(2nB')$ in size. Hence small savings on B', can translate in to big savings in the actual size of the search region.

In Chapter 2 we will adopt a different approach to the infinite descent:

1. We will give an algorithm which will allow us, in most cases, to calulate a sharper upper bound for the quantity $h(P) - \hat{h}(P)$.

2. We will show how a basis of a submodule of the torsion-free part of $E(K)$, having full rank, can be enlarged efficiently to a basis for $E(K)$.

The algorithm for infinite descent we will give uses both of these ingredients, and involves searching much smaller regions than the above.

1.3 Applications of Computing the Mordell-Weil Group

As noted already, we will be concerned with the problem of determining a basis for the Mordell-Weil group of an elliptic curve. We hope to convince the reader that this is an interesting and engaging problem in itself. It is however appropriate to describe some of the applications of computing the Mordell-Weil group of an elliptic curve [4]:

1.3.1 Describing Rational Solutions to Elliptic Diophantine Equations.

Many diophantine problems are equivalent to computing the Mordell-Weil group of an elliptic curve, or showing that an elliptic curve has rank at least 1. This includes many geometrical problems. A well-known example is the so called 'congruent number problem' (see [Kob]): An integer n is said to be a 'congruent number' if it is the area of some right-angled triangle with rational sides. It turns out that n is congruent if and only if the elliptic curve

$$Y^2 = X(X^2 - n^2)$$

[4] It should be noted that what is presented here is necessarily a random sample, and that some parts are perhaps out of date. This is because of the vastness of the topic and the modest knowledge of the author. Nice references here are [Guy], [Mord].

has rank at least 1, and that to find a particular right-angled triangle with area n it necessary to find a point of infinite order on this curve. This can be a highly non-trivial problem. For example, as a corollary to the calculations in Section (4.7) it turns out that the simplest right-angled triangle having area 2833 has base and perpendicular

$$\frac{5334745291350384}{709516254613385}, \quad \frac{20100595493197l9705}{2667372645675192},$$

and has hypotenuse

$$\frac{1426240910614742861472434930476897}{18925442492176578988382456449 20}.$$

1.3.2 Integral Points on Elliptic Diophantine Equations.

Many diophantine problems are equivalent to computing all the integral points on a model of an elliptic curve. For elliptic curves in standard minimal Weierstrass form there is now a practical algorithm for performing this using elliptic logarithms (see [GPZ], [Smart], [Str, Tz], and for a generalization to number fields [Sm, Ste]). These algorithms require the computation of the Mordell-Weil group beforehand. This method has been extended (see [Tz]) to finding the integral points on elliptic curves of the form

$$Y^2 = f(X)$$

where $f(X) \in \mathbb{Z}[X]$ is a quartic polynomial. [5]

As an example, we mention Ljunggren's infamous equation (see [Mord] page 271)

$$Y^2 = 2X^4 - 1.$$

Ljunggren had shown that the only integral solutions of this are $(\pm 1, \pm 1)$ and $(\pm 13, \pm 239)$. However his method was exceedingly complicated, especially for such an innocuous looking equation, and Mordell had wished if only a simpler proof could be found. It is striking to note that this can now be resolved by a

[5] See [Str, We] for an example which arises from the theory of Radon Transforms.

couple of computer programs. Letting $y = 2XY$ and $x = 2X^2$ we find that

$$E : y^2 = x(x^2 - 2). \tag{1.15}$$

Using `mwrank` (see page 19) we find that the Mordell-Weil group of this elliptic curve is $\langle(0,0)\rangle \bigoplus \langle(-1,1)\rangle$. We thank N. Smart for computing the integral points on this equation using his own implementation of the elliptic logarithm method mentioned above. After a minute or so the program output that the only integral points on E are

$$(0,0),\ (-1,\pm1),\ (2,\pm2),\ (338,\pm6214).$$

It easily follows that the only integral solutions to Ljunggren's equation are the ones he gave.

1.3.3 Rational Points on Certain Curves of Genus > 1.

Given a curve of genus > 1, we can occasionally cover an elliptic curve by this curve, and then use the Mordell-Weil group of the elliptic curve to obtain information about its rational points. We give one example of a method essentially due to Dem'Janenko (see [Ca5]). Suppose we wanted to determine the rational points on

$$X^4 + Y^4 = 2Z^4. \tag{1.16}$$

If (X, Y, Z) is a non-trivial solutions of the equation (1.16) then we may assume that $X,\ Y,\ Z$ are coprime integers. It easily follows that

$$P_1 = \left(\frac{2Z^2}{X^2}, \frac{2Y^2Z}{X^3}\right),\ P_2 = \left(\frac{2Z^2}{Y^2}, \frac{2X^2Z}{Y^3}\right),$$

are rational points on the elliptic curve (1.15). We note that the naive heights of P_1, and P_2 are equal: it is clear that $2Z^2 > X^2,\ Y^2$. Hence it follows that the difference between their canonical heights is bounded. Using Silverman's estimates for the difference between the logarithmic and canonical heights we get

$$|\hat{h}(P_1) - \hat{h}(P_2)| \leq 9.6988.$$

We recall that the Mordell-Weil group of (1.15) is

$$\langle (0,0)\rangle \bigoplus \langle (-1,1)\rangle$$

where $(0,0)$ is a point of order 2 and $(-1,1)$ is of infinite order. Write

$$P_i = m_i(-1,1) + n_i(0,0)$$

where $m_i \in \mathbb{Z}$ and $n_i \in \{0,1\}$ $(i = 1,2)$. It follows that

$$|m_1^2 - m_2^2| \leq \frac{9.6988}{\hat{h}((-1,1))} = 15.93.$$

From this we deduce that either $m_1 = m_2$ or $|m_1|+|m_2| \leq 15$. It is now a simple matter to check that the only (non-trivial) solutions to (1.16) are $(\pm 1, \pm 1, \pm 1)$.

1.3.4 Rational Points on Certain Surfaces.

It is possible to describe certain surfaces by a parametric family of elliptic curves. In this case one can obtain information about the rational points on the surface by studying the Mordell-Weil groups of these elliptic curves. For example, in [SwD2], Swinnerton-Dyer uses this idea to show that the rational points on the variety

$$X^4 + Y^4 = Z^4 + W^4$$

are dense (with respect to the Euclidean topology).

An impressive recent success for this method is the counterexample by N. D. Elkies (see [Elkies]) to a conjecture of Euler that there are no solutions in positive integers to

$$A^4 + B^4 + C^4 = D^4. \tag{1.17}$$

Elkies parametrized this as a pencil of curves of genus 1. By finding the simplest curve in the pencil which is everywhere locally soluble and checking that it has a solution, he found a solution [6] to Euler's (1.17):

$$2682440^4 + 15365639^4 + 18796760^4 = 20615673^4 \quad .$$

He also showed that the rational points are dense in the real locus of

$$r^4 + s^4 + t^4 = 1.$$

1.4 Computer Packages

In preparing the examples in this thesis, we have found it useful to use some computer packages and programs which we list below.

1.4.1 `mwrank` and `findinf`

These are programs written by J. Cremona for elliptic curves defined over $\mathbb{Q}$. `mwrank` is an implementaion of the Birch and Swinnerton-Dyer method of 2-descent ([Bi, SwD] and [Cre] pages 68-76). It also attempts an infinite descent via the traditional method explained on page 9.

The 2-descent step is remarkably successful for curves of small discriminant [7]. The infinite descent is not so successful for reasons explained on page 14.

`findinf` is a program for searching for points up to a given logarithmic height on an elliptic curve using a quadratic sieve method.

1.4.2 `Pari/GP`

We have found this package very useful for number-theoretic computations. It has many functions for doing arithmetic on elliptic curves, including elliptic

[6] Clearly this solution could not be easily found by a naive computer search. The smallest solution

$$95800^4 + 217519^4 + 414560^4 = 422481^4$$

was later found by Roger Frye - using Elkies' ideas- in a search which took 100 hours of computer time.

[7] Though of course there is no guarantee of finding rational points on all the everywhere locally soluble homogeneous spaces.

logarithms, and canonical height computations. Moreover it provides tools for dealing with modular arithmetic, algebraic numbers, p-adic numbers.

All the programming we did was done using this package.

Chapter 2

The Infinite Descent

The contents of this chapter have been accepted for publication by the Rocky Mountain Journal of Mathematics.

2.1 The bound on the difference $h(P) - \hat{h}(P)$

2.1.1 Preliminaries

Let E be an elliptic curve given by the Weierstrass equation

$$E \ : \quad Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6 \tag{2.1}$$

where $a_1, \ldots, a_6$ are in the ring of integers $\mathcal{O}_K$ of a number field K. In this section we shall give an algorithm for obtaining an upper bound for the quantity $h(P)-\hat{h}(P)$. This is based on the traditional method of estimating the difference $h(2P)-4h(P)$. Generally speaking, when this has been done in the past, it relied on the use of elimination theory, which leads to poor upper bounds. The method we shall give bypasses elimination theory using explicit calculations over some local completions of K.

Apart from Silverman's Theorem 1.1.3, there are other results which give bounds on the quantity $h(P) - \hat{h}(P)$, most notably in [Zim] and [Dem]. The reason why we make specific comparisons only with Silverman's theorem is that this is currently the most widely used and quoted in the literature.

As our method is very different from Silverman's method for obtaining his estimate (1.1.3), we have no easy way of deciding a priori which should give the smaller bound. We can only note that, in practice, we have found that our method gives much smaller bounds most of the time, or exceptionally bounds which are slightly better. For example, a straightforward application of Silverman's Theorem 1.1.3 for the curve

$$Y^2 + Y = X^3 - 7X + 6$$

gives

$$h(P) - \hat{h}(P) \leq 5.4.$$

In [BGZ] Buhler, Gross and Zagier derive that

$$h(P) - \hat{h}(P) \leq 0 \quad \text{for all } P \in E(\mathbb{Q}),$$

and we get this also by applying our Theorem 2.1.1. Needless to say, here our method gave a much better bound than Silverman's. In contrast to this, for the curve

$$Y^2 = X(X^2 - p^2)$$

where p is prime and > 2, Silverman's theorem gives

$$h(P) - \hat{h}(P) \leq \log(p) + 4.505$$

and our Theorem 2.1.1 gives

$$h(P) - \hat{h}(P) \leq \log(p) + 0.347 \quad \text{for all } P \in E(\mathbb{Q}).$$

Here for small primes p our bound looks much better and for large p it looks roughly the same as Silverman's. However, even here, the extra work we had to do to get our bound was worthwhile, since to search for all rational points on the curve of canonical height $\leq B$, the size of the search region if we apply our bound is roughly

$$1.682p^{1.5}\exp(1.5B),$$

and if we apply Silverman's bound it is roughly

$$860.488 p^{1.5} \exp(1.5B).$$

Accordingly, we believe, that the small amount of work that goes into obtaining our bound will usually be amply rewarded by the time saved through searching smaller regions.

We employ some standard notation to do with number fields and elliptic curves. Given a number field K we let M_K be the set of all valuations on K. We write M_K^0 and M_K^∞ for the sets of non-archimedean and archimedean valuations on K respectively. For an elliptic curve E given by a Weierstrass equation of the form (2.1) we define some associated constants (see [Si2] page 46):

$$\begin{aligned} b_2 &= a_1^2 + 4a_2, \\ b_4 &= 2a_4 + a_1 a_3, \\ b_6 &= a_3^2 + 4a_6, \\ b_8 &= a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2, \\ \Delta &= -b_2^2 b_8 - 8b_4^3 - 27b_6^2 + 9b_2 b_4 b_6. \end{aligned} \tag{2.2}$$

Let

$$\begin{aligned} f(X) &= 4X^3 + b_2 X^2 + 2b_4 X + b_6 \\ g(X) &= X^4 - b_4 X^2 - 2b_6 X - b_8. \end{aligned} \tag{2.3}$$

It will be seen that the polynomials $f,\ g$ arise in the duplication formula for a point on the curve E and a little study of these polynomials essentially gives us our required bound for $h(P) - \hat{h}(P)$.

As usual, we denote the residue field of a completion K_v with respect an non-archimedean prime v by k_v, and we denote the canonical map $K_v \to k_v \cup \{\infty\}$ by $x \to \bar{x}$. We let π be a prime element for v (i.e. $\pi \in K_v$ such that $v(\pi) = 1$).

Lemma 2.1.1 *Suppose that v is a non-archimedean valuation on K and $P = (x, y) \in E(K_v)$ is such that its reduction $\bar{P} = (\bar{x}, \bar{y}) \in E(k_v)$ is non-singular. Then*

$$\max\{|f(x)|_v, |g(x)|_v\} = \max\{1, |x|_v\}^4 .$$

Proof. If $|x|_v > 1$ then $|f(x)|_v \leq |x|_v^3$ and $|g(x)|_v = |x|_v^4$ and in this case the conclusion is obvious.

Hence we can suppose that $|x|_v \leq 1$. Now we are required to prove that

$$\max\{|f(x)|_v, |g(x)|_v\} = 1$$

Hence it is enough to show that when $f(x) \equiv 0 \pmod{\pi}$ and $g(x) \equiv 0 \pmod{\pi}$ then $\bar{P}$ is singular on $E(k_v)$.

By a change of variable which is non-singular modulo π, we may suppose that $(x, y) = (0, 0)$. Now the condition for $(0, 0)$ to be on the Weierstrass equation is that $a_6 = 0$. Moreover, since $f(0) \equiv g(0) \equiv 0 \pmod{\pi}$ we get that $b_6 \equiv b_8 \equiv 0 \pmod{\pi}$. Hence from the formulae for $b_6,\ b_8$ we get that $a_3 \equiv a_4 \equiv 0 \pmod{\pi}$. This is a sufficient condition for $(0, 0)$ to be singular on $E(k_v)$. □

Here is some more notation which we will find useful:

$$\begin{aligned} f'(X') &= X'^4 f(\tfrac{1}{X'}) \\ g'(X') &= X'^4 g(\tfrac{1}{X'}). \end{aligned} \tag{2.4}$$

Further let, for each $v \in M_K$,

$$D_v = \left\{X \in K_v : |X|_v \leq 1 \textit{ and } f(X) \in {K_v}^2\right\}$$

$$D'_v = \left\{X' \in K_v : |X'|_v \leq 1 \textit{ and if } X' \neq 0 \textit{ then } f\left(\frac{1}{X'}\right) \in {K_v}^2\right\}$$

.

Lemma 2.1.2 *Define constants d_v, d'_v by*

1. $d_v = \inf_{X \in D_v} \max\{|f(X)|_v, |g(X)|_v\}$,
2. $d'_v = \inf_{X' \in D'_v} \max\{|f'(X')|_v, |g'(X')|_v\}$.

Then, d_v, d'_v are non-zero.

Proof. We begin by noting that the sets D_v, D'_v, are compact subsets of K_v (with respect to the v-adic topology), and hence the infimums d_v, d'_v must be attained. If say d_v was zero then there would exist $X_1 \in D_v$ such that $f(X_1) = g(X_1) = 0$. However, from [Si3] p347 we have that

$$\text{Resultant}(f, g) = \text{Resultant}(f', g') = \Delta^2$$

where Δ is the discriminant of the elliptic curve E. Accordingly, as this cannot be zero, $d_v \neq 0$. Similarly $d'_v \neq 0$.

□

If E is minimal at some non-archimedean valuation v then we define

$$c_v = [E(K_v) : E^0(K_v)].$$

i.e. c_v is the Tamagawa index at v.

Lemma 2.1.3 *Let, for any valuation v on K,*

$$\epsilon_v{}^{-1} = \inf_{(X,Y)\in E(K_v)} \frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} \tag{2.5}$$

Then

1. *ϵ_v exists. (i.e. the quantity on the right exists and is non-zero). Moreover $\epsilon_v{}^{-1} = \min(d_v, d'_v)$.*

2. *$\epsilon_v \geq 1$.*

3. *If v is non-archimedean, E is minimal at v, and the local Tamagawa index $c_v = 1$, then $\epsilon_v = 1$.*

4. *If v is non-archimedean, then $\epsilon_v = d_v{}^{-1}$ where d_v is as defined in Lemma (2.1.2).*

5. *If v is non-archimedean, and*

$$\lfloor \frac{v(4\Delta)}{2} \rfloor = n,$$

then $\epsilon_v \leq |\pi|_v^{-2n}$ (where $\lfloor\ \ \rfloor$ denote the integer part of a number).

Proof. Suppose $(X,Y) \in E(K_v)$. Then by a standard manipulation of the Weierstrass equation (2.1) we get

$$(2Y + a_1X + a_3)^2 = f(X) \tag{2.6}$$

Hence, if $|X|_v \leq 1$ then $X \in D_v$ and

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} = \max(|f(X)|_v, |g(X)|_v).$$

If $|X|_v \geq 1$ then $X' = X^{-1} \in D'_v$ and

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} = \max(|f'(X')|_v, |g'(X')|_v).$$

Hence it is clear that the quantity on the right of (2.5) exists and is equal to $\min(d_v, d'_v)$, and so is non-zero (by Lemma (2.1.2)). This proves the first part of the above.

For the second part we note that we may take $(X,Y) \in E(K_v)$ to be arbitrarily close to 0. Hence X is unbounded with respect to the metric $|\ |_v$ and so

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4}$$

is arbitrarily close to 1. It follows that ${\epsilon_v}^{-1} \leq 1$, and hence that $\epsilon_v \geq 1$, as required for part 2.

Part 3 is clear from Lemma (2.1.1).

For part 4 we note that if v is non-archimedean and $|X|_v > 1$ then by the proof of Lemma (2.1.1),

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} = 1,$$

and if $|X|_v \leq 1$ then

$$\frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4} \leq 1,$$

so by the definition of ϵ_v we get

$${\epsilon_v}^{-1} \leq \inf_{(X,Y) \in E(K_v),\ |X|_v \leq 1} \max(|f(X)|_v, |g(X)|_v)$$

which immediately gives part 4.

Let us now prove part 5. Let n be as defined in the Lemma. Suppose that

$$\inf_{X \in D_v} \max(|f(X)|_v, |g(X)|_v) \leq |\pi|_v^{2n+1}$$

and it is sufficient to derive a contradiction. If this were the case then there would exist $(X, Y) \in E(K_v)$, with

$$f(X) \equiv g(X) \equiv 0 \pmod{\pi^{2n+1}}.$$

But from equation (2.6) we must deduce that $f(X) \equiv 0 \pmod{\pi^{2n+2}}$. We now invoke the following identity:

$$4g(X) = (6X^2 + b_2X + b_4)^2 - (8X + b_2)f(X). \tag{2.7}$$

This is easily verified. It follows that $(6X^2 + b_2X + b_4)^2 \equiv 0 \pmod{\pi^{2n+2}}$. Finally we use congruence

$$[48X^2 + 8b_2X + (-b_2^2 + 32b_4)](6X^2 + b_2X + b_4)^2 \equiv -4\Delta \pmod{f(X)} \tag{2.8}$$

in $\mathbb{Z}[X, a_1, \ldots, a_6]$. This is straight forward but rather tedious to verify (it is a slightly more general form of the congruence in page 51 of [Ca1]). We can now conclude that π^{2n+2} divides 4Δ as required. □

For a non-archimedean valuation v, we let (as usual) $E^0(K_v)$ be the subgroup of points on $E(K_v)$ with non-singular reduction modulo π. It is useful to define $\mu_v = \mu_v(E)$ as follows:

1. if v is archimedean, then $\mu_v = \frac{1}{3}$,

2. if v is non-archimedean and E is not minimal at v, then $\mu_v = \frac{1}{3}$,

3. if v is non-archimedean and E is minimal at v, then

$$\mu_v = \begin{cases} 0 & if\ [E(K_v) : E^0(K_v)] = 1 \\ 1/4 & if\ E(K_v)/E^0(K_v) \cong \mathbb{Z}/2\mathbb{Z}\ or\ \left(\mathbb{Z}/2\mathbb{Z}\right)^2 \\ \left(1 - \frac{1}{4^\alpha}\right)/3 & if\ E(K_v)/E^0(K_v) \cong \mathbb{Z}/2^\alpha\mathbb{Z}\ where\ \alpha \geq 1 \\ 1/3 & if\ [E(K_v) : E^0(K_v)]\ is\ not\ a\ power\ of\ 2. \end{cases}$$

Here we recall that for non-archimedean υ at which E is minimal, the group $E(K_\upsilon)/E^0(K_\upsilon)$ is either cyclic or is equal to $(\mathbb{Z}/2\mathbb{Z})^2$ (see for example Theorem VII.6.1 on page 183 of [Si2]). Hence the above definition for υ covers all the possible cases.

We are now ready to state our main Theorem on the bound $h - \hat{h}$.

Theorem 2.1.1 *Let M_K be a complete set of inequivalent valuations on K. For each $\upsilon \in M_K$, let $n_\upsilon = [K_\upsilon : \mathbb{Q}_\upsilon]$. Define a function*

$$\epsilon : \; M_K \times E(K) \to \mathbb{R}_{\geq 1} \tag{2.9}$$

by

$$\epsilon(\upsilon, P) = \begin{cases} 1 & \text{if } \upsilon \in M_K^0, \ E \text{ is minimal at } \upsilon, \text{ and } P \in E^0(K_\upsilon) \\ \epsilon_\upsilon & \text{otherwise.} \end{cases} \tag{2.10}$$

Then for all $P \in E(K)$ we have

$$\begin{aligned} h(P) - \hat{h}(P) \quad &\leq \quad \tfrac{1}{[K:\mathbb{Q}]} \left(\textstyle\sum_{\upsilon \in M_K} \mu_\upsilon n_\upsilon \log(\epsilon(\upsilon, P)) \right) \\ &\leq \quad \tfrac{1}{[K:\mathbb{Q}]} \left(\textstyle\sum_{\upsilon \in M_K} \mu_\upsilon n_\upsilon \log(\epsilon_\upsilon) \right). \end{aligned} \tag{2.11}$$

We note here that if υ is non-archimedean, E is minimal at υ, and the Tamagawa index $c_\upsilon = 1$, then by the definition for μ_υ above, and Lemma (2.1.3) we have that $\mu_\upsilon = \log(\epsilon(\upsilon, P)) = \log(\epsilon_\upsilon) = 0$. Hence only finitely many terms in the above sums are non-zero.

Proof. We begin by noting that for all $P \in E(K), \ \upsilon \in M_K$,

$$\max(|f(X)|_\upsilon, |g(X)|_\upsilon) \geq \epsilon(\upsilon, P)^{-1} \max(1, |X|_\upsilon)^4 \tag{2.12}$$

using the definition of ϵ_υ on page 25, and the definition of $\epsilon(\upsilon, P)$ above, and Lemma (2.1.1).

Now if $P = (X, Y) \in E(K)$ then by the duplication formula (see [Si2] p59) the x-coordinate of $2P$ is $g(X)/f(X)$. Hence using the product definition for naive heights and Lemma (2.1.1) above we get

$$\begin{aligned} H_K(2P) &= \prod\nolimits_{\upsilon\in M_K} \max\left\{|f(X)|_\upsilon, |g(X)|_\upsilon\right\}^{n_\upsilon} \\ &\geq \prod\nolimits_{\upsilon\in M_K} \left(\epsilon(\upsilon,P)^{-1} \max\left\{1, |X|_\upsilon\right\}^4\right)^{n_\upsilon} \qquad (2.13) \\ &= \left(\prod\nolimits_{\upsilon\in M_K} \epsilon(\upsilon,P)^{-n_\upsilon}\right) H_K(P)^4. \end{aligned}$$

Recall that

$$h(P) = \frac{1}{[K:\mathbb{Q}]}\log(H_K(P))$$

and so

$$h(2P) - 4h(P) \geq \frac{1}{[K:\mathbb{Q}]}\left(\sum_{\upsilon\in M_K} n_\upsilon \log(\epsilon(\upsilon,P)^{-1})\right).$$

Rearranging, we get

$$h(P) \leq \frac{1}{4}h(2P) + \frac{1}{4[K:\mathbb{Q}]}\left(\sum_{\upsilon\in M_K} n_\upsilon \log(\epsilon(\upsilon,P))\right).$$

Using

$$\hat{h}(P) = \lim_{n\to\infty} 4^{-n}h(2^nP)$$

we get

$$h(P) \leq \frac{1}{[K:\mathbb{Q}]}\left(\sum_{\upsilon\in M_K} n_\upsilon \left(\sum_{n=1}^{\infty}\frac{1}{4^n}\log(\epsilon(\upsilon,2^nP))\right)\right) + \hat{h}(P).$$

However, from the definition of the function ϵ we find that

$$\log(\epsilon(\upsilon,2^nP)) = \begin{cases} 0 & \upsilon\in M_K^0,\ E \text{ is minimal at } \upsilon, \text{ and } 2^nP\in E^0(K_\upsilon), \\ \log(\epsilon_\upsilon) & \text{otherwise.} \end{cases}$$

It is now an easy matter to show that for all $\upsilon \in M_K$,

$$\sum_{n=1}^{\infty}\frac{1}{4^n}\log(\epsilon(\upsilon,2^nP)) \leq \mu_\upsilon \log(\epsilon(\upsilon,P))$$

where $\mu_\upsilon$is as defined above. This completes the proof. □

It is apparent from our Theorem above that to get an upper bound on $h - \hat{h}$, all that remains is to calculate the values ϵ_v at the finitely many valuations for which μ_v is not zero: to recall these are the cases when either v is archimedean (i.e. where $K_v = \mathbb{R}$ or $\mathbb{C}$), or where v is non-archimedean but E is not minimal at v, or it is minimal but the Tamagawa index $c_v \neq 1$.

We give separate algorithms for calculating $\epsilon_v = \min(d_v, d'_v)^{-1}$ for three different cases:

- $K_v = \mathbb{R}$
- $K_v = \mathbb{C}$
- v is non-archimedean.

2.1.2 v is Real

Suppose that $K_v = \mathbb{R}$. Note that there exists $\sigma \in Gal(K/\mathbb{Q})$ such that $K^\sigma \subset \mathbb{R}$ and for all $x \in K,\ |x|_v = |x^\sigma|$ where $|\ |$ is the ordinary absolute value. Hence, by replacing $f,\ g,\ f',\ g'$ by $f^\sigma,\ g^\sigma,\ f'^\sigma,\ g'^\sigma$ if necessary, we can assume $f,\ g,\ f',\ g'$ are all real polynomials. Now the problem is reduced to finding

$$d_v = \inf_{X \in D_v} \max\left\{|f(X)|_v, |g(X)|_v\right\},$$

$$d'_v = \inf_{X' \in D'_v} \max\left\{|f'(X')|_v, |g'(X')|_v\right\},$$

where

$$D_v = \{X \in \mathbb{R} : |X| \leq 1 \ \mathit{and}\ f(X) \geq 0\}$$

and

$$D'_v = \left\{X' \in \mathbb{R} : |X'| \leq 1 \ \mathit{and\ either}\ X' = 0 \ \mathit{or}\ f(\frac{1}{X'}) \geq 0\right\}$$

are clearly finite unions of intervals. Finally we use the following elementary lemma.

Lemma 2.1.4 *If $f,\ g$ are continuous real functions and I is an interval then the infimum of the continuous function* $\max\{|f(X)|, |g(X)|\}$ *over the interval I is attained at one of the following points*

(i) *an end point of I,*

(ii) *at one of the roots of f, g, $f+g$, $f-g$ in the interval I,*

(iii) *at a turning point of one of the functions f, g.*

Proof. We simply note that at any point in I not listed in (i) or (ii), the function $\max\{|f(X)|, |g(X)|\}$ is equal to one of $\pm f$, $\pm g$ and its infimum must be a local supremum or infimum of f, or g. □

Hence, to calculate d_v, we write D_v as a union of intervals (I) and calculate the infimum of $\max\{|f(X)|, |g(X)|\}$ over each interval separately using the above Lemma, and then d_v will be the minimum of these (finitely many) infima. Similarly we calculate d'_v, and then $\epsilon_v = \min(d_v, d'_v)^{-1}$.

2.1.3 v is Complex

Suppose that $K_v = \mathbb{C}$. In the same way as the real case, we can if necessary replace f, g, f', g' by appropriate conjugates so that

$$d_v = \inf_{X \in D_v} \max\{|f(X)|_v, |g(X)|_v\},$$

$$d'_v = \inf_{X' \in D'_v} \max\{|f'(X')|_v, |g'(X')|_v\},$$

where $D_v = D'_v = D = \{z \in \mathbb{C} : |z| \leq 1\}$ is the closed unit disc. We make use of the following Lemma.

Lemma 2.1.5 *let f and g be as above. Then the continuous function $k : \mathbb{C} \to \mathbb{R}_{>0}$ defined by*

$$k(z) = \max\{|f(z)|, |g(z)|\}$$

attains its infimum over D at a point z_0 satisfying **either**

1. $|z_0| = 1$ *(i.e. it is on the boundary of D),* **or**

2. $|f(z_0)| = |g(z_0)|$.

Proof. For each $\rho \in \mathbb{C}$ there are, counting multiplicities, 4 solutions to the equation $f(X) = \rho g(X)$. In fact by Cardano's formulae, there exist 4 functions $\phi_1, \ldots, \phi_4 : \mathbb{C} \to \mathbb{C}$ such that $\phi_1(\rho), \ldots, \phi_4(\rho)$ are solutions to $f(X) = \rho g(X)$.

Let

$$S = \{\rho \in \mathbb{C} : |\rho| = 1\}.$$

It follows that each $\phi_i(S)$ is a path in $\mathbb{C}$. We note that for all $z \in \mathbb{C}$, $|f(z)| = |g(z)|$ if and only if there exist $\rho \in S$ such that $f(z) = \rho g(z)$ and hence if and only if $z \in \phi_i(S)$ for some i.

Now the paths $\phi_1(S), \ldots, \phi_4(S)$ divide the unit disc D into finitely many connected regions $U_1, \ldots, U_n$. Consider a region U_j; denote the interior of U_j by $\mathrm{int}(U_j)$ and its closure by $\overline{U}_j$. We note that that the intersection of $\mathrm{int}(U_j)$ and $\phi_i(S)$ is empty for $i = 1, \ldots, 4$. Hence, by the connectedness of U_j, we get that either $|f| > |g|$ or $|g| > |f|$ on all of $\mathrm{int}(U_j)$. Suppose, without loss of generality, that $|f| > |g|$ on all of $\mathrm{int}(U_j)$. Then $k(z) = |f(z)|$ for all $z \in \overline{U}_j$. It is easy to see that f is never zero on $\overline{U}_j$: if f is zero at some point of $\overline{U}_j$, then g is also zero at that point, contradicting Lemma (2.1.2). Let $w(z) = \frac{1}{f(z)}$. Then w is holomorphic on $\mathrm{int}(U_j)$ and continuous on $\overline{U}_j$ and so by the Maximum Modulus Theorem of Complex Analysis (see [Pr] p76), it attains its maximum modulus over $\overline{U}_j$ on the boundary $\overline{U}_j \backslash \mathrm{int}(U_j)$. Hence $k(z) = |f(z)|$ attains its infimum over $\overline{U}_j$ on the boundary $\overline{U}_j \backslash \mathrm{int}(U_j)$. But each of these boundaries is a subset of $S \cup \phi_1(S) \cup \ldots \cup \phi_4(S)$. Since the $\overline{U}_j$ cover D we get that k attains its infimum over D on $S \cup \phi_1(S) \cup \ldots \cup \phi_4(S)$. This is the statement of the theorem. □

It is plain that the Lemma is true for f', g', instead of f, g. Now it is necessary to estimate $\inf \{|f|, |g|\}$ over the boundary S, and over the sections of the paths $\phi_i(S)$ inside the unit disc D. We will use the following naive method. Fix some $n \geq 2$ (this should be roughly 1 more than the number of significant digits we want to determine d_v to). Let $\theta_j = 10^{-n} j$ for $j = 1, \ldots, 10^n$. For each θ_j we solve (numerically) the equation

$$f(X) = e^{2\pi\theta_j} g(X),$$

and let

$$\kappa_j = \min \left\{ \max(|f(e^{2\pi\theta_j})|, |g(e^{2\pi\theta_j})|) \right\} \cup \left\{ |f(X)| : X \in D \textit{ and } f(X) = e^{2\pi\theta_j} g(X) \right\}.$$

Finally, we take $d_v = \min(\kappa_j)$. Similarly, we estimate d'_v, and take $\epsilon_v = \min(d_v, d'_v)^{-1}$.

Of course, this method is crude, and great improvements must be possible, but we will not do this.

2.1.4 v is Non-Archimedean

In this section we want to calculate

$$\epsilon_v{}^{-1} = \inf_{(X,Y)\in E(K_v)} \frac{\max(|f(X)|_v, |g(X)|_v)}{\max(1, |X|_v)^4}$$

for non-archimedean v. We note by Lemma (2.1.1), that if the reduction of the curve $E(k_v)$ is non-singular then $\epsilon_v = 1$. Hence, we can assume that E has bad reduction at v, and calculate the infimum over the points of $E(K_v)$ which have singular reduction modulo v. To do this we define the following sequence of sets:

We define U_i for $i = 1, 2, \ldots$, to be the set of all $X \pmod{\pi^{2i}}$ satisfying

1. $f(X) \equiv 0 \pmod{\pi^{2i}}$,
2. $g(X) \equiv 0 \pmod{\pi^{2i-1}}$, and
3. there exists $X_0 \in K_v$ such that $X \equiv X_0 \pmod{\pi^{2i}}$ and $f(X_0) \in K_v{}^2$.

And we define V_i for $i = 1, 2, \ldots$, to be the set of all $X \pmod{\pi^{2i}}$ satisfying

1. $f(X) \equiv g(X) \equiv 0 \pmod{\pi^{2i}}$,
2. there exists $X_0 \in K_v$ such that $X \equiv X_0 \pmod{\pi^{2i}}$ and $f(X_0) \in K_v{}^2$.

Lemma 2.1.6 *1. Suppose $v(2) = 0$. If $i \geq 1$ and $U_i \neq \emptyset$, then $V_i = U_i$ and $\pi^{2i} \mid \Delta$.*

2. Suppose $v(2) = e > 0$. If $U_i \neq \emptyset$ or $V_i \neq \emptyset$, then $\pi^{2i} \mid 4\Delta$.

Proof. We recall the identity and the congruence we used in the proof of Lemma (2.1.3) (on page 27)

$$4g(X) = (6X^2 + b_2X + b_4)^2 - (8X + b_2)f(X). \tag{2.14}$$

$$[48X^2 + 8b_2X + (-b_2^2 + 32b_4)](6X^2 + b_2X + b_4)^2 \equiv -4\Delta \pmod{f(X)}. \tag{2.15}$$

It follows from the first that if $\upsilon(2) = 0$, and $X \in U_i$, then

$$(6X^2 + b_2X + b_4)^2 \equiv 0 \pmod{\pi^{2i}}$$

and so $\pi^{2i} \mid g(X)$ and so $X \in V_i$. Further, by the congruence, $\pi^{2i} \mid \Delta$, and this completes the proof of the first part. The proof of the second part is similar. □

Corollary 2.1.1 *If $\upsilon(2) = 0$ and $U_1 = \emptyset$ then $\epsilon_\upsilon = 1$. If $U_j \neq \emptyset$ and $U_{j+1} = \emptyset$ then $\epsilon_\upsilon = {|\pi|_\upsilon}^{-2j}$.*

Hence if $\upsilon(2) = 0$ then we compute (U_i) explicitly for $i = 1, 2, \ldots$ until we reach the empty set. Then the value of ϵ_υ is given by the above corollary. Here in calculating the (U_i), it is needed to be able to test, given $X \pmod{\pi^{2i}}$, if there exists $X_0 \in K_\upsilon$ such that $X \equiv X_0 \pmod{\pi^{2i}}$ and $f(X_0) \in {K_\upsilon}^2$. For this the algorithm given in the Appendix A can be used.

Corollary 2.1.2 *Suppose $\upsilon(2) \neq 0$*

1. *If $U_1 = \emptyset$ then $\epsilon_\upsilon = 1$.*
2. *If $U_j \neq \emptyset$ and $V_j = \emptyset$ then $\epsilon_\upsilon = {|\pi|_\upsilon}^{-(2j-1)}$.*
3. *If $V_j \neq \emptyset$ and $U_{j+1} = \emptyset$ then $\epsilon_\upsilon = {|\pi|_\upsilon}^{-2j}$.*

Hence if $\upsilon(2) \neq 0$, then we compute (U_j) and (V_j) explicitly until one of them is empty. Then we compute ϵ_υ from the above corollary.

2.1.5 The Height Modulo Torsion

As will be seen in the examples, curves where the bound obtained by Theorem (2.1.1) is small tend to be those where the Tamagawa indices are trivial at the larger primes which divide the discriminant. This is often not the case where the torsion group is non-trivial. However the following Theorem will show us how to exploit the torsion group in order to reduce the bound obtained.

Theorem 2.1.2 *Under the notation and hypotheses of Theorem (2.1.1), let $\upsilon_1, \ldots, \upsilon_n$ be the (finitely many) valuations in M_K where the quantities $\mu_\upsilon \log(\epsilon_\upsilon)$*

are non-zero. Suppose (for some $m \leq n$) that $\upsilon_1, \ldots, \upsilon_m$ are non-archimedean valuations such that E is minimal at each of them, and there exists a subgroup $H \leq \mathrm{Tor}(E(K))$ such that H surjects onto $E(K_{\upsilon_i})/E^0(K_{\upsilon_i})$ (via the natural map) for $1 \leq i \leq m$. Then for each $P \in E(K)$, there exists $T \in H$ such that

$$\begin{aligned} h(P+T) - \hat{h}(P) &\leq \tfrac{1}{[K:\mathbb{Q}]} \left(\tfrac{|H|-1}{|H|}\right) \left(\textstyle\sum_{i=1}^{m} \mu_\upsilon n_\upsilon \log(\epsilon_\upsilon)\right) \\ &+ \tfrac{1}{[K:\mathbb{Q}]} \left(\textstyle\sum_{i=m+1}^{n} \mu_\upsilon n_\upsilon \log(\epsilon_\upsilon)\right). \end{aligned} \tag{2.16}$$

Proof. Let

$$H = \{T_1, \ldots, T_k\}.$$

Given any $P \in E(K)$, and $1 \leq i \leq m$ we must have at least one of $P + T_j$ has good reduction at υ_i. Hence, using Theorem (2.1.1), we get that

$$\textstyle\sum_{j=1}^{k} h(P+T_j) - \hat{h}(P) = \sum_{j=1}^{k} h(P+T_j) - \hat{h}(P+T_j)$$

$$\leq \tfrac{1}{[K:\mathbb{Q}]} \left(\textstyle\sum_{i=1}^{n} \mu_{\upsilon_i} n_{\upsilon_i} \sum_{j=1}^{k} \log(\epsilon(\upsilon_i, P+T_j))\right) \tag{2.17}$$

$$\leq \tfrac{k-1}{[K:\mathbb{Q}]} \left(\textstyle\sum_{i=1}^{m} \mu_{\upsilon_i} n_{\upsilon_i} \epsilon_{\upsilon_i}\right) + \tfrac{k}{[K:\mathbb{Q}]} \left(\textstyle\sum_{i=m+1}^{n} \mu_{\upsilon_i} n_{\upsilon_i} \epsilon_{\upsilon_i}\right)$$

Hence, for one of the T_j we must have that

$$k(h(P+T_j) - \hat{h}(P)) \leq \frac{k-1}{[K:\mathbb{Q}]} \left(\sum_{i=1}^{m} \mu_{\upsilon_i} n_{\upsilon_i} \epsilon_{\upsilon_i}\right) + \frac{k}{[K:\mathbb{Q}]} \left(\sum_{i=m+1}^{n} \mu_{\upsilon_i} n_{\upsilon_i} \epsilon_{\upsilon_i}\right)$$

which gives us the statement of the Theorem □

2.1.6 Examples

Example 2.1.1

$$E: \ Y^2 = X^3 - 73705X - 7526231 \tag{2.18}$$

We find that the equation is minimal and that its discriminant is

$$\Delta = 1155136043932048 = 2^4 \times 199 \times 362793983647$$

as a product of prime factors. Hence the Tamagawa indices will be 1*, except possibly at* 2*, and so from the definition on page 27, all the* $\mu_p = 0$ *except possibly for* $p = 2$*, or* $p = \infty$*. Using* `Pari/GP` *we find that the Tamagawa index at* 2 *is* 3*. Hence* $\mu_2 = \mu_\infty = \frac{1}{3}$*. To use Theorem (2.1.1) it remains to calculate* ϵ_2 *and* ϵ_∞*.*

We find that

$$f = 4x^3 - 294820x - 30104924 = 4(x^3 - 73705x - 7526231)$$

and

$$g = x^4 + 147410x^2 + 60209848x + 5432427025.$$

Now if $g \equiv 0 \pmod 2$ *then* x *is odd. But clearly, if* x *is odd then* $|f|_2 = 1/4$*, and* $|g|_2 \leq 1/4$*. Moreover,* $(-137, -1) \in E(\mathbb{Q}) \subseteq E(\mathbb{Q}_2)$ *and* $|f(-137)| = |g(-137)| = 1/4$*. Hence* $\epsilon_2 = 4$*.*

In computing ϵ_∞ *we find*

$$D_\infty = \emptyset$$

and

$$D'_\infty = [-0.007299, -0.005691] \cup [0, 0.003198].$$

Using Lemma (2.1.4) we find $\epsilon_\infty = 2.939442$*. Applying Lemma (2.1.1) we get*

$$h(P) - \hat{h}(P) \leq 0.8215047. \tag{2.19}$$

for all $P \in E(\mathbb{Q})$*.*

Here we note that Silverman's Theorem (1.1.3) gives a bound

$$h(P) - \hat{h}(P) \leq 13.0242$$

Example 2.1.2 *We begin with a curve of Mestre (quoted on page 234 of [Si2])*

$$E: \; Y^2 + Y = X^3 - 6349808647X + 193146346911036 \tag{2.20}$$

The discriminant of this curve is

$$\Delta = 60259 \times 550469 \times 11241887 \times 722983930261$$

as a product of primes. Since it is not divisible by any squares we must have that all constants $\mu_p = 0$ *for all finite primes* p*. By definition* $\mu_\infty = \frac{1}{3}$ *and it remains to determine* ϵ_∞*. Hence we write* D_∞*, and* D'_∞ *as unions of intervals as described on page 30 :*

$$D_\infty = [-1, 1]$$

and

$$D'_\infty = [-1\ ,\ \ -1.08780 \times 10^{-5}] \cup [0\ ,\ 2.02512 \times 10^{-5}] \cup [2.35024 \times 10^{-5}\ ,\ 1].$$

Hence we find that $d_\infty \approx 4 \times 10^{19}$ *and* $d'_\infty = 0.1289169$*. So* $\epsilon_\infty = 7.75693$ *and using Theorem (2.1.1) we get*

$$h(P) - \hat{h}(P) \leq \mu_\infty \log(\epsilon_\infty) = 0.68286 \tag{2.21}$$

for all points $P \in E(\mathbb{Q})$*. We note here that Silverman's theorem (1.1.3) gives an upper bound of* 21.7782 *instead* 0.68286*.*

It is apparent in the last two examples that the reason why the bound for $h(P) - \hat{h}(P)$ is so small is that all or almost all of the Tamagawa indices were 1. Here is an example where this is not the case:

Example 2.1.3 *We compute the bound for the following curve which is given by Thomas Kretschmer in [Kret] (page 633)*

$$Y^2 + XY = X^3 - 5818216808130X + 5401285759982786436 \tag{2.22}$$

The model given here is minimal and the discriminant is

$$\Delta = 2^6 \times 3^8 \times 7^2 \times 11^2 \times 29^2 \times 31^2 \times 41^2 \times 47^2 \times 277891391058913$$

We compute the following table

p	c_p	μ_p	ϵ_p
2	6	1/3	2^6
3	8	21/64	3^8
7	2	1/4	7^2
11	2	1/4	11^2
29	2	1/4	29^2
31	2	1/4	31^2
41	2	1/4	41^2
47	2	1/4	47^2
∞	-	1/3	518.48024

Hence we get

$$h(P) - \hat{h}(P) \leq 15.70819.$$

In comparison Silverman's bound is 27.5866*.*

Here we note that although our bound is much smaller than Silverman's it is still somewhat large for the purpose of the infinite descent (see the continuation of this example on page 51). However we note that the reduction of the point of order 2

$$Q = [1402932, -701466]$$

is singular at the primes 7, 11, 29, 31, 41, 47*. Hence using Theorem (2.1.2) we get that for all points* $P \in E(\mathbb{Q})$ *there is a* $T \in \{0,\ Q\}$ *such that*

$$h(P+T) - \hat{h}(P) \leq 11.03099$$

2.2 The Canonical Height and Results from the Geometry of Numbers

It is worth recalling at the outset of this section, that in the case when the elliptic curve E has rank 1 over the number field K, the infinite descent can be performed in a much easier way than that described in the introduction. This is well known: suppose $P \in E(K)$ has infinite order and and let us say that P

generates $E(K)/2E(K)$. Then, modulo torsion, $P = nQ$ where $n \geq 1$, and Q generates the free part of $E(K)$. Since P generates $E(K)/2E(K)$, n cannot be even and hence $n = 1$ or $n \geq 3$. If $n \geq 3$ then

$$\hat{h}(Q) \leq \frac{1}{9}\hat{h}(P)$$

and so, if P is not the generator of the free part of $E(K)$, we will find a generator in a much smaller region than that given by Zagier's Theorem (1.1.2).

In this section we develop a general technique for the infinite descent which is analogous to the reduction of the bound for the rank 1 case given above. The inspiration for much of this comes from Manin's Theorem (see [Ge, Zi]). There it is shown, using the Geometry of Numbers, how an upper bound for the regulator gives an upper bound for heights of generators of a sublattice of full rank. Below, we shall use the Geometry of numbers to show how given a basis for a sublattice of full rank, and a little extra information, we get an upper bound on the index.

We shall employ the language of lattices. Following [Ge, Zi] we define $\hat{E}(K) = E(K)/\mathrm{Tor}(E(K))$, where $\mathrm{Tor}(E(K))$ is the torsion of $E(K)$. Suppose that $P_1, \ldots, P_r$ generate a sublattice of $\hat{E}(K)$ of full rank (for example $P_1, \ldots, P_r$ could be a basis of $\hat{E}(K)/m\hat{E}(K)$ for some $m \geq 1$). Suppose that this sublattice had index n. If $n = 1$, then of course, $P_1, \ldots, P_r$ is a basis for $\hat{E}(K)$, and we can easily recover a basis for $E(K)$. We will define the height pairing matrix of $P_1, \ldots, P_r$ as follows:

$$H(P_1, \ldots, P_r) = (\langle P_i, P_j \rangle)_{i,j=1,\ldots,r} \tag{2.23}$$

where for all P, Q in $E(K)$

$$\langle P, Q \rangle =: \frac{1}{2}(\hat{h}(P+Q) - \hat{h}(P) - \hat{h}(Q)) \tag{2.24}$$

Let $R(P_1, \ldots, P_r)$ be the determinant of the height pairing matrix $H(P_1, \ldots, P_r)$. If R is the regulator of $E(K)$ it follows that

$$R = \frac{1}{n^2} R(P_1, \ldots, P_r) \tag{2.25}$$

We recall that the regulator is roughly of the same order of magnitude as the product of the canonical heights of some basis for $\hat{E}(K)$ (See, for example, the proof of Manin's theorem in [Ge, Zi]). Hence if the index n was very large we would expect (by virtue of (2.25)) there to be points of $\hat{E}(K)-\{0\}$ of very small canonical height. We make this idea precise. Roughly it tells us that if there are no points of $\hat{E}(K)-\{0\}$ of height smaller than some lower bound, then we can get an upper bound for the index n and hence reduce the infinite descent to checking the index of $P_1,\ldots,P_r$ in $\hat{E}(K)$. We make use of the following Lemma from the Geometry of Numbers.

Lemma 2.2.1 *(Hermite, Minkowski and others) Suppose*

$$f(\mathbf{x})=\sum_{i,j=1}^{r} f_{ij}x_ix_j \tag{2.26}$$

where (f_{ij}) is a symmetric positive definite matrix with determinant

$$D=\det(f_{ij})>0. \tag{2.27}$$

Then there exists a positive constant γ_r such that

$$\inf_{\mathbf{m}\neq 0 \ \textit{integral}} f(\mathbf{m}) \leq \gamma_r D^{\frac{1}{r}} \tag{2.28}$$

Moreover we can take

$$\begin{array}{llll} \gamma_1^1=1, & \gamma_2^2=\frac{4}{3}, & \gamma_3^3=2, & \gamma_4^4=4, \\ \gamma_5^5=8, & \gamma_6^6=\frac{64}{3}, & \gamma_7^7=64, & \gamma_8^8=2^8 \end{array} \tag{2.29}$$

and for $r\geq 9$

$$\gamma_r=\left(\frac{4}{\pi}\right)\Gamma\left(\frac{r}{2}+1\right)^{\frac{2}{r}} \tag{2.30}$$

Proof. The Lemma with constant $\gamma_r=\left(\frac{4}{3}\right)^{\frac{(r-1)}{2}}$ was originally due to Hermite. The formula (2.28) with γ_r given for all r by (2.30) is the formula for the 'first Minima' in Minkowski's Second Theorem (see [Ca2] p260, and [Sieg1] p26 for the formula (2.30)). The constants $\gamma_1,\ldots,\gamma_8$ given above are, for $1\leq r\leq 8$, the smallest constants which make the Lemma valid (See [Ca3] p332).

I am unaware if the smallest possible values of γ_r have been determined for any $r \geq 9$. □

Lemma 2.2.2 *Let E be an elliptic curve defined over a number field K. Let R be the regulator of $E(K)$. If the rank r is ≥ 1 then there exists a point Q in $E(K)$ of infinite order such that*

$$\hat{h}(Q) \leq \gamma_r R^{\frac{1}{r}} \tag{2.31}$$

Proof. Suppose $Q_1, \ldots, Q_r$ is a basis for $\hat{E}(K)$. If $Q = \sum_{i=1}^{r} m_i Q_i$ then

$$\hat{h}(Q) = \sum_{i,j=1}^{r} m_i m_j < Q_i, Q_j > . \tag{2.32}$$

Recall that the height pairing matrix $H(Q_1, \ldots, Q_r) = (< Q_i, Q_j >)$ is symmetric positive definite, and its determinant is R, the regulator of $E(K)$. It follows from Lemma (2.28) that there exist an $\mathbf{m} \neq \mathbf{0}$ integral such that

$$\hat{h}(Q) = \left(\sum_{i,j=1}^{r} m_i m_j < Q_i, Q_j > \right) \quad \leq \gamma_r R^{\frac{1}{r}}. \tag{2.33}$$

Since $Q_1, \ldots, Q_r$ is a basis for $\hat{E}(K)$ and $\mathbf{m} \neq \mathbf{0}$, Q must have infinite order, and the Lemma now follows. □

We now combine the above with the observation (2.25) to deduce the following theorem.

Theorem 2.2.1 *Let E be an Elliptic curve defined over a number field K. Suppose that $E(K)$ contains no point Q of infinite order with canonical height $\hat{h}(Q) \leq \lambda$ where λ is some positive real number. Suppose that $P_1, \ldots, P_r$ generate a sublattice of $\hat{E}(K)$ of full rank $r \geq 1$. Then the index n of the span of $P_1, \ldots, P_r$ in $\hat{E}(K)$ satisfies*

$$n \leq R(P_1, \ldots, P_r)^{\frac{1}{2}} \left(\frac{\gamma_r}{\lambda} \right)^{\frac{r}{2}} \tag{2.34}$$

where $R(P_1, \ldots, P_r)$ is the determinant of the height pairing matrix and

$$\begin{gathered} \gamma_1^1 = 1, \quad \gamma_2^2 = \tfrac{4}{3}, \quad \gamma_3^3 = 2, \quad \gamma_4^4 = 4, \\ \gamma_5^5 = 8, \quad \gamma_6^6 = \tfrac{64}{3}, \quad \gamma_7^7 = 64, \quad \gamma_8^8 = 2^8 \end{gathered} \tag{2.35}$$

and for $r \geq 9$

$$\gamma_r = \left(\frac{4}{\pi}\right) \Gamma\left(\frac{r}{2}+1\right)^{\frac{2}{r}} \tag{2.36}$$

Proof. By Lemma (2.2.2), if R is the regulator of $E(K)$ then there exists Q in $E(K)$ of infinite order such that

$$\hat{h}(Q) \leq \gamma_r R^{\frac{1}{r}}.$$

It follows that

$$\lambda \leq \gamma_r R^{\frac{1}{r}}.$$

But $R = \frac{1}{n^2} R(P_1, \ldots, P_r)$. Hence

$$\lambda^r \leq \frac{\gamma_r^r R(P_1, \ldots, P_r)}{n^2}.$$

Rearranging, we get the required inequality

$$n \leq R(P_1, \ldots, P_r)^{\frac{1}{2}} \left(\frac{\gamma_r}{\lambda}\right)^{\frac{r}{2}}.$$

□

2.3 A Sub-lattice Enlargement Procedure

Suppose we are given $P_1, \ldots, P_r$ which is a basis for a sublattice of $\hat{E}(K)$ of full rank. By the methods of the previous section, we can establish an upper bound for n, the index of this sublattice in $\hat{E}(K)$. If $n < 2$, then it is clear that $P_1, \ldots, P_r$ is a basis for $\hat{E}(K)$ and the infinite descent is finished.

Suppose now that the method of the previous section gave us a bound $n \leq \alpha$ for some $\alpha \geq 2$. Here it is necessary to check, for each prime $p \leq \alpha$ whether or not the index n is divisible by p. Equivalently, we must determine if there exist $a_1, \ldots, a_r \in \mathbb{Z}$, not all divisible by p. such that

$$\sum a_i P_i = pQ \tag{2.37}$$

for some $Q \in \hat{E}(K)$.

It is clear that in checking this we can assume that $|a_i| \leq p/2$. This leaves us with a finite number of equations of type (2.37) to solve. We explain how these may be solved later. However, as these equations can be many, it is useful to start with some sieving. In practice, we have found the sieving described below to be very effective.

2.3.1 Sieving

In the notation of above, given a prime $p \leq \alpha$, we let $P_{r+1}, \ldots, P_{r+s}$ be a basis for $\mathrm{Tor}(E(K))/p\mathrm{Tor}(E(K))$, where $\mathrm{Tor}(E(K))$ is the torsion subgroup of $E(K)$ (and so typically $s = 0$). We let

$$V_p = \left\{ \bar{\mathbf{a}} \in {\mathbb{F}_p}^{r+s} : \; if \; \mathbf{a} \in \mathbb{Z}^{r+s} \; and \; \mathbf{a} \equiv \bar{\mathbf{a}} \pmod{p} \; then \; \sum_{i=1}^{r+s} a_i P_i \in pE(K) \right\}$$

It is clear the V_p is an $\mathbb{F}_p$-linear subspace of ${\mathbb{F}_p}^{r+s}$ and that the index n is divisible by p if and only if $V_p \neq \{0\}$.

Suppose that $v \in M_K^0$ is a prime such that:

1. E has good reduction at v,

2. $|E(k_v)|$ is divisible by p but not by p^2.

Write $|E(k_v)| = lp$ where p does not divide l.

We let π be a uniformizer at v and compute $P_i' \equiv lP_i \pmod{\pi}$. If $P_i' \equiv 0 \pmod{\pi}$ for $i = 1, \ldots, r+s$, then the sieving modulo π, will give us nothing and we should start with another $v \in M_K$ satisfying the 2 conditions above. However, suppose, say that P_1' is not $0 \pmod{\pi}$. We note that the subgroup $lE(k_v)$ of $E(k_v)$ is cyclic of order p, and contains $P_1', \ldots, P_{r+s}'$; in particular $P_1' \pmod{\pi}$ generates $lE(k_v)$. By computing the multiples of $P_i' \pmod{\pi}$, we determine m_i such that $P_i' \equiv m_i P_1' \pmod{\pi}$. Hence, if $(\bar{a}_1, \ldots, \bar{a}_{r+s}) \in V_p$, we must have that

$$\sum m_i \bar{a}_i = 0 \tag{2.38}$$

in $\mathbb{F}_p$. This gives us a relation that must be satisfied by the vectors in V_p. If we were to compute $r+s$ independent relations by this method, then $V_p = \{0\}$, and the index would not be divisible by p.

At the very least, our hope is that by sieving modulo a few of these prime π, we have reduced V_p to being in a much smaller subspace of ${\mathbb{F}_p}^{r+s}$, and so we have considerably reduced the number of equations of type (2.37) to be checked.

Our method of sieving has an obvious gap, which is to find $\upsilon \in M_K$, for which $|E(k_\upsilon)|$ is divisible by p but not p^2. At least the second assumption is not always attainable (for example if $\mathrm{Tor}(E(K))$ had a subgroup of order p^2). So we note that the assumption that p^2 does not divide $|E(k_\upsilon)|$ can be easily circumvented after determining the structure of the p-Sylow subgroup of $E(k_\upsilon)$, as the reader may readily verify. However, the assumption that p divides $|E(k_\upsilon)|$ is essential to the idea of the sieving.

If primes $\upsilon \in M_K$ satisfying the conditions above exist, we hope to uncover some by computing sufficiently many $|E(k_\upsilon)|$. If $K = \mathbb{Q}$, then there exist efficient methods of computing $|E(\mathbb{F}_q)|$ for primes q, and judging from [Cohen] (pages 396-398), these have become very impressive.

2.3.2 Solving the Equation $P = pQ$

If the sieving described above has not been entirely successful in proving that $V_p = \{0\}$, then it will leave us with a subspace V_p' of ${\mathbb{F}_p}^{r+s}$, containing V_p (V_p' is simply the set of all solutions to the equations (2.38)). Here it is useful to take a projective subset of V_p', which we denote by S_p; we will let S_p be a subset of $\mathbb{Z}^{r+s} \backslash \{0\}$ with the following properties

1. if $(b_1, \ldots, b_{r+s}) \in S_p$, then $|b_i| \leq (p-1)/2$ unless $p = 2$ in which case $b_i = 0$ *or* 1,

2. for every $(\bar{a}_1, \ldots, \bar{a}_{r+s}) \in V_p \backslash \{0\}$, there exists exactly one $(b_1, \ldots, b_{r+s}) \in S_p$ such that $(\bar{a}_1, \ldots, \bar{a}_{r+s}) \equiv \beta(b_1, \ldots, b_{r+s}) \pmod{p}$ for some $\beta \in \mathbb{F}_p$.

It is clear that all that remains is to check, for all $(b_1, \ldots, b_{r+s}) \in S_p$, if

$$\sum_{i=1}^{r+s} b_i P_i = pQ \tag{2.39}$$

for some $Q \in E(K)$.

For each $(b_1, \ldots, b_{r+s}) \in S_p$, the equation (2.39) has exactly p^2 solutions in $E(\mathbb{C})$, and it is not at all difficult to find these p^2 possible $Q = (x, y) \in E(\mathbb{C})$ with $x,\ y \in \mathbb{C}$ computed as accurately as is desired using elliptic logarithms (see [Cohen]). This leaves us with the problem of deciding, given a sufficiently accurate computation of $x,\ y \in \mathbb{C}$, whether or not these are in our number field K. We make use of the following Lemma.

Lemma 2.3.1 *Suppose the elliptic curve E is given by Weierstrass equation (2.1) with $a_1, \ldots, a_6 \in \mathcal{O}_K$, and suppose that $P = nQ$, where $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ are on $E(K)\setminus\{0\}$. If $v \in M_K^0$ and $v(x_2) < 0$ then $v(x_1) \leq v(x_2)$.*

Moreover, if $c \in \mathcal{O}_K$ is such that $cx_1 \in \mathcal{O}_K$, then $cx_2 \in \mathcal{O}_K$.

Proof. Let E' be the minimal Weierstrass equation at v, and let $(x', y') \in E'(K_v)$ correspond to coordinates $(x, y) \in E(K_v)$. Then by [Si2] p172, there exists $u,\ r,\ t,\ s \in \mathcal{O}_v$ such that

$$x = u^2x' + r$$

$$y = u^3y' + u^2sx' + t.$$

If $v(x) < 0$ then $v(x') = v(x) - 2v(u)$, where $v(u) \geq 0$. Hence it is sufficient to assume that $v(x_2') < 0$ and show that $v(x_1') \leq v(x_2')$.

Let $v(x_2') = -2m$, where $m \in \mathbb{Z}$ (as is well known, $v(x_2') < 0$ implies that $3v(x_2') = 2v(y_2')$ and hence that $v(x_2')$ is even). Then the subset

$$E_m'(K_v) = \{(x', y') \in E'(K_v) : v(x') \leq -2m\} \cup \{0\}$$

is a subgroup of $E'(K_v)$ (see for example [Si2], p187). Hence $P' \in E_m'(K_v)$ and $v(x_1') \leq -2m = v(x_2')$.

This concludes the proof of the first part of the Lemma. The second part is now obvious. □

Hence given $(b_1, \ldots, b_{r+s}) \in S_p$, we calculate $P = (x_1, y_1) = \sum b_i P_i$, and find a $c \in \mathcal{O}_K$ such that $cx_1 \in \mathcal{O}_K$. If $P = pQ$, with $Q = (x_2, y_2) \in \mathcal{O}_K$, then $cx_2 \in \mathcal{O}_K$ by the above Lemma. So if we compute the p^2 values x_2

accurately enough [1] we can determine if any of the cx_2 is expressible as a $\mathbb{Z}$-linear combination of any $\mathbb{Z}$-basis for $\mathcal{O}_K$, using an LLL-based algorithm such as the one given on page 100 of [Cohen]. (Of course, if $K = \mathbb{Q}$, then we can be much more down to earth. We simply calculate the x_2s accurately enough to see if any of cx_2 is an integer to many decimal places.) If any cx_2 seems to equal an element $a \in \mathcal{O}_K$, then we can substitute a/c for x in the equation for E and ask if there is a solution $y \in K$.

If we have found that none of the equations (2.39) is soluble with $Q \in E(K)$, then we have proven that the index is not divisible by p, and we can proceed to the next prime until we reach α, our upper bound for the index. However, if we find that $\sum b_i P_i = pQ$ with $Q \in E(K)$, then there is a $1 \leq j \leq r$, such that p does not divide b_j. Here we replace P_j by Q. The index of the sublattice generated by the new $P_1, \dots, P_r$ in $\hat{E}(K)$ is $\leq \alpha/p$. In any case, we continue until we get to show that the index is 1.

2.4 Examples

Example 2.4.1 *Here we return to our Example on page 35*

$$E: \; Y^2 = X^3 - 73705X - 7526231.$$

We recall that we established

$$h(P) - \hat{h}(P) \leq 0.8215 \tag{2.40}$$

for all $P \in E(\mathbb{Q})$. It is easy to show that this curve has no torsion. Using Cremona's program `mwrank`*, we found that the 2-part of the Tate-Shafarevich group is trivial, that the rank is 4, and that a basis for $E(\mathbb{Q})/2E(\mathbb{Q})$ is*

$$P_1 = (-137, -1), \; P_2 = (-157, -419), \; P_3 = (-175, -113), \; P_4 = (413, -5699);$$

[1] Here, if K has a real embedding, then it is useful to replace K with a real conjugate field at the beginning of the computation, and so reject all the values of x_2 which are not real (taking into account that floating-point arithmetic a real number is one with a very small imaginary part!).

this the program did in approximately 1.5 *minutes.*

The determinant of the height pairing matrix of $P_1, \ldots, P_4$ *is* 248.987*. We search for points of logarithmic height* ≤ 5 *using Cremona's program* `findinf`*. The search takes a few seconds and turns up only one point:* $P_1 = (-137, -1)$*. This has canonical height* 4.41996*. We note that had there been any point of canonical height* ≤ 4.1*, then its logarithmic height would have been* $\leq 4.1 + 0.8215 < 5$ *and would have been uncovered by the search. Hence there are no points of canonical height* ≤ 4.1*. Using Theorem (2.2.1) we find that the index of the span of* $P_1, \ldots, P_4$ *is* ≤ 1.88 *. Hence we have found the Mordell-Weil group.*

Next we compare our method to that outlined in the introduction. We recall that if $(X, Y) \in E(\mathbb{Q})$*, then we can write* $X = x/z^2$ *where* $x,\ z \in \mathbb{Z}$*. Hence to search up to logarithmic height* 5*, our search region on* $x,\ z$ *is*

$$-148 \leq x \leq 148, \quad 1 \leq z \leq 12.$$

We note that had we used Zagier's (1.1.2) on page 10, we would be required to enumerate all the points on $E(\mathbb{Q})$ *of canonical height* ≤ 13.5831*. If we combine this with our estimate (2.40) above, we must list all points with logarithmic height* 14.4046*. The corresponding search region is*

$$-1802346 \leq x \leq 1802346, \quad 1 \leq z \leq 1321.$$

To search this region is possible using a well written program such as `findinf` *mentioned above, but this would take a few hours on a work station.*

Moreover we note that if we had to use Silverman's bound on the difference $h(P) - \hat{h}(P)$ *as well as Zagier's Lemma we would have to search for all points on* $E(\mathbb{Q})$ *with logarithmic height* ≤ 26.6073*. Then the search region would be*

$$-359255618029 \leq x \leq 359255618029, \quad 1 \leq z \leq 599379.$$

Finally, at the suggestion of Dr Cremona, we compute the following table to give another illustration of how effective our bound (2.40) is.

P	$h(P)$	$\hat{h}(P)$	$h(P)-\hat{h}(P)$
P_1	4.9199809	4.4199587	0.50002214
P_2	5.0562458	4.4416097	0.61463607
P_3	5.1647859	4.4605122	0.70427372
P_4	6.0234476	5.8817481	0.14169942

Example 2.4.2 *We return here to Mestre's curve:*

$$E:\ Y^2+Y=X^3-6349808647X+193146346911036 \tag{2.41}$$

We recall that on page 36 we proved that

$$h(P)-\hat{h}(P)\leq 0.682862 \tag{2.42}$$

for all points $P\in E(\mathbb{Q})$. Mestre (see [Mestre]) has shown that this curve has rank at least 12 *and has given* 12 *independent points (Mestre in fact gave a non-minimal model of the curve, and the equation (2.41) which we will work with is the minimal model). Moreover he has shown that the standard conjectures* [2] *imply that the rank is* 12*. Here we will not take on the task of determining the rank unconditionally* [3]*; we will simply assume that the rank is* 12*, and obtain a basis from the points given by Mestre. Here is a list of the points that Mestre gave (after applying the change of variable which takes the points onto our minimal model (2.41)):*

$$\begin{array}{lll}
P_1=[49421,\ 200114], & P_2=[49493,\ 333458], & P_3=[49513,\ 362258],\\
P_4=[49632,\ 502899], & P_5=[49667,\ 538049], & P_6=[49797,\ 654674],\\
P_7=[49899,\ 735713], & P_8=[50012,\ 818375], & P_9=[50165,\ 921837],\\
P_{10}=[50215,\ 954017], & P_{11}=[50823,\ 1305633], & P_{12}=[51108,\ 1454591].
\end{array}$$

[2] The Birch and Swinnerton-Dyer conjecture, the Taniyama-Weil conjecture, and a suitable Riemann hypothesis.

[3] Here `mwrank` would take too long. In the absence of 2-torsion, `mwrank` uses the algorithm for 2-descent described in [Bi, SwD] and in [Cre] pages 68-76. In this algorithm the size of the search region for the homogeneous spaces is roughly proportional to the square root of the discriminant of the elliptic curve. In cases where the discriminant is very large, such as that for Mestre's curve above, the algorithm is no longer practical. Unfortunately there does not seem to be any unconditional algorithm suited for determining Mordell-Weil groups of curves of large discriminant and no torsion.

Here we proceeded with the sieving first. We used `Pari/GP`*, which calculates* $|E(\mathbb{F}_q)|$ *for prime* q *using the Shanks-Mestre algorithm (see [Cohen] page 397). We found that it took roughly 1 second to compute* $|E(\mathbb{F}_q)|$ *for the first 200 primes* q *(i.e. for all the primes* ≤ 1223*). We wrote a program which does the following: for each prime* $2 \leq p \leq 11$ *it lists all the primes* $q \leq 1223$ *for which* $|E(\mathbb{F}_q)|$ *is divisible by* p *but not* p^2 *as recommended by our sieving algorithm on page 43. Next, for each prime* q *satisfying these conditions, it computes a relation modulo* p*, which must be satisfied by the vectors in* V_p *as defined on page 43 using the idea described there; if it finds* 12 *independent relations then the rank of* V_p *is* 0 *and the index is not divisible by* p*. For each of the primes* p*, the program continues computing relations until the rank of the relations is 12 or until there are no more prime* $q \leq 1223$ *satisfying the conditions described. The program took roughly 25 seconds to run and output that for all the primes* $p \leq 11$ *the rank of relations found is* 12 *except for* $p = 2$ *where the rank was* 10*. We note that there are* 47 *primes* q *in the above range satisfying the criterion that* 2 *divides* $|E(\mathbb{F}_q)|$ *but* 4 *does not. Hence it seems very probable that the index is divisible by 2. Calculating the kernel of the relations obtained we get that*

$$V_2' = span\,\{(1,0,0,1,1,0,1,0,1,1,0,0),(1,0,0,1,1,0,0,1,0,0,0,1)\} \pmod 2.$$

Hence we want to test if any of the 3 *points* $P_1+P_4+P_5+P_7+P_9+P_{10}$*,* $P_1+P_4+P_5+P_8+P_{12}$*,* $P_7+P_8+P_9+P_{10}+P_{12}$ *is 2-divisible in* $E(\mathbb{Q})$*. Using* `Pari/GP` *we calculate the periods of* E *and the 2-division points of the first* 2 *points. We get for each one a division point which is integral to* 50 *decimal places. We checked that these give us integral points on the curve. We replace our old* P_7*, and* P_8 *with these two new points:*

$$\begin{aligned} P_7 &= [38756, -2294721] \\ P_8 &= [208314, 88938858], \end{aligned}$$

thus gaining index 4*.*

We repeat the sieving for $p = 2$*. This time the rank of relations obtained for* $p = 2$ *is* 11*. We find that if the index is still divisible by 2 then* $P_3 + P_5 + P_6 +$

$P_8 + P_{10} + P_{11} + P_{12}$ must be 2-divisible in $E(\mathbb{Q})$. Here none of the 2-division points were integral and we used Lemma (2.3.1) to recover a rational 2-division point. This becomes our new P_3:

$$P_3 = \left[\frac{2739835340}{5041}, \frac{141949849330392}{357911}\right].$$

Repeating the sieving described for $p = 2$ we find get that the rank of relations obtained is 12, and hence the index of the span of our new $P_1, \ldots, P_{12}$ is not divisible by 2. Moreover, this index is not divisible by any prime $3 \leq p \leq 11$ since the index of the span of the original points was not.

We return to the sieving again. We calculate $|E(\mathbb{F}_q)|$ for the first 2500 primes q (i.e. all the primes $q \leq 22307$), and we extend our range for the prime p to all the primes ≤ 200. It took `Pari/GP` *roughly 25 seconds to compute all the $|E(\mathbb{F}_q)|$ for all the primes $q \leq 22307$. Our program this time took about 10 minutes to stop. In each case the rank of relations computed was 12 except for $p = 167,\ 179,\ 191$ where the ranks were respectively 8, 10, 10. Hence if the index of the span of our new $P_1, \ldots, P_{12}$ is not 1, then it must be ≥ 167.*

The determinant of the height matrix of $P_1, \ldots, P_{12}$ is

$$R(P_1, \ldots, P_{12}) = 586593208.77747$$

and computing γ_{12} we get 3.81181 according to formula (2.36) . Hence Theorem (2.2.1) gives us that if there are no rational points on E with canonical height $\leq \lambda$ then the index of the span of $P_1, \ldots, P_{12}$ in $E(\mathbb{Q})$ satisfies:

$$n \leq \frac{74295365.4988}{\lambda^6}.$$

Using this inequality we find that if there were no points of canonical height ≤ 8.73 then the index would be ≤ 166.9 and we would be finished. Using the inequality (2.42) we see that we need to find all points of logarithmic height ≤ 9.41. We used Cremona's program `findinf` *and found none in that range of canonical height ≤ 8.73 (the program took roughly 5 minutes to list all the points of logarithmic height ≤ 9.41). Hence the points listed below form a basis assuming that the rank (as predicted by the Birch and Swinnerton-Dyer conjecture) is*

12*:*

$$\begin{array}{lll}
P_1 = [49421,\ 200114], & P_2 = [49493,\ 333458], & P_3 = \left[\frac{2739835340}{5041}, \frac{1419498493330392}{357911}\right], \\
P_4 = [49632,\ 502899], & P_5 = [49667,\ 538049], & P_6 = [49797,\ 654674], \\
P_7 = [38756,\ -2294721], & P_8 = [208314,\ 88938858], & P_9 = [50165,\ 921837], \\
P_{10} = [50215,\ 954017], & P_{11} = [50823,\ 1305633], & P_{12} = [51108,\ 1454591].
\end{array}$$

Example 2.4.3 *Here we return to the curve*

$$Y^2 + XY = X^3 - 5818216808130X + 5401285759982786436 \tag{2.43}$$

In [Kret] Kretchmer gave this as a curve of (exact) rank 8 *with torsion of order* 2*, but did not give the points he found on the curve. We used Cremona's program* `mwrank` *and it gave a basis for* $E(\mathbb{Q})/2E(\mathbb{Q})$*:*

$$\begin{array}{ll}
P_1 = [1410240,\ -29977314], & P_2 = [1704648,\ -661672482], \\
P_3 = [1421184,\ -55353570], & P_4 = [259761720/125,\ -189069355038/125], \\
P_5 = [4740024,\ 9180268266], & P_6 = [975216,\ 808674546], \\
P_7 = [7028688,\ -17659711842], & P_8 = [3418038804/289,\ 195936026213238/4913], \\
Q = [1402932,\ -701466], &
\end{array}$$

where $P_1, \ldots, P_8$ *are of infinite order and* Q *is a point of order* 2*. Here it is easy to show that there are no other torsion points. It remains to complete the infinite descent.*

Of course the index of the span of the points above is not divisible by 2 *since the points are independent modulo* $2E(\mathbb{Q})$*. Sieving (as in the above example) with roughly* 200 *primes (here we excluded all the primes of bad reduction), we were able to show that the index of the span of the given points is not divisible by* 5, 7, 11, 13 *and detected a possibly 3-divisible linear combination of the points. We found*

$$P_4 - P_5 - P_6 - P_7 + P_8 = 3\ [-2623596, -1613325930]$$

and hence replacing P_8 *by*

$$P_8 = [-2623596, -1613325930]$$

we reduce the index by 3*. Repeating the sieving we found that the new index is not divisible by* 3*. Now we continued the sieving using* 15000 *primes* q *and our program proved that the index is not divisible by any prime* p *less than* 500 *(this took roughly 30 minutes).*

The determinant of the height pairing matrix of the new $P_1, \ldots, P_8$ *is* 184808.298*. Using Theorem (2.2.1) it is now sufficient to show that there are no points of canonical height* ≤ 1.96 *whence it would follow that the index is* 1*. Here we recall that we proved that*

$$h(P) - \hat{h}(P) \leq 15.70819.$$

and so that to check that there are no points of canonical height ≤ 1.96 *using this it would be necessary to uncover all the points of logarithmic height* ≤ 17.67*. We expect that this computation would take roughly* 10 *days. However we also proved that for any point* P *there is a point* T *which is either* 0 *or* Q *such that*

$$h(P+T) - \hat{h}(P) \leq 11.03099 \tag{2.44}$$

Now it is sufficient to enumerate all the points of logarithmic height ≤ 13 *and check that none have canonical height* ≤ 1.96*. We did this in roughly* 45 *minutes using* `findinf`*. Hence it follows that*

$$\begin{array}{ll}
P_1 = [1410240, \ -29977314], & P_2 = [1704648, \ -661672482], \\
P_3 = [1421184, \ -55353570], & P_4 = [259761720/125, \ -189069355038/125], \\
P_5 = [4740024, \ 9180268266], & P_6 = [975216, \ 808674546], \\
P_7 = [7028688, \ -17659711842], & P_8 = [-2623596, \ -1613325930] \\
Q = [1402932, \ -701466], &
\end{array}$$

is a basis for $E(\mathbb{Q})$*.*

Finally we would like to point out that we were able to obtain the bound (2.44) using the fact that the torsion group surjects onto $E(\mathbb{Q}_p)/E^0(\mathbb{Q}_p)$ *for most of the primes where the Tamagawa index is not* 1*. Since this will not be be the case for most curves we would like to illustrate a third method which can be used to complete the infinite descent when the bound for* $h(P) - \hat{h}(P)$ *is too large. We note that for all the non-archimedean primes except* 2 *and* 3*, the Tamagawa* *index is either* 1 *or* 2 *(see the table on page (2.1.3)). In any case, if* $P \in E(\mathbb{Q})$ *was of infinite order, and had canonical height* ≤ 1.96*, then* $2P$ *will have canonical height* ≤ 7.84 *and will have good reduction at all the non-archimedean primes except possibly at* 2 *or* 3*. Hence, in the notation of Theorem (2.1.1) we have*

$$\epsilon(p, 2P) = 1$$

for all primes $p \neq 2,\ 3,\ \infty$ *and*

$$\epsilon(p, 2P) \leq \epsilon_p$$

for $p = 2,\ 3,\ \infty$*. Using the values of* ϵ_p *given in the table on page (2.1.3) for the primes* $p = 2,\ 3,\ \infty$ *and Theorem (2.1.1) we get*

$$h(2P) - \hat{h}(2P) \leq 6.39956.$$

Hence to uncover $2P$ *we need to find all points of logarithmic height* ≤ 14.24 *and this would not take much longer than the search we have already done. Finally we would have to test each point found with canonical height* ≤ 7.84 *to see if it is twice a point.*

Chapter 3

Computing the 2-Selmer Group of an Elliptic Curve

When trying to compute the Mordell-Weil group of an elliptic curve one normally first computes the 2-Selmer group. This is a group which contains a subgroup isomorphic to $E(\mathbb{Q})/2E(\mathbb{Q})$. Whilst computing the 2-Selmer group is certainly an effective procedure there is no known effective procedure for computing the subgroup isomorphic to $E(\mathbb{Q})/2E(\mathbb{Q})$. However all is not lost as at least the 2-Selmer group gives one an upper bound on the rank of the elliptic curve. We set

$$L_D(\alpha, \beta) = (e^{(\log D)^{\alpha}(\log\log D)^{1-\alpha}})^{\beta+o(1)}.$$

This is a function which interpolates between polynomial time, $\alpha = 0$, and exponential time, $\alpha = 1$. In this chapter we show the complexity of computing the 2-Selmer group is $O(L_D(0.5, c_1))$ where D denotes the absolute discriminant of the elliptic curve.

Let E be our elliptic curve given by

$$E \;:\; Y^2 = X^3 + AX + B.$$

We shall assume that the elliptic curve has no points of order 2 defined over $\mathbb{Q}$. This is certainly the most difficult case for finding the 2-Selmer group. The modern method of computing the 2-Selmer group in this case goes back to the paper of Birch and Swinnerton-Dyer, [Bi, SwD]. In their method a search is carried out for the quartics which represent the homogeneous spaces given their invariants. As we noted in the Introduction (page 12), this method is certainly fast for small values of D, however it is not hard to see that its complexity is at least $O(\sqrt{|D|})$, [Bi, SwD][Page 11]. In the present chapter we shall show how the "old-fashioned" technique which is the basis for Weil's proof of the Mordell-Weil Theorem combined with a method derived from a paper of Brumer and Kramer, [Brum, Kra], will determine the 2-Selmer group in our stated time. Our complexity is therefore much better than the Birch and Swinnerton-Dyer algorithm.

As we have pointed out already, this chapter is based on joint work with Dr N. Smart.

We let S denote the set of primes dividing $2D$, we note that this has cardinality $O(\log D)$. Let K denote the number field generated by θ where $\theta^3 + A\theta + B = 0$. We will let R denote the set of primes of K lying above those in S as well as the infinite primes. As usual we let $K(R,2)$ denote the group of all elements of K^*/K^{*2} such that by adjoining a square root of an element of $K(R,2)$ to K one obtains an extension of K unramified outside R. Equivalently we have

$$K(R,2) = \{\alpha \in K^*/K^{*2} : \mathrm{ord}_{\wp}(\alpha) \equiv 0 \pmod 2 \quad \text{if} \quad \wp \notin R.\}. \tag{3.1}$$

One can show (see for example Exercise 10.9 on page 320 of [Si2]), that $K(R,2)$ contains the 2-Selmer group. We first find $K(R,2)$ and then reduce it to the 2-Selmer group.

3.1 The Method of Brumer and Kramer

For each prime $p \in S \cup \{\infty\}$ we define

$$K_p = \mathbb{Q}_p[T]/(f(T)) = \mathbb{Q}_p(t)$$

where $(f(T))$ is the ideal in $\mathbb{Q}_p[T]$ generated by $f(T) = T^3 + AT + B$, and $t = T + (f(T))$. Just as in the classical case of the 2-descent over $\mathbb{Q}$ we have an embedding

$$E(\mathbb{Q}_p)/2E(\mathbb{Q}_p) \to K_p^*/K_p^{*2} \tag{3.2}$$

with the usual definition of K_p^*/K_p^{*2}. Here, for each prime p we have the following diagram

$$\begin{tikzcd}
0 \arrow[r] & E(\mathbb{Q})/2E(\mathbb{Q}) \arrow[r, "X-t"] \arrow[d] & K(R,2) \arrow[d] \\
0 \arrow[r] & E(\mathbb{Q}_p)/2E(\mathbb{Q}_p) \arrow[r, "X-t"] & K_p^*/K_p^{*2}.
\end{tikzcd} \tag{3.3}$$

We denote the natural map from $K(R,2)$ to K_p^*/K_p^{*2} by σ.

For each prime $p \in S \cup \{\infty\}$ we let U_p be the image of $E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)$ in K_p^*/K_p^{*2} under the mapping (3.2). In [Brum, Kra] Brumer and Kramer showed that the Selmer group is the maximal subgroup of $K(R,2)$ whose image under the natural map σ is contained in U_p for all primes $p \in S \cup \{\infty\}$. Ostensibly, to use this method for the computation of the 2-Selmer group, one must first calculate $E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)$ for each prime $p \in S \cup \{\infty\}$. However, we have found this mildly troublesome, and indeed what is really needed is to compute the images U_p. We note that the size of K_p^*/K_p^{*2} is bounded for all primes p and all (cubic) polynomials f.

To determine U_p it is sufficient to take each element of K_p^*/K_p^{*2} which has norm in $\mathbb{Q}_p^{*2}$ and determine whether or not it is in U_p. As in the classical case (see page 65) this leads to a homogeneous space as the intersection of 2 quadric surfaces, and here all that is required is to check their solubility over the local field $\mathbb{Q}_p$. This can be done by the polynomial time algorithm given in Chapters 5 and 6 [1]; these algorithms are non-constructive (they do not give points on the homogeneous space but simply determine whether or not they have a point defined over $\mathbb{Q}_p$, which is all that is needed here).

[1] It is to be noted here that any homogeneous spaces (over $\mathbb{Q}_p$) arising here will be the intersection of 2 quadric surfaces, one of which is singular. Hence if $p \neq 2, \infty$ then we can use our polynomial time algorithm in Section 5.3

3.2 Finding $K(R,2)$

It will be seen later that determining $K(R,2)$ is useful for our higher descents which we describe in the next chapter. But there K will not necessarily be a cubic field. We should also point out that the above method of Brumer and Kramer has been applied to computing the Mordell-Weil group of Jacobians of hyperelliptic curves of higher genus by Schaefer [Scha], where again the field K is not necessarily cubic. Hence for the purpose of this section we will assume that K is a general number field with absolute discriminant D, and R is a finite set of valuations on K which includes all those at infinity. It should be noted however, that some of the complexity analysis is valid only for the case where $|R| = O(\log(D))$. We shall assume that we are given an integral basis for the maximal order of K and generators for the unit and class groups. To determine this information will take time $O(L_D(0.5, c_2))$ as computing a basis for the maximal order can be done in time $O(L_D(1/3, c_3))$, [Bu, Len], and computing the unit and class groups can be done in time $O(L_D(0.5, c_2))$, [Buch] assuming GRH and a certain conjecture about the number of reduced smooth ideals of a number field. The class group Cl_K is then presented as a set of ideals $\mathfrak{c}_1, \ldots, \mathfrak{c}_g$ and integers s_i with $s_{i-1}|s_i$, such that, if for an ideal $\mathfrak{a}$ we denote by $\overline{\mathfrak{a}}$ the image of $\mathfrak{a}$ in the class group, we have

$$Cl_K \cong \langle\overline{\mathfrak{c}_1}\rangle \times \ldots \times \langle\overline{\mathfrak{c}_g}\rangle ,$$

with $\langle\overline{\mathfrak{c}_i}\rangle \cong \mathbb{Z}/s_i\mathbb{Z}$. We denote by $\eta_1, \ldots, \eta_r$ a set of r fundamental units for K. Given an ideal of K then using the basis of the relation lattice which was used in computing the class group one can determine whether the ideal is principal and if so compute a generator in time $O(L_D(0.5, c_4))$ (see [Buch]). We note that in general one cannot write down the elements we require in polynomial time when we express them in standard representation so throughout we assume all elements are in a compact representation, see [Thiel]. We now give the algorithm to compute $K(R,2)$ as a product of cyclic groups of order 2. Let the finite prime ideals in R be denoted $\wp_1, \ldots, \wp_t$.

Suppose $\alpha \in K(R,2)$. Then by the definition (3.1) above $(\alpha) = \mathfrak{a}\mathfrak{b}^2$ where

$\mathfrak{a}|(2D)$. Let $\mathcal{F}$ be the group of fractional ideals. We have a homomorphism

$$\phi : K(R,2) \to \mathcal{F}/\mathcal{F}^2$$

given by $\alpha \to (\alpha)\mathcal{F}^2$. Clearly the image of ϕ is contained in the group

$$H_1 = \left\langle \wp_1 \mathcal{F}^2 \right\rangle \times \ldots \times \left\langle \wp_n \mathcal{F}^2 \right\rangle.$$

Let

$$H_2 = \left\{ \mathfrak{d}\mathcal{F}^2 \in H_1 : \mathfrak{d}\mathcal{F}^2 = (\gamma)\mathcal{F}^2 \text{ for some } \ \gamma \in K^* \right\}.$$

Clearly $Im(\phi) = H_2$. We want to show how to calculate H_2 and then how to refine it to obtain $K(R,2)$ as a product of cyclic groups of order 2. We assume that for each $\wp_i$ that we can write

$$\overline{\wp_j} = \prod_{i=1}^{g} \overline{\mathfrak{c}_i}^{b_{ij}};$$

this can be done by the method in [Buch] in time $O(L_D(0.5, c_4))$. Suppose $\mathfrak{d}\mathcal{F}^2 \in H_2$, then we can take $\mathfrak{d} = \prod_{j=1}^{n} {\wp_j}^{a_j}$. Hence

$$\overline{\mathfrak{d}} = \prod_{i=1}^{g} \overline{\mathfrak{c}_i}^{e_i},$$

where $e_i = \sum_{j=1}^{n} a_j b_{ij}$. Suppose that $s_1, \ldots, s_k$ are odd, and $s_{k+1}, \ldots, s_g$ are even. Then $\mathfrak{d}\mathcal{F}^2$ lies in H_2 if and only if $\sum_{j=1}^{n} a_j b_{ij} \equiv 0 \pmod 2$ for $i = k+1, \ldots, g$.

By a computing an $\mathbb{F}_2$-basis for the subspace of the vectors $(a_1, \ldots, a_n)$ in $\mathbb{F}_2^n$ which satisfy the congruences above, we get a basis for H_2. Further we may replace the representative of each element of this basis by one which is a principal ideal as follows: Suppose $\mathfrak{d}$ is such a representative which we want to replace by a principal ideal. By construction of this basis we know $\mathfrak{d}$ as a product of the $\wp_i$ and hence we can write $\overline{\mathfrak{d}} = \prod \overline{\mathfrak{c}_i}^{u_i}$ where $u_{k+1}, \ldots, u_g$ are even. Now since $s_1, \ldots, s_k$ are odd we can find $t_1, \ldots, t_k$ such that $u_i + 2t_i \equiv 0 \pmod{s_i}$ for $i = 1, \ldots, k$. We take $t_j = -u_j/2$ for $j = k+1, \ldots, g$. Hence we have that

$$\mathfrak{d} \prod_{i=1}^{g} {\mathfrak{c}_i}^{2t_i} = (\alpha)$$

for some $\alpha \in K^*$. This α can be computed in time $O(L_D(0.5, c_4))$ as we stated above. Hence we can write

$$H_2 = \left\langle (\alpha_1)\mathcal{F}^2 \right\rangle \times \ldots \times \left\langle (\alpha_n)\mathcal{F}^2 \right\rangle$$

for some $\alpha_1, \ldots, \alpha_n \in K^*$.

Lemma 3.2.1 *Let $\mathfrak{b}_1, \ldots, \mathfrak{b}_l$ be an $\mathbb{F}_2$-basis for $Cl[2]$. Write ${\mathfrak{b}_i}^2 = (\beta_i)$. Then*

$$\alpha_1 {K^*}^2, \ldots, \alpha_n {K^*}^2,\ \beta_1 {K^*}^2, \ldots, \beta_l {K^*}^2,\ \eta_1 {K^*}^2, \ldots, \eta_r {K^*}^2,\ \eta_{r+1} {K^*}^2 \qquad (3.4)$$

is a basis for $K(R, 2)$, where $\eta_1, \ldots, \eta_r$ is a system of fundamental units for K, and we take η_{r+1} a generator for the roots of unity.

Proof. It is clear that the elements of the list above generate $K(R, 2)$. What remains is to show that these are independent. Suppose that

$$\prod_{i=1}^{n} {\alpha_i}^{a_i} \prod_{i=1}^{l} {\beta_i}^{b_i} \prod_{i=1}^{r+1} {\eta_i}^{c_i} \in {K^*}^2$$

where the a's, b's, c's, are in $\{0, 1\}$. Then $\prod ((\alpha_i)\mathcal{F}^2)^{a_i} = (1)\mathcal{F}^2$ which implies that $a_i = 0$ for $i = 1, \ldots, n$. Hence we can now assume that

$$\prod_{i=1}^{l} {\beta_i}^{b_i} \prod_{i=1}^{r+1} {\eta_i}^{c_i} \in {K^*}^2.$$

Hence $\prod {\mathfrak{b}_i}^{2b_i} = (\epsilon)^2$ where $\epsilon \in K^*$, i.e. $\prod {\mathfrak{b}_i}^{b_i} = (\epsilon)$, so $b_i = 0$. The result now follows. □

Lemma 3.2.2 *The complexity of finding $K(R, 2)$ as a product of cyclic groups of order 2 is given by $O(L_D(0.5, c_1))$.*

Proof. We note that the number of ideals $\wp_i$ dividing $(2D)$ is $O([K : \mathbb{Q}] \log D)$. The number of elements in a basis of $Cl[2]$ is $O(\log(h_K)) = O(\log(D))$. Hence the number of ideals that we need to check to be principal is a polynomial function in $\log D$. As we stated earlier for each ideal this can be done in time $O(L_D(0.5, c_4))$ by an algorithm which will also produce a generator of any principal ideal found. The desired complexity then follows. □

3.3 Computing The 2-Selmer Group

We return now to the special case where K is a cubic field arising from our elliptic curve E. For each element of $K(R,2)$ we eliminate those elements which do not lie in the image of σ in diagram (3.3) for all $p \in S \cup \{\infty\}$. Suppose we know that the Selmer group is a subgroup of some group

$$\langle k_1 \rangle \times \ldots \times \langle k_v \rangle \leq K(R,2)$$

where the $\langle k_i \rangle$ are cyclic groups of order 2 (it is understood that the k_i are in fact $k_i K^{*2}$). Consider any prime $p \in S \cup \{\infty\}$; recall that we denoted the image of the map

$$E(\mathbb{Q}_p)/2E(\mathbb{Q}_p) \to K_p^*/K_p^{*2} \tag{3.5}$$

by U_p. To determine the Selmer group we want to determine the maximal subgroup of $\langle k_1 \rangle \times \ldots \times \langle k_v \rangle$ whose image under σ is in U_p for all primes p; obviously we need only consider those primes which divide $2D$ and the infinite prime. This idea we find explained in [Brum, Kra] or [Scha] as we have already stated.

Lemma 3.3.1 *The image of an element of $K(R,2)$ under σ can be checked to lie in U_p in polynomial time.*

Proof. Suppose $X^3 + AX + B$ has three roots in $\mathbb{Q}_p$ and $p > 2$; then

$$U_p \leq {\mathbb{Q}_p}^*/{\mathbb{Q}_p}^{*2} \times {\mathbb{Q}_p}^*/{\mathbb{Q}_p}^{*2} \times {\mathbb{Q}_p}^*/{\mathbb{Q}_p}^{*2}.$$

There are at most four elements of ${\mathbb{Q}_p}^*/{\mathbb{Q}_p}^{*2}$ and $|U_p|$ has order $O(1)$.

We therefore have $O(1)$ tests to perform as to whether an element of $\mathbb{Q}_p$ is a p-adic square. This can certainly be done in polynomial time. The other cases are similar. □

For $i = 1, \ldots, v$, we define the subgroup S_i of $\langle k_1 \rangle \times \ldots \times \langle k_i \rangle$ to be the maximal subgroup of $\langle k_1 \rangle \times \ldots \times \langle k_i \rangle$ whose image under σ is in U_p. We let

$$b_1, \ldots, b_{j_i} \in \langle k_1 \rangle \times \ldots \times \langle k_i \rangle$$

be such that

$$H_i := \langle b_1 S_i \rangle \times \ldots \times \langle b_{j_i} S_i \rangle = \left(\langle k_1 \rangle \times \ldots \times \langle k_i \rangle \right) / S_i.$$

Notice that $|H_i| = O(1)$. This is because $|E(\mathbb{Q}_p)/2E(\mathbb{Q}_p)| = O(1)$. Hence if there were too many b_j then there would exist a relation of the form

$$\sigma({b_1}^{s_1}) \ldots \sigma({b_{j_i}}^{s_{j_i}}) = \text{identity of } K_p^*/K_p^{*2}$$

where the $s_j \in \{0, 1\}$ and not all $s_j = 0$. But certainly the identity is in the image of the map (3.2). Hence ${b_1}^{s_1} \ldots {b_{j_i}}^{s_{j_i}}$ is in S_i giving a contradiction. Hence as we claimed $|H_i| = O(1)$.

Now we determine the S_i and H_i recursively. To determine S_1 simply check if the image of k_1 is in U_p. If it is then $S_1 \cong \langle k_1 \rangle$ and $H_1 \cong \{S_1\}$. If it is not then $S_1 \cong \{\text{identity}\}$, and $H_1 \cong \langle k_1 S_1 \rangle$.

Suppose we have determined S_i and the H_i. To determine S_{i+1} and H_{i+1} we check if

$$\sigma({b_1}^{s_1}) \ldots \sigma({b_{j_i}}^{s_{j_i}}) \sigma(k_{i+1}) \tag{3.6}$$

is in U_p for any $s_j = 0$ or 1. If none of these are in U_p then $S_{i+1} = S_i$, and

$$H_{i+1} = \langle b_1 S_{i+1} \rangle \times \ldots \times \langle b_{j_i} S_{i+1} \rangle \times \langle k_{i+1} S_{i+1} \rangle .$$

If, on the other hand, the expression (3.6) is in U_p for some choice of $s_j = 0$ or 1 (there can be at most one such choice), then

$$S_{i+1} \cong S_i \times \langle {b_1}^{s_1} \ldots {b_{j_i}}^{s_{j_i}} k_{i+1} \rangle$$

and

$$H_{i+1} \cong \langle b_1 S_{i+1} \rangle \times \ldots \times \langle b_{j_i} S_{i+1} \rangle .$$

The number of choices of b_j that we have is $O(1)$ as $|H_i| = O(1)$. Hence we can determine S_k as a product of cyclic groups all of order 2. The time to do this is then polynomial in $\log D$ via Lemma 3.3.1

Now to determine the Selmer group, we start with $K(R, 2)$ expressed as a product of cyclic groups. For our primes $p_1, \ldots, p_r$ dividing $2D$ we start with p_1

and we determine as above the maximal subgroup $V_{p_1} \leq K(R,2)$ whose image under $\sigma = \sigma_{p_1}$ which is contained in U_{p_1}. Our construction will give us V_{p_1} as a product of cyclic groups of order 2. This will certainly contain the Selmer group. We now discard $K(R,2)$ and find the maximal subgroup of V_{p_1} whose image under σ_{p_2} is contained in U_{p_2}. Doing this recursively we arrive at the Selmer group as soon as we have carried out the above construction for all of $p_1, \ldots, p_r$ and also the infinite prime.

If we have $K(R,2)$ as a product of cyclic groups of order 2 then we will find the Selmer group in polynomial time. Hence the total complexity is given by the complexity of finding $K(R,2)$.

Chapter 4

Descents on the Intersections of 2 Quadrics

4.1 Introduction

Let us briefly review the progress made so far with the problem of computing the Mordell-Weil group of an elliptic curve. We have shown, jointly with N. Smart, how the 2-Selmer group may be computed efficiently (at least in theory) for an elliptic curve defined over $\mathbb{Q}$. We have also given practical methods for performing the infinite descent. We want to show how to determine the coset representatives for $E(\mathbb{Q})/2E(\mathbb{Q})$ once we have computed the Selmer group. Equivalently, we want to determine which of the homogeneous spaces representing the 2-Selmer group has a rational point on it. This is an unsolved problem, and one which we have failed to solve.

Hopefully, a computer search will find rational points on all of the homogeneous spaces representing the 2-Selmer group and this would give us the coset representatives for $E(\mathbb{Q})/2E(\mathbb{Q})$. Occasionally we will not be able to find rational points on some of the homogeneous spaces representing the 2-Selmer group. Here we have two possibilities:

1. The homogeneous space, though everywhere locally soluble, has no global points. Unfortunately, there is no local-to-global principle for curves of genus 1 (see, say [Ca1] pages 85-88).

2. The homogeneous space, has global (i.e. rational) points but these are too large to be found by a naive computer search.

Here, in order to try to find out if the homogeneous space has points on it we will use what is referred to as 'higher descents'. It is appropriate here to explain what is meant by a descent (Compare this to [Ca4] p205): Given a homogeneous space D we construct other curves $D_1, \ldots, D_n$ and rational maps (which are also defined over $\mathbb{Q}$) $\phi_i : D_i \to D$ (of degree > 1) such that for all $P \in D(\mathbb{Q})$ there would exist $Q \in D_i(\mathbb{Q})$ for some i such that $\phi_i(Q) = P$. We will sometimes refer to D as the 'parent', and the D_i as the 'descendants'. [1]

Of course, if we discover that none of the D_i is everywhere locally soluble, then D cannot have a rational point and we would be finished. If a search reveals a rational point Q on one of the D_i then $\phi_i(Q)$ is a rational point on D and we are finished.

Our hope is that the method of constructing the curves D_i and the maps ϕ_i will be recursive. Since $\deg(\phi_i) > 1$, we expect that if there is a rational point P on $D(\mathbb{Q})$ and if $P = \phi_i(Q)$, $Q \in D_i(\mathbb{Q})$, then the (logarithmic) height of Q will be smaller than that of P and hence it will be easier to find a rational point on D_i.

Repeating this process, and rejecting at each stage the D_i which are not everywhere locally soluble, we hope to finally arrive at some curve D_i which has a small rational point or prove that D does not have a rational point.

The basic idea behind such methods is known and has been used to treat particular families of elliptic curves (see for example [Brem, Ca], [Brem], [Str, Top]). However, our exposition will be in a more general setting, allowing the meth-

[1] In the descents which we will study in this chapter, our descendant curves D_i, will always be intersections of two quadric surfaces in $\mathbb{P}^3$. Our D will always be a curve genus 1, and it follows from Theorem B.0.3 that the intersection of the 2 quadric surfaces is in fact transverse, and that D_i is a curve of genus 1.

ods to be better understood and applied. Moreover, we will give new faster algorithms for testing the curves D_i for local solubility.

It should be clear that there is no guarantee for such a method to succeed in deciding whether a particular homogeneous space D has a rational point or not (see [Brem, Bue] and our Chapter (4.8)). But such methods are often effective, and at the very least we hope to have minimized the possibility of failure.

In theory, it happens to be no extra trouble for us to consider these descents over a general number field K, and we will do this. However, we will always assume that we can determine the class group and fundamental units of the number fields we use.

4.2 The Homogeneous Spaces for the 2-Descent

Suppose E is an elliptic curve defined over the number field K and given by the equation

$$y^2 = f(x) \tag{4.1}$$

where $f(x) = x^3 + Ax + B$ is a polynomial over K with no repeated roots. Let L be the algebra

$$L = K[T]/(f(T)),$$

and let Θ be the image of T under the natural homomorphism

$$K[T] \to L.$$

We recall that we have an group homomorphism (see [Ca1] page 66 or [Ca6] page 31)

$$\alpha : E(K) \to L^*/L^{*2} \tag{4.2}$$

given explicitly by

$$P = (x, y) \to (x - \Theta)L^{*2}. \tag{4.3}$$

This homomorphism has as its kernel $2E(K)$. Moreover, the image of this homomorphism is contained in the 2-Selmer group which we regard as a (finite) subgroup of L^*/L^{*2}. We assume that we have already determined the 2-Selmer group which we will denote by S. For each element of S we would like to know if this element can be expressed as $\alpha(P)$ for some $P \in E(K)$. If $s_1, \ldots, s_n$ are the elements of S which can be so expressed, and if say $s_i {L^*}^2 = \alpha(P_i)$ where the $P_i \in E(K)$ then $P_1, \ldots, P_n$ is a complete set of coset representatives for $E(K)/2E(K)$.

Hence we would like to know, for each $s \in S$ if it is possible for us to have

$$(x - \Theta) = s\epsilon^2 \tag{4.4}$$

for some $x \in K$ and $\epsilon \in L$, and if so determine the x (and ϵ) explicitly. Now any such ϵ can be written in the form

$$\epsilon = u_1 + u_2\Theta + u_3\Theta^2$$

where $u_1,\ u_2,\ u_3 \in K$. Substituting in equation (4.4) and comparing coefficients of $1,\ \Theta,\ \Theta^2$ we get

$$\left.\begin{aligned} Q_1(u_1, u_2, u_3) &= x \\ Q_2(u_1, u_2, u_3) &= -1 \\ Q_3(u_1, u_2, u_3) &= 0. \end{aligned}\right\} \tag{4.5}$$

Here we would solve our problem for the particular s if and only if we can find a simultaneous solution to the last two equations above. Thus our 'homogeneous space' is an intersection of 2 quadric surfaces in four (homogeneous) variables:

$$\left.\begin{aligned} Q_2(u_1, u_2, u_3) &= -{u_4}^2 \\ Q_3(u_1, u_2, u_3) &= 0. \end{aligned}\right\} \tag{4.6}$$

This will be a curve of genus 1 and the intersection of the 2 quadric surfaces in (4.6) is a transverse intersection by Theorem B.0.3. It is on this curve that we look for a solution, and it is this curve which is the starting point for our higher descents. We note for future reference that that the second of these quadrics is singular.

4.3 'Coprimality' in number fields

When dealing with homogeneous equations over the rationals (in say n variables), one often make the simplifying assumption that all the solutions are given by integer n-tuples which are not all divisible by a common prime factor. Of course, this useful device does not extend without modification to homogeneous equations defined over number fields, due to the failure of unique factorization over number fields in general. Our purpose here is to show how this method may be modified as follows: instead of insisting that there be no common prime ideal factors at all, we demand that there are no common prime ideal factors from outside a certain finite pre-determined set.

Theorem 4.3.1 *Let $\mathfrak{a}_1, \ldots, \mathfrak{a}_k$ be a set of ideals whose image in the ideal class group of K generates the ideal class group, and let $S = \{\wp_1, \ldots, \wp_m\}$ be the set of prime ideals which divide any of the $\mathfrak{a}_i$. If $(a_1, \ldots, a_n) \in K^n \backslash \mathbf{0}$ then there exists $u \in K^*$ such that $(ua_1, \ldots, ua_n) \in \mathcal{O}_K^n \backslash \mathbf{0}$ and if $\wp$ is a prime ideal dividing all the ua_i then $\wp \in S$.*

Proof. It is obvious that the images of the $\wp_i$ in the ideal class group also generate the ideal class group. By scaling, we may assume that $(a_1, \ldots, a_n) \in \mathcal{O}_K^n \backslash \mathbf{0}$. Suppose we are given $\wp$, a prime ideal which divides all of the a_i but is not contained in S. Let

$$r = \min_{i=1,\ldots,n} \{\mathrm{ord}_{\wp}(a_i)\} \geq 1.$$

Then there exists non-negative integers $\alpha_1, \ldots, \alpha_m$ such that

$$\left(\prod_{j=1}^{m} {\wp_j}^{\alpha_j}\right) \wp^r = (b)$$

for some $b \in \mathcal{O}_K$. Also, clearly, there exist $\beta_1, \ldots, \beta_m$ such that $\beta_j \geq \alpha_j$ for $j = 1, \ldots, m$ and

$$\prod_{j=1}^{m} {\wp_j}^{\beta_j} = (c).$$

Hence the n-tuple $\left(\frac{a_1 c}{b}, \ldots, \frac{a_n c}{b}\right)$ is proportional to $(a_1, \ldots, a_n)$ and is not divisible by $\wp$. Furthermore, if $\wp'$ is a prime ideal not in $S \cup \{\wp\}$ then clearly

$$\min_{i=1,\ldots,n} \left\{\mathrm{ord}_{\wp'}(\frac{a_i c}{b})\right\} = \min_{i=1,\ldots,n} \{\mathrm{ord}_{\wp'}(a_i)\}.$$

Proceeding recursively in this fashion, it follows that there is a vector in ${\mathcal{O}_K}^n$ which is proportional to our original $(a_1, \ldots, a_n)$ and which is not divisible by any common prime ideal which is not in S. □

It is clear that our set S depends on our choice of generators for the ideal class group and so is not unique. However, we will still find it convenient to think of it as fixed at the outset and will denote it by S_K. If $(a_1, \ldots, a_n) \in \mathcal{O}_K^n$ we will say that $a_1, \ldots, a_n$ are coprime outside S_K if no prime ideal $\wp$ in S_K divides all of the a_i.

4.4 Diagonalization

For this section we will assume that $A,\ B$ are $n \times n$ matrices with coefficients in our number field K. We write $\overline{K}$ for the algebraic closure over K. The following is a trivial modification of a standard result for Hermitian matrices.

Lemma 4.4.1 *Suppose $\mathbf{x}_1,\ \mathbf{x}_2 \in \overline{K}^n$ which satisfy*

$$(A - \lambda_1 B)\mathbf{x}_1 = (A - \lambda_2 B)\mathbf{x}_2 = \mathbf{0}$$

for $\lambda_1,\ \lambda_2 \in \overline{K}$. If $\lambda_1 \neq \lambda_2$ then

$$\mathbf{x}_1^t A \mathbf{x}_2 = \mathbf{x}_1^t B \mathbf{x}_2 = 0.$$

Proof. Note that $A\mathbf{x}_2 = \lambda_2 B \mathbf{x}_2$ and $\mathbf{x}_1^t A = \lambda_1 \mathbf{x}_1^t B$. Hence

$$\lambda_1 \mathbf{x}_1^t B \mathbf{x}_2 = \mathbf{x}_1^t A \mathbf{x}_2 = \lambda_2 \mathbf{x}_1^t B \mathbf{x}_2.$$

If $\lambda_1 \neq \lambda_2$ then the result follows. □

Theorem 4.4.1 *Suppose that A and B are $n \times n$ symmetric matrices defined over a number field K such that*

$$F(X, Y) = \det(XA - YB)$$

has distinct roots over K.

If $F(X,Y)$ factorizes as

$$F(X,Y) = F_1(X,Y)\dots F_r(X,Y)$$

where $F_1,\dots,F_r$ are homogeneous and irreducible over K, then there exists a non-singular matrix $P \in GL(K,n)$ such that

$$P^tAP = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}, \; P^tBP = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_r \end{pmatrix} \tag{4.7}$$

where A_i and B_i are symmetric matrices of size $\deg(F_i)$*, and*

$$\det(XA_i - YB_i) = c_iF_i(X,Y)$$

with $c_i \in K$ satisfying $\prod c_i = 1$.

Proof. By replacing A by $A + \epsilon B$ for some $\epsilon \in K$, we may assume that $\det(A) \neq 0$. Suppose λ_j is a root of $F_j(X,1)$. Then there exists $\mathbf{v}_j \in \overline{K}^n$ such that $(A - \lambda_jB)\mathbf{v}_j = \mathbf{0}$. Let $\mathbf{v}_j^{(i)}$ $(i = 1,\dots,\deg(F_j))$ be the conjugates of $\mathbf{v}_j$. Using Lemma 5.8.1, page 40 of [Si2] we know that there are $\mathbf{w}_j^{(i)} \in K^n$ $(i = 1,\dots,\deg(F_j))$ which span the same $\overline{K}$-linear subspace of $\overline{K}^n$ as the $\mathbf{v}_j^{(i)}$. Now let P be the matrix with columns

$$\mathbf{w}_1^{(1)},\dots,\mathbf{w}_1^{(\deg(F_1))},\dots,\mathbf{w}_r^{(1)},\dots,\mathbf{w}_r^{(\deg(F_r))}.$$

It is easy to see, using Lemma 4.4.1 that P is the matrix required by the Theorem. □

4.5 Parametrization of the Singular Combinations

As is noted before our intersections of pairs of quadric surfaces, will be curves of genus 1, must be a transverse intersections, and the singular combinations of the corresponding pencils will all have rank=3. Hence a singular combination is a curve of genus 0. Now since we would have checked our intersection of 2 quadrics for everywhere local solubility we know that this singular combination is everywhere locally soluble [2]. Hence this singular combination must have a global solution and if we were to find one such solution then all others will be given parametrically. It is this parametrization which will enable us to perform our descents in Section 4.6.

Of course, after a non-singular change of variable, we may take our singular combination to be of the form

$$aX^2 + bY^2 + cZ^2 = 0 \tag{4.8}$$

where $a,\ b,\ c \in \mathcal{O}_K \backslash \{0\}$. In [Sieg2] Siegel gives a region for the triple (X, Y, Z) which is guaranteed to contain a solution. We omit giving this because of its complexity. If $K = \mathbb{Q}$ then we have the following Theorem of Holzer

Theorem 4.5.1 *The solvable equation $ax^2+by^2+cz^2 = 0$ taken in its canonical form with $a > 0,\ b > 0,\ c < 0$ has a non-trivial solution with*

$$|x| \leq |bc|^{\frac{1}{2}}, \quad |y| \leq |ca|^{\frac{1}{2}}, \quad |z| \leq |ab|^{\frac{1}{2}}.$$

Proof. See [Holzer], or [Mord] page 47. □

Once we have one non-trivial solution of the Equation (4.8) we can parametrize all others as is well-known.

Theorem 4.5.2 *Suppose $x_0,\ y_0,\ z_0 \in K$ is a non-trivial solution of equation (4.8) then there exist binary quadratic forms $q_1,\ q_2,\ q_3$ defined over K*

[2] Here the assumption that the intersection is transverse is critical. For in this case our equations can be taken of the form

$$q_1(x_1, x_2, x_3) = 0,\ q_2(x_1, x_2, x_3) = {x_4}^2,$$

and if this has a non-trivial solution over any (say local) field then not all of x_1, x_2, x_3 will be zero, and hence $q_1(x_1, x_2, x_3) = 0$ has a non-trivial solution. This of course is not the case for the rational curve

$$x^2 + y^2 + z^2 = 0, x^2 + 2y^2 = zw$$

which has the solution $(0, 0, 0, 1)$, the singular combination $x^2+y^2+z^2 = 0$ has no non-trivial solutions over either of $\mathbb{R}$, or $\mathbb{Q}_2$.

such that $x,\ y,\ z \in K$ is a solution if and only if it is a K-rational multiple of $q_1(u,v),\ q_2(u,v),\ q_3(u,v)$ for some $u,\ v \in K$.

Proof. See [Mord] page 48 which gives $q_1,\ q_2,\ q_3$ explicitly in terms of $a,\ b,\ c$ and $x_0,\ y_0,\ z_0$. □

It follows from the above that one can always constructively obtain a parametric solution to any singular combination of an everywhere locally soluble transverse intersection of 2 quadric surfaces in 4 variables.

4.6 Descents

We will explain 3 methods of 'descent' on the (transverse) intersection of 2 quadrics

$$\left.\begin{array}{r}\mathbf{x}^t A\mathbf{x} = 0\\ \mathbf{x}^t B\mathbf{x} = 0\end{array}\right\} \tag{4.9}$$

Write $F(X,Y) = \det(XA - YB)$. Each method will require that F splits over K in a certain way. Since we are performing the descents to determine if the homogeneous space (4.9) has rational points, we will assume that (4.9) is everywhere locally soluble.

4.6.1 F has at least 2 roots defined over K

Suppose F has at least 2 roots defined over K. By Theorem (4.4.1) we know that by performing a simultaneous change of variable we can take

$$A = \left(\begin{array}{cc|c} a_1 & & \\ & a_2 & \smash{\raisebox{1ex}{0}} \\ \hline \multicolumn{2}{c|}{0} & A_1 \end{array}\right)$$

and

$$B = \left(\begin{array}{c|c} \begin{matrix} b_1 & \\ & b_2 \end{matrix} & 0 \\ \hline 0 & B_1 \end{array}\right).$$

By again taking suitable linear combinations we can assume that in the above $a_1 = 1$, $a_2 = 0$, $b_1 = 0$, $b_2 = 1$. Hence our original equations will have become of the form

$$\left.\begin{aligned} x_1^2 &= Q_1(x_3, x_4), \\ x_2^2 &= Q_2(x_3, x_4) \end{aligned}\right\} \tag{4.10}$$

where $Q_1,\ Q_2$ are quadratic forms with coefficients in K. Since our original pair was everywhere locally soluble, both $x_1^2 = Q_1(x_3, x_4)$, and $x_2^2 = Q_2(x_3, x_4)$ are everywhere locally soluble, and hence have parametric solutions of the form

$$\left.\begin{aligned} x_1 : x_3 : x_4 &= p_1(u_1, u_2) : p_3(u_1, u_2) : p_4(u_1, u_4) \\ x_2 : x_3 : x_4 &= q_2(v_1, v_2) : q_3(v_1, v_2) : q_4(v_1, v_2) \end{aligned}\right\} \tag{4.11}$$

where the p_i and the q_i are binary quadratic forms with coefficients in K.

We construct a set S_K of prime ideals as in Section 4.3. Clearly we may assume that the $p_i,\ q_j$ have coefficients in $\mathcal{O}_K$. Moreover we may assume that the pairs $u_1,\ u_2$ and $v_1,\ v_2$ are defined over $\mathcal{O}_K$, and that each pair is not divisible by any prime ideal not in S_K. Now, we will have a solution to the pair (4.10) if we can find $\alpha \in K^*$ such that

$$\left.\begin{aligned} p_3(u_1, u_2) &= \alpha q_3(v_1, v_2) \\ p_4(u_1, u_2) &= \alpha q_4(v_1, v_2) \end{aligned}\right\} \tag{4.12}$$

has a solution. We note that if $\wp$ is an ideal such that $\mathrm{ord}_\wp(\alpha) > 0$ then either $\wp \in S_K$ or $\wp$ divides the resultant of p_3, p_4, which is necessarily non-zero [3]. Similarly if $\mathrm{ord}_\wp(\alpha) < 0$ then $\wp$ is either in S_K or $\wp$ divides the resultant of

[3] Suppose that the resultant of $p_3,\ p_4$ is 0. Then $p_3,\ p_4$ (when regarded as homogeneous polynomials) have a common factor. Since $p_1,\ p_3,\ p_4$ satisfies $p_1^2 = Q_1(p_3, p_4)$, this common factor must divide p_1. If this common factor is quadratic, then there is precisely one solution to $x_1^2 = Q_1(x_3, x_4)$ in $\mathbb{P}^2$, and this is impossible. Hence $p_i(u_1, u_2) = l(u_1, u_2)m_i(u_1, u_2)$, where

q_3, q_4. Now dropping our requirement that $v_1,\ v_2$ are in $\mathcal{O}_K$, we see that α matters only up to squares in K^*. Hence we may, by the method used for determining $K(R, 2)$ in Chapter 3, determine all possible α modulo ${K^*}^2$. Doing this we will have finitely many curves (4.12) defined over K. The process of deriving these curves from our original curve (4.9) is a 'descent' according to the paradigm of our introduction to this chapter (page 64). The explicit maps from the 'descendants' (4.12) to the 'parent' (4.10) are given explicitly by (4.11).

4.6.2 F is the Product of 2 Irreducible Quadratic Factors

Suppose $F(X, Y) = \det(XA - YB)$ is the product of 2 irreducible factors defined over K. Then, by Theorem 4.4.1, after a non-singular change of variable defined over K we can assume

$$A = \left(\begin{array}{c|c} A_1 & \mathbf{0} \\ \hline \mathbf{0} & A_2 \end{array}\right)$$

and

$$B = \left(\begin{array}{c|c} B_1 & \mathbf{0} \\ \hline \mathbf{0} & B_2 \end{array}\right)$$

where $A_1,\ A_2,\ B_1,\ B_2$ are 2×2 matrices over K and

$$\det(XA - YB) = \det(XA_1 - YB_1)\det(XA_2 - YB_2).$$

Hence $\det(XA_1 - YB_1)$ and $\det(XA_2 - YB_2)$ are irreducible. Hence we may re-write our original equations 4.9 in the form

$$\left.\begin{array}{c} \mathbf{y}^t A_1 \mathbf{y} = -\mathbf{z}^t A_2 \mathbf{z} \\ \mathbf{y}^t B_1 \mathbf{y} = -\mathbf{z}^t B_2 \mathbf{z} \end{array}\right\} \tag{4.13}$$

where

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \ \mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

$l,\ m_1,\ m_3,\ m_4$ are linear, and m_3, m_4 are linearly independent. By a non-singular change of variable, we can assume that $m_2(u_1, u_2) = u_1,\ m_3(u_1, u_2) = u_2$. Hence $m_1(u_1, u_2)^2 = Q_1(u_1, u_2)$. So Q_1 is a square polynomial, and it follows that $x_1^2 - Q_1(x_3, x_4)$ has rank=2, contradicting our assumption that (4.9) is a transverse intersection of quadric surfaces.

Let λ be a root of $\det(A_2 - YB_2)$, so λ is defined over a quadratic extension of K. Then $A_2 - \lambda B_2$ has rank=1. So there exists $\delta \in K(\lambda)$ such that

$$\mathbf{y}^t(A_1 - \lambda B_1)\mathbf{y} = \delta\zeta^2. \tag{4.14}$$

Since our original pair of equations (4.9) was soluble over all localizations of K, it follows that this equation regarded as curve of genus 0 over $K(\lambda)$ must have solutions over all localizations of $K(\lambda)$. By the results of Section 4.5 there exists binary quadratic forms $q_1,\ q_2,\ q_3$ defined over $K(\lambda)$ such that the triple $y_1,\ y_2, \zeta$ is a solution to (4.14) over $K(\lambda)$ if and only if

$$y_1 : y_2 : \zeta = q_1(\psi_1, \psi_2) : q_2(\psi_1, \psi_2) : q_3(\psi_1, \psi_2). \tag{4.15}$$

By comparing the coefficients of $1,\ \lambda$ in (4.14), we see that there exists a solution to (4.13) defined over K if and only if there exists $y_1,\ y_2 \in K$ and $\alpha \in K(\lambda)^*$, $\psi_1,\ \psi_2 \in K(\lambda)$ (not both zero) such that

$$\begin{aligned} y_1 &= \alpha q_1(\psi_1, \psi_2) \\ y_2 &= \alpha q_2(\psi_1, \psi_2). \end{aligned} \tag{4.16}$$

Suppose for the moment that we have a fixed α. Now $\psi_1 = \beta_1 + \lambda\beta_2$, and $\psi_2 = \beta_3 + \lambda\beta_4$ for some $\beta_1, \ldots, \beta_4 \in K$, and not all zero. Then expanding $\alpha q_1(\psi_1, \psi_2)$ and $\alpha q_2(\psi_1, \psi_2)$ and comparing the coefficients of λ in (4.16) we get a pair of homogeneous quadratic equations in 4 variables (defined over K) which we want to solve:

$$\begin{aligned} Q_1(\beta_1, \ldots, \beta_4) &= 0 \\ Q_2(\beta_1, \ldots, \beta_4) &= 0. \end{aligned}$$

Now this pair depends on the choice of α and we need to show that α can be taken from a finite set whose elements can be effectively enumerated. This is what we shall do. We note that we may assume the following

1. By scaling we can assume that $A_1,\ A_2,\ B_1,\ B_2$ all have entries in $\mathcal{O}_K$.

2. If $y_1,\ y_2,\ z_1,\ z_2$ is a solution to (4.13) then we may assume that they are in $\mathcal{O}_K$ and that the 4-tuple (y_1, y_2, z_1, z_2) is not divisible by any prime ideal not in S_K.

3. $q_1,\ q_2$ are defined over $\mathcal{O}_{K(\lambda)}$, and $\psi_1,\ \psi_2$ are in $\mathcal{O}_{K(\lambda)}$ and are not both divisible by any prime ideal not in $S_{K(\lambda)}$.

Now if $\wp$ is a prime ideal of $K(\lambda)$ such that $\mathrm{ord}_\wp(\alpha) \leq -1$ then it is easy to see from (4.16) and the above assumptions that either $\wp$ divides the resultant of $q_1,\ q_2$ or $\wp$ is in $S_{K(\lambda)}$. Moreover if $\wp$ is a prime ideal of $K(\lambda)$ such that $\mathrm{ord}_\wp(\alpha) \geq 1$ then $\wp$ will divide both $y_1,\ y_2$ and hence $\wp$ will divide one of the ideals in $\left\{\wp'\mathcal{O}_{K(\lambda)} : \wp' \in S_K\right\}$. Hence it is clear that $\mathrm{ord}_\wp(\alpha) = 0$ for all prime ideal $\wp$ not in some finite set, and that this set can be determined. We now drop our requirement that $\psi_1,\ \psi_2$ are in $\mathcal{O}_{K(\lambda)}$, and so we see from (4.16) that α matters only up to squares in $K(\lambda)^*$. By the methods of Chapter 3 we may determine all possible α modulo $K(\lambda)^{*2}$. This shows that the α may be taken from a finite set which can be enumerated.

4.6.3 F has exactly one root defined over K

Suppose $F(X,Y) = \det(XA - YB)$ has exactly one rational root. By a non-singular change of coordinates (defined over K) and then taking appropriate linear combinations, we may assume that

$$A = \left(\begin{array}{c|c} A_1 & \mathbf{0} \\ \hline \mathbf{0} & 0 \end{array}\right)$$

$$B = \left(\begin{array}{c|c} B_1 & \mathbf{0} \\ \hline \mathbf{0} & 1 \end{array}\right),$$

where $A_1,\ B_1$ are 3×3 matrices over K and $\det(XA_1 - YB_1)$ is irreducible. Hence we may rewrite our original equations (4.9) in the form

$$\left.\begin{array}{rcl} \mathbf{z}^t A_1 \mathbf{z} & = & 0 \\ \mathbf{z}^t B_1 \mathbf{z} & = & y^2 \end{array}\right\} \qquad (4.17)$$

where $\mathbf{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$. As usual we parametrize the solutions to $\mathbf{z}^t A_1 \mathbf{z} = 0$ by

$$z_1 : z_2 : z_3 = q_1(X_1, X_2) : q_2(X_1, X_2) : q_3(X_1, X_2), \tag{4.18}$$

Suppose that $z_i = \alpha q_i(X_1, X_2)$ for some $\alpha \in K^*$. Substituting this into $\mathbf{z}^t B_1 \mathbf{z} = y^2$ we see that

$$G(X_1, X_2) = Y^2 \tag{4.19}$$

where $Y = \alpha^{-1} y$ and G is a binary quartic form with coefficients in K. It is easy to see that (4.17) has a solution defined over K if and only if (4.19) has a solution defined over K. Hence if G has a root defined over K, (4.17) has a solution defined over K.

Lemma 4.6.1 *IF G has no roots defined over K then it is irreducible (i.e. it is not a product of* 2 *irreducible quadratic factors).*

Proof. Suppose G is a product of two irreducible quadratic factors. Let L' be the splitting field of G. Let λ be a root of the irreducible cubic polynomial $\det(B_1 - XA_1)$. Now $[K(\lambda) : K] = 3$ and $[L' : L] = 2$ or 4, and so $\lambda \notin L'$.

Now $\mathbf{z}^t(B_1 - \lambda A_1)\mathbf{z}$ has rank $= 2$ and so

$$\mathbf{z}^t(B_1 - \lambda A_1)\mathbf{z} = \beta m(z_1, z_2, z_3)^2 + \gamma n(z_1, z_2, z_3)^2$$

where $\beta,\ \gamma \in K(\lambda)^*$ and $m,\ n$ are linear forms with coefficients in $K(\lambda)$. Let $L'' = L'(\sqrt{-\beta/\gamma})$. Clearly $\lambda \notin L''$, since again $[L'' : K]$ is a power of 2.

Then

$$\mathbf{z}^t(B_1 - \lambda A_1)\mathbf{z} = M(z_1, z_2, z_3)N(z_1, z_2, z_3) \tag{4.20}$$

where $M,\ N$ are linear forms with coefficients in $L''(\lambda)$. But if we let $z_i = q_i(X_1, X_2)$ then $\mathbf{z}^t B_1 \mathbf{z} = G(X_1, X_2)$ and $\mathbf{z}^t A_1 \mathbf{z} = 0$. Hence

$$G(X_1, X_2) = M(q_1, q_2, q_3)N(q_1, q_2, q_3),$$

where $M(q_1, q_2, q_3)$ and $N(q_1, q_2, q_3)$ will be quadratic factors of $G(X_1, X_2)$. But the roots of G are contained in L' and hence in L''. So there is some

$\delta \in L''(\lambda)\backslash \{0\}$ such that $\delta M(q_1, q_2, q_3)$ and $\delta^{-1}N(q_1, q_2, q_3)$ are in $L''[X_1, X_2]$. Write

$$\delta M(z_1, z_2, z_3) = M_1(z_1, z_2, z_3) + \lambda M_2(z_1, z_2, z_3) + \lambda^2 M_3(z_1, z_2, z_3)$$

where $M_1,\ M_2,\ M_3$ are linear forms with coefficients in L''. Hence $M_2(q_1, q_2, q_3)$ is (identically) zero, and so (since $q_1 : q_2 : q_3$ is a parametrization of the zeros of $\mathbf{z}^t A_1 \mathbf{z}$,) the (projective) variety $M_2(z_1, z_2, z_3) = 0$ contains the locus of zeros of the non-singular conic $\mathbf{z}^t A_1 \mathbf{z} = 0$. It follows that $M_2(z_1, z_2, z_3)$ is (identically) zero. Similarly $M_3 = 0$ and hence $\delta M(z_1, z_2, z_3)$ has coefficients in L''. This is also true for $\delta^{-1}N(z_1, z_2, z_3)$ by the same argument. Hence by (4.20) it follows the coefficients of $\mathbf{z}^t(B_1 - \lambda A_1)\mathbf{z}$ are in L'', which contradicts $\lambda \notin L''$. This completes the proof [4]. □

We have already stated that if G has a root in K then (4.19) and hence (4.17) has a point defined over K. Hence we will assume that G has no roots over K, and by the above Lemma it will follow that G is irreducible. We now return to the descents. Rewrite the equation (4.19) in the form

$$aY^2 = F(X_1, X_2) \tag{4.21}$$

where $a \in \mathcal{O}_K$ and F is irreducible monic of degree 4 with coefficients in $\mathcal{O}_K$ [5]. Let Θ be a root of $F(X_1, 1)$ and let $L = K(\Theta)$. It is clear that if (X_1, X_2, Y) is a solution to (4.21) then we can assume that $X_1,\ X_2$ are coprime outside a certain predetermined set of prime ideals S_K.

Lemma 4.6.2 *If (X_1, X_2, Y) is a solution to the equation (4.21) and X_1, X_2 are coprime outside S_K, then we can write*

$$(X_1 - \Theta X_2)\mathcal{O}_L = \mathfrak{a}\mathfrak{b}^2 \tag{4.22}$$

where $\mathfrak{a}$, $\mathfrak{b}$ are ideals of L, and

[4] In essence we have shown that $\det(B_1 - XA_1)$ is the resolvent cubic of G.

[5] Clearly by multiplying the equation (4.19) by a suitable element in ${\mathcal{O}_K}^2$ we can assume that G has coefficients in $\mathcal{O}_K$. Now if a is the leading coefficient of $G(X_1, X_2)$, then we simply multiply (4.19) by a^3, replace Y by $a^{-2}Y$, and X_1 by $a^{-1}X_1$.

1. *$\mathfrak{a}$ is square-free,*

2. $\mathrm{Norm}_{L/K}(\mathfrak{a}) \in aK^{*2}$,

3. *If $\wp$ is a prime ideal of L and $\wp|\mathfrak{a}$ then either*

   (a) *$\wp|a$, or*

   (b) *$\wp|\Delta(F)$, where $\Delta(F)$ is the discriminant of F, or*

   (c) *$\wp|\mathfrak{q}\mathcal{O}_L$ where $\mathfrak{q} \in S_K$.*

Proof. It is clear that we can write $(X_1 - \Theta X_2)\mathcal{O}_L$ in the form (4.22), where $\mathfrak{a}$ satisfies conditions 1 and 2 of the Lemma. Suppose $\wp$ is a prime ideal of L, such that $\wp|\mathfrak{a}$, and $\wp$ does not divide $a,\ \Delta(F)$. Let L' be the splitting field of F over K, so that $L' \supseteq L \supseteq K$. Since $\wp$ does not divide $\Delta(F)$ we see that $\wp$ does not ramify over L'. Hence if $\mathfrak{q}|\wp\mathcal{O}_{L'}$, where $\mathfrak{q}$ is a prime ideal of L', then

$$\mathrm{ord}_{\mathfrak{q}}(X_1 - \Theta X_2)\mathcal{O}_{L'} = \mathrm{ord}_{\wp}(X_1 - \Theta X_2)\mathcal{O}_L$$

which is odd. But

$$\mathrm{ord}_{\mathfrak{q}} \prod_{i=1}^{4}(X_1 - \Theta_i X_2) = \mathrm{ord}_{\mathfrak{q}} F(X_1, X_2) = \mathrm{ord}_{\mathfrak{q}} aY^2$$

is even, since $\mathfrak{q}$ does not divide a. So $\mathfrak{q}|(X_1 - \Theta' X_2)\mathcal{O}'_L$ where Θ' is a root of $F(X_1, 1)$ which does not equal Θ. Hence $\mathfrak{q}|(\Theta - \Theta')X_1$ and $\mathfrak{q}|(\Theta - \Theta')X_2$. But $\mathfrak{q}$ does not divide $(\Theta - \Theta')$ since otherwise $\mathfrak{q}|\Delta(F)$ which would imply $\wp|\Delta(F)$. Hence $\mathfrak{q}|X_1,\ X_2$. Since this is true for all ideals $\mathfrak{q}$ of L' dividing $\wp\mathcal{O}_{L'}$, it follows that $\wp|X_1,\ X_2$. Hence the conclusion follows. □

Suppose that (X_1, X_2, Y) is a solution to (4.21), and $X_1,\ X_2$ are coprime outside S_K. We will write

$$X_1 - \Theta X_2 = \epsilon\gamma^2 \tag{4.23}$$

where $\epsilon,\ \gamma \in L^*$. As usual, our ϵ matters only up to squares in L^*. From the above Lemma, we see that our ϵ is supported, modulo square ideals, by the (finitely many) prime ideals specified in condition 3 of the Lemma. Hence, by the method in Chapter 3, we can list a complete set of representatives of L^*

modulo L^{*2} which are supported by these prime ideals. Of these, we keep only those whose Norm is in aK^{*2}; these will be our required ϵs. For any fixed ϵ, if we write $\gamma = \sum_{i=1}^{4} v_i \Theta^{(i-1)}$, $v_i \in K^*$ and compare coefficients of $1, \ldots, \Theta^3$ in (4.23), we will get

$$\left.\begin{array}{rcl} Q_1(v_1, \ldots, v_4) & = & X_1 \\ Q_2(v_1, \ldots, v_4) & = & X_2 \\ Q_3(v_1, \ldots, v_4) & = & 0 \\ Q_4(v_1, \ldots, v_4) & = & 0 \end{array}\right\} \tag{4.24}$$

where the Q_i are quadratic forms with coefficients in K. Now

$$\left.\begin{array}{rcl} Q_3(v_1, \ldots, v_4) & = & 0 \\ Q_4(v_1, \ldots, v_4) & = & 0 \end{array}\right\} \tag{4.25}$$

defines an intersection of 2 quadrics as is required.

4.7 Examples

We have applied the method of the previous section to obtain generators on the congruent number curve

$$E : Y^2 = X(X^2 - p^2) \tag{4.26}$$

for primes

$$p = 257,\ 313,\ 353,\ 1201,\ 1217,\ 1249,\ 1321,\ 2113,\ 2273,\ 2777,\ 2833,\ 2953.$$

these primes are all congruent to $1 \pmod 8$, and it is easy to show that the 2-Selmer rank will always be 2 (see below). In [Serf], all the integers $n \leq 3000$ for which the rank of $Y^2 = X(X^2 - n^2)$ is 2 are predicted. For most of these, Mordell-Weil generators were found in [Ge, Zi]. However some values were omitted (presumably because the generators were too large for the method). The primes above are all the primes $p \equiv 1 \pmod 8$ for which generators are not given in [Ge, Zi]. We shall omit most of the details, as they are quite tedious.

As usual, we have a map

$$E(\mathbb{Q}) \to \mathbb{Q}^*/\mathbb{Q}^{*2} \times \mathbb{Q}^*/\mathbb{Q}^{*2} \times \mathbb{Q}^*/\mathbb{Q}^{*2}$$

given by [6]

$$(X,Y) \to [X\mathbb{Q}^{*2}, (X-p)\mathbb{Q}^{*2}, (X+p)\mathbb{Q}^{*2}].$$

It is not hard to show that the 2-Selmer group, regarded as a subgroup of $\mathbb{Q}^*/\mathbb{Q}^{*2} \times \mathbb{Q}^*/\mathbb{Q}^{*2} \times \mathbb{Q}^*/\mathbb{Q}^{*2}$ is generated by $[p,1,p]$, $[p,2p,2]$, and the images of the points of order 2. Hence if we can find points on the homogeneous spaces corresponding to these 2 elements of the 2-Selmer group, we will be able to write down generators for $E(\mathbb{Q})/2E(\mathbb{Q})$.

1. The homogeneous space corresponding to $[p,1,p]$ is

   $$\left.\begin{array}{rcl} U^2 - V^2 &=& pA^2 \\ U^2 + V^2 &=& B^2 \end{array}\right\}$$

   A rational point on this homogeneous space gives a corresponding rational point on E: $(pU^2/V^2, p^2ABU/V^3)$. After 2 descents starting from this homogeneous space we arrived at the equation

   $$e_1^4 + 8e_1^3 f_1 + 12e_1^2 f_1^2 + 16e_1 f_1^3 + 4f_1^4 = pg_1^2. \tag{4.27}$$

   A rational point on this gives P_1 on E with X-coordinate

   $$X = p\left(\frac{e_1^4 + 4e_1^3 f_1 + 4e_1^2 f_1^2 + 8e_1 f_1^3 + 4f_1^4}{4e_1 f_1(e_1^2 + 2e_1 f_1 + 2f_1^2)}\right)^2.$$

   After a small search for points on (4.27) we have found a point for each p in our list and computed the corresponding generator P_1 on E. The information is contained in the table below.

2. Similarly, after 2 descents on the homogeneous space corresponding to $[p,2p,2]$ we arrive at

   $$e_2^4 - 8e_2^3 f_2 + 18e_2^2 f_2^2 + 8e_2 f_2^3 + f_2^4 = pg_2^2. \tag{4.28}$$

   which gives generator P_2 on E with X-coordinate

   $$X = p\left(\frac{e_1^4 - 4e_1^3 f_1 + 10e_1^2 f_1^2 + 4e_1 f_1^3 + f_1^4}{e_1^4 - 4e_1^3 f_1 - 6e_1^2 f_1^2 + 4e_1 f_1^3 + f_1^4}\right)^2.$$

[6] For (X,Y) a point of order 2 the definition must be adjusted as on page 67 of [Ca1]

Again, after a short search for points on (4.28), we found a point with small coordinates for each p in our list.

p	e_1	f_1	$\hat{h}(P_1)$	e_2	f_2	$\hat{h}(P_2)$	$R(P_1, P_2)$
257	1	2	10.243	13	2	19.340	168.055
253	121	4	38.629	34	23	30.979	177.561
313	194	3	39.493	-18	103	35.825	193.933
1201	151	6	40.455	-56	57	36.927	247.108
1217	-29	4	25.829	8	1	15.628	316.657
1249	27	52	36.378	206	45	41.421	443.022
1321	11	3	21.274	9	2	16.394	285.289
2113	-1	6	16.437	-24	31	30.879	490.464
2273	8	1	14.848	1346	751	59.533	785.932
2777	22	73	35.506	11	2	17.919	621.627
2833	164	3	38.172	-117	82	41.802	1147.840
2953	537	29	50.718	-261	184	48.236	848.351

It should be clear from the height of the points listed that not all could be found by a naive computer search for points [7] on E.

4.8 Local to Global- A Counter Example

As we saw, given an intersection of two quadrics which defines a curve of genus 1, if there exist a singular combination of the quadrics which defined over the ground number field, then it is possible to perform a descent arriving at other curves of genus 1. It was therefore clear a priori, that it is unreasonable that a local-to-global principle should exist for such curves. This is because, even though the original curve is everywhere locally soluble, it is possible that none of its descendents are. This leaves us with the following question: Given an

[7] Experience with `findinf` shows that it is not very feasible to search for points whose logarithmic height is much greater that 15.

intersection of two quadrics

$$\left.\begin{array}{r}\mathbf{x}^t A\mathbf{x} = 0\\ \mathbf{x}^t B\mathbf{x} = 0\end{array}\right\} \tag{4.29}$$

where A, B, are symmetric 4×4 matrices with entries over a number field K, and suppose that $F(X,Y) = \det(XA - YB)$ is irreducible over K, and that the pair (4.29) is everywhere locally soluble. Is it the case that the pair must have a non-trivial solution defined over K? Infact, in the literature of Curves of Genus 1 I know of no counterexamples to local-to-global principle which cannot be demonstrated by constructing a complete covering of rational [8] descendents of the original curve, and showing that these are all insoluble on local grounds. Here we are faced with the task of constructing such a counterexample with no obvious method of constructing a covering of rational descendants. However there are well-known counterexamples where (by an argument due to Lind, see [Ca4] page 284) it is possible to show that there are no global points, without constructing any coverings , even though in such cases it was always easy to disprove global solubility by constructing a complete set of rational coverings.

In our counterexample to the question posed we will mimic Lind's argument for disproving global solubility, and use our results from Chapters 5, and 6 to prove everywhere local solubility.

Example 4.8.1

$$\left.\begin{array}{lcl}-2x^2 + 34x(z+w) + y^2 - 17z^2 & = & 0\\ -17x^2 + 3y^2 + 4yz + w^2 & = & 0\end{array}\right\} \tag{4.30}$$

Here, the relevant $\det(XA - YB)$ *is irreducible, as required by our question. We will show that (4.30) is everywhere locally soluble but has no non-trivial rational points.*

Let us prove first that the pair (4.30) have no common solutions over $\mathbb{Q}$*. Suppose that a non-trivial rational solution* (x, y, z, w) *exists. We may suppose that* x, y, z, w *are integers and that they have no common divisor.*

[8] Here, by rational, we mean defined over the ground field.

It is easy to see that $x \neq 0$. *If* $17|x$ *then by the first equation* $y^2 - 17z^2 \equiv 0 \pmod{17^2}$ *(since* $17^2|x^2$ *and* $17^2|34x$*). It would follow that* $17|y$ *and* $17|z$. *From the second equation we get that* $17|w$. *This contradiction shows that* 17 *does not divide* x.

Suppose p is an odd prime dividing x. *Reducing the first equation of (4.30) modulo p , we get that* $y^2 \equiv 17z^2 \pmod{p}$. *It is straightforward to see that if* $p|y$ *or* $p|z$ *then* $p|x,\ y,\ z,\ w$, *giving us a contradiction. Hence we deduce that* 17 *is a quadratic residue modulo p. By the Law of Quadratic Reciprocity it follows that p is a quadratic residue modulo* 17. *This is true of all odd primes that divide* x. *But* -1, *and* 2 *are quadratic residues modulo* 17. *Hence we can write*

$$x \equiv x_1^2 \pmod{17}.$$

Similarly to the above, it is easy to show that 17 *does not divide y, and that there exists* y_1 *such that*

$$y \equiv y_1^2 \pmod{17}.$$

Since 17 *does not divide* $x,\ y$ *we get that* 17 *does not divide* $x_1,\ y_1$ *either.*

Now reduce the first equation of (4.30) modulo 17. *It follows that*

$$-2x^2 + y^2 \equiv 0 \pmod{17}$$

and so

$$2x_1^4 \equiv y_1^4 \pmod{17}.$$

But 2 *is not a quartic residue modulo* 17. *This gives us a contradiction. Hence no rational solution to the pair (4.30) exists.*

Let us now prove that the pair (4.30) has solutions everywhere locally. We can write (4.30) in the form of (4.29) where

$$A = \begin{pmatrix} -2 & 0 & 17 & 17 \\ 0 & 1 & 0 & 0 \\ 17 & 0 & -17 & 0 \\ 17 & 0 & 0 & 0 \end{pmatrix}$$

and

$$B = \begin{pmatrix} -17 & 0 & 0 & 0 \\ 0 & 3 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Let $F(\lambda,\mu) = \det(\lambda A - \mu B)$*. The discriminant of* F *is*

$$-11\ 327\ 320\ 899\ 466\ 789\ 632\ 139.$$

The prime factorization of this is

$$17^3 \times 2\ 305\ 581\ 294\ 416\ 199\ 803.$$

We denote the second prime by p_1*. It follows from our Theorem 5.2.1 that to prove everywhere local solubility it sufficient to prove that the pair (4.30) has solutions over* $\mathbb{R}$*,* $\mathbb{Q}_2$*,* $\mathbb{Q}_{17}$*,* $\mathbb{Q}_{p_1}$*.*

1. Over $\mathbb{R}$. *Here we note that the roots of* $F(1,\mu)$ *are roughly* 0.34, 13.29, $-0.38 \pm 3.79i$*. Hence* F *is not totally real and so we must have solubility in* $\mathbb{R}$ *by Lemma (6.2.2) on page 97.*

2. Over $\mathbb{Q}_2$. *We recall that any* 2*-adic number which is congruent to* 1 *modulo* 8 *is a* 2*-adic square. Hence* 17 *is a* 2*-adic square. Moreover, it is easy to show that the (two)* 2*-adic square-roots of* 17 *are congruent to* 1, -1 *modulo* 8 *respectively. Let* $\alpha \in \mathbb{Q}_2$ *satisfying* $\alpha^2 = 17$ *and* $\alpha \equiv 1 \pmod 8$*. Then* $-3\alpha^2 - 4\alpha \equiv 1 \pmod 8$*. Hence there exists* $\beta \in \mathbb{Q}_2$ *satisfying*

$$\beta^2 = -3\alpha^2 - 4\alpha.$$

   It follows that $(0,\alpha,1,\beta)$ *is a* 2*-adic solution to the pair (4.30).*

3. Over $\mathbb{Q}_{17}$. *Here it is sufficient to observe that* $(1,6,4,0) \pmod{17}$ *is a non-singular point on the reduction of the pair (4.30) modulo* 17*, and hence by Theorem (5.2.1) it must lift to a non-trivial solution on* $\mathbb{Q}_{17}$*.*

4. Over $\mathbb{Q}_{p_1}$ *Here* $F(\lambda,\mu)$ *has only one double root modulo* p_1*. This is*

$$1 : 393\ 077\ 095\ 592\ 234\ 641.$$

Write this as $1 : \mu_1$*. Then* $A + \mu_1 B \pmod{p_1}$ *is, up to scalar multiples, the only linear combination of* $A,\ B \pmod{p_1}$ *which might have rank* ≤ 2*. Using* `Pari/GP`*, we find that its rank* $= 3$*. This means that every non-trivial linear combination of* $A,\ B \pmod{p_1}$ *has rank* ≥ 3*. It is now easy to use the proof of Lemma B.0.4 to construct a non-singular point on the pair (4.30) modulo* p_1*, and hence this must have a non-trivial point defined over* $\mathbb{Q}_{p_1}$*.*

Chapter 5

Local Solubility I: Over Non-Archimedean Completions

5.1 Introduction

Let K be a number field, and let A, B be 4×4 symmetric matrices with entries in K such that $\det(XA - YB)$ has distinct roots (i.e. the combinant [1] $\partial(A, B) \neq 0$). Our goal in this chapter is to give algorithms for determining the solubility of

$$\left.\begin{aligned} \mathbf{x}^t A \mathbf{x} = 0 \\ \mathbf{x}^t B \mathbf{x} = 0 \end{aligned}\right\} \tag{5.1}$$

over the non-archimedean completions of K. Our notation for this chapter is as follows:

K	a number field
M_K^0	a full set of inequivalent non-archimedean valuations on K
$\mathcal{O}$	the set of integers of K

[1] See page 107.

υ	a non-archimedean valuation on K
$\mathcal{O}_\upsilon$	$:= \{x \in K_\upsilon : \upsilon(x) \geq 0\}$
π	a prime element for υ (i.e. $\pi \in K_\upsilon$ such that $\upsilon(\pi) = 1$)
k_υ	residue field associated with K_υ
q	the number of elements in k_υ
$\mathbb{P}_\pi$	$:= \{(x, y) : x,\ y \in \mathcal{O}_\upsilon \text{ and } \min(\upsilon(x), \upsilon(y)) = 0\}$

Further, when working with a fixed valuation υ, we let $\mathcal{O}_\upsilon \to k_\upsilon$ be the natural map sending $x \in \mathcal{O}_\upsilon$ to $\overline{x} \in k_\upsilon$. Similarly, given vectors $\mathbf{v}$ and matrices C defined over $\mathcal{O}_\upsilon$, we denote by $\overline{\mathbf{v}}$ and $\overline{C}$ to be their reductions in the obvious way.

We can assume without loss of generality that A and B have entries in $\mathcal{O}$ and hence that $\partial(A, B)$ is in $\mathcal{O}$.

The first algorithm we will give, relies on searching for points on

$$\left.\begin{aligned} \mathbf{x}^t A\mathbf{x} \equiv 0 \\ \mathbf{x}^t B\mathbf{x} \equiv 0 \end{aligned}\right\} mod(\pi) \tag{5.2}$$

for πs corresponding to a finite pre-determined set of υs, and then attempting to lift the points found to points modulo powers of π until it is certain that they will lift to points defined over ${\mathcal{O}_\upsilon}^4$. We need two pieces of information:

1. For which of the infinitely many $\upsilon \in M_K^0$ is it necessary to do this?

2. Modulo which power of the corresponding π is it sufficient to find a solution, to be sure that this solution will lift?

The second algorithm we give assumes that $A,\ B$ have a singular combination defined over $\mathcal{O}_\upsilon$.

5.2 Algorithm I

Theorem 5.2.1 *Suppose $A,\ B$ are 4×4 symmetric matrices with entries in $\mathcal{O}_\upsilon$ such that $\partial(A, B) \neq 0$. We have*

1. *If $\upsilon(2\partial(A,B)) = 0$ then*
$$\left.\begin{aligned} \mathbf{x}^t A\mathbf{x} = 0 \\ \mathbf{x}^t B\mathbf{x} = 0 \end{aligned}\right\} \tag{5.3}$$
has a non-trivial solution over $\mathcal{O}_\Upsilon$.

2. *Suppose that there exists $\mathbf{x}_0 \in {\mathcal{O}_\upsilon}^4 \backslash \pi {\mathcal{O}_\upsilon}^4$ such that*
$$\mathbf{x}_0^t A\mathbf{x}_0 \equiv \mathbf{x}_0^t B\mathbf{x}_0 \equiv 0 \pmod{\pi^{2\delta+1}}$$
and there is no pair $(\lambda,\mu) \in \mathbb{P}_\pi$ such that $2(\lambda A\mathbf{x}_0 - \mu B\mathbf{x}_0) \equiv \mathbf{0} \pmod{\pi^{\delta+1}}$. Then there exists $\mathbf{x} \in {\mathcal{O}_\upsilon}^4$ such that $\mathbf{x} \equiv \mathbf{x}_0 \pmod{\pi^{\delta+1}}$ and $\mathbf{x}$ is a non-trivial solution to the pair of equations (5.3).

Proof. For the first part it is sufficient to note that if $\upsilon(2\partial(A,B)) = 0$ then $\mathbf{x}^t\overline{A}\mathbf{x} \equiv \mathbf{x}^t\overline{B}\mathbf{x} \equiv 0 \pmod{\pi}$ has genus 1 by Theorem B.0.2; it follows then from [Ca4] page 205 that there is a non-trivial solution to (5.3). The second part is a special case of Theorem 5.21 on page 64 of [Gre]. □

Thus it is clear that to test local solubility at the archimedean places, it sufficient to check solubility over K_υ only for those $\upsilon \in M_K^0$ for which $\upsilon(2\partial(A,B)) \neq 0$. For any such υ, we can do this using the following algorithm.

Algorithm 5.2.1 *We write down a complete set of coset representatives of ${\mathcal{O}_\upsilon}^4/\pi{\mathcal{O}_\upsilon}^4$. For any of these (other than the one contained in $\pi{\mathcal{O}_\upsilon}^4$) we check if it gives a solution to $\mathbf{x}^tA\mathbf{x} \equiv \mathbf{x}^tB\mathbf{x} \equiv 0 \pmod{\pi}$. If there are none which give a solution to this, then (5.3) has no solution over K_υ and we can stop. If there are some, and for one of them we can establish that it lifts to a point on (5.3) by the above Theorem then we can stop. If not then we will be left with $\mathbf{x}_1,\ldots,\mathbf{x}_n$ satisfying $\mathbf{x}^tA\mathbf{x} \equiv \mathbf{x}^tB\mathbf{x} \equiv 0 \pmod{\pi}$ and for each there exists $(\lambda,\mu) \in \mathbb{P}_\pi$ such that $2(\lambda A\mathbf{x} - \mu B\mathbf{x}) \equiv \mathbf{0} \pmod{\pi}$.*

Suppose now that after r steps we are left with a set of $\mathbf{x}_1,\ldots,\mathbf{x}_n$ (not necessarily the same $\mathbf{x}_i$ as before) satisfying $\mathbf{x}^tA\mathbf{x} \equiv \mathbf{x}^tB\mathbf{x} \equiv 0 \pmod{\pi^{2r+1}}$ and for each there exists $(\lambda,\mu) \in \mathbb{P}_\pi$ such that $2(\lambda A\mathbf{x} - \mu B\mathbf{x}) \equiv \mathbf{0} \pmod{\pi^{r+1}}$. Then for each $i = 1,\ldots,n$, we write a complete set of representatives of $\mathcal{O}_\upsilon/\pi^{2r+3}\mathcal{O}_\upsilon$ which are congruent to $\mathbf{x}_i$ modulo π^{2r+1}. If none of these are on $\mathbf{x}^tA\mathbf{x} \equiv \mathbf{x}^tB\mathbf{x} \equiv$

$0 \pmod{\pi^{2r+3}}$ *then we go to the next* i *(if there are none for all the* i*s then 5.3 has no solutions and we can stop). If there are some on* $\mathbf{x}^t A\mathbf{x} \equiv \mathbf{x}^t B\mathbf{x} \equiv 0 \pmod{\pi^{2r+3}}$*, which using the above Theorem will lift to points on (5.3) then we can stop. So we may suppose that pooling our* $\mathbf{x}$*s that we get for each* i *we end up with a new list* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *all satisfying* $\mathbf{x}^t A\mathbf{x} \equiv \mathbf{x}^t B\mathbf{x} \equiv 0 \pmod{\pi^{2r+3}}$ *and for each there exists a pair* $(\lambda, \mu) \in \mathbb{P}_\pi$ *such that* $2(\lambda A\mathbf{x} - \mu B\mathbf{x}) \equiv \mathbf{0} \pmod{\pi^{r+2}}$*. The following Lemma shows that we must eventually stop.*

Lemma 5.2.1 *Suppose that there exists* $\mathbf{x}_1 \in {\mathcal{O}_v}^4$ *such that*

$$\mathbf{x}_1 A\mathbf{x}_1 \equiv \mathbf{x}_1 B\mathbf{x}_1 \equiv 0 \pmod{\pi^\alpha}$$

and there exists $(\lambda : \mu) \in \mathbb{P}_\pi$ *such that* $(\lambda A\mathbf{x}_1 - \mu B\mathbf{x}_1) \equiv \mathbf{0} \pmod{\pi^\beta}$*, then* $\min(\alpha, \beta) \leq v(\partial(A,\ B))$*.*

Proof. Let $\gamma = \min(\alpha,\ \beta)$. Choose $\mathbf{x}_2,\ \mathbf{x}_3,\ \mathbf{x}_4 \in {\mathcal{O}_v}^4$ such that $\mathbf{x}_1,\ \mathbf{x}_2,\ \mathbf{x}_3,\ \mathbf{x}_4$ are linearly independent modulo π. Let T be the 4×4 matrix with columns $\mathbf{x}_1,\ \mathbf{x}_2,\ \mathbf{x}_3,\ \mathbf{x}_4$. Further, choose $(\lambda' : \mu') \in \mathbb{P}_\pi$ such that $\lambda\mu' - \lambda'\mu \not\equiv 0 \pmod{\pi}$. Write

$$C = T^t(\lambda A - \mu B)T,\ D = T^t(\lambda' A - \mu' B)T.$$

Then by Theorem B.0.4 we have that $v(\partial(C, D)) = v(\partial(A, B))$. Now note that

$$C \equiv \left(\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & C_1 \end{array}\right) \pmod{\pi^\gamma}$$

where C_1 is a 3×3 matrix with entries in $\mathcal{O}_v$. Also

$$D \equiv \left(\begin{array}{c|c} \mathbf{0} & \mathbf{v^t} \\ \hline \mathbf{v} & D_1 \end{array}\right) \pmod{\pi^\gamma}$$

where D_1 is a 3×3 matrix with entries in $\mathcal{O}_v$, and $\mathbf{v} \in {\mathcal{O}_v}^3$. It is now easily seen that the coefficients of X^4 and X^3Y in $G(X, Y) = \det(XC - YD)$ are congruent to 0 modulo π^γ. By considering the formula for the discriminant of G in terms of its coefficients, we see that $\pi^\gamma | \partial(C,\ D)$. This completes the proof.

□

5.3 Algorithm II: F has a rational root over $\mathcal{O}_v$

If $F(X,Y) = \det(XA - YB)$ has a root defined over $\mathcal{O}_v$, then by parametrizing a singular combination of our equation

$$\left.\begin{aligned} \mathbf{x}^t A\mathbf{x} = 0 \\ \mathbf{x}^t B\mathbf{x} = 0 \end{aligned}\right\} \tag{5.4}$$

we can reduce the testing of solubility of (5.4) over K_v, to the problem of testing solubility over K_v of an equation of the form

$$Y^2 = g(X), \tag{5.5}$$

where $g(X) \in K_v[X]$ has degree 4 and no repeated roots. In our algorithm here, instead of Hensel's Lemma we intend to use techniques such as finding the roots of polynomials over finite fields, which will give us an overall polynomial time complexity [2]. We shall restrict ourselves to the case where $v(2) = 0$, or equivalently where the residue field has odd characteristic.

5.3.1 Parametrizing the Singular Combination

As in the case over a number field (see page 76), to get to an equation of the form (5.5), it sufficient to find a single non-trivial solution of the singular combination. By a change of variable defined over K_v we may assume that our singular combination is of the form

$$aX^2 + bY^2 + cZ^2 = 0 \tag{5.6}$$

[2] The polynomials we wish to solve all have degree at most 4, and hence are soluble. Hence computing the roots is reduced to extracting pure roots of elements of finite fields. This problem is soluble in probablistic polynomial time, or alternatively in deterministic polynomial time assuming the Generalized Riemann Hypothesis (see [Cohen] pages 31-34 and page 37).

where $a,\ b,\ c \in \mathcal{O}_\upsilon$, and $\upsilon(a) = \upsilon(b) = 0$ and $\upsilon(c) = 0$ or 1. If $\upsilon(c) = 1$, then $-ab^{-1}$ must be a square in $\mathcal{O}_\upsilon$, otherwise (5.4) does not have a solution over K_υ and we may stop. So if $\alpha^2 = -ab^{-1}$ then $(1\ ,\alpha\ ,0)$ is a non-trivial solution to (5.6), and we are finished. If $\upsilon(c) = 0$, then heuristically, for 50% of pairs (x, y), $-c^{-1}(ax^2 + by^2)$ is a square in $\mathcal{O}_\upsilon$. Thus we assume that we can arrive at a solution in $O(1)$ steps.

5.3.2 Local Solubility Testing for $Y^2 = g(X)$

We recall that $g \in \mathcal{O}_\upsilon[X]$ has degree 4, and non-zero discriminant, and that the characterstic of the residue field k_υ is odd. We write q for the number of elements in the residue field k_υ. When considering solutions to

$$Y^2 = g(X)$$

we shall include those at infinity; thus this curve has a pair of points at infinity if and only if the leading coefficient of g is square in $\mathcal{O}_\upsilon$. If f is a polynomial in $\mathcal{O}_\upsilon[X]$, we write $\overline{f}$ for the image of f under the map $\mathcal{O}_\upsilon[X] \to k_\upsilon[X]$ induced by the natural map $\mathcal{O}_\upsilon \to k_\upsilon$. If $\deg f = 4$ but $\deg \overline{f} \leq 3$ we shall say that $\overline{f}$ has a root at infinity; if $\deg \overline{f} \leq 2$ we shall say that $\overline{f}$ has a multiple root at infinity. These conventions should be borne in mind in what follows.

Lemma 5.3.1 *Suppose the curve*

$$C : aY^2 = f(X) \tag{5.7}$$

is given with $f(X) \in \mathcal{O}_\upsilon[X]$, $a \in \mathcal{O}_\upsilon$. Let $x_1,\ y_1 \in \mathcal{O}_\upsilon$ such that

$$ay_1^2 \equiv f(x_1) \pmod{\pi}.$$

Then there exists $x,\ y \in \mathcal{O}_\upsilon$ with $x \equiv x_1,\ y \equiv y_1 \pmod{\pi}$, such that

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

except possibly when

$$ay_1 \equiv f'(x_1) \equiv 0 \pmod{\pi}.$$

Proof. The conclusion follows by applying Hensel's Lemma to the polynomial

$$G_1(X) = f(X) - ay_1^2$$

in the case $f'(x_1) \not\equiv 0 \pmod{\pi}$, and to the polynomial

$$G_2(Y) = aY^2 - f(x_1)$$

in the case $ay_1 \not\equiv 0 \pmod{\pi}$. A suitable version of Hensel's Lemma is given on page 49 of [Ca7]. □

Lemma 5.3.2 *Suppose that $f(X) \in \mathcal{O}_v[X]$ such that $\deg f = 4$ and $\deg \overline{f} = 3$ or 4. Suppose $\overline{f}(X)$ has no repeated factors. Then the equation*

$$Y^2 = f(X)$$

has solutions over K_v. [3]

Proof. Under the hypotheses of the Lemma, the equation

$$Y^2 = \overline{f}(X)$$

is a curve of genus 1 defined over k_v. It follows (see [Ca1] page 119) that it has at least one point defined over k_v. Again, since $\overline{f}$ does not have repeated factors, we can use Lemma 5.3.1 to show that this solution lifts to one defined over K_v. □

Lemma 5.3.3 *Suppose $f(X) \in \mathcal{O}_v[X]$ such that $1 \leq \deg \overline{f} \leq 4$. Suppose that $\overline{f} = \overline{g}^2\overline{h}$ where $\deg \overline{g} \geq 0$, $\deg \overline{h} \geq 1$ and $\overline{h}$ is a square-free polynomial. Then the equation*

$$Y^2 = f(X)$$

has solutions in K_v.

[3] There is nothing new here: see [Ca4] page 205 (where the term elliptic curve really means a curve of genus 1).

Proof. The curve $Y^2 = \overline{h}(X)$ has genus 0, and hence has $q+1$ points defined over k_v. Of these at most 2 are at infinity. Further, there is at most 1 root of g. If this root is x_0 say, then there are at most 2 points on $Y^2 = \overline{h}(X)$ whose x-coordinate is x_0. Hence if $q \geq 5$ then $Y^2 = \overline{h}(X)$ has at least one point $(x_1, y_1) \in k_v^2$ with $x_1 \not\equiv x_0$. Then the point $(x_1,\ y_1\overline{g}(x_1))$ lifts to a point on $Y^2 = f(X)$ by Lemma 5.3.1. For the case $q = 3$ the Lemma can be established by a lengthy but straightforward case-by-case check which we omit. □

The following theorems follow easily from the above Lemmas.

Theorem 5.3.1 *Suppose $\overline{f} \not\equiv 0$. If*

$$Y^2 = f(X) \tag{5.8}$$

has no points over K_v then

$$\overline{f} \equiv \alpha \overline{g}^2$$

where $\overline{g}(X) \in k_v[X]$ and $\alpha \in {k_v}^* \backslash {k_v}^{*2}$.

Proof. The only case that remains to be checked is that if $\overline{f} \not\equiv 0$ and $\overline{f} \equiv \overline{g}^2$ then (5.8) has a solution over K_v. For this it is sufficient to choose any x_0 such that $g(x_0) \not\equiv 0 \pmod{\pi}$, and then note that $(x_0, g(x_0))$ lifts by Lemma 5.3.1. □

Theorem 5.3.2 *Suppose $f(X) \in \mathcal{O}_v[X]$ such that $\overline{f} \not\equiv 0 \pmod{\pi}$, and $\deg \overline{f} \leq 4$. Then*

$$\pi Y^2 = f(X)$$

has a solution in K_v if $\overline{f}$ has a root defined k_v which is not a repeated root.

Algorithm 5.3.1 *Testing*

$$Y^2 = f(X) \tag{5.9}$$

for solubility over K_v, where $f(X) \in \mathcal{O}_v[X]$, $\deg f = 4$, and the discriminant of f is non-zero.

Step I *If $\overline{f} \equiv 0 \pmod{\pi}$, then go to Step II. Now suppose $\overline{f} \not\equiv 0$. Check if $\overline{f} = \overline{a}\overline{g}^2$ for some $\overline{g} \in k_v[X]$, and $\overline{a} \in k_v$. If this is not the case, or if $\overline{a} \in k_v^{*2}$*

then we have local solubility by the above theorems and we can stop. Hence we can assume that $\overline{f} = \overline{a}\overline{g}^2$, *and* $\overline{a} \notin k_v^{*2}$. *So any solution* $(X_0,\ Y_0) \in {\mathcal{O}_v}^2$ *to (5.9) must satisfy* $Y_0 \equiv 0$ *and* $\overline{g}(X_0) \equiv 0$. *Now* $\overline{g}$ *has at most two solutions* $\overline{\epsilon_1},\ \overline{\epsilon_2} \pmod{\pi}$*; if* $\overline{g}$ *has no solutions in* k_v *then (5.9) has no solutions in* $\mathcal{O}_v$ *and we can stop. Hence*

$$Y_0 = \pi Y_1 \textit{ and } X_0 = \pi X_1 + \epsilon_i$$

where $Y_1,\ X_1 \in \mathcal{O}_v$. *Choose* $a \in \mathcal{O}_v$ *and* $g \in \mathcal{O}_v[X]$ *such that the images of* a *and* g *under* $\mathcal{O}_v \to k_v$ *are* $\overline{a}$ *and* $\overline{g}$. *Then* $f = ag^2 + \pi h$ *where* h *has coefficients in* $\mathcal{O}_v$. *Since* $\pi^2 | Y_0^2 = f(X_0)$ *and* $\pi | g(X_0)$, *we get that* $\pi | h(X_0)$. *Hence if neither of* ϵ_1 *and* ϵ_2 *is a root of* $\overline{h}$ *then (5.9) is not soluble and we can stop. If say* ϵ_i *is a root of* $\overline{h}$ *then* π *divides the trailing coefficient of* $h(\pi X + \epsilon_i)$. *So we will get at most 2 equations of the form*

$$Y^2 = f_i(X)$$

where $f_i(X) = \frac{1}{\pi^2} f(\pi X + \epsilon_i) \in \mathcal{O}_v[X]$. *It is now necessary and sufficient that one of these should have solutions in* $\mathcal{O}_v$, *and we use Step I again with* f_i *instead of* f.

Step II *Here* f *is divisible by* π. *If* f *is divisible by* π^2 *then we can replace* f *by* $\frac{1}{\pi^2} f$ *and go to Step I. So suppose that* $f_1 = \frac{1}{\pi} f \not\equiv 0 \pmod{\pi}$. *We see that we want to determine if*

$$\pi Y_1^2 = f_1(X)$$

has solutions in $\mathcal{O}_v$. *If* f_1 *has no roots in* k_v *then (5.9) is not soluble and we can stop. If* f_1 *has a root which is not a repeated root then (5.9) is soluble and we can stop. Suppose that* f_1 *has repeated roots* ϵ_i *where* $i = 1$, *or* $i = 1,\ 2$. *Then it is necessary and sufficient to determine if either of*

$$Y_1^2 = \frac{1}{\pi} f_1(\pi X_1 + \epsilon_i)$$

is soluble, and $\frac{1}{\pi} f_1(\pi X_1 + \epsilon_i) \in \mathcal{O}_v[X]$. *So we use Step I again.*

Lemma 5.3.4 *Suppose $r = \upsilon(\partial g)$ where ∂g is the discriminant of g. In the above algorithm, if we are still undecided after $r+1$ steps, then the equation (5.9) has a solution defined over K_v and we can stop.*

Proof. It is clear that after r steps, we may write down a $Z \in \mathcal{O}_v$, such that $f(Z) \equiv \pi^{2(r+1)}$. By [Ca7] page 52, f has a root in $\mathcal{O}_v$. This immediately implies that (5.9) has a solution defined over K_v. $\square$

Chapter 6

Local Solubility II: Over Archimedean Completions

6.1 Introduction

Let K be a number field, and A, B be $n \times n$ symmetric matrices with entries in O_K, the ring of integers of K. Suppose further that $F(X,Y) = \det(XA - YB)$ is non-zero and does not have any repeated roots. We want to determine the local solubility of

$$\left.\begin{aligned} \mathbf{x}^t A\mathbf{x} = 0 \\ \mathbf{x}^t B\mathbf{x} = 0 \end{aligned}\right\} \tag{6.1}$$

over all completions of K isomorphic to $\mathbb{R}$. If $\sigma_1, \ldots, \sigma_n : K \hookrightarrow \mathbb{R}$ are the real embeddings of K, then this is equivalent to determining if, for each i,

$$\left.\begin{aligned} \mathbf{x}^t A^{\sigma_i}\mathbf{x} = 0 \\ \mathbf{x}^t B^{\sigma_i}\mathbf{x} = 0 \end{aligned}\right\} \tag{6.2}$$

has a non-trivial solution over $\mathbb{R}$.

Without loss of generality, we will assume for the rest of this chapter that $\sigma K = K \subseteq \mathbb{R}$, and hence that A, B are $n \times n$ real matrices.

Further, as $\det(XA - YB)$ is non-zero, by taking appropriate linear combinations of A and B (if necessary), we can assume that $\det A$ and $\det B$ are non-zero. Hence $F(\lambda) = \det(A - \lambda B)$ is a real polynomial of degree n with distinct roots.

6.2 Reducing to Totally Real $F(\lambda)$

The following lemma of Swinnerton-Dyer allows us to get a better grip on the problem.

Lemma 6.2.1 *(Swinnerton-Dyer) Let $f,\ g$ be homogeneous real quadratic forms; the manifold $f = g = 0$ contains non-zero real points if and only if the quadratic form $\lambda f - \mu g$ is not definite for all real λ, μ.*

Proof. This is part of Lemma 1 of [SwD1]. □

We are now ready for a simplification:

Lemma 6.2.2 *Suppose that $F(\lambda) = \det(A - \lambda B)$ has a non-real root. Then*

$$\left.\begin{aligned} \mathbf{x}^t A \mathbf{x} = 0 \\ \mathbf{x}^t B \mathbf{x} = 0 \end{aligned}\right\} \tag{6.3}$$

has a non-trivial solution over $\mathbb{R}$.

Proof. Recall first our assumption above that $\det A$ and $\det B$ are non-zero.

Suppose for a contradiction that the pair of equations (6.3) has no non-trivial real solutions. By Lemma (6.2.1) above, there exists a real linear combination of A and B which is a positive definite matrix. Without loss of generality, we may suppose that this is B, and that $\det A \neq 0$. From Linear Algebra we know that there exists a non-singular real matrix P such that $P^t B P = I$, the identity $n \times n$ matrix. Note that $P^t A P$ is a real symmetric matrix, and hence must have only real eigenvalues. Hence the solutions to $\det((P^t A P) - \lambda I)$, or equivalently those of $\det(A - \lambda B)$ are all real. This gives our desired contradiction. □

6.3 Results on the Totally Real Case

By Lemma (6.2.2), we may restrict our attention to the case where $F(X,Y) = \det(XA - YB)$ has n real roots. Hence by the next Lemma, the two matrices A, B are simultaneously diagonalizable over $\mathbb{R}$. Naturally, it is much easier to ask if there is a definite linear combination of two matrices when they are diagonal.

Lemma 6.3.1 *Suppose that* $\det A, \det B$ *are non-zero, and that* $\det(A - YB)$ *is a polynomial of degree n, which has n real roots* $\lambda_1, \ldots, \lambda_n$ *say. Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be non-trivial vectors in* $\mathbb{R}^{\ltimes}$ *such that*

$$(A - \lambda_i B)\mathbf{x}_i = \mathbf{0}. \tag{6.4}$$

Let $P = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$*, the* $n \times n$ *matrix with the* $\mathbf{x}_i$ *as its columns. Then* $P \in GL_n(\mathbb{R})$ *and*

$$P^t A P = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \quad , \quad P^t B P = \begin{pmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_n \end{pmatrix} \tag{6.5}$$

where $\alpha_i = \lambda_i {\mathbf{x}_i}^t B \mathbf{x}_i$, $\beta_i = {\mathbf{x}_i}^t B \mathbf{x}_i$.

Proof. This is straightforward (cf Lemma 4.4.1 and Theorem 4.4.1). □

Lemma 6.3.2 *Under the hypotheses and notation of Lemma (6.3.1), the pair of equations*

$$\left.\begin{aligned} \mathbf{x}^t A \mathbf{x} = 0 \\ \mathbf{x}^t B \mathbf{x} = 0 \end{aligned}\right\} \tag{6.6}$$

has a non-trivial real solution if and only if there do not exist real λ^*, μ^* *(not both zero) such that the real numbers* $\mu^*\alpha_i - \lambda^*\beta_i$ *all have the same sign.*

Proof. This is immediate from Lemmas (6.2.1), (6.3.1). □

Lemma 6.3.3 *Under the hypotheses and notation of Lemma (6.3.1), the pair of equations*

$$\left.\begin{aligned} \mathbf{x}^t A \mathbf{x} = 0 \\ \mathbf{x}^t B \mathbf{x} = 0 \end{aligned}\right\} \tag{6.7}$$

has no non-trivial real solution if and only if there exists λ_j, one of the roots of $F(\lambda) = det(A - \lambda B)$, such that $A - \lambda_j B$ is semi-definite.

Proof. Suppose first that the pair of equations (6.7) has no non-trivial real solution. By Lemma (6.3.3) above, there exist real λ^*, μ^* such that $\mu^* \alpha_i - \lambda^* \beta_i$ all have the same sign. If $\mu^* = 0$ then we can replace it by a very small non-zero real number and still have that $\mu^* \alpha_i - \lambda^* \beta_i$ all have the same sign. Hence, we will assume that $\mu \neq 0$. By dividing by μ^*, we see that there is a real λ^{**} such that $\alpha_i - \lambda^{**} \beta_i$ all have the same sign. Let λ_j be the root of $F(\lambda)$ which is closest to λ^{**}.We note that as we vary λ along the real line, none of the $\alpha_i - \lambda \beta_i$ change sign until we cross a root of $\prod(\alpha_i - \lambda \beta_i) = F(\lambda)$. Since λ_j is the closest root of $F(\lambda)$ to λ^{**}, it follows that $\alpha_i - \lambda_j \beta_i$ $i \neq j$ all have the same sign and that, of course, $\alpha_j - \lambda_j \beta_j = 0$. Hence $A - \lambda_j B$ is semi-definite, as required.

Conversely, suppose that $A - \lambda_j B$ is semi-definite, where λ_j is a root of $F(\lambda)$. Write

$$A = \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} \quad , \quad B = \begin{pmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_n \end{pmatrix} \tag{6.8}$$

as in Lemma (6.3.1). Recall that the alphas and betas are all non-zero, since by assumption $\det A, \det B \neq 0$. $A - \lambda_j B$ is semi-definite and so all the $\alpha_i - \lambda_j \beta_i$ are all of the same sign except $\alpha_j - \lambda_j \beta_j = 0$. Note $\alpha_j - (\lambda_j + \epsilon)\beta_j = -\epsilon \beta_j$; hence, since $\beta_j \neq 0$, by choosing ϵ small enough and with appropriate sign, we will have that $\alpha_i - (\lambda_j + \epsilon)\beta_i$ are all of the same sign. Hence $A - (\lambda_j + \epsilon)B$ is definite and the lemma follows. □

Theorem 6.3.1 *Under the notation and hypotheses of Lemma (6.3.1),the pair of equations*

$$\left.\begin{aligned} \mathbf{x}^t A \mathbf{x} = 0 \\ \mathbf{x}^t B \mathbf{x} = 0 \end{aligned}\right\} \tag{6.9}$$

has a non-trivial solution in $\mathbb{R}$ if and only if, for each λ_j, the real numbers $\alpha_i - \lambda_j \beta_i$ $(i \neq j)$ do not all have the same sign.

Proof. Immediate from Lemma (6.3.3). □

6.4 The Algorithm

We can now present our algorithm for determining the solubility of the pair

$$\left.\begin{aligned}\mathbf{x}^t A\mathbf{x} = 0\\ \mathbf{x}^t B\mathbf{x} = 0\end{aligned}\right\} \qquad (6.10)$$

over $\mathbb{R}$, under the assumption that $\det(XA - YB)$ is non-zero, and has distinct roots.

Algorithm 6.4.1 *Since* $\det(XA - YB)$ *is non-zero, we can choose K-linear combinations of A and B which are non-singular and hence assume that* $\det(A)$ *and* $\det(B)$ *are non-zero. Now check if all the roots of* $\det(A - \lambda B)$ *are real. If there is a non-real root then the pair (6.10) has a non-trivial solution over* $\mathbb{R}$ *and we can stop. Otherwise let* $\lambda_1, \ldots, \lambda_n$ *be the roots of* $F(\lambda) = \det(A - \lambda B)$ *and let* $\mathbf{x}_i \in K(\lambda_i)^n - \{\mathbf{0}\}$ *be solutions to* $(A - \lambda_i B)\mathbf{x}_i = \mathbf{0}$ *(this is simply solving linear equations). Define* $\beta_i = \mathbf{x}_i{}^t B\mathbf{x}_i$ *and* $\alpha_i = \lambda_i \mathbf{x}_i{}^t B\mathbf{x}_i$*. It is clear that the* λ_i, β_i *and* α_i *can be calculated to arbitrary accuracy in* $\mathbb{R}$*. Check if, for each* $1 \leq j \leq n$*, the numbers* $\alpha_i - \lambda_j\beta_i$ $(1 \leq i \leq n, i \neq j)$ *do not all have the same sign. These are non-zero, and so it is easy to determine their signs. If for some* j*,* $\alpha_i - \lambda_j\beta_i$ $i \neq j$ *have the same sign then (6.10) has only the trivial solution. Otherwise, it has a non-trivial solution.*

6.5 A Special Case for Two Quadrics in Four Variables

We record in passing the following theorem, which says that if $n = 4$, $K = \mathbb{Q}$ and $F(X, Y) = \det(XA - YB)$ has 4 distinct roots, all in $\mathbb{Q}$, then solubility over $\mathbb{R}$ is guaranteed once solubility over $\mathbb{Q}_p$ has been checked for all (finite) primes p.

Theorem 6.5.1 *Suppose that A, B are symmetric* 4×4 *matrices with rational entries such that* $F(X, Y) = \det(XA - YB)$ *can be factorized completely over*

$\mathbb{Q}$. *Then, if*

$$\left.\begin{aligned}\mathbf{x}^t A\mathbf{x} = 0\\ \mathbf{x}^t B\mathbf{x} = 0\end{aligned}\right\} \tag{6.11}$$

has non-trivial solutions in $\mathbb{Q}_p$ *for all finite primes p, it has a non-trivial solution in* $\mathbb{R}$.

Proof. Suppose (6.11) has non-trivial solutions in $\mathbb{Q}_p$ for all finite primes p, and it has no non-trivial solution in $\mathbb{R}$. As before, we may assume that A and B are non-singular. Hence by Lemma (6.3.3), there exists λ_j, a root of $\det(A - \lambda B)$ such that $A - \lambda_j B$ is semi-definite. Now after a simultaneous diagonalization over $\mathbb{Q}$ (see Theorem 4.4.1), we can assume that

$$\left.\begin{array}{lcccccccc}\mathbf{x}^t A\mathbf{x} & = & \alpha_1 x_1^2 & + & \alpha_2 x_2^2 & + & \alpha_3 x_3^2 & + & \alpha_4 x_4^2\\ \mathbf{x}^t(A - \lambda_j B)\mathbf{x} & = & \gamma_1 x_1^2 & + & \gamma_2 x_2^2 & + & \gamma_3 x_3^2 & & \end{array}\right\} \tag{6.12}$$

where $\gamma_1, \gamma_2, \gamma_3$ have the same sign. Recall that this has non-trivial solutions over $\mathbb{Q}_p$, for all finite primes p. We will show that the second equation, as an equation in 3 variables, has a non-trivial solutions over all completions $\mathbb{Q}_p$ $(p \neq \infty)$. If $(x_1, x_2, x_3, x_4) \in {\mathbb{Q}_p}^4 - \{\mathbf{0}\}$ and solves the pair (6.12) then we cannot have $x_1 = x_2 = x_3 = 0$. For otherwise $\alpha_4 {x_4}^2 = 0$ in $\mathbb{Q}_p$ and so either $\alpha_4 = 0$ or $x_4 = 0$. Hence either $\det(A) = 0$ or $(x_1, x_2, x_3, x_4) = \mathbf{0}$. This contradiction shows that the equation

$$\gamma_1 x_1^2 + \gamma_2 x_2^2 + \gamma_3 x_3^2 = 0 \tag{6.13}$$

is solvable at all the finite primes. By the well known lemma below, this must also have a solution in the reals. This contradicts the fact that $\gamma_1, \gamma_2, \gamma_3$ share the same sign, and so we are finished. □

Lemma 6.5.1 *for any conic over* $\mathbb{Q}$, *the number of primes p (including* ∞*) for which there is not a point over* $\mathbb{Q}_p$ *is even.*

Proof. See [Ca2] page 46. □

Appendix A

Hensel Lifting for $Y^2 = g(X)$

Let K be a field complete with respect to a non-archimedean valuation υ, such that the corresponding residue field is finite. Let $\mathcal{O}$ be the ring of valuation integers, and let π be a uniformizer for υ. Suppose $g(X)$ is a non-zero polynomial with coefficients in $\mathcal{O}$, and has non-zero discriminant. In this Appendix we consider the following problem: Given $x_0 \in \mathcal{O}$ and $\epsilon \geq 0$, does there exist $x \in \mathcal{O}$ such that $g(x)$ is a square in $\mathcal{O}$ and $\upsilon(x - x_0) \geq \epsilon$?

The question arises from our method of computing the non-archimedean contribution to the upper bound for $h - \hat{h}$ in Chapter 2. This question is considered in [Bi, SwD] and [Cre] for the case where $K = \mathbb{Q}_p$ for some prime p, and g is a polynomial of degree 4, though the details for our general case are not any more difficult. The following Lemma, is a direct generalization of Lemmas 6 and 7 of [Bi, SwD], and the details of their proof carry over without any changes, and thus a proof is omitted.

Lemma A.0.2 *Suppose $\upsilon(2) = e \geq 0$. Suppose that $x_0 \in \mathcal{O}$, and let*

$$\upsilon(g(x_0)) = \lambda, \ \upsilon(g'(x_0)) = \mu.$$

Then there exists $x \in \mathcal{O}$, with $g(x)$ a square in $\mathcal{O}$, and $\upsilon(x - x_0) \geq \epsilon$ if

1. *$g(x_0)$ is a square in $\mathcal{O}$, or*

2. *$\lambda - \mu \geq \epsilon > \mu$, or*

3. *λ is even, and there exists i such that $1 \leq i \leq 2e$, $\lambda = \mu + \epsilon - i$, $\epsilon > \mu$, and $\pi^{-\lambda} g(x_0) \equiv 1 \pmod{\pi^i}$.*

There may exist $x \in \mathcal{O}$, with $g(x)$ a square in $\mathcal{O}$, and $\upsilon(x - x_0) \geq \epsilon$ if

1. *$\mu \geq \epsilon$ and $\lambda \geq 2\epsilon$, or*

2. *$\mu \geq \epsilon$ and $\lambda = 2\epsilon - 2i$ where $1 \leq i \leq e$.*

There does not exist $x \in \mathcal{O}$, with $g(x)$ a square in $\mathcal{O}$, and $\upsilon(x - x_0) \geq \epsilon$ in any other case.

Now suppose that $x_0 \in \mathcal{O}$, and that we want to know if there exists $x \in \mathcal{O}$ such that $g(x)$ is a square in $\mathcal{O}$ and $\upsilon(x - x_0)\epsilon$. If we use the above Lemma we will be able to decide this question unless $\mu_0 \geq \epsilon$, and $\lambda_0 \geq 2\epsilon - 2e$, where $\lambda_0 = \upsilon(g(x_0))$, and $\mu = \upsilon(g'(x_0))$. Suppose that this is the case. Let $\alpha_1, \ldots, \alpha_q$ be a complete set of coset representatives for $\mathcal{O}/\pi\mathcal{O}$. Now it is sufficient to determine, if for some j, there exists $x \in \mathcal{O}$ such that $g(x)$ is a square in $\mathcal{O}$ and $\upsilon(x - x_1) \geq \epsilon + 1$ where $x_1 = x_0 + \alpha_j \pi^\epsilon$. For any fixed j, we use the above Lemma. If we are still undecided, then $\mu_1 \geq \epsilon + 1$, and $\lambda_1 \geq 2(\epsilon + 1) - 2e$, where $\lambda_1 = \upsilon(g(x_1))$ and $\mu_1 = \upsilon(g'(x_1))$. Continuing recursively in the obvious manner, if our question remains undecided forever, then we will have constructed a sequence $(x_k) \in \mathcal{O}$ $(k = 0,\ 1, \ldots)$ such that $\upsilon(g(x_k)) \geq \epsilon + k$, and $\upsilon(g'(x_k)) \geq 2(\epsilon + k) - 2e$. Since K is complete, and the discriminant of g is non-zero, we arrive at a contradiction.

Appendix B

The Geometry of the Intersection of two Quadrics

Suppose A and B are linearly independent 4×4 matrices with entries in a ground field K, which has characteristic $\neq 2$. We shall say that two distinct quadric surfaces (in $\mathbb{P}^3$)

$$H_1 : \mathbf{x}^t A \mathbf{x} = 0$$

and

$$H_2 : \mathbf{x}^t B \mathbf{x} = 0$$

intersect transversely if $F(X, Y) = \det(XA - YB)$ is non-zero and has distinct roots. We assume throughout that our ground field K has characteristic $\neq 2$.

Theorem B.0.2 *The intersection of the 2 distinct quadric surfaces H_1, and H_2 is an (irreducible) curve of genus 1 if and only if it is transverse. Moreover, if the intersection is not transverse, then it has a (Zariski) component which is a curve of genus 0.*

Proof. Suppose that H_1, and H_2 intersect transversely. Then by Proposition 22.38 on page 304 of [Harris], the intersection

$$\left.\begin{array}{r}\mathbf{x}^t A\mathbf{x} = 0\\ \mathbf{x}^t B\mathbf{x} = 0\end{array}\right\} \tag{B.1}$$

is a curve of genus 1. The rest of the theorem follows from the two Lemmas below. □

Lemma B.0.3 *If* $\det(XA - YB) = 0$ *identically, then (B.1) must have a Zariski component which is a curve of genus* 0*.*

Proof. Let y be transcendental over K. Now $\det(A - yB) = 0$ and so there exists a vector $\mathbf{v}(y) \in K[y]^4\backslash\{\mathbf{0}\}$ such that

$$(A - yB)\mathbf{v}(y) = \mathbf{0}.$$

Moreover, we may assume that the elements of $\mathbf{v}(y)$ are coprime as polynomials in y. Now let $y_1,\ y_2$ be independent transcendental elements over K. By the usual argument (cf. Lemma 4.4.1) $y_1 \neq y_2$ implies that

$$\mathbf{v}(y_1)^t A\mathbf{v}(y_2) = \mathbf{v}(y_1)^t B\mathbf{v}(y_2) = 0.$$

Now substitute $y_1 = y_2 = y$ in the above. So we have

$$\mathbf{v}(y)^t A\mathbf{v}(y) = \mathbf{v}(y)^t B\mathbf{v}(y) = 0.$$

If $\mathbf{v}(y)$ is a non-constant vector then the conclusion follows. Suppose that $\mathbf{v}(y) = \mathbf{v} \in K^4\backslash\{\mathbf{0}\}$. So $(A - yB)\mathbf{v} = \mathbf{0}$, which implies that $A\mathbf{v} = B\mathbf{v} = \mathbf{0}$. In this case it is easy to show now that the intersection $H_1 \cap H_2$ is a collection of straight lines (in $\mathbb{P}^3$). □

Lemma B.0.4 *Suppose* $\det(XA{-}YB) \neq 0$ *but it has a multiple root. Then (B.1) has a component which is a curve of genus* 0*.*

Proof. Without loss of generality we may assume that Y^2 divides $\det(XA - YB)$. Clearly the rank of A is at most 3. If A has rank 1 or 2, then let

$L(x_1, \ldots, x_4)$ be a linear form dividing $\mathbf{x}^t A\mathbf{x}$. Then the intersection of $L(x_1, \ldots, x_4) = 0$ and $\mathbf{x}^t B\mathbf{x} = 0$ is either a pair of lines or a plane conic. In either case the conclusion of the lemma follows.

Hence suppose that the rank of A is 3. By a non-singular change of variable we may assume that

$$A = \left(\begin{array}{c|c} A_1 & \mathbf{0} \\ \hline \mathbf{0} & 0 \end{array}\right).$$

Write

$$\mathbf{x}^t B\mathbf{x} = q(x_1, x_2, x_3) + x_4 l(x_1, x_2, x_3) + bx_4^2$$

where q is quadratic, l is linear and b is constant. It follows that the coefficient of X^3Y in $\det(XA - YB)$ is $b \det(A_1)$. Since $Y^2 | \det(XA - YB)$ we have that $b = 0$. So (B.1) is birational to the conic

$$(x_1, x_2, x_3)A_1(x_1, x_2, x_3)^t = 0$$

via the map

$$x_4 = \frac{-q(x_1, x_2, x_3)}{l(x_1, x_2, x_3)}.$$

□

This completes the proof.

Theorem B.0.3 *Suppose that the characteristic of K is 0. Suppose that C is a curve of genus 1, and D is the intersection of 2 distinct quadric surfaces H_1 and H_2 (in $\mathbb{P}^3$) as above. If there is a non-constant morphism $\phi : D \to C$ then the intersection is transverse and D is a curve of genus 1.*

Proof. If the intersection is not transverse, then D contains a component D_1 which is a curve of genus 0. Restricting ϕ to D_1 we see that we have a non-constant morphism from a curve of genus 0 into a curve of genus 1. This is impossible (see exercise 2.8 on page 43 of [Si2]). □

B.0.1 The Combinant

Define $\partial(A, B) = \mathrm{disc}(\det(XA - YB))$. We need(ed) the following theorem.

Theorem B.0.4 *If $a,\ b,\ c,\ d$ are elements of K and P is a 4×4 matrix then write $C = P^t(aA - bB)P$ and $D = P^t(cA - dB)P$. We have*

$$\partial(C,\ D) = (ad - bc)^{12}(\det(P))^{12}\partial(A,\ B).$$

Proof. See [Bi, Le, Mu] page 112. □

Bibliography

[Bi, SwD] B.J. Birch and H.P.F. Swinnerton-Dyer, *Notes on Elliptic Curves I*, J. Reine Angew. Math. **212** (1963), 7-25.

[Bi, Le, Mu] B.J. Birch, D.J. Lewis, T. G. Murphy, *Simultaneous Quadratic Forms*, Am .J. Math. 84 (1962), 110-112.

[BGZ] J.P. Buhler, B.H. Gross, and D.B. Zagier, *On the Conjecture of Birch and Swinnerton-Dyer for an elliptic curve of rank 3*, math. Comp. **44** (1985), 473–481.

[Brem] A. Bremner, *On the Equation* $Y^2 = X(X^2+p)$, in “Number Theory and Applications” (R.A. Mollin, ed.), Kluwer, Dordrecht, 1989, 3–23.

[Brem, Bue] A. Bremner and D. Buell, *Three Points of Great Height on Elliptic Curves*, Math.Comp. **61** (1993), 111-115.

[Brem, Ca] A. Bremner and J.W.S. Cassels, *On the Equation* $Y^2 = X(X^2+p)$, Math. Comp. **42** (1984), 257–264.

[Brum, Kra] A. Brumer and K. Kramer, *The Rank of Elliptic Curves*, Duke Math. J., **44** (1977), 715–743.

[Buch] J. Buchmann, *A Subexponential Algorithm for the Determination of Class Groups and Regulators of Algebraic Number Fields*, Séminaire de théorie des nombres, Paris (1988-1989), 28–41.

[Bu, Len] J. Buchmann and H.W. Lenstra Jnr, *Approximating Rings of Integers in Number Fields*, To Appear.

[Ca1] J.W.S. Cassels, *Lectures on Elliptic Curves*, LMS Student Texts, Cambridge University Press, 1991.

[Ca2] J.W.S. Cassels, *Rational Quadratic Forms*, LMS Monographs, Academic Press (London), 1978.

[Ca3] J.W.S. Cassels, *Introduction to the Geometry of Numbers*, Springer-Verlag, 1959.

[Ca4] J.W.S. Cassels, *Survey Article: Diophantine Equations with Special Reference to Elliptic Curves*, J.L.M.S. *41* (1966), 193-291.

[Ca5] J.W.S. Cassels, *On a theorem of Dem'Janenko*, J. London Math. Soc. **43** (Part I) (1968), 67-70.

[Ca6] J.W.S. Cassels, *The Mordell-Weil Group of Curves of Genus 2*, .

[Ca7] J.W.S. Cassels, *Local Fields*, LMS Student Texts, Cambridge University Press, 1986.

[Cohen] H. Cohen, *A Course in Computational Algebraic Number Theory*, GTM 138, Springer-Verlag, 1993.

[Cre] J.E. Cremona, *Algorithms for Modular Elliptic Curves*, Cambridge University Press, 1992.

[Dem] V. A. Dem'Janenko, *An estimate of the remainder term in Tate's formula*, Mat. Zametki **3** (1968), 271–278. (Russian)

[Elkies] N. D. Elkies, *On* $A^4 + B^4 + C^4 = D^4$, Math. Comp. **51** (1988), 825-835.

[Ge, Zi] J. Gebel and H.G. Zimmer, *Computing the Mordell-Weil group of an elliptic curve over* Q, in "Elliptic Curves and related Topics", ed. H.Kisilevsky and M. Ram Murty, CRM Proceedings and Lecture Notes Volume 4, American Mathematical Society 1994.

[GPZ] J. Gebel, A. Pethő, and H.G. Zimmer, *Computing integral points on elliptic curves*, To Appear Acta. Arith.

[Gre] M. J. Greenberg, *Lectures on Forms in Many Variable*, W. A. Benjamin, 1969.

[Guy] R. K. Guy, *Unsolved Problems in Number Theory*, Springer-Verlag, 1981.

[Harris] J. Harris, *Algebraic Geometry*, GTM 133, Springer-Verlag, 1992.

[Hart] R. Hartshorne, *Algebraic Geometry*, GTM 52, Springer-Verlag, 1977.

[Holzer] L. Holzer, *Minimal solutions of diophantine equations*, Can. J. Math. **11** (1950), 238-244.

[Hus] D. Husemoller, *Elliptic Curves*, Springer-Verlag, 1987.

[Kob] N. Koblitz, *Introduction to Elliptic Curves and Modular Forms*, GTM 97, Springer-Verlag, 1984.

[Kret] T.J. Kretschmer, *Construction of Elliptic Curves with Large Rank*, Math. Comp. **46** (1986), 627-635.

[Mazur] B. Mazur, *On the Passage from Local to Global in Number Theory*, Bulletin (new series) of the AMS **29** (1993), 14-50.

[Mestre] J-F Mestre, *Construction d'une courbe elliptique de rang* ≥ 12, C.R. Acad. Sc. Paris **295** (1982), 643-644.

[Mord] L. J. Mordell, *Diophantine Equations*, Academic Press, 1969.

[Pari] C. Batut, D. Bernardi, H. Cohen, M. Olivier, *User's Guide to PARI-GP* (version 1.39), 1995.

[Pr] H.A. Priestley, *Introduction to Complex Analysis*, Oxford University Press, 1985.

[Scha] E.F. Schaefer, *2-descent on the Jacobians of Hyperelliptic Curves*, To Appear: Journal of Number Theory.

[Serf] P. Serf, *Congruent Numbers and Elliptic Curves*, in “Computational Number Theory”, ed. A. Pethő, M. Pohst, H. Williams, H.Zimmer , Walter de Gruyter, 1991.

[Sieg1] C. L. Siegel, *Lectures on the Geometry of Numbers*, Springer-Verlag, 1988.

[Sieg2] C. L. Siegel, *Normen Algebraischer Zahlen*, Nachr. Akad. Wiss Gottingen, Math.-Phys. Kl.

[Si1] J.H. Silverman, *The Difference between the Weil Height and the Canonical Height on Elliptic Curves*, Math. Comp. **55** (1990), 723-743.

[Si2] J.H. Silverman, *The Arithmetic of Elliptic Curves*, GTM 106, Springer-Verlag, 1986.

[Si3] J.H. Silverman, *Computing heights on Elliptic Curves*,Math. Comp. **51** (1988), 339-358.

[Sik] S. Siksek, *Infinite Descents on Elliptic Curves*, to appear in Rocky Mountain Journal.

[Si, Sm] S. Siksek and N. P. Smart, *On the Complexity of Computing the 2-Selmer Group of and Elliptic Curve*, submitted to the Journal of Number Theory.

[Smart] N. P. Smart, *S-integral points on elliptic curves*, Proc. Camb. Phil. Soc. **116** (1994), 391–399.

[Sm, Ste] N.P. Smart and N.M. Stephens, *Integral points on Elliptic Curves over Number Fields*, (to appear).

[Str, We] R.J. Stroeker and B. M. M. de Weger, *On a quartic diophantine equation*, Report 9371/B, Econometric Institute, Erasmus University Rotterdam, 1993.

[Str, Top] R.J. Stroeker and J. Top, *On the Equation* $Y^2 = (X+p)(X^2+p^2)$, Rocky Mountain J. Math. (1994).

[Str, Tz] R.J. Stroeker and N. Tzanakis, *Solving elliptic diophantine equations by estimating linear forms in elliptic logarithms*, Acta. Arith. **67** (1994), 177–196.

[SwD1] H. P. F. Swinnerton-Dyer, *Rational Zeros of two Quadratic Forms*, Acta. Arith. **9** (1964)

[SwD2] H. P. F. Swinnerton-Dyer, $A^4+B^4 = C^4+D^4$ *revisited*, J. London Math. Soc. **43** (Part I) (1968), 149-151.

[Thiel] C. Thiel, *Under the assumption of the Generalized Riemann Hypothesis verifying the class number belongs to* $\mathcal{NP} \cap$ *co*-$\mathcal{NP}$, In *ANTS-1 : Algorithmic Number Theory*, Eds L.M. Adleman and M-D. Huang, Springer-Verlag (Berlin), Lecture Notes In Computer Science No. 877, (1994),234–247.

[Trop] A.M. Tropper, *Linear Algebra*, Van Nostrand Reinhold, 1969.

[Tz] N. Tzanakis, to appear.

[Zim] H. G. Zimmer, *On the difference of the Weil Height and the Neron-Tate height*, Math. Z. **147** (1976), 35–51.