Université catholique de Louvain
Faculté des Sciences

MATH2MA
Master of Science in Mathematics

Continued Fractions and their relation to Integral Binary Quadratic Forms

Author:
Jim Barthel

Advisor :
Jean-Pierre Tignol

Reading Committee :
Pierre-Emmanuel Caprace
Luc Haine

2017-2018

Abstract

Continued fractions and integral binary quadratic forms belong to the classical objects of number theory. At first, both theories have been studied independently. Short time later, mathematicians discovered their natural connection. Unfortunately, there is not a unique way to combine both theories and neither is there a general consent about which approach should be considered. Most of the time, the approaches are chosen depending on the problem to solve.

The aim of this Master thesis is first to present one such approach in the modern mathematical language by giving formal definitions and second to highlight that there is indeed a natural connection between integral binary quadratic forms and continued fractions. This fact will be used to prove some classical results of continued fractions. Finally, those results lead to a characterization of irrational numbers with purely periodic continued fractions. By the natural connection of both domains, the given non-standard approach develops simultaneously some well-known properties of integral binary quadratic forms. In particular the equivalence of integral binary quadratic forms will be studied and cycles of strongly reduced forms will be discovered. Ultimately, all the main results will be combined to deduce in a first instance when a given integral binary quadratic form and its negative are equivalent and to obtain in a second step a well-justified algorithm to compute the class number for a given positive non-square discriminant.
Key words: Continued fractions, integral binary quadratic forms, purely periodic, equivalence, strongly reduced, class number.

Word of thanks

After an intensive period of almost ten months, today is finally the day when I'm finishing my Master thesis by writing this note of thanks.

In the first place, I would like to thank my advisor Professor Jean-Pierre Tignol from the Université catholique de Louvain (UCL). Professor Tignol's patient guidance and constant encouragement to continue the study of the theory given in this Master thesis pushed me to investigate the theory beyond the expected point. In particular, I would also like to thank him for all the advices, constructive comments, and critics he has provided throughout my time as his student. Furthermore, I immensely appreciated that he agreed to change the subject of my Master thesis when I found several inconsistencies in the main sources and that he strengthened my suspicions. Moreover, I acknowledge every single minor correction he suggested for this written report and his instructions on the general shape of a scientific document. Last but not least, I am extremely grateful that he was so dedicated to my project and managed to devote many insightful hours for my queries. I have been extremely lucky to have a world-class supervisor whose office was always open whenever I ran into a trouble spot, who cared so much about my work and who responded to my questions so promptly.

Besides my supervisor, I would like to thank the rest of my Master thesis committee Professor Pierre-Emmanuel Caprace and Professor Luc Haine for reading and evaluating this report and for all the questions and comments they might provide.

I must also express my very profound gratitude to my parents, family and my girlfriend's family for providing me with their unfailing support, continuous encouragement and financial resources throughout all my years of study and through the process of researching and writing this thesis.

Finally, I would like to thank my girlfriend for all the forgiven hours that I invested in this Master thesis, my study of Mathematics, my time-outs to read scientific papers in general and for her powerful mental support throughout my studies.

This accomplishment would not have been possible without all of you. Thank you.

Contents

1 Introduction

Although they have been known a long time ago, it was only in 1695 that John Wallis (1616-1703) published in his book *Opera Mathematica* the theoretical framework of so called continued fractions. Those number theoretical tools convinced the mathematicians not only because of their own surprising properties but also because of their properties to approximate real numbers, to solve a specific Diophantine equation and to study integral binary quadratic forms.

Integral binary quadratic forms are special polynomials in two variables. They have been extensively studied by Gauss (1777-1855) and Lagrange (1736-1813). As a result of their connection to quadratic fields, integral binary quadratic forms are still of great interest in modern mathematics. In particular integral binary quadratic forms were used in the study of lattices of quadratic fields, the ideals of quadratic fields and zeta functions of quadratic fields.

Both theories described previously seem to be independent from each other and indeed, as the classical approach shows, they can be studied independently. However, as already the founders and pioneers of those domains guessed, they can be studied simultaneously. This combination of both theories allows to state and prove several results in a natural and elegant manner. However, since there are several different approaches to continued fractions, as well as to quadratic forms, this combination has to be well chosen to work out fine.

This Master thesis is split into three parts which are subdivided into several chapters.

The first part defines continued fraction in an algorithmic way which shows that there is a firm difference between finite and infinite continued fractions. Then it goes on to prove that each continued fraction is identified with a unique real number which is summarized in the final conclusion given by theorem 3.27. However the main work is hidden in proposition 3.19 where the connection of real projective points and continued fractions is established and in proposition 3.21 where it is proven that any infinite continued fraction converges to an irrational number. The first part ends by studying the equivalence of real projective points through continued fractions in further detail and by developing in proposition 4.5 a product decomposition of matrices with non-negative entries and ± 1 determinant which will become an important tool in part III chapter 12.

The second part starts by defining integral binary quadratic forms but will quickly consider indefinite integral binary quadratic forms only. Next, the equivalence of those objects and the class number will be defined. Actually, the class number of indefinite integral binary quadratic forms which is the number of equivalence classes of indefinite integral binary quadratic forms consists in the main interest of this Master thesis and hence the rest of the document will study its properties. In this regard, the second part will ultimately conclude some rough upper bounds of the class number in section 8.3. To achieve this task, the idea is to show that any integral binary quadratic form is equivalent to some integral binary quadratic form of a special form and that there are only finitely many such special integral binary quadratic forms. Those special integral binary quadratic forms are exactly the strongly reduced integral binary quadratic forms that are defined in chapter 6. Furthermore, two easy verifications to determine whether a given integral binary quadratic form is strongly reduced or not will also be developed in this chapter. The reduction method used to obtain the main result that any integral binary quadratic form is equivalent to a strongly reduced one (theorem 8.4) relies on so called derived forms which are defined in chapter 7. Additionally some of their properties are already developed in this chapter as well as an efficient computation method of

the derived form of a strongly reduced integral binary quadratic form (c.f. proposition 7.10). Let us highlight the fact that from chapter 6 on, the definitions given in this document will, in order to keep up the natural path of deductions, slightly differ from the standard definitions found in the literature.

The third and final part will finally combine explicitly both theories to obtain several results in both domains. The first application of the theory of integral binary quadratic forms onto continued fractions is in chapter 11 where in theorem 11.4 the connection of strongly reduced quadratic surds and purely periodic continued fractions is outlined. The proof of this result needs an important conclusion about strongly reduced integral binary quadratic forms which is developed in chapter 10. This result states that by consecutively taking the derived form of a given strongly reduced quadratic form, one ends up at the initial form and so produces a cycle of strongly reduced integral binary quadratic forms (this can be found in theorem 10.1). Then some special continued fractions will be developed for whose development we need so called associate forms. They will reappear in chapter 12 when one of the main results of part III is proven, namely that two strongly reduced integral binary quadratic forms are equivalent if and only if they belong to the same cycle or associate cycles (c.f theorems 12.7 and 12.8). Up to this point the main work of the Master thesis was to collect and rewrite the results in a modern mathematical language. This bibliographic work will finally be used in chapter 13 to obtain a method to study cycles of integral binary quadratic forms from a different perspective and to to obtain several innovative results. These results are new as to the knowledge of the author. At last, these results will justify a basic algorithm to compute the class number of integral binary quadratic forms.

Each chapter is accompanied with several bibliographic references which are given as precisely as possible and such that the main sources and the complementing sources are well distinguished. Since several definitions and results have been slightly modified, the individual parts of each chapter are not referenced on their own.

Part I
Continued fractions

The first part is divided into chapter 2, 3 and 4. The main goal of this part is to introduce continued fractions, to define the group action of $GL(2,\mathbb{Z})$ onto continued fractions when they are the first coefficient of a (normalized) real projective point, to show the identification of continued fractions with real numbers and to develop the equivalence of real projective points through continued fractions which consists in a handy tool to work with later on.

In chapter 2 a recursive algorithmic process will be defined which attributes to each real number a particular fraction. Due to this expansion, continued fractions will be defined as abstract mathematical objects that generalize this expansion.

In chapter 3, it is shown that there is a one-to-one correspondence between real numbers and continued fractions. In a first place, the correspondence between finite continued fractions and rational numbers is established in section 3.2 and then the same is done for infinite continued fractions and irrational numbers in section 3.3. However, this proof is more challenging and thus several preparations need to be developed. The method used to prove this result relies on a special equivalence relation of the real projective line caused by the action of the general linear group of order two. Actually, one observes in proposition 3.19 that certain actions of matrices of the general linear group of order two onto real projective points produce formal continued fractions. Using this proposition, it is shown in proposition 3.21 that any infinite continued fraction converges to an irrational number. At last, theorem 3.27 concludes the one-to-one correspondence.

Chapter 4 can be seen at this point as a complementing chapter. In fact, it analyzes the equivalence of real projective points in further detail and ends at the conclusion that two (normalized) real projective points are equivalent if and only if the continued fractions of their first entries are both finite (c.f. proposition 4.1) or both infinite and equal after finitely many partial quotients (c.f. proposition 4.8). However, the main work of this chapter is given by the product decomposition of matrices of $GL(2,\mathbb{Z})$. This is developed in lemma 4.4, proposition 4.5 and corollary 4.7. This decomposition will play a crucial role in chapter 12.

2 Continued fractions

Sources: [1, p.3-4], [2, p.1-3], [3, p.325-327], [4, p.5-8]
Other approaches: [5, p.6], [6, p.3-9], [7, p.8-9]

Continued fractions can be introduced through different approaches. Some approaches arise naturally from number theoretical facts but most of those approaches treat the continued fractions of rational numbers only. One such approach is given in appendix A.

We are going to use in this chapter a modern algorithmic approach to define continued fractions.

The definition of continued fractions

In order to formalize the idea of continued fractions, let us first formalize some basic number theoretical notions.

Definition 2.1. The *integral part* of $x \in \mathbb{R}$ is the unique integer $\lfloor x \rfloor$ such that $x-1 < \lfloor x \rfloor \leq x$.

Example. $\left\lfloor \frac{5}{2} \right\rfloor = 2$, $\left\lfloor -\frac{16}{5} \right\rfloor = -4$, $\left\lfloor \frac{1}{2} \right\rfloor = 0$ and $\lfloor 1 \rfloor = 1$.

Definition 2.2. The *derived number* of $x \in \mathbb{R} \setminus \mathbb{Z}$ is the unique number ∂x that satisfies $x = \lfloor x \rfloor + \frac{1}{\partial x}$ or equivalently $\partial x = \frac{1}{x - \lfloor x \rfloor}$ (hence $\partial x > 1$).
The *derived number* of $x \in \mathbb{Z} \cup \{\infty\}$ is set by convention to be $\partial x = \infty$.
We use the notation $\partial^0 x = x$, $\partial^1 x = \partial x$ and $\partial^k x = \partial(\partial^{k-1} x)$ for all $k \in \mathbb{N}$ and call $\partial^k x$ the *k-th derived number* of x.

Example. $\partial\left(\frac{5}{2}\right) = 2$, $\partial^2\left(\frac{5}{2}\right) = \partial 2 = \infty$ and $\partial\left(-\frac{16}{5}\right) = \frac{5}{4}$.

The modern mathematical background of continued fractions can be reduced to the following algorithm.
Let $x \in \mathbb{R}$ and set $\partial^0 x = x = x_0$. We define a recursive process by the following two rules:

1. If $x_i \notin \mathbb{Z}$, then $x_{i+1} = \partial x_i$.
2. If $x_i \in \mathbb{Z}$, then the process ends (i.e. x_k does not exist for $k > i$).

We can observe that if x_n exists for some $n \in \mathbb{N}$, then x_k exists for all $0 \leq k \leq n$ and hence $x_k \notin \mathbb{Z}$ for all $0 \leq k \leq n-1$. Furthermore, the first rule implies that in this case $x_n = \partial^n x$. Hence, if x_n exists, then the following system of equations is satisfied:

$$\begin{aligned} x_0 = \partial^0 x &= \lfloor x_0 \rfloor + \frac{1}{x_1}; \\ x_1 = \partial^1 x &= \lfloor x_1 \rfloor + \frac{1}{x_2}; \\ &\vdots \\ x_{n-1} = \partial^{n-1} x &= \lfloor x_{n-1} \rfloor + \frac{1}{x_n}. \end{aligned}$$

If we insert the expression of the last equation into the second last equation and then the resulting equation into the third last equation and so on, we obtain

$$x = \lfloor x \rfloor + \cfrac{1}{\lfloor \partial x \rfloor + \cfrac{1}{\ddots \cdot \lfloor \partial^{n-1} x \rfloor + \frac{1}{\partial^n x}}}. \tag{1}$$

This leads us to the first main definition.

Definition 2.3. A *(simple, integral) partial continued fraction* of order j is any display of the form

$$r_0 + \cfrac{1}{r_1 + \cfrac{1}{r_2 + \cfrac{1}{\ddots + \frac{1}{r_{j-1} + \frac{1}{C}}}}}$$

where $r_0 \in \mathbb{Z}$, $r_i \in \mathbb{N} \setminus \{0\}$ for all $i \in \{1, ..., j-1\}$ and $C \in]1, +\infty[$. One calls r_i the *i-th partial quotient* of the (partial) continued fraction and C the *j-th complete quotient* of the partial continued fraction.

Consider the right hand side of (1). Since the integral part of a real number is an integer, $\lfloor x \rfloor, \lfloor \partial x \rfloor, ..., \left\lfloor \partial^{n-1} x \right\rfloor$ are integers. Since $x_k = \partial^k x \notin \mathbb{Z}$ for all $0 \leq k \leq n-1$, we deduce that $x_k > 1$ for all $1 \leq k \leq n$ by definition (2.2). Hence $\partial^n x \in]1, +\infty[$ and $\left\lfloor \partial^k x \right\rfloor \geq 1$ for all $1 \leq k \leq n-1$. Hence the right hand side of (1) is a partial continued fraction and is called the *partial continued fraction expansion of order n* of x. Notice also that the n-th complete quotient of the continued fraction expansion of x corresponds to the n-th derived number of x.

Example. The continued fraction expansion of order 1 of $\frac{21}{15}$ is given by

$$\frac{21}{15} = 1 + \frac{1}{\frac{5}{2}}.$$

We deduce that $\frac{5}{2}$ is the first complete quotient of $\frac{21}{15}$ and 1 is the 0-th partial quotient of $\frac{21}{15}$.

Since the above notation of partial continued fractions is quite cumbersome, we will use the so called bracket notation

$$x = [r_0, r_1, ..., r_{n-1}, x_n] = r_0 + \cfrac{1}{r_1 + \cfrac{1}{r_2 + \cfrac{1}{\ddots + \frac{1}{r_{n-1} + \frac{1}{x_n}}}}}.$$

Example. $\frac{21}{15} = 1 + \frac{1}{2 + \frac{1}{2}} = [1, 2, 2]$.

Our construction method uses a recursive process to compute a partial continued fraction expansion of a real number x with a real complete quotient that is strictly greater than 1. However, we want to obtain a quantity that involves integers only. Thus, we could tempt to replace the n-th complete quotient by another partial continued fraction expansion with some real complete quotient that is strictly greater than 1, as we did for the real number x in the first place. If we do so, we will elongate the partial continued fraction expansion of x and obtain a partial continued fraction expansion of order k of x where $k > n$. Since the construction method is based on a recursive process which is defined by two if-rules, we have to consider two cases:

1. Either, the k-th complete quotient $x_k \in \mathbb{Z}$ and so the process ends. Then $\lfloor x_k \rfloor = x_k$ which implies that $x_k = r_k$ is a partial quotient. Then we obtain a so called *finite continued fraction*.

   Definition 2.4. A *finite continued fraction* is any display of the form

   $$[r_0, r_1, r_2, ..., r_j] = r_0 + \cfrac{1}{r_1 + \cfrac{1}{r_2 + \cfrac{1}{\ddots + \frac{1}{r_j}}}}$$

   where $r_0 \in \mathbb{Z}$, $r_i \in \mathbb{N} \setminus \{0\}$ for all $i \in \{1, ..., j-1\}$ and $r_j \in \mathbb{N} \setminus \{0, 1\}$.

   Example.

   $$[1, 2, 2] = 1 + \frac{1}{2 + \frac{1}{2}}.$$

2. Or, for all $k \in \mathbb{N}$ the k-th complete quotient $x_k \notin \mathbb{Z}$. Then the process will never end and we obtain a so called *infinite continued fraction*.

   Definition 2.5. An *infinite continued fraction* is any display of the form

   $$[r_0, r_1, r_2, ...] = r_0 + \cfrac{1}{r_1 + \cfrac{1}{r_2 + \cfrac{1}{r_3 + \cfrac{1}{\ddots}}}}$$

   where $r_0 \in \mathbb{Z}$, $r_i \in \mathbb{N} \setminus \{0\}$ for all $i \in \mathbb{N} \setminus \{0\}$.

   Example.

   $$[3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, ...] \ (= \pi).$$

Remark 2.6. *Through history, there have been a lot of different approaches to continued fractions. That's why in today's literature, one can still find different approaches and even different definitions of continued fractions. We are going to highlight some of them in this remark.*

First of all, one can find general continued fractions which are displayed by

$$a_0 + \cfrac{b_0}{a_1 + \cfrac{b_1}{a_2 + \cfrac{b_2}{\ddots}}}$$

and involve two general sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$. Then, one can consider simple continued fractions by setting $b_k = 1$ for all $k \in \mathbb{N}$. Actually, the sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ in general continued fractions can be defined in any number field and hence the theory of general continued fractions has a wider spectrum than the theory that will be presented in this document. However the general theory is not yet well understood and has still a lot of inconsistencies. Because of the difficulties of general continued fractions, the biggest part of literature about continued fractions treats simple continued fractions only.

A big separation of the different studies of simple continued fractions is due to the choice of the considered number field. By setting $a_n \in \mathbb{Z} \cup \{\infty\}$, one considers formal (simple, integral) continued fractions. Several authors use these hypotheses on the partial quotients of continued

fractions. Indeed, in some fields of mathematics (like for example knot theory), it is easier to accept the value of the partial quotients to be negative and 0 *and* ∞ *with the convention that* $\frac{1}{0} = \infty$ *and* $\frac{1}{\infty} = 0$*. A big disadvantage of the theory of formal continued fractions is that there may be several different continued fractions that denote the same real number.*

Example. $\frac{21}{15} = 1 + \frac{1}{2+\frac{1}{2}} = 1 + \frac{1}{2+\frac{1}{2+\frac{1}{\infty}}} = 2 + \frac{1}{-1+\frac{1}{-1+\frac{1}{-2}}} = ...$

In order to bypass the plurality of continued fractions that denote the same real number, one can consider the sequence $(a_n)_{n\in\mathbb{N}}$ *to be defined by* $a_0 \in \mathbb{Z}$*,* $a_n \geq 1$ *as it is the case in definition 2.3, 2.4 and 2.5. In section 3, we will see that under our hypotheses continued fractions and real numbers can be identified one-by-one.*

On the other hand, the uncomplicated character of formal continued fractions simplifies several proofs. Fortunately, every formal continued fraction with finitely many negative, 0 *or* ∞ *partial quotients denotes a real number or* ∞ *and can be converted into a continued fraction (except those who denote* ∞*). Hence, there shall not be a problem in using formal continued fractions, in this document too. We are going to use our hypotheses as the main hypotheses and emphasize the change to formal continued fractions each time we use them. Nevertheless, the use of formal continued fractions will be restricted to some proofs only and one should note that one could have given the same results and proofs without relying on formal continued fractions (by accepting more complicated and longer arguments).*

3 The set of continued fractions is equal to the set of real numbers

Sources: [1, p.4-7], [2, p.3-20], [3, p.327-340], [4, p.8-28,51-76]
Other approaches: [5, p.7-8],[6, p.30-69, p.191-200], [7, p.9-19]

In chapter 2, we defined continued fractions independently from real numbers. Since we do not know any properties of continued fractions yet, we can imagine them as abstract mathematical objects. In order to quantify the set of continued fractions, we should find a bijection between this set and a known algebraic set. The considered set will be the set of real numbers.

Actually continued fractions are not only in one-to-one correspondence with real numbers, but each continued fraction is equal to a real number. In the interest of proving this fact, we are going to get a bit more abstract by defining the value of a continued fraction in 3.1. Then, in 3.2, we prove that the value of a finite continued fraction is a rational number and that each rational number is the value of exactly one finite continued fraction. Unfortunately at this point, we do not know if a rational number might also be the value of some infinite continued fractions. Thus, we are going to show that the value of any infinite continued fraction is an irrational number in section 3.3. This result will finally allow us to conclude in section 3.4 that continued fractions and real numbers are identified one-by-one.

3.1 The continued fraction expansion of a real number and the value of a continued fraction

The continued fraction expansion of a real number

In the previous chapter, we discussed briefly what we mean by the continued fraction expansion of a real number. Nonetheless, let us repeat this here.

Using a recursive process, we gave in chapter 2 a general method to construct a partial continued fraction for any real number (c.f. formula (1)). By including this process into a possibly infinite algorithm, we deduced that we can obtain a continued fraction for any real number.

Definition 3.1. Let $x \in \mathbb{R}$. The unique continued fraction obtained by the maybe infinite process described above is called the *continued fraction expansion of* x.

Proposition 3.2. *Any real number x has a continued fraction expansion.*

Example. $\frac{5}{2} = [2, 2]$, $-\frac{16}{5} = [-4, 1, 4]$, $8 = [8]$, $\sqrt{2} = [1, 2, 2, 2, 2, 2, 2, 2, ...]$,
$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, ...]$ and $\frac{1+\sqrt{5}}{2} = [1, 1, 1, 1, 1, 1, 1, ...]$.

The value of a continued fraction

Since the rational number field $\mathbb{Q}$ is closed under addition, multiplication and division by non-zero elements and since for any finite continued fraction $[r_0, r_1, r_2, ..., r_n]$, the partial quotients $r_1, r_2, ..., r_n$ are all strictly positive, we deduce that a finite continued fraction is actually equal to a rational number. We would like to obtain a similar statement for infinite continued fractions. Thus, consider the following definition.

Definition 3.3. Let $[r_0, r_1, ..., r_n]$ be a finite continued fraction, then, the *value* of $[r_0, r_1, ..., r_n]$ is defined to be

$$r_0 + \cfrac{1}{r_1 + \cfrac{1}{r_2 + \cfrac{1}{\ddots + \frac{1}{r_n}}}}.$$

Let $[r_0, r_1, ..., r_n, ...]$ be an infinite continued fraction. If the rational number sequence

$$r_0, \quad r_0 + \frac{1}{r_1}, \quad r_0 + \cfrac{1}{r_1 + \frac{1}{r_2}}, \quad ...$$

converges, then, the *value* of $[r_0, r_1, ..., r_n, ...]$ is defined to be the limit value of this sequence. If this sequence does not converge, then the *value* of $[r_0, r_1, ..., r_n, ...]$ is defined to be ∞.

Corollary 3.4. *The value of a continued fraction is unique.*

In this logic, the proof that real numbers are identified one-by-one with real numbers is reduced to the proof that the value of any continued fraction is a real number and that any real number is the value of exactly one continued fraction, namely its continued fraction expansion.

3.2 The value of a finite continued fraction is a rational number and each rational number is the value of exactly one finite continued fraction

Let $x, y \in \mathbb{Z}$ with $y > 0$ and consider the rational number $\frac{x}{y}$. Then our recursive algorithm to compute a continued fraction expansion of $\frac{x}{y}$ corresponds to the Euclidean division algorithm. Actually, each step of the Euclidean division algorithm will produce one partial quotient of the continued fraction expansion of $\frac{x}{y}$ (c.f. appendix A). Since the Euclidean division algorithm ends after a finite number of steps, the resulting continued fraction expansion will be finite.

Proposition 3.5. *Any rational number has a finite continued fraction expansion.*

Example. Using the Euclidean algorithm (or equivalently our construction to obtain (1)), we obtain the following continued fraction expansions:

$$\frac{5}{6} = [0, 1, 5] \qquad \text{and} \qquad \frac{10101}{103} = [98, 14, 1, 2, 2]$$

Hence, any rational number R is the value of at least one continued fraction and at least one of all the continued fractions with value R is finite. Now, we should prove that R is equal to exactly one continued fraction, namely its continued fraction expansion. In this regard, we tackle the problem from the other side and observe that by definition of the value of a finite continued fraction the following proposition holds.

Proposition 3.6. *The value of a finite continued fraction is unique and rational.*

Example. Using the operations of $\mathbb{Q}$, we can simplify the value of a finite continued fraction:

$$[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] = \frac{2789144166880}{1946194117057} \qquad \text{and} \qquad [2, 3, 2, 3, 2] = \frac{126}{55}$$

Next, we would like to prove that there is only one finite continued fraction that is equal to a given rational number, namely its continued fraction expansion. To do so, consider the following lemma.

Lemma 3.7. *Let $[r_0, C]$ and $[r'_0, C']$ be two partial continued fractions. Then $[r_0, C] = [r'_0, C']$ if and only if $r_0 = r'_0$ and $C = C'$.*

Proof. We have $[r_0, C] = [r'_0, C']$ if and only if

$$0 = r_0 + \frac{1}{C} - r'_0 - \frac{1}{C'} = (r_0 - r'_0) + \left(\frac{1}{C} - \frac{1}{C'}\right). \tag{2}$$

Since $0 < \frac{1}{C}, \frac{1}{C'} < 1$ by the definition of partial continued fractions, we deduce that $-1 < \frac{1}{C} - \frac{1}{C'} < 1$. Since $r_0, r'_0 \in \mathbb{Z}$, equation (2) holds if and only if $r_0 = r'_0$ and $C = C'$. □

If the value of $[r'_0, r'_1, ..., r'_n]$ is the rational number R and if $[r_0, r_1, ..., r_m]$ is the continued fraction expansion of R, then by definition of the value of finite continued fractions $[r'_0, r'_1, ..., r'_n] = R$ and by construction $[r_0, r_1, ..., r_m] = R$. Hence$[r'_0, r'_1, ..., r'_n] = [r_0, r_1, ..., r_m]$. If one of the two integers n, m is zero and the other is strictly positive, we get a contradiction since a continued fraction with one partial quotient only is an integer, whereas continued fraction with at least two partial quotients can be shown not to be an integer. Using this, the previous lemma and induction on the number of partial quotients, we deduce the following corollary.

Corollary 3.8. *The continued fraction expansion of a rational number is the unique finite continued fraction whose value is equal to this rational number.*

In order to conclude that each rational number is the value of exactly one continued fraction, namely its continued fraction expansion (which is finite), we need to exclude the case that the value of an infinite continued fraction is equal to a rational number.

3.3 The value of an infinite continued fraction is an irrational number and each irrational number is the value of exactly one (infinite) continued fraction

If an irrational number is the value of a continued fraction, then this continued fraction is infinite.

By proposition 3.2 any irrational number has a continued fraction expansion. Assume by contradiction that an irrational number Ir is the value a finite continued fraction. Since by proposition 3.6, the value of any finite continued fraction is unique and rational, we get a contradiction by the fact that the real number Ir cannot be both rational and irrational. Hence we conclude the following result.

Proposition 3.9. *If an irrational number is the value of a continued fraction, then this continued fraction is infinite.*

Proposition 3.10. *Any irrational number has an infinite continued fraction expansion.*

Example. Using our construction (c.f. formula (1)), we obtain the following continued fraction expansions:

$$\sqrt{2} = [1, 2, 2, 2, 2, 2, 2, 2, ...] \qquad \text{and} \qquad \pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, ...]$$

In order to grant continued fractions the status of real numbers, the search of a formal proof which confirms that the value of any continued fraction is equal to a real number (and not ∞) was probably one of the most difficult tasks. Whereas the statement for rational numbers and finite continued fractions is an easy exercise as is shown above, the corresponding statement for infinite continued fractions has to be developed cautiously. Today, there are different ways to do so. We are going to carry out a proof that uses the general linear group and a certain equivalence relation of points of the real projective line.

3.3.1 The general linear group of order two over the integers, the real projective line and their connection to continued fractions

Before we fulfill the main task of this section, let us recall the definition of the general linear group of order two over the integers and the real projective line and let us show how the general linear group of order two over the integers acts on the real projective line.

The general linear group of order two over the integers and the real projective line

Throughout the whole document, we are going to use a special type of 2-by-2 matrices to perform computations and actions on different sets. These matrices are exactly the elements of the general linear group of order two over the integers.

Definition 3.11. The *general linear group of order two over the integers* $GL(2,\mathbb{Z})$ is the group of 2×2 matrices $S=\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ where $(a,b,c,d)\in\mathbb{Z}^4$ that satisfy $\det(S)=ad-bc=\pm 1$. In other words $GL(2,\mathbb{Z})=\left\{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) : ad-bc=\pm 1\right\}$.

Example. Some elements of $GL(2,\mathbb{Z})$ are $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -1 & 0 \\ -1 & -1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -1 & 2 \\ 3 & -5 \end{smallmatrix}\right)$.

As we will see in the second part of this section, the real projective line consists in a handy tool to study continued fractions. In order to avoid any confusions, let us remind the definition and computation rules of the real projective line.

Definition 3.12. Let $(x,y),(w,z)\in\mathbb{R}^2$ and define the equivalence relation $(x,y)\sim(w,z)$ if and only if $(x,y)=(tw,tz)$ for some $t\in\mathbb{R}\setminus\{0\}$.
The real projective line $P^1(\mathbb{R})$ is the set of equivalence classes in $\mathbb{R}^2\setminus\{(0,0)\}$ of the equivalence $\sim$. Elements of $P^1(\mathbb{R})$ (i.e. the equivalence classes of $\sim$) are denoted by $(x:y)$ where (x,y) is a representative of the equivalence class. In other words

$$P^1(\mathbb{R})=\left\{(x:y):(x,y)\in\mathbb{R}^2\setminus\{(0,0)\}\right\}=\left(\mathbb{R}^2\setminus\{(0,0)\}\right)\Big/_{\{(x,y)\sim(\lambda x,\lambda y)\}}.$$

Let $(x_1:x_2),(y_1:y_2)\in P^1(\mathbb{R})$. Then, if $x_2\neq 0$, we have $(x_1:x_2)=\left(\frac{x_1}{x_2}:1\right)$. Whence the set of couples $(x:1)$ is in bijection with the real affine line $\mathbb{R}$. If $x_2=0$, then $(x_1:0)=(1:0)$ and we call this element ∞. Thus $P^1(\mathbb{R})\cong\mathbb{R}\cup\{\infty\}$.

Furthermore, we define the operations $+,-$, juxtaposition and $(\quad)^{-1}$ on the real projective line as follows:

1. $(x_1:x_2)+(y_1:y_2)=(x_1y_2+y_1x_2:x_2y_2)$ (if at least one of the two points is not ∞) and thus the neutral element for addition is $(0:1)$. Moreover $(x:1)+\infty=\infty$ and $-(x:1)=(-x:1)$. Furthermore $\infty+\infty=\infty$ and ∞ does not have an additive inverse.

2. $(x_1 : x_2)(y_1 : y_2) = (x_1y_1 : x_2y_2)$ (not defined for $(0 : 1)\infty$) and thus the neutral element for addition is $(1 : 1)$. Moreover $(x : 1)\infty = \infty$ (if $x \neq 0$), $\infty\infty = \infty$ and $(x : 1)^{-1} = (\frac{1}{x} : 1)$ (if $x \neq 0$). Furthermore ∞ does not have a multiplicative inverse.

Example. $(1 : 1)$ and $(2 : 1)$ denote two different elements of the real projective line, whereas $(1 : 1)$ and $(2 : 2)$ denote the same.
Furthermore $(-2 : 1)$ is the additive inverse of $(2 : 1)$ and $(\frac{1}{2} : 1)$ is its multiplicative inverse.

Remark 3.13. *There may be multiple representatives of a single equivalence class (called points) of the real projective line for example* $(1 : 1)$ *and* $(2 : 2)$*, however this does not mean that their individual coordinates (or respective points in the Euclidean plane) are identified* $(2 \neq 1)$*. However, if one coordinate is non-zero and equal in both points, then both points are equal if and only if the second coordinate is also identified. In particular, this implies that* $(x : 1) = (y : 1)$ *if and only if* $x = y$*.*

The group action of $GL(2, \mathbb{Z})$ on $P^1(\mathbb{R})$ and the resulting equivalence relation

Let us study the interplay of the above defined sets, discover some of their properties and develop some computation rules.

Proposition 3.14. $GL(2, \mathbb{Z})$ *acts on* $P^1(\mathbb{R})$ *by*

$$S \star (x : y) = (ax + by : cx + dy)$$

for all $S = \pm \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$.

Remark 3.15. *Since* $(-x : -y) = (x : y)$ *for any point* $(x : y)$ *on the real projective line, we deduce that for any matrix* $S \in GL(2, \mathbb{Z})$*, we have that* $S \star (x : y) = (-S) \star (x : y)$ *for any* $(x : y) \in P^1(\mathbb{R})$.

By the properties of the projective line, we can summarize the action of $GL(2, \mathbb{Z})$ on $P^1(\mathbb{R})$ by the following computation rules. For all $S = \pm \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$, we have

$$\begin{aligned} S \star (x : 1) &= \left(\frac{ax + b}{cx + d} : 1\right) \quad \text{if } x \neq -\frac{d}{c} \\ S \star (-\frac{d}{c} : 1) &= \infty \\ S \star \infty &= \left(\frac{a}{c} : 1\right) \quad \text{if } c \neq 0 \\ S \star \infty &= \infty \quad \text{if } c = 0 \end{aligned}$$

Example. Let $S = \pm \left(\begin{smallmatrix} 1 & 2 \\ -3 & 5 \end{smallmatrix}\right) \in Gl(2, \mathbb{Z})$. Then $S \star (1 : 1) = \left(\frac{3}{2} : 1\right)$, $S \star \left(\frac{5}{3} : 1\right) = \infty$ and $S \star \infty = \left(-\frac{1}{3} : 1\right)$.

Notice that if $S = \pm \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$ with $c = 0$, then $a, d = \pm 1$. By the last equation above, we know that the matrices $\pm \left(\begin{smallmatrix} \pm 1 & b \\ 0 & 1 \end{smallmatrix}\right)$ where $b \in \mathbb{Z}$ stabilize ∞. Actually, one easily shows that those are the only matrices that stabilize ∞.

Proposition 3.16. *The stabilizer of* ∞ *in* $GL(2, \mathbb{Z})$ *is* $\Gamma_\infty = \left\{ \begin{pmatrix} \pm 1 & n \\ 0 & \pm 1 \end{pmatrix} \in GL(2, \mathbb{Z}) \middle| n \in \mathbb{Z} \right\}$ *where the sign of the first entry does not depend on* n*.*

We define an equivalence relation on the real projective line via the action of $GL(2,\mathbb{Z})$.

Definition 3.17. We say that two points X and Y in $P^1(\mathbb{R})$ are *equivalent* if there exists $S \in GL(2,\mathbb{Z})$ such that $X = S \star Y$.

Example. Since $(2:1) = \pm \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \star (1:1)$, $(2:1)$ and $(1:1)$ are equivalent. Moreover, since $\pm \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \star (n-1:1) = (n:1)$ for all $n \in \mathbb{Z}$, we deduce by transitivity that $(n:1)$ is equivalent to $(1:1)$ for all $n \in \mathbb{Z}$.

The connection of $GL(2,\mathbb{Z})$ and $P^1(\mathbb{R})$ to continued fractions

We are going to use the above group action to prove one of the most important results of this document. To this end, let us get back to the notions of chapter 2. Let us begin with a preliminary observation that connects chapter 2 and chapter 3.

Remark 3.18. *Let $x \in \mathbb{R} \setminus \mathbb{Z}$. Then $(x:1)$ is equivalent to $(\partial x:1)$ where ∂x is the derived number of x. Indeed, consider the matrix $T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$ with $r = \lfloor x \rfloor$. Then $T(r) \star (\partial x:1) = (r + \frac{1}{\partial x}:1) = (x:1)$.*

By transitivity, we conclude that if $\partial^n x$ exists for some $n \in \mathbb{N}$, then $(x:1)$ is equivalent to $(\partial^n x:1)$. More precisely, we deduce that $(x:1) = D_n \star (\partial^n x:1)$ where $D_n = T(r_0)T(r_1)\dots T(r_{n-1})$ with $r_i = \lfloor \partial^i x \rfloor$. Using the notation $x_n = \partial^n x$, we deduce that $([r_0, r_1, ..., r_{n-1}, x_n]:1) = T(r_0)\dots T(r_{n-1}) \star (x_n:1)$.

This seemingly coincidental observation will be one of the strongest arguments in future proofs. Therefore we should study the given equivalence under this kind of matrices in further detail.

Proposition 3.19. *Let $(r_n)_{n\geq 0}$ be a sequence of integers and set*

$$T(r_n) = \begin{pmatrix} r_n & 1 \\ 1 & 0 \end{pmatrix} \qquad (n \geq 0);$$

$$D_n = T(r_0)\dots T(r_{n-1}) \qquad (n \geq 1).$$

1. *For $(x:1), (x':1) \in P^1(\mathbb{R})$ with $x, x' \neq 0$ the following two are equivalent:*
   - *(a) $(x:1) = D_n \star (x':1)$;*
   - *(b) $x = [r_0, r_1, ..., r_{n-1}, x']$.*

   Notice that $[r_0, r_1, ..., r_{n-1}, x']$ may denotes a formal continued fraction since we do not suppose that $r_k > 1$ for all $1 \leq k \leq n-1$, nor that $x' > 0$.

2. *We have*

$$D_n = \begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix}$$

   where the sequences $(p_n)_{n\geq -1}$ and $(q_n)_{n\geq -1}$ are defined by the recurrence relations:
   - *(a) $p_{-1} = 1$ and $p_0 = r_0$;*
   - *(b) $p_n = r_n p_{n-1} + p_{n-2}$;*
   - *(c) $q_{-1} = 0$ and $q_0 = 1$;*

(d) $q_n = r_n q_{n-1} + q_{n-2}$.

Moreover we have $\det(D_n) = p_{n-1}q_{n-2} - q_{n-1}p_{n-2} = (-1)^n$ *and so* $D_n \in GL(2,\mathbb{Z})$ *for all* $n \in \mathbb{N}$.

Proof. 1. We prove the equivalence of (a) and (b) by induction on n.
For $n = 1$ one observes that if $(x:1) = D_1.(x':1)$, then

$$(x:1) = D_1 \star (x':1) = T(r_0) \star (x':1) = (r_0 + \frac{1}{x'}:1) = ([r_0, x']:1)$$

and so $x = [r_0, x']$.
On the other hand, if $x = [r_0, x']$, then

$$(x:1) = ([r_0, x']:1) = (r_0 + \frac{1}{x'}:1) = T(r_0) \star (x':1) = D_1 \star (x':1).$$

Next, assume that the claim holds for $n-1 \geq 0$, then we will prove it for n. Indeed, if $(x:1) = D_n \star (x':1)$, then $(x:1) = D_{n-1}T(r_{n-1}) \star (x':1) = D_{n-1} \star \left(r_{n-1} + \frac{1}{x'}:1\right)$. Thus by induction $x = [r_0, r_1, ..., r_{n-2}, r_{n-1} + \frac{1}{x'}] = [r_0, r_1, ..., r_{n-1}, x']$. Conversely, assume $x = [r_0, r_1, ..., r_{n-1}, x']$, then $x = [r_0, r_1, ..., r_{n-2}, r_{n-1} + \frac{1}{x'}]$. Thus by induction $(x:1) = D_{n-1} \star \left(r_{n-1} + \frac{1}{x'}:1\right) = D_{n-1}T(r_{n-1}) \star (x':1) = D_n \star (x':0)$.

2. For all $n \geq 1$, set

$$D_n = \begin{pmatrix} p_{n-1} & p'_{n-1} \\ q_{n-1} & q'_{n-1} \end{pmatrix}.$$

Then the relation $D_{n+1} = D_n T(r_n)$ yields:

$$\begin{aligned} p_n &= r_n p_{n-1} + p'_{n-1} \quad \text{and} \quad p'_n = p_{n-1}; \\ q_n &= r_n q_{n-1} + q'_{n-1} \quad \text{and} \quad q'_n = q_{n-1}. \end{aligned}$$

This gives us the claimed formula of D_n for all $n \geq 1$ and the three term recurrence relations of the sequences (p_n) and (q_n). The initial values of the sequences are given by $D_1 = T(r_0)$. By noticing that $\det(T(r_i)) = -1$ for all $i \geq 0$ and by the multiplicative property of the determinant, the determinant formula follows.

□

The second part of proposition 3.19 describes the matrix D_n in terms of recurrence relations which depend on the sequence $(r_n)_{n\geq 0}$. This knowledge will simplify future computations. Since proposition 3.19 is one of the most important results of this document and since we are going to use the established formulas and notations several times hereinafter, we reserve the following notations $T(r_n)$, D_n, p_n and q_n for the objects described in proposition 3.19.

Corollary 3.20. *For any finite sequence of integers* $r_0, ..., r_n$ *we have* $[r_0, ..., r_n] = \frac{p_n}{q_n}$.

Proof. The claim holds for $n = 0$ since $[r_0] = r_0 = \frac{r_0}{1} = \frac{p_0}{q_0}$.
Moreover, for $n \geq 1$, we observe that

$$([r_0, ..., r_n]:1) = D_n \star (r_n:1) = \left(\frac{p_{n-1}r_n + p_{n-2}}{q_{n-1}r_n + q_{n-2}}:1\right) = \left(\frac{p_n}{q_n}:1\right).$$

□

3.3.2 The value of an infinite continued fractions is an irrational number

We are finally able to prove the main result, namely that the value of every infinite continued fraction is an irrational number.

Proposition 3.21. *Take an infinite sequence of integers $(r_n)_{n\geq 0}$ with $r_n \geq 1$ for $n \geq 1$ and set*

$$\zeta_n = [r_0, ..., r_n] = \frac{p_n}{q_n}.$$

Then the sequence of finite continued fractions (respectively rational numbers) $(\zeta_n)_{n\in\mathbb{N}} = \left(\frac{p_n}{q_n}\right)_{n\in\mathbb{N}}$ converges and defines an irrational number $\zeta = \lim_{n\to\infty} \zeta_n$. Thus, the value of $[r_0, r_1, ..., r_n, ...]$ is equal to ζ.

Moreover

$$\left|\zeta - \frac{p_n}{q_n}\right| \leq \frac{1}{q_n q_{n+1}} < \frac{1}{q_n^2}.$$

Proof. We separate the proof into four parts:

1. First we prove by induction (using the determinant formula from the second part of proposition 3.19) that

   $$\zeta_n = \frac{p_n}{q_n} = r_0 + \frac{(-1)^2}{q_0 q_1} + ... + \frac{(-1)^{n+1}}{q_{n-1} q_n}$$

   for all $n \geq 1$.
   Indeed, for $n = 1$ we have

   $$\zeta_1 = \frac{p_1}{q_1} = \frac{p_1 q_0 - p_0 q_1}{q_1 q_0} + \frac{p_0}{q_0} = \frac{(-1)^2}{q_1 q_0} + r_0.$$

   Next assume that the claim holds for $n - 1 \geq 1$, then

   $$\zeta_n = \frac{p_n}{q_n} = \frac{p_n q_{n-1} - q_n p_{n-1}}{q_n q_{n-1}} + \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n+1}}{q_n q_{n-1}} + \frac{(-1)^n}{q_{n-1} q_{n-2}} + ... + \frac{(-1)^2}{q_1 q_0} + r_0$$

   where we used induction on $\frac{p_{n-1}}{q_{n-1}}$ to obtain the last equality.

2. Next, we show by using the first part of the proof that the sequence of rational numbers (finite continued fractions) $(\zeta_n)_{n\in\mathbb{N}}$ converges.
   Since by proposition 3.19 the sequence $(q_n)_{n\geq -1}$ satisfies $q_{-1} = 0$, $q_0 = 1$ and $q_n = r_n q_{n-1} + q_{n-2}$ for all $n \geq 1$ and since $r_i \geq 1$ for all $i \geq 0$, we conclude that $q_{n+1} > q_n \geq n$ for all $n \geq 1$. From this we take the following conclusions:

   (a) The sub-sequence $(\zeta_{2n})_{n\in\mathbb{N}}$ is strictly increasing. Indeed, for all $n \in \mathbb{N}$, we have $\zeta_{2n} < \zeta_{2n+2}$ since

   $$\zeta_{2n} - \zeta_{2n+2} = -\frac{(-1)^{2n+2}}{q_{2n} q_{2n+1}} - \frac{(-1)^{2n+3}}{q_{2n+1} q_{2n+2}} = \frac{-q_{2n+2} + q_{2n}}{q_{2n} q_{2n+1} q_{2n+2}} < 0.$$

   (b) The sub-sequence $(\zeta_{2n+1})_{n\in\mathbb{N}}$ is strictly decreasing. Indeed, for all $n \in \mathbb{N}$, we have $\zeta_{2n+1} > \zeta_{2n+3}$ since

   $$\zeta_{2n+1} - \zeta_{2n+3} = -\frac{(-1)^{2n+3}}{q_{2n+1} q_{2n+2}} - \frac{(-1)^{2n+4}}{q_{2n+2} q_{2n+3}} = \frac{q_{2n+3} - q_{2n+1}}{q_{2n+1} q_{2n+2} q_{2n+3}} > 0.$$

(c) For any $n \in \mathbb{N}$, we have $\zeta_{2n} < \zeta_{2n+1}$ since

$$\zeta_{2n} - \zeta_{2n+1} = -\frac{(-1)^{2n+2}}{q_{2n}q_{2n+1}} < 0.$$

By (a),(b) and (c), we conclude that for all $n, m \in \mathbb{N}$, we have

$$\zeta_{2n} < \zeta_{2n+2m} < \zeta_{2n+2m+1} < \zeta_{2m+1}$$

and thus

$$\zeta_0 < \zeta_2 < \zeta_4 < ... < \zeta_5 < \zeta_3 < \zeta_1.$$

Since the sub-sequence $(\zeta_{2n})_{n\in\mathbb{N}}$ is strictly increasing and bounded above by ζ_1, the monotone convergence theorem implies that the sub-sequence $(\zeta_{2n})_{n\in\mathbb{N}}$ converges to some real number x.

Since the sub-sequence $(\zeta_{2n+1})_{n\in\mathbb{N}}$ is strictly decreasing and bounded below by ζ_0, the monotone convergence theorem implies that the sub-sequence $(\zeta_{2n+1})_{n\in\mathbb{N}}$ converges to some real number y.

Let $\epsilon > 0$. Since $q_{n+1} > q_n \geq n$ for all $n \in \mathbb{N}$, (c) implies that there exists $N_1 \in \mathbb{N}$ such that for all $n > N_1$, we have $|\zeta_{2n} - \zeta_{2n+1}| < \frac{1}{(2n_1)^2} < \frac{\epsilon}{3}$. Furthermore, since $(\zeta_{2n})_{n\in\mathbb{N}}$ converges to x and $(\zeta_{2n+1})_{n\in\mathbb{N}}$ converges to y, there exist $N_2, N_3 \in \mathbb{N}$ such that for all $m > N_2$ we have $|x - \zeta_{2m}| < \frac{\epsilon}{3}$ and such that for all $k > N_3$ we have $|\zeta_{2k+1} - y| < \frac{\epsilon}{3}$. Let $N = \max\{N_1, N_2, N_3\}$. Then for any $n > N$, we have

$$|x - y| \leq |x - \zeta_{2n}| + |\zeta_{2n} - \zeta_{2n+1}| + |\zeta_{2n+1} - y| < \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we deduce that $x = y$.

Hence the sequence $(\zeta_n)_{n\in\mathbb{N}}$ converges to some real number ζ which is then by definition the value of the infinite continued fraction $[r_0, r_1, r_2, ...]$.

3. Let us compute the speed of convergence of the sequence $(\zeta_n)_{n\geq 0}$ to its limit ζ.

   By the second part of the proof, any infinite continued fraction $[r'_0, r'_1, r'_2, ...]$ converges to an irrational number η which verifies $r'_0 < \eta < r'_0 + \frac{1}{r'_1}$.

   Let $\eta_n = [r_{n+1}, r_{n+2}, ...]$ with $n \geq 1$ be the n-th remainder of the infinite continued fraction $\zeta = [r_0, r_1, r_2, ...]$. Since $r_{n+1} \geq 1$, we deduce that $\eta_n \geq 1$.

   Observe that $(\zeta : 1) = ([r_0, ..., r_n, \eta_n] : 1) = D_{n+1} \star (\eta_n : 1)$ and so

$$\begin{aligned}
\left(\zeta - \frac{p_n}{q_n} : 1\right) = (\zeta : 1) - \left(\frac{p_n}{q_n} : 1\right) &= \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \star (\eta_n : 1) - \left(\frac{p_n}{q_n} : 1\right) \\
&= \left(\frac{p_n\eta_n + p_{n-1}}{q_n\eta_n + q_{n-1}} : 1\right) - \left(\frac{p_n}{q_n} : 1\right) \\
&= \left(\frac{q_n(p_n\eta_n + p_{n-1}) - p_n(q_n\eta_n + q_{n-1})}{q_n(q_n\eta_n + q_{n-1})} : 1\right) \\
&= \left(\frac{q_np_{n-1} - p_nq_{n-1}}{q_n(q_n\eta_n + q_{n-1})} : 1\right) \\
&= \left(\frac{(-1)^{n+2}}{q_n(\eta_nq_n + q_{n-1})} : 1\right).
\end{aligned}$$

Thus

$$0 < \left|\zeta - \frac{p_n}{q_n}\right| = \left|\frac{(-1)^{n+2}}{q_n(\eta_n q_n + q_{n-1})}\right| < \frac{1}{q_n^2} \leq \frac{1}{n^2}.$$

4. At last, let us show that ζ is irrational.

   Observe that by the third part of the proof $0 < \left|\zeta - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2}$ for all $n \geq 0$ and so the inequality $0 < \left|\zeta - \frac{p}{q}\right| < \frac{1}{q^2}$ with unknowns $p, q \in \mathbb{Z}$ such that $q > 0$ has infinitely many solutions.

   Assume by contradiction that ζ is rational. Then $\zeta = \frac{a}{b}$ with $a \in \mathbb{Z}$, $b \in \mathbb{N} \setminus \{0\}$ and $\gcd(a, b) = 1$. Then consider again the system of inequalities

   $$0 < \left|\frac{a}{b} - \frac{p}{q}\right| < \frac{1}{q^2} \tag{3}$$

   with unknowns $p, q \in \mathbb{Z}$ such that $q > 0$. Observe that if $0 < \left|\frac{a}{b} - \frac{p}{q}\right| = \left|\frac{aq-pb}{bq}\right|$, we have that $\frac{1}{bq} \leq \left|\frac{a}{b} - \frac{p}{q}\right|$ since $b > 0$, $q > 0$ and $0 < |aq - pb|$. We claim that the system of inequalities (3) has only a finite number of solutions for p and q.

   (a) If $\frac{1}{bq} \leq \frac{a}{b} - \frac{p}{q} < \frac{1}{q^2}$, then by the first inequality $0 \leq \frac{aq-bp-1}{bq}$ and by the second inequality $0 < \frac{b-aq^2+pbq}{bq^2}$. By taking the sum of those two positive fractions, we obtain $0 < \frac{b-q}{bq^2}$ and so $b > q > 0$ and thus there are only $b-1$ possible integer values for q. Moreover, for each fixed q there is a finite number of integer values for p such that $\frac{1}{bq} \leq \frac{a}{b} - \frac{p}{q} < \frac{1}{q^2}$ or equivalently such that $-q\left(\frac{1}{bq} - \frac{a}{b}\right) \geq p > -q\left(\frac{1}{q^2} - \frac{a}{b}\right)$.

   (b) If $\frac{1}{bq} \leq -\left(\frac{a}{b} - \frac{p}{q}\right) < \frac{1}{q^2}$, then by the first inequality $0 \leq \frac{-aq+bp-1}{bq}$ and by the second inequality $0 < \frac{b+aq^2-pbq}{bq^2}$. By taking the sum of those two positive fractions, we obtain $0 < \frac{b-q}{bq^2}$ and so $b > q > 0$ and thus there are only $b-1$ possible integer values for q. Moreover, for each fixed q there is a finite number of integer values for p such that $\frac{1}{bq} \leq -\left(\frac{a}{b} - \frac{p}{q}\right) < \frac{1}{q^2}$ or equivalently such that $-q\left(-\frac{1}{bq} - \frac{a}{b}\right) \leq p < -q\left(-\frac{1}{q^2} - \frac{a}{b}\right)$.

   Hence, the system $0 < \left|\frac{a}{b} - \frac{p}{q}\right| < \frac{1}{q^2}$ has only a finite number of solutions for p and q.

   However, this contradicts our initial observation of the fourth part of the proof. Hence ζ is an irrational number.

□

In particular, since any infinite continued fraction $[r_0, r_1, r_2, ...]$ can be approximated by an infinite sequence of finite continued fractions $\zeta_n = [r_0, r_1, ..., r_n]$, proposition 3.21 yields that the value of any infinite continued fraction is an irrational number.

Corollary 3.22. *The value of an infinite continued fraction is an irrational number.*

Example. The value of the infinite continued fraction $[1, 2, 2, 2, 2, 2, 2, 2, 2, ...]$ is $\sqrt{2}$ and the value of $[1, 2, 1, 2, 1, 2, ...]$ is $\frac{1+\sqrt{3}}{2}$.

By proposition 3.10, any irrational number has an infinite continued fraction expansion. By proposition 3.9, we know that if an irrational number is the value of a continued fraction, then this continued fraction is infinite. Assume that an irrational number is the value of multiple continued fractions, then they all are infinite. Furthermore, by using lemma 3.7 and induction, one can easily prove that all those continued fractions are equal.

Proposition 3.23. *The continued fraction expansion of an irrational number is the unique continued fraction whose value is equal to this rational number.*

3.4 Conclusions

Finally, we would like to conclude the one-to-one correspondence between the real number set and the set of continued fractions.

Actually, without proposition 3.19 (or a similar result), one cannot state any such result since one does not know if infinite continued fractions converge to real numbers at all and hence one cannot exclude the case that the value of a given continued fraction is ∞.

By corollary 3.8, we know that the continued fraction expansion of a rational number is the unique finite continued fraction whose value is equal to this rational number. By corollary 3.22, we can now exclude the case that the value of an infinite continued fraction is equal to a rational number. Hence, we conclude that the continued fraction expansion of a rational number is the only continued fraction whose value is equal to this rational number. Furthermore, since any finite continued fraction is equal to a rational number, we deduce the following proposition.

Proposition 3.24. *The set of finite continued fractions is equal to the set of rational numbers.*

Since the value of any infinite continued fraction is an irrational number, we can naturally embed the set of infinite continued fractions into the real number set by setting that an infinite continued fraction is equal to its value. Then proposition 3.23 implies the following proposition.

Proposition 3.25. *The set of infinite continued fractions is equal to the set of irrational numbers.*

The last two results show the one-to-one correspondence between the real number set and the set of continued fractions and how the set of continued fractions is identified with the set of real numbers.

We can finally summarize all the previous results in the following proposition.

Proposition 3.26.
1. *Any real number is the value of a unique continued fraction which is finite if and only if the number is rational.*
2. *The value of any continued fraction is a real number which is rational if and only if the continued fraction is finite.*

Hence, we conclude that our method to generate the continued fraction expansion produces the only possible continued fraction of a given real number. Furthermore, all the previous results lead to the following conclusion.

Theorem 3.27. *The set of continued fractions is equal to the set of real numbers.*

Hence, in the preceding continued fractions and their values and real numbers and their continued fraction expansions will be interchanged without further notifications and we will drop the logical and linguistic difference between those four objects.

4 An analysis of the equivalence relation of real projective points with the help of continued fractions

Source: [1, p.8-12]
Other approaches: [8, p.350-352]

The following chapter is a complementing chapter and introduces some notations that will be needed later on in the document. We use this chapter to study, in addition to the preparations for future computations, the equivalence relation of the real projective line in further detail.

The equivalence relation described in sub-section 3.3.1 partitions the real projective line into equivalence classes. Since we observed that $P^1(\mathbb{R}) = \{(x:1) : x \in \mathbb{R}\} \cup \{\infty\}$, we are going to study the equivalence relation on ∞ and points of the form $(x:1)$ with $x \in \mathbb{R}$.

We are going to study the equivalence relations between points of the form $(x:1)$ where $x \in \mathbb{R}$ through an analysis of the continued fraction of their first component. This will finally lead to the conclusion that we can reduce the search for equivalence classes to the unit interval.

Let us first discover the complete equivalence class of the special element ∞ and then develop the other equivalence classes.

4.1 The equivalence class of ∞

The following result will gather all rational numbers and show their membership to the same equivalence class.

Proposition 4.1. *Any point $(x:1)$ with $x \in \mathbb{Q}$ is equivalent to ∞. Furthermore, those are the only points that are equivalent to ∞.*

Proof. Let $r \in \mathbb{Q}$. Then $r = \frac{x}{y}$ for some $x, y \in \mathbb{Z}$ with $y > 0$. Since $\frac{x}{y}$ is rational, its continued fraction is finite. Let $[r_0, ..., r_{m-1}]$ be the continued fraction of $r = \frac{x}{y}$. Then $(\frac{x}{y} : 1) = ([r_0, ..., r_{m-1}] : 1)$ and so by the first part of proposition 3.19

$$\left(\frac{x}{y} : 1\right) = T(r_0) \star ([r_1, ..., r_{m-1}] : 1) = ... = T(r_0)T(r_1)...T(r_{m-2}) \star (r_{m-1} : 1) = D_m \star \infty.$$

Since $D_m \in GL(2, \mathbb{Z})$, $(r:1)$ is equivalent to ∞. Furthermore, by the computation rules in subsection 3.3.1 of the action of $GL(2, \mathbb{Z})$ onto $P^1(\mathbb{R})$, we know that for any $S = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$, we have $S \star \infty = (\frac{a}{c} : 1)$ (if $c \neq 0$) or $S \star \infty = \infty$ (if $c = 0$). Hence ∞ can only be equivalent to real projective points of the form $(x:1)$ where x is rational. □

By transitivity, we conclude the next corollary.

Corollary 4.2. *The equivalence class of ∞ is given by $\{(r:1) : r \in \mathbb{Q}\} \cup \{\infty\}$.*

Since rational points are identified with finite continued fractions, we deduce the following corollary.

Corollary 4.3. *The equivalence class of ∞ is given by*
$\{([r_0, ..., r_n] : 1) : n \in \mathbb{N}, r_0 \in \mathbb{Z}, r_i \geq 1$ for all $i = 1, ..., n$ and $r_n > 1\} \cup \{\infty\}$.

Example. By the previous proposition, the two points $(-\frac{3}{2} : 1)$ and $(\frac{21}{15} : 1)$ are equivalent to ∞ on $P^1(\mathbb{R})$. Thus in particular the two points $([-2, 2] : 1)$ and $([1, 2, 2] : 1)$ are equivalent on $P^1(\mathbb{R})$.

4.2 The equivalence classes of the remaining points

The analysis of the equivalence of the remaining points is not as straightforward as the preceding analysis. To study the equivalence relation of real projective points of the form $(x:1)$ where x is irrational, we need to develop some properties of the group $GL(2,\mathbb{Z})$.

Product decomposition of elements of $GL(2,\mathbb{Z})$

In order to have a better control over the elements of $GL(2,\mathbb{Z})$, let us exhibit a set of generators.

Let us denote the subgroup $\left\{\pm\left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right);\pm\left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)\right\}$ of $GL(2,\mathbb{Z})$ by Δ. Then Δ is generated by $\pm\left(\begin{smallmatrix}-1&0\\0&1\end{smallmatrix}\right)$.

Lemma 4.4. *For any $S=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in GL(2,\mathbb{Z})$, there exist $Z_1,Z_2\in\Delta$ such that $S=Z_1\begin{pmatrix}|a|&|b|\\|c|&|d|\end{pmatrix}Z_2$.*

Proof. Since any element of Δ is its own inverse, we deduce that there exist $Z_1,Z_2\in\Delta$ such that $S=Z_1\begin{pmatrix}|a|&|b|\\|c|&|d|\end{pmatrix}Z_2$ if and only if $Z_1SZ_2=\begin{pmatrix}|a|&|b|\\|c|&|d|\end{pmatrix}$. Hence it is sufficient to prove that there exist $Z_1,Z_2\in\Delta$ such that $Z_1SZ_2=\begin{pmatrix}|a|&|b|\\|c|&|d|\end{pmatrix}$. In order to prove this, we observe that

$$\begin{pmatrix}\epsilon_1&0\\0&\epsilon_2\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}\epsilon_1'&0\\0&\epsilon_2'\end{pmatrix}=\begin{pmatrix}\epsilon_1\epsilon_1'a&\epsilon_1\epsilon_2'b\\\epsilon_2\epsilon_1'c&\epsilon_2\epsilon_2'd\end{pmatrix}$$

for any $\epsilon_1,\epsilon_2,\epsilon_1',\epsilon_2'\in\{-1,1\}$.

If at least one of the integers a,b,c,d is equal to zero, it is easy to check that there exist $\epsilon_1,\epsilon_2,\epsilon_1',\epsilon_2'\in\{-1,1\}$ such that $\begin{pmatrix}\epsilon_1\epsilon_1'a&\epsilon_1\epsilon_2'b\\\epsilon_2\epsilon_1'c&\epsilon_2\epsilon_2'd\end{pmatrix}=\begin{pmatrix}|a|&|b|\\|c|&|d|\end{pmatrix}$.

If none of the integers a,b,c,d is equal to zero, then, we observe that since $ad-bc=\pm1$, the two products ad and bc have equal sign. Thus either two or four entries have the same sign[1]. Then, the lemma can be easily proven by direct case by case computations. □

Hence any element of $GL(2,\mathbb{Z})$ is given by a matrix with non-negative entries which is multiplied on the left and right by elements of Δ.

Set

$$W=\begin{pmatrix}0&1\\1&0\end{pmatrix}\qquad\text{and}\qquad U=\begin{pmatrix}1&1\\0&1\end{pmatrix}.$$

Then

$$U^r=\begin{pmatrix}1&r\\0&1\end{pmatrix},\quad U^rW=T(r)=\begin{pmatrix}r&1\\1&0\end{pmatrix},\quad WU^r=\begin{pmatrix}0&1\\1&r\end{pmatrix},\quad WU^rW=\begin{pmatrix}1&0\\r&1\end{pmatrix}.$$

Proposition 4.5. *Let $S=\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\in GL(2,\mathbb{Z})$ with $a,b,c,d\geq0$.*

Then S can be written in a unique manner as a product

$$S=WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W$$

where $m\in\mathbb{N}$, $r_0,r_m\geq0$ and if $m\geq2$ then $r_i\geq1$ for all $i=1,...,m-1$. We are going to refer to this product as the product decomposition of S.

[1] The sign distribution of a matrix in $GL(2,\mathbb{Z})$ with non-zero entries, up to identification, can only take the forms $\pm\left(\begin{smallmatrix}+&+\\+&+\end{smallmatrix}\right),\pm\left(\begin{smallmatrix}-&-\\+&+\end{smallmatrix}\right),\pm\left(\begin{smallmatrix}-&+\\+&-\end{smallmatrix}\right),\pm\left(\begin{smallmatrix}-&+\\-&+\end{smallmatrix}\right)$ where $+$ denotes a positive entry (i.e. >0) and $-$ a negative (i.e. <0)

Proof. We prove the existence of such a product by strong induction on b.

1. If $b = 0$, then since $\det(S) = \pm 1$ and since $a, b, c, d \geq 0$, this implies that
$$S = \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} = WU^cW.$$

2. If $b = 1$. Then we have to consider two sub-cases.
   (a) If $\det(S) = ad - bc = -1$, then $c = ad + 1$ and so
   $$S = \begin{pmatrix} a & 1 \\ ad+1 & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ d & 1 \end{pmatrix} \begin{pmatrix} a & 1 \\ 1 & 0 \end{pmatrix} = WU^dWU^aW.$$
   (b) If $\det(S) = ad - bc = 1$, then $c = ad - 1$ and so
   $$S = \begin{pmatrix} a & 1 \\ ad-1 & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ d-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a-1 & 1 \\ 1 & 0 \end{pmatrix} = WU^{d-1}WUWU^{a-1}W.$$

3. Assume that the claim holds true for all $b \in \{0, 1, ..., n-1\}$ for some $n - 1 \geq 1$ and let $b = n > 1$. We distinguish again two sub-cases.
   (a) If $0 \leq a < b$. Then
   $$S = SWW = \left(\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right) W.$$
   Since $b, a, d, c \geq 0$ and since $a < b$, by strong induction, we conclude that $\left(\begin{smallmatrix} b & a \\ d & c \end{smallmatrix}\right) = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W$ for some $m \in \mathbb{N}$, $r_0, r_m \geq 0$ and if $m \geq 2$ then $r_i \geq 1$ for all $i = 1, ..., m-1$. Hence, if $r_m \neq 0$, then
   $$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}WU^0W,$$
   if $r_m = 0$ and $m \geq 1$, then
   $$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}W$$
   and if $r_m = 0$ and $m = 0$, then
   $$S = WU^0WU^0W.$$
   (b) If $a \geq b > 1$, then $a = \mu b + \rho$ with $0 \leq \rho < b$. Since $\det(S) = ad - bc = \pm 1$, we deduce that $\gcd(a, b) = 1$ and that in particular a is not a multiple of b. Thus $\rho \neq 0$. Set $\sigma = c - \mu d$. Then
   $$S = \begin{pmatrix} b & \rho \\ d & \sigma \end{pmatrix} \begin{pmatrix} \mu & 1 \\ 1 & 0 \end{pmatrix}.$$
   Since $S \in GL(2, \mathbb{Z})$ and since $\left(\begin{smallmatrix} \mu & 1 \\ 1 & 0 \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$, we conclude that also $\left(\begin{smallmatrix} b & \rho \\ d & \sigma \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$.
   Assume by contradiction that $\sigma < 0$. Then $b\sigma - \rho d \leq b\sigma$ since $0 < \rho$ and $d \leq 0$. Then, $b\sigma \leq -b < -1$ since $b > 1$ and $\sigma \leq -1$. Hence $\det\left(\begin{smallmatrix} b & \rho \\ d & \sigma \end{smallmatrix}\right) = b\sigma - \rho d < -1$ which contradicts the fact that $\left(\begin{smallmatrix} b & \rho \\ d & \sigma \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$. Hence $\sigma \geq 0$.

Since $b, \rho, d, \sigma \geq 0$ and since $\rho < b$, by strong induction we conclude that $\left(\begin{smallmatrix} b & \rho \\ d & \sigma \end{smallmatrix}\right) = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W$ for some $m \in \mathbb{N}$, $r_0, r_m \geq 0$ and if $m \geq 2$ then $r_i \geq 1$ for all $i = 1, ..., m-1$. Hence, if $r_m \neq 0$, then

$$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}WU^{\mu}W,$$

if $r_m = 0$ and $m \geq 1$, then

$$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}+\mu}W$$

and if $r_m = 0$ and $m = 0$, then

$$S = WU^0WU^{\mu}W.$$

Let us now prove uniqueness of this product. Let

$$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W = WU^{r'_0}WU^{r'_1}W...U^{r'_{n-1}}WU^{r'_n}W$$

for some $m, n \in \mathbb{N}$, $r_0, r_m, r'_0, r'_n \geq 0$ and if $m \geq 2$ then $r_i \geq 1$ for all $i = 1, ..., m-1$ and if $n \geq 2$ then $r'_i \geq 1$ for all $i = 1, ..., n-1$.

1. If $n = m = 0$, the claim is trivial since $WU^{r_0}W = WU^{r'_0}W$ implies $r_0 = r'_0$ and so the two products are equal.

2. Assume by contradiction that one of the two non-negative integers n, m is zero and the other is strictly positive. Assume without loss of generality that $n = 0$ and $m > 0$. Then, we deduce that

   $$WSW = U^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m} = U^{r'_0}.$$

   Thus

   $$U^{r_0}WU^{r_1}W...U^{r_{m-1}}W = U^{r'_0}(U^{r_m})^{-1} = U^{r'_0 - r_m}.$$

   If $m = 1$, this yields that $U^{r_0}W = U^{r'_0 - r_m}$. This implies that $\left(\begin{smallmatrix} r_0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & r'_0 - r_m \\ 0 & 1 \end{smallmatrix}\right)$ which is clearly a contradiction.

   If $m > 1$, we observe that $U^rW = T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Thus, we can use proposition 3.19 to rewrite $U^{r_0}WU^{r_1}W...U^{r_{m-1}}W = D_m = \left(\begin{smallmatrix} p_{m-1} & p_{m-2} \\ q_{m-1} & q_{m-2} \end{smallmatrix}\right)$.

   Since $r_0 \geq 0$ and $r_1, ..., r_{m-1} > 0$, the second part of proposition 3.19 implies that $q_i \geq 1$ for all $i \in \{0, 1, ..., m-1\}$ hence in particular $q_{m-1} \geq 1$ which gives a contradiction since the lower left entry of $U^{r'_0 - r_m}$ is 0.

3. If $n, m \geq 1$, we deduce that

   $$WSW = U^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m} = U^{r'_0}WU^{r'_1}W...U^{r'_{n-1}}WU^{r'_n}.$$

   This implies that

   $$U^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m} \star \infty = U^{r'_0}WU^{r'_1}W...U^{r'_{n-1}}WU^{r'_n} \star \infty.$$

   Since by proposition 3.16, U^{r_m} and $U^{r'_n}$ belong to the stabilizer of ∞, we deduce that

   $$U^{r_0}WU^{r_1}W...U^{r_{m-1}}W \star \infty = U^{r'_0}WU^{r'_1}W...U^{r'_{n-1}}W \star \infty.$$

Since $U^r W = T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$, by the first part of proposition 3.19, we deduce that

$$([r_0, r_1, ..., r_{m-1}] : 1) = ([r'_0, r'_1, ..., r'_{n-1}] : 1).$$

Hence $[r_0, r_1, ..., r_{m-1}] = [r'_0, r'_1, ..., r'_{n-1}]$. Since by construction $[r_0, r_1, ..., r_{m-1}]$ and $[r'_0, r'_1, ..., r'_{n-1}]$ are finite continued fractions that are equal, we deduce by lemma 3.7 that $m - 1 = n - 1$ (and hence $m = n$) and that $r_i = r'_i$ for all $i \in \{0, 1, ..., m-1\}$. By considering again our initial assumption

$$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W = WU^{r'_0}WU^{r'_1}W...U^{r'_{n-1}}WU^{r'_n}W,$$

we deduce from the last conclusion that $(WU^{r_0}WU^{r_1}W...U^{r_{m-1}}W)^{-1}SW = U^{r_m} = U^{r'_n}$. Hence, both products are equal.

□

The last proposition does not only state that each matrix in $GL(2, \mathbb{Z})$ with non-negative entries has a product decomposition, but its proof gives also a general method to construct this product decomposition.

Remark 4.6. *Let $S = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$ with $a, b, c, d \geq 0$. Assume*

$$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W.$$

If $r_m \neq 0$, then

$$\left(\begin{smallmatrix} c & d \\ a & b \end{smallmatrix}\right) = U^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W$$

yields that the continued fraction of $\frac{c}{a} = [r_0, r_1, ..., r_m]$ since

$$(\frac{c}{a} : 1) = \left(\begin{smallmatrix} c & d \\ a & b \end{smallmatrix}\right) \star \infty = U^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W \star \infty = ([r_0, r_1, ..., r_m] : 1)$$

.

Similarly, since

$$\left(\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right) = S^T = (WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W)^T = WU^{r_m}WU^{r_{m-1}}W...U^{r_1}WU^{r_0}W,$$

we deduce that if $r_0 \neq 0$, then $\frac{b}{a} = [r_m, r_{m-1}, ..., r_1, r_0]$.

Combining lemma 4.4 and proposition 4.5 gives the following corollary.

Corollary 4.7. *Any $S \in GL(2, \mathbb{Z})$ can be written as a product*

$$S = Z_1WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}WZ_2$$

where $Z_1, Z_2 \in \Delta$, $m \in \mathbb{N}$, $r_0, r_m \geq 0$ and if $m \geq 2$ then $r_i \geq 1$ for all $i = 1, ..., m-1$. This product is unique up to multiplication to the left and to the right by elements of Δ.

Equivalence of irrational real projective points and the relation between their continued fractions

We proved in the last subsection that each matrix $S \in GL(2,\mathbb{Z})$ has a product decomposition with finitely many terms. This result will now be used to compute the equivalence classes of $P^1(\mathbb{R})$. To understand intuitively what we are going to prove, we will briefly explain a new point of view. Actually, we know that two points $(x:1)$ and $(y:1)$ of $P^1(\mathbb{R})$ are equivalent if there exist $S \in GL(2,\mathbb{Z})$ such that $S \star (x:1) = (y:1)$. Instead of considering two different points $(x:1)$ and $(y:1)$ and analyzing whether they are equivalent or not, we can take one point $(x:1)$ and study the action of any matrix $S \in GL(2,\mathbb{Z})$ on this point. In this way, we find all points that are equivalent to $(x:1)$. Moreover, instead of considering S to act on $(x:1)$ we can consider its product decomposition. Hence, instead of considering the action of each possible element of $GL(2,\mathbb{Z})$, we can simply study the action of the generators at those elements. This intuitive idea will be used in the proof of the next result.

Proposition 4.8. *Let x and y be two irrational numbers. Then $(x:1)$ and $(y:1)$ are equivalent in $P^1(\mathbb{R})$ if and only if*

$$x = [r_0, r_1, ..., r_p, u_0, u_1, ...],$$
$$y = [s_0, s_1, ..., s_q, u_0, u_1, ...],$$

where the sequence of partial quotients of the continued fraction of x after the p-th partial quotient is the same as the sequence of partial quotients of the continued fraction of y after the q-th partial quotient.

Proof. First assume that

$$x = [r_0, r_1, ..., r_p, u_0, u_1, ...],$$
$$y = [s_0, s_1, ..., s_q, u_0, u_1, ...].$$

Set

$$z = [u_0, u_1, ...],$$
$$X = \begin{pmatrix} r_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} r_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} r_p & 1 \\ 1 & 0 \end{pmatrix},$$
$$Y = \begin{pmatrix} s_0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} s_1 & 1 \\ 1 & 0 \end{pmatrix}\cdots\begin{pmatrix} s_q & 1 \\ 1 & 0 \end{pmatrix}.$$

From the first part of proposition 3.19, we deduce that $(x:1) = X\star(z:1)$ and $(y:1) = Y\star(z:1)$. Since $X \in GL(2,\mathbb{Z})$, X has an inverse $X^{-1} \in GL(2,\mathbb{Z})$. Then $X^{-1} \star (x:1) = (z:1)$. Thus $YX^{-1} \star (x:1) = (y:1)$ and $YX^{-1} \in GL(2,\mathbb{Z})$. Thus $(x:1)$ and $(y:1)$ are equivalent.

On the other hand, let $(x:1)$ and $(y:1)$ be equivalent with $x \in \mathbb{R}\setminus\mathbb{Q}$. Then there exists $S \in GL(2,\mathbb{Z})$ such that $(y:1) = S\star(x:1)$. From corollary 4.7, we know that S has a product decomposition $S = Z_1WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}WZ_2$ for some $m \in \mathbb{N}$.

Since x is irrational, theorem 3.27 yields that x has an infinite continued fraction. Let $x = [r'_0, r'_1, r'_2, ...]$. Then we are going to show that $y = [s_0, s_1, ..., s_{m'}, r'_{n'}, r'_{n'+1}, r'_{n'+2}, ...]$ for some $m', n' \in \mathbb{N}$. Since $S = Z_1WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}WZ_2$, it is sufficient to prove that

for any infinite continued fraction $[r_0, r_1, r_2, ...]$ and any $X \in \{Z_1, Z_2, U^{r_0}, U^{r_1}, ..., U^{r_m}, W\}$ we have $X \star ([r_0, r_1, r_2, ...] : 1) = ([s_0, s_1, ..., s_{n_0}, r_n, r_{n+1}, ...] : 1)$ for some $n_0, n \in \mathbb{N}$ $(n > 0)$ and for some $s_0 \in \mathbb{Z}$ and $s_i \geq 1$ for all $i \in \{1, ..., n_0\}$. If this holds, then the action of each matrix in $\{Z_1, Z_2, U^{r_0}, U^{r_1}, ..., U^{r_m}, W\}$ changes only the beginning of a continued fraction and so does then any finite product of them. In order to prove the claim, we are going to carry out a case-by-case proof depending on the value of X.

Consider first $X = Z_1$ (and so at the same time $X = Z_2$). By construction $Z_1 \in \Delta$. Furthermore, we know that Δ is generated by $\pm \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and that the action of S and $-S$ is the same for any $S \in GL(2, \mathbb{Z})$. Thus we consider the action of $X = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ only. There are two sub-cases to consider.
If $r_1 > 1$, we have

$$\begin{aligned}
\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \star ([r_0, r_1, r_2, ...] : 1) &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_1 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\
&= \begin{pmatrix} -r_0 r_1 - 1 & -r_0 \\ r_1 & 1 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\
&= \begin{pmatrix} -r_0 - 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_1 - 1 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\
&= ([-r_0 - 1, 1, r_1 - 1, r_2, r_3, ...] : 1).
\end{aligned}$$

If $r_1 = 1$, then the previous computation can be simplified to

$$\begin{aligned}
\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \star ([r_0, 1, r_2, ...] : 1) &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\
&= \begin{pmatrix} -r_0 - 1 & -r_0 \\ 1 & 1 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\
&= \begin{pmatrix} -r_0 - 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\
&= \begin{pmatrix} -r_0 - 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_2 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_3, r_4, ...] : 1) \\
&= \begin{pmatrix} -r_0 - 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_2 + 1 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_3, r_4, ...] : 1) \\
&= ([-r_0 - 1, r_2 + 1, r_3, r_4, ...] : 1).
\end{aligned}$$

By using the general formula $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \star (x : 1) = \left(\frac{ax+b}{cx+d} : 1\right)$ (since $x \neq -\frac{d}{c}$ since x is irrational), we observe that the action of $\left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ sends $(x : 1)$ to $(-x : 1)$ for any irrational x and so we obtain a formula to compute the continued fraction of the additive inverse of an irrational number from its continued fraction.

Next consider $X = U = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

$$\begin{aligned}
X \star ([r_0, r_1, r_2, ...] : 1) &= \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} r_0 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_1, r_2, ...] : 1) \\
&= \begin{pmatrix} r_0 + 1 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_1, r_2, ...] : 1) \\
&= ([r_0 + 1, r_1, r_2, ...] : 1)
\end{aligned}$$

By using the general formula $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \star (x:1) = \left(\frac{ax+b}{cx+d} : 1\right)$ (since $x \neq -\frac{d}{c}$ since x is irrational), we observe that the action of U sends $(x:1)$ to $(x+1:1)$. In the same fashion, one can show that $U^k \star ([r_0, r_1, r_2, ...] : 1) = ([r_0 + k, r_1, r_2, ...] : 1)$ for any $k \in \mathbb{Z}$ and that the action of U^k sends $(x:1)$ to $(x+k:1)$ for any $k \in \mathbb{Z}$. We deduce the effect of the addition of an integer to an infinite continued fraction.

At last consider $X = W = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. We have to distinguish three sub-cases.
If $r_0 > 0$ then

$$\begin{aligned} X \star ([r_0, r_1, r_2, ...] : 1) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_0 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_1, r_2, ...] : 1) \\ &= ([0, r_0, r_1, ...] : 1). \end{aligned}$$

If $r_0 = 0$, then

$$\begin{aligned} X \star ([r_0, r_1, r_2, ...] : 1) &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_1 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\ &= \begin{pmatrix} r_1 & 1 \\ 1 & 0 \end{pmatrix} \star ([r_2, r_3, ...] : 1) \\ &= ([r_1, r_2, ...] : 1). \end{aligned}$$

If $r_0 < 0$, then we have

$$\begin{aligned} X \star ([r_0, r_1, r_2, ...] : 1) &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}^2 W \star ([r_0, r_1, r_2, ...] : 1) \\ &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} W \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \star ([r_0, r_1, r_2 ...] : 1) \\ &= - \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} W \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \star ([r_0, r_1, r_2 ...] : 1) \\ &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} W \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \star ([r_0, r_1, r_2 ...] : 1) \\ &= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \star \left(W \star \left(\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \star ([r_0, r_1, r_2 ...] : 1) \right) \right) \end{aligned}$$

We can use the case $X = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ to perform the rightmost action which changes the first partial quotients only and which sets the 0-th partial quotient to be $-r_0 - 1 \geq 0$. Then we can use one of the two first sub-cases of $X = W$ to compute the action of $X = W$ and at last we can use again the case $X = \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ to perform the leftmost action. Since the three actions change the beginning of the continued fraction only, so does their product.
By using the general formula $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \star (x:1) = \left(\frac{ax+b}{cx+d} : 1\right)$ (since $x \neq -\frac{d}{c}$ since x is irrational), we observe that the action of W sends $(x:1)$ to $(\frac{1}{x} : 1)$. In particular, this gives a general method to compute the continued fraction of the multiplicative inverse of an irrational number from its continued fraction. □

Example. Let us check the if part of the proposition on an easy example. Consider $\sqrt{2}$. Let $\left(\begin{smallmatrix} 1 & 2 \\ 1 & 3 \end{smallmatrix}\right) \in GL(2, \mathbb{Z})$. Then $\left(\begin{smallmatrix} 1 & 2 \\ 1 & 3 \end{smallmatrix}\right) \star (\sqrt{2} : 1) = \left(\frac{4+\sqrt{2}}{7} : 1\right)$. Thus $(\sqrt{2} : 1)$ and $\left(\frac{4+\sqrt{2}}{7} : 1\right)$ are

equivalent in $P^1(\mathbb{R})$ and hence their continued fractions should end in the same sequence of partial quotients. Indeed, the continuous fraction of $\sqrt{2} = [1, 2, 2, 2, 2, 2, ...]$ and the continuous fraction of $\frac{4+\sqrt{2}}{7} = [0, 2, 1, 2, 2, 2, 2, 2, ...]$.

Remark 4.9. *Notice that, under some assumptions, proposition 4.8 could also be used to obtain the equivalence class of ∞.*

Indeed, by considering $\emptyset$ as the sequence of partial quotients with no partial quotients inside, one can also conclude that finite continued fractions end up with the same sequence after discarding finitely many terms.

On the other hand, one can use exactly the same proof as the one given above to prove the other direction (except that this time $x = -\frac{d}{c}$ might be possible and has to be excluded). This implies that we can also use all the computation rules that we concluded for infinite continued fractions on finite continued fractions.

Remark 4.10. *We showed during the proof of proposition 4.8 that for irrational number x, the real projective points $(x : 1), (-x : 1), (\frac{1}{x} : 1)$ and $(x + n : 1)$ $(n \in \mathbb{Z})$ are equivalent. Hence their continued fractions end up with the same sequence of partial quotients.*

The last example verified the proposition on an easy example, namely on the square root of an integer. In part II, we are going to develop some computation rules to construct the continued fraction of square roots of rational numbers in general. After we do so, the last example seems then trivial. Unfortunately a similar construction does not exist yet for other irrational numbers and that's the strong characteristic of proposition 4.8. Indeed, one can apply its conclusions to any kind of numbers for example transcendental numbers like π and conclude that $\pi, -\pi, \frac{1}{\pi}$ have continued fractions that end up, after some initial partial quotients that may be different, with the same sequence of partial quotients. This is rather difficult to check since we can only use simple algorithms that rely mostly on the Euclidean division algorithm to construct the partial quotients and hence do not end when they compute the continued fraction of irrational numbers.

By proposition 4.5 and proposition 4.8, we deduce that in order to find all the equivalence classes of $P^1(\mathbb{R})$ it is sufficient to find all equivalence classes with representative $(x : 1)$ where $x \in [0, 1] \setminus \mathbb{Q}$ and to include the equivalence class of ∞. This actually amounts to finding all possible different endings of infinite continued fractions. Using Cantor's diagonalization method, one can show that there are uncountably many such endings and hence uncountably many equivalence classes in $P^1(\mathbb{R})$.

Part II
Integral binary quadratic forms

The second part is divided into chapter 5, 6, 7 and 8. The main goal of this part is to introduce integral binary quadratic forms, to define their equivalence, to define the class number and to prove that the class number is finite for positive discriminants. To do so, a certain reduction method is chosen where each integral binary quadratic form is shown to be equivalent to a strongly reduced one.

Chapter 5 consists in an introductory chapter where all the basic definitions of integral binary quadratic forms are set and their equivalence relation is described. Let us highlight that we do only consider indefinite integral binary quadratic forms from this point on.

In chapter 6 the strongly reduced integral binary quadratic forms of a given discriminant are defined. The set of strongly reduced integral binary quadratic forms of a given discriminant will be shown to have finitely many elements and to consist in a complete set of representatives for the equivalence classes of integral binary quadratic forms of this discriminant. Furthermore, the chapter gives two easy verification methods to check if a given integral binary quadratic forms is strongly reduced or not (c.f. proposition 6.10 and lemma 6.11).

In chapter 7 the reduction method is prepared. The chosen method relies on derived forms and not as is common in the literature on GAUSS' right neighbor forms. One strong result and motivation to consider derived forms to perform computations is given in proposition 7.10, where an effective method to compute the derived form with respect to the first root of the corresponding polynomial of a given strongly reduced quadratic form is given.

Finally, chapter 8 concludes the finiteness of the class number for positive discriminants. More precisely, in section 8.1 the result is proven for positive square discriminants. Then in section 8.2 the result is proven for positive non-square discriminants. To this end, the reduction operator given by the derived forms that have been developed in chapter 7, as well as the properties of the strongly reduced integral binary quadratic forms that have been developed in chapter 6 are combined. At last, some rough upper bounds for the class number of positive discriminants will be discovered in section 8.3.

5 Integral binary quadratic forms

Sources: [1, p.1-8], [8, p.318-321], [9, p.1-12], [3, p.150-152,155-159], [10, p.9,12-15,21-23], [11, art. 153-154,157-159]
Other approaches: [12, p.354-356]
Another source which is particularly difficult to decompose into different chapters is given by LAGRANGE's *Recherches d'Arithmétiques* which can be found in [7, 3rd volume p. 695-795]. This source develops the classical approach (which is modernized by WEIL) which is nearest to our development and shall be considered as complementing source for the second and third part of this document.

Integral binary quadratic forms belong to the classical mathematical objects of number theory. Several mathematicians studied their properties and in 1801 GAUSS published in [11] a general method to completely classify them.

Using the strong results that GAUSS developed, mathematicians found a connection between continued fractions and integral binary quadratic forms and proceeded to deduce properties of continued fractions from the known theory about integral binary quadratic forms.

Although this is a successful way to study both theories, we are not going to follow this path. We are going to fusion both theories so that we obtain results in both fields in a simple, elegant and modern manner. This will finally lead to a classification of integral binary quadratic forms and a basic algorithm to obtain this classification.

Since several definitions throughout the rest of the document are non-standard, we would like to warn the reader to not confuse them with the (gaussian) standard definitions.

The definition of integral binary quadratic forms

Definition 5.1. An *integral binary quadratic form* is a homogeneous polynomial in two variables of degree two with integral coefficients

$$Q(X,Y) = aX^2 + bXY + cY^2 \in \mathbb{Z}[X,Y].$$

Its *discriminant* is $D = b^2 - 4ac$ (and satisfies $D \equiv 0, 1 \mod 4$).

Since we are going to work with integral binary quadratic forms only, there should be no confusion in calling them quadratic forms in the rest of the document.

Set

$$A = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix} \text{ and } V = \begin{pmatrix} X \\ Y \end{pmatrix}.$$

Then

$$\frac{1}{2} V^T A V = Q(X,Y)$$

and $\det(A) = -D$. We say that the matrix A *represents* the quadratic form $Q(X,Y)$.

In order to shorten the notation and for better perception of the coefficients of the quadratic forms, one usually uses the notation

$$\langle a, b, c \rangle_{X,Y} = aX^2 + bXY + cY^2.$$

Quadratic forms can be distinguished by the value of their discriminant.

Definition 5.2. A quadratic form $Q(X,Y) = aX^2 + bXY + cY^2$ with discriminant D is called:

1. *indefinite* if $D > 0$;
2. *semi-definite* if $D = 0$;
3. *definite* if $D < 0$.

Example. $Q(X,Y) = 2X^2 + 3XY + Y^2 = \langle 2,3,1\rangle_{X,Y}$ is indefinite whereas $Q(X,Y) = \langle 1,0,1\rangle_{X,Y}$ is definite.

Although the general theory about quadratic forms is quite interesting, we are going to consider indefinite quadratic forms only. For more general developments, we refer the reader to [9, p.13-20], [3, p.170-175], [10, p.85-105], [12, p.356], [11, art. 171-181,215-221].

Definition 5.3. The set of all quadratic forms of discriminant D is denoted by $Quad(D)$.

Example. $Q(X,Y) = 2X^2 + 3XY + Y^2 \in Quad(1)$.

Since the sets $Quad(D)$ partition the indefinite quadratic forms, they are a great tool to classify them. In this logic, we are going to study those sets in detail.

$GL(2,\mathbb{Z})$ acts on $Quad(D)$

Proposition 5.4. *$GL(2,\mathbb{Z})$ acts on $Quad(D)$ by*

$$Q(X,Y) \circ S = \frac{1}{2} V^T A V \circ S = \frac{1}{2} V^T S^T A S V$$

for all $S = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$.

Remark 5.5. *Actually the action of $S \in GL(2,\mathbb{Z})$ on $Q(X,Y) \in Quad(D)$ can be seen as a linear change of the variables X and Y. Indeed, one observes that if $S = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$, then we have*

$$\begin{aligned} Q(X,Y) \circ S &= \frac{1}{2} V^T S^T A S V \\ &= \frac{1}{2} (X,Y) \left(\begin{smallmatrix} p & r \\ q & s \end{smallmatrix}\right) A \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \left(\begin{smallmatrix} X \\ Y \end{smallmatrix}\right) \\ &= \frac{1}{2} (pX + qY, rX + sY) A \left(\begin{smallmatrix} pX+qY \\ rX+sY \end{smallmatrix}\right) \\ &= Q(pX + qY, rX + sY). \end{aligned}$$

Remark 5.6. *Notice that for all $GL(2,\mathbb{Z})$ we have $Q(X,Y) \circ S = Q(X,Y) \circ (-S)$.*

Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ with representative matrix $A = \left(\begin{smallmatrix} 2a & b \\ b & 2c \end{smallmatrix}\right)$ and $S = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$. Then the matrix that represents $Q(X,Y) \circ S$ is given by

$$\begin{aligned} S^T A S &= \begin{pmatrix} p & q \\ r & s \end{pmatrix}^T \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} \\ &= \begin{pmatrix} 2ap^2 + 2bpr + 2cr^2 & 2apq + b(ps+qr) + 2crs \\ 2apq + b(ps+qr) + 2crs & 2aq^2 + 2bqs + 2cs^2 \end{pmatrix} \end{aligned}$$

Thus $\langle a', b', c' \rangle_{X,Y} = Q(X,Y) \circ S$ is given by

$$\begin{cases} a' &= & Q(p,r) \\ b' &= & (q\ s) \left(\begin{smallmatrix} 2a & b \\ b & 2c \end{smallmatrix}\right) \left(\begin{smallmatrix} p \\ r \end{smallmatrix}\right) \\ c' &= & Q(q,s) \end{cases}$$

We deduce the following proposition.

Proposition 5.7. *If $Q(X,Y) = \langle a,b,c \rangle_{X,Y} \in Quad(D)$ and $S = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$, then*

$$Q(X,Y) \circ S = \langle Q(p,r),\ (q,s) \left(\begin{smallmatrix} 2a & b \\ b & 2c \end{smallmatrix}\right) \left(\begin{smallmatrix} p \\ r \end{smallmatrix}\right),\ Q(q,s) \rangle_{X,Y}$$

Equivalence of quadratic forms and the class number

Through the above group action of $GL(2,\mathbb{Z})$, we can define an equivalence relation on quadratic forms as follows.

Definition 5.8. Two quadratic forms $Q(X,Y) \in Quad(D)$ and $Q'(X,Y) \in Quad(D)$ are said to be *equivalent (in the wide sense) in $Quad(D)$* if $Q(X,Y) = Q'(X,Y) \circ S$ for some $S \in GL(2,\mathbb{Z})$ (i.e. $\det(S) = \pm 1$).
More precisely, we say that $Q'(X,Y)$ and $Q(X,Y) = Q'(X,Y) \circ S$ are

1. *properly equivalent* if $\det(S) = 1$ and
2. *improperly equivalent* if $\det(S) = -1$.

Example. Since $\langle 1,3,3 \rangle_{X,Y} = \langle 1,1,3 \rangle_{X,Y} \circ \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, we deduce that $X^2+XY+3Y^2$ is equivalent (in the wide sense) and properly equivalent to $X^2+3XY+3Y^2$.
Since $\langle 1,-1,6 \rangle_{X,Y} = \langle 1,1,3 \rangle_{X,Y} \circ \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, we deduce that $X^2+XY+3Y^2$ is equivalent (in the wide sense) and improperly equivalent to $X^2-XY+6Y^2$.

Remark 5.9. *Observe that a quadratic form $Q(X,Y) \in Quad(D)$ can be properly and improperly equivalent to another quadratic form $Q'(X,Y)$ or itself.*

Example. Since $Q(X,Y) = \langle 1,-2,1 \rangle_{X,Y}$ satisfies $Q(X,Y) \circ \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) = \langle 1,-2,1 \rangle_{X,Y}$ and $Q(X,Y) \circ \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right) = \langle 1,-2,1 \rangle_{X,Y}$, we deduce that $Q(X,Y)$ is properly and improperly equivalent to itself.

Let $H(D)$ denote the quotient space $Quad(D)/GL(2,\mathbb{Z})$ (i.e. the set of equivalence classes in the wide sense) and let us define the *class number* of $Quad(D)$ (denoted by $h(D)$) to be the number of equivalence classes in $Quad(D)$.

Remark 5.10. *Observe that equivalence (in the wide sense) and proper equivalence are equivalence relations in $Quad(D)$, whereas improper equivalence is not an equivalence relation since transitivity is not fulfilled.*

In the following, we are going to say that two quadratic forms are equivalent when we mean equivalent in the wide sense.

Before we close this chapter, let us consider two general examples of equivalent forms which will reappear later on.

Example. Let $Q(X,Y) = aX^2+bXY+cY^2$ and let $Q'(X,Y) = cX^2+bXY+aY^2$. Since $Q(X,Y) = Q'(Y,X) \circ \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $Q(X,Y)$ is equivalent to $Q(Y,X)$ (thus quadratic forms are equivalent under the exchange of their variables).

Example. Let $Q(X,Y) = aX^2 + bXY + cY^2$ and $Q'(X,Y) = aX^2 - bXY + cY^2$. Since $Q'(X,Y) = Q(X,Y) \circ \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $Q(X,Y)$ is equivalent to $Q'(X,Y)$.

Let us highlight the fact that quadratic forms that are equivalent share a lot of properties for example they represent the same numbers. Hence, in order to study indefinite quadratic forms, it is sufficient to study their equivalence classes. In other words, it is sufficient to analyze the sets $H(D)$ for all $D \in \mathbb{N} \setminus \{0\}$.

6 Quadratic surds and reduced quadratic forms

```
Sources:  [1, p.12-15], [8, p.321-322,353], [3, p.159-160], [4, p.96-104],
          [10, p.79-82,107-111]
Other approaches:  [9, p.21-22]
```

Chapter note: Assume throughout this chapter that $D \in \mathbb{N} \setminus \mathbb{N}^2$.

In order to find the equivalence classes of $Quad(D)$, we would like to reduce the set of all quadratic forms of discriminant D to a smaller set of quadratic forms which includes at least one representative for each equivalence class of $Quad(D)$. In the second part of theorem 8.4 in chapter 8, we are going to see that such a set is given by the set of all strongly reduced quadratic forms of discriminant D. Before we can prove this result, we have to define what it means for a quadratic form to be strongly reduced. Since there are again several definitions of strongly reducedness that can be found in the literature, where most of them are given synthetically by a certain comparison of the coefficients that define the quadratic form, we would like to give a more natural definition and show how this can be seen by comparing the coefficients. That's why we start this section by defining so called quadratic surds which will be shown to partially define quadratic forms.

6.1 The definition of quadratic surds

Definition 6.1. A *quadratic surd* (or *quadratic irrationality*) is a real irrational number which is a root of a quadratic polynomial with integral coefficients.

Example. $\sqrt{2}$ is a quadratic surd since it is the root of $X^2 - 2$ and irrational.
$\frac{7}{19}$ is not a quadratic surd since it is not irrational.
π and $\sqrt[3]{2}$ are not quadratic surds since they are not roots of a quadratic equation with integral coefficients.

Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ and consider the polynomial $Q(T,1) = aT^2 + bT + c \in \mathbb{Z}[T]$. Since $D \in \mathbb{N} \setminus \mathbb{N}^2$, the roots of $Q(T,1)$ are distinct and given by the two quadratic surds

$$x^+ = \frac{-b+\sqrt{D}}{2a} = \frac{-2c}{b+\sqrt{D}} \quad \text{and} \quad x^- = \frac{-b-\sqrt{D}}{2a} = \frac{-2c}{b-\sqrt{D}}.$$

We define $\chi_1 = \max\{x^+, x^-\}$ to be the *first root* of $Q(T,1)$ and $\chi_2 = \min\{x^+, x^-\}$ to be the *second root* of $Q(T,1)$. One can easily verify that if $a > 0$ then $\chi_1 = x^+$ and $\chi_2 = x^-$ and if $a < 0$ then $\chi_1 = x^-$ and $\chi_2 = x^+$. In particular, we have $\{x^+, x^-\} = \{\chi_1, \chi_2\}$.

Example. Consider $Q(X,Y) = X^2 - 2Y^2 \in Quad(8)$, then the first root of $Q(T,1) = T^2 - 2$ is $\chi_1 = x^+ = \sqrt{2}$ and the second root of $Q(T,1)$ is $\chi_2 = x^- = -\sqrt{2}$.

Proposition 6.2. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y}$ be a quadratic form and let $S = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$. Then the roots of $Q'(X,Y) = Q(X,Y) \circ S = \langle a',b',c'\rangle_{X,Y}$ are*

$$\frac{-b' + \det(S)\sqrt{D}}{2a'} \quad \textit{and} \quad \frac{-b' - \det(S)\sqrt{D}}{2a'}.$$

Consider the roots x^+ and x^- of $Q(T,1)$ where $Q(X,Y) \in Quad(D)$ $(0 < D \notin \mathbb{N}^2)$. Let $S \in GL(2,\mathbb{Z})$ and let $Q'(X,Y) = Q(X,Y) \circ S = \langle a', b', c' \rangle_{X,Y}$, then a brute force computation shows that

$$S^{-1} \star (x^+ : 1) = \left(\frac{-b' + \det(S)\sqrt{D}}{2a'} : 1 \right) \quad \text{and} \quad S^{-1} \star (x^- : 1) = \left(\frac{-b' - \det(S)\sqrt{D}}{2a'} : 1 \right).$$

Hence, the above conclusion gives us a method to generate the roots of a polynomial given by a quadratic form in terms of the roots of a polynomial given by a quadratic form that is equivalent to the first one. We verify this with an example.

Example. Let $Q(X,Y) = \langle 1, 2, -4 \rangle_{X,Y} \in Quad(20)$ and let $S = \left(\begin{smallmatrix} -1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$. Then $Q_1(X,Y) = Q(X,Y) \circ S = \langle 1, -4, -1 \rangle_{X,Y}$. The roots of $Q(T,1)$ are $x^+ = -1 + \sqrt{5}$ and $x^- = -1 - \sqrt{5}$. The roots of $Q_1(T,1)$ are $x_1^+ = 2 + \sqrt{5}$ and $x_1^- = 2 - \sqrt{5}$. Furthermore the inverse of S is given by $S^{-1} = \left(\begin{smallmatrix} -1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = S$. Thus $S^{-1} \star (x^+ : 1) = S^{-1} \star (-1 + \sqrt{5} : 1) = \left(\frac{-(-1+\sqrt{5})+1}{1} : 1 \right) = (2 - \sqrt{5} : 1) = (x_1^- : 1)$ and similarly $S^{-1} \star (x^- : 1) = (x_1^+ : 1)$.

Notice that the above conclusion states that if the two quadratic forms $Q(X,Y)$ and $Q^{\triangleright}(X,Y)$ are equivalent, then the points $(\chi_1 : 1)$ and $(\chi_1^{\triangleright} : 1)$ or the points $(\chi_1 : 1)$ and $(\chi_2^{\triangleright} : 1)$ are equivalent in $P^1(\mathbb{R})$ where χ_1 denotes the first root of $Q(T,1)$, $\chi_1^{\triangleright}$ the first and $\chi_2^{\triangleright}$ the second root of $Q^{\triangleright}(T,1)$.

Unfortunately the converse does not hold. One can prove that the two forms $Q(X,Y) = \langle 2, -2, -1 \rangle_{X,Y}$ and $-Q(X,Y)$ are not equivalent (c.f. appendix D). However the first root of $Q(T,1)$ is equal to the first root of $-Q(T,1)$ and thus trivially equivalent to itself.

Corollary 6.3. *If x is a quadratic surd and $(y : 1) = S \star (x : 1)$ where $S \in GL(2,\mathbb{Z})$, then y is a quadratic surd.*

6.2 Weakly and strongly reduced quadratic surds

Definition 6.4. Let $x = a + b\sqrt{D}$ with $a, b \in \mathbb{Q}$ and $D \in \mathbb{N} \setminus \mathbb{N}^2$. Then the *conjugate* of x is the number $x' = a - b\sqrt{D}$.

Observe that the function $x \mapsto x'$ is the non-trivial automorphism of $\mathbb{Q}(\sqrt{D})$. Hence the computation rules of $x \mapsto x'$ are well studied and will be used without further precisions hereinafter.

Definition 6.5. We say that a quadratic surd x is *weakly reduced* (in the sense of Lagrange) if $x > 0$ and $x' < 0$ and that x is *strongly reduced* if $x > 1$ and $-1 < x' < 0$.

Example. Since the conjugate of $1+\sqrt{2}$ is $1-\sqrt{2}$ and since $1+\sqrt{2} > 1$ and $-1 < 1-\sqrt{2} < 0$, the quadratic surd $1 + \sqrt{2}$ is strongly reduced.

Let $Q(X,Y) = \langle a, b, c \rangle_{X,Y}$ be a quadratic form with positive non-square discriminant D. Let x be a root of $Q(T,1)$, then we know that x is a quadratic surd. Let $x = x_1 + x_2\sqrt{D}$ for some $x_1, x_2 \in \mathbb{Q}$. Since $Q(x,1) = 0$, we deduce that $ax_1^2 + ax_2^2 D + bx_1 + c = 0$ and $2ax_1x_2 + bx_2 = 0$. Then $x' = x_1 - x_2\sqrt{D}$ is also a root of $Q(T,1)$ since $Q(x',1) = ax_1^2 - 2ax_1x_2\sqrt{D} + ax_2D + bx_1 - bx_2\sqrt{D} + c = 0$. In particular this shows that $\{x, x'\} = \{x^+, x^-\} = \{\chi_1, \chi_2\}$.
Hence, we deduce that the first root of a polynomial of second degree can be weakly or strongly reduced, whereas the second root cannot be weakly or strongly reduced since $\chi_1 > \chi_2$.

Lemma 6.6. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y}$ and $Q^{\triangleright}(X,Y) = \langle a_{\triangleright}, b_{\triangleright}, c_{\triangleright}\rangle_{X,Y}$ be two quadratic forms. Let $D \in \mathbb{N}\backslash\mathbb{N}^2$, $D_{\triangleright} \in \mathbb{N}\backslash\mathbb{N}^2$ be the discriminants of $Q(X,Y)$ and $Q^{\triangleright}(X,Y)$ respectively. Let χ_1 be the first root of $Q(T,1)$ and let $\chi_1^{\triangleright}$ be the first root of $Q^{\triangleright}(T,1)$.*

1. *If the sign of a and the sign of $a_{\triangleright}$ are equal, then $Q(X,Y) = Q^{\triangleright}(X,Y)$ if and only if $D = D_{\triangleright}$ and $\chi_1 = \chi_1^{\triangleright}$.*
2. *If the sign of a and the sign of $a_{\triangleright}$ are opposite, then $Q(X,Y) = -Q^{\triangleright}(X,Y)$ if and only if $D = D_{\triangleright}$ and $\chi_1 = \chi_1^{\triangleright}$.*

Proof. If $Q(X,Y) = \pm Q^{\triangleright}(X,Y)$, then their corresponding polynomials have the same roots and their discriminant are equal.

Conversely, if $D = D_{\triangleright}$ and $\chi_1 = \chi_1^{\triangleright} = \frac{-b+\sqrt{D_1}}{2a}$, then since $D \in \mathbb{N} \setminus \mathbb{N}^2$, the second root of $Q(T,1)$ and $Q^{\triangleright}(T,1)$ is given by $(\chi_1^{\triangleright})'$. Hence $Q^{\triangleright}(X,Y) = kQ(X,Y)$ for some $k \in \mathbb{Z}$. Since both quadratic forms have the same discriminant, this yields that $Q^{\triangleright}(X,Y) = \pm Q(X,Y)$. The claim follows by comparing the signs of a and $a_{\triangleright}$. □

6.3 Reduced quadratic forms

Definition 6.7. Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$. Let χ_1 be the first root of $Q(T,1)$. We call $Q(X,Y)$ *strongly reduced* (respectively *weakly reduced*) if χ_1 is strongly reduced (respectively weakly reduced).

Example. Let $Q(X,Y) = X^2 - 2XY - Y^2$. Then the first root of $Q(T,1)$ is $\chi_1 = x^+ = 1+\sqrt{2}$. Since $1+\sqrt{2}$ is strongly reduced, so is then $Q(X,Y)$.

For any $Q(X,Y) \in Quad(D)$, the polynomials $Q(T,1)$ and $-Q(T,1)$ have the same roots. Thus, we deduce the following proposition.

Proposition 6.8. *$Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ is strongly reduced if and only if $-Q(X,Y) = \langle -a,-b,-c\rangle_{X,Y}$ is strongly reduced.*

Unfortunately, the direct method to check whether a quadratic form is weakly reduced, strongly reduced or none of these demands a lot of computational work. In particular, one has to determine the first root of the corresponding polynomial and check whether it is weakly or strongly reduced. Fortunately, one can bypass these computations and deduce the reducedness of a quadratic form by a direct comparison of its coefficients.

Lemma 6.9. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$. Then $Q(X,Y)$ is weakly reduced if and only if $ac < 0$.*

Proof. Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ and observe that the two roots of $Q(T,1)$ are given by $x^+ = \frac{-b+\sqrt{D}}{2a} = \frac{-2c}{b+\sqrt{D}}$ and $x^- = \frac{-b-\sqrt{D}}{2a} = \frac{-2c}{b-\sqrt{D}}$. Let χ_1 be the first and χ_2 be the second root of $Q(T,1)$. Then $\chi_1\chi_2 = \frac{c}{a}$.

By definition $Q(X,Y)$ is weakly reduced if and only if $\chi_1\chi_2 < 0$. Since $\chi_1\chi_2 = \frac{c}{a}$, we deduce that $Q(X,Y)$ is weakly reduced if and only if $\frac{c}{a} < 0$ which holds if and only if $ac < 0$. □

Proposition 6.10. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$. The following are equivalent:*

1. $Q(X,Y)$ *is strongly reduced.*
2. $a > 0$, $-\sqrt{D} < b < 0$, $c < 0$ *and* $b + \sqrt{D} < 2a < -b + \sqrt{D}$
   or
   $a < 0$, $0 < b < \sqrt{D}$, $c > 0$ *and* $-b + \sqrt{D} < -2a < b + \sqrt{D}$.
3. $a > 0$, $-\sqrt{D} < b < 0$, $c < 0$ *and* $b + \sqrt{D} < -2c < -b + \sqrt{D}$
   or
   $a < 0$, $0 < b < \sqrt{D}$, $c > 0$ *and* $-b + \sqrt{D} < 2c < b + \sqrt{D}$.

The conditions above are called Gauss conditions.

Proof. Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ and observe that the two roots of $Q(T,1)$ are given by $x^+ = \frac{-b+\sqrt{D}}{2a} = \frac{-2c}{b+\sqrt{D}}$ and $x^- = \frac{-b-\sqrt{D}}{2a} = \frac{-2c}{b-\sqrt{D}}$. Let χ_1 be the first and χ_2 be the second root of $Q(T,1)$. Then $\chi_1\chi_2 = \frac{c}{a}$.

1. Assume first that $Q(X,Y)$ is strongly reduced reduced. Since $Q(X,Y)$ is weakly reduced, lemma 6.9 implies that $ac < 0$. Hence $D = b^2 - 4ac > b^2$ and so $0 \leq |b| < \sqrt{D}$.
   (a) If $a > 0$, then $\chi_1 = x^+$, $\chi_2 = x^-$ and $c < 0$. Since $Q(X,Y)$ is strongly reduced, x^+ is strongly reduced and so $1 < x^+$ and $-1 < x^- < 0$. Firstly, $1 < x^+$ implies that $2a < -b + \sqrt{D}$ and $b + \sqrt{D} < -2c$ and secondly $-1 < x^- < 0$ implies that $b + \sqrt{D} < 2a$ and $-2c < -b + \sqrt{D}$. We observe in particular that $b + \sqrt{D} < 2a < -b + \sqrt{D}$ and hence $b < -b$ and so $-\sqrt{D} < b < 0$.
   (b) If $a < 0$, then $\chi_1 = x^-$, $\chi_2 = x^+$ and $c > 0$. Since $Q(X,Y)$ is strongly reduced, x^- is strongly reduced and so $1 < x^-$ and $-1 < x^+ < 0$. Firstly, $1 < x^-$ implies that $-2a < b + \sqrt{D}$ and $-b + \sqrt{D} > 2c$ and secondly $-1 < x^+ < 0$ implies that $-b + \sqrt{D} < -2a$ and $2c < b + \sqrt{D}$. We observe in particular that $-b + \sqrt{D} < -2a < b + \sqrt{D}$ and hence $-b < b$ and so $0 < b < \sqrt{D}$.
2. Assume next that the conditions in *(2.)* hold.
   (a) If $a > 0$, then $\chi_1 = x^+$ and $\chi_2 = x^-$. Since $0 < 2a < -b + \sqrt{D}$, we deduce that $x^+ > 1$ and since $0 < b + \sqrt{D} < 2a$ we deduce that $-1 < x^- < 0$.
   (b) If $a < 0$, then $\chi_1 = x^-$ and $\chi_2 = x^+$. Since $0 < -2a < b + \sqrt{D}$, we deduce that $x^- > 1$ and since $0 < -b + \sqrt{D} < -2a$ we deduce that $-1 < x^+ < 0$.
3. Finally assume that the conditions in *(3.)* hold.
   (a) If $a > 0$, then $\chi_1 = x^+$ and $\chi_2 = x^-$. Since $0 < -2c < -b + \sqrt{D}$, we deduce that $0 > x^- > -1$ and since $b + \sqrt{D} < -2c$ we deduce that $x^- > 1$.
   (b) If $a < 0$, then $\chi_1 = x^-$ and $\chi_2 = x^+$. Since $0 < 2c < b + \sqrt{D}$, we deduce that $0 > x^+ > -1$ and since $0 < -b + \sqrt{D} < 2c$ we deduce that $1 < x^-$.

□

Example. Let $Q(X,Y) = X^2 - 4XY - 13Y^2 \in Quad(68)$. Since $1 > 0$ and $-13 < 0$, the quadratic form $Q(X,Y)$ is weakly reduced. However, since $\sqrt{68} - 4 > 4 > 2$, the Gauss conditions are not satisfied and so the quadratic form $Q(X,Y)$ is not strongly reduced.

Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$. Let χ_1 and χ_2 the roots of $Q(T,1)$. Then an easy computation shows that $\chi_1 + \chi_2 = -\frac{b}{a}$ and $\chi_1\chi_2 = \frac{c}{a}$. Furthermore, we deduce that

$$Q(1,0)Q(0,1) = ac = a^2\frac{c}{a} = a^2\chi_1\chi_2$$
$$Q(1,0)Q(1,1) = a(a+b+c) = a^2\left(1+\frac{b}{a}+\frac{c}{a}\right) = a^2(1-(\chi_1+\chi_2)+\chi_1\chi_2) = a^2(1-\chi_1)(1-\chi_2)$$
$$Q(1,0)Q(-1,1) = a(a-b+c) = a^2\left(1-\frac{b}{a}+\frac{c}{a}\right) = a^2(1+(\chi_1+\chi_2)+\chi_1\chi_2) = a^2(1+\chi_1)(1+\chi_2)$$

Lemma 6.11. *$Q(X,Y) = \langle a,b,c\rangle_{X,Y}$ is strongly reduced if and only if*

$$a > 0, \quad c < 0, \quad a+b+c < 0, \quad a-b+c > 0$$

or

$$a < 0, \quad c > 0, \quad a+b+c > 0, \quad a-b+c < 0.$$

Proof. Let $Q(X,Y)$ be strongly reduced and let χ_1 be the first and let χ_2 be the second root of $Q(T,1)$.
Suppose $a > 0$ then $c < 0$ by the Gauss conditions. Furthermore since $\chi_1 > 1$ and $\chi_2 < 0$, the above observation yields firstly that $0 > a^2(1-\chi_1)(1-\chi_2) = a(a+b+c)$ and thus $a+b+c < 0$ and secondly that $0 < a^2(1+\chi_1)(1+\chi_2) = a(a-b+c)$ and thus $a-b+c > 0$.
Suppose $a < 0$ then $c > 0$ by the Gauss conditions. Furthermore since $\chi_1 > 1$ and $\chi_2 < 0$, the above observation yields firstly that $0 > a^2(1-\chi_1)(1-\chi_2) = a(a+b+c)$ and thus $a+b+c > 0$ and secondly that $0 < a^2(1+\chi_1)(1+\chi_2) = a(a-b+c)$ and thus $a-b+c < 0$.

Conversely, assume that the conditions hold.
If $a > 0$. Then since $a > 0$ and $a+b+c < 0$, we have $0 > a(a+b+c) = a^2(1-\chi_1)(1-\chi_2)$ and thus (since $\chi_1 > \chi_2$) we have $1-\chi_1 < 0 < 1-\chi_2$ whence $\chi_1 > 1$. Moreover $a > 0$ and $a-b+c > 0$ implies $0 < a(a-b+c) = a^2(1+\chi_1)(1+\chi_2)$ whence $\chi_2 > -1$. Last of all since $a > 0$ and $c < 0$, we have $0 > ac = a^2\chi_1\chi_2$ and so $\chi_2 < 0$. Hence $Q(X,Y)$ is strongly reduced.
If $a < 0$. Then since $a < 0$ and $a+b+c > 0$, we have $0 > a(a+b+c) = a^2(1-\chi_1)(1-\chi_2)$ and thus (since $\chi_1 > \chi_2$) we have $1-\chi_1 < 0 < 1-\chi_2$ whence $\chi_1 > 1$. Moreover $a < 0$ and $a-b+c < 0$ implies $0 < a(a-b+c) = a^2(1+\chi_1)(1+\chi_2)$ whence $\chi_2 > -1$. Last of all since $a < 0$ and $c > 0$, we have $0 > ac = a^2\chi_1\chi_2$ and so $\chi_2 < 0$. Hence $Q(X,Y)$ is strongly reduced. □

Example. By the first example of this section, we know that $Q(X,Y) = \langle 1,-2,-1\rangle_{X,Y}$ is strongly reduced. Moreover, $1 > 0$, $-1 < 0$, $1-2-1 = -2 < 0$ and $1+2-1 = 4 > 0$, which verifies the lemma.

7 Derived forms

Sources: [1, p.15], [8, p.353-354], [10, p.112-113]
Other approaches: [9, p.22-23], [11, art.160]

Chapter note: Assume throughout this chapter that $D \in \mathbb{N} \setminus \mathbb{N}^2$.

In order to prove that the set of all strongly reduced forms of discriminant D includes at least one representative for each equivalence class of $Quad(D)$, we need to prove that each quadratic form of discriminant D is equivalent to a strongly reduced one.

To do so, one can tempt to exhibit a chain of equivalent quadratic forms which ends up at a strongly reduced quadratic form. Any such approach needs a so called *reduction* method that defines the chain of equivalent quadratic forms. For this purpose, GAUSS defined the *right neighbor* of a quadratic form and *right neighbor chains*. However this method is not suited for our construction. That's why we are going to consider so called *derived forms* and *cycles of derived forms*.

7.1 The derived form of a quadratic form

Definition 7.1. Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$. Let x be any root of $Q(T,1)$. Then the *(first) derived form of $Q(X,Y)$ with respect to x* is

$$\partial_x Q(X,Y) = \partial_x^1 Q(X,Y) = Q(X,Y) \circ T(\lfloor x \rfloor), \quad \text{where} \quad T(r) = \begin{pmatrix} r & 1 \\ 1 & 0 \end{pmatrix}.$$

We observe that $\partial_x Q(X,Y)$ and $Q(X,Y)$ are improperly equivalent and by using proposition 5.7, we get the following result.

Proposition 7.2. *Let $Q(X,Y) = \langle a,b,c \rangle_{X,Y} \in Quad(D)$ and let x be any root of $Q(T,1)$. Then $\partial_x Q(X,Y) = \langle a_1, b_1, c_1 \rangle_{X,Y}$ is given by:*

$$\begin{cases} a_1 = & a \lfloor x \rfloor^2 + b \lfloor x \rfloor + c \\ b_1 = & b + 2a \lfloor x \rfloor \\ c_1 = & a \end{cases}$$

Example. Let $Q(X,Y) = \langle 1, -4, -13 \rangle_{X,Y} \in Quad(68)$. Then, $\chi_1 = x^+ = 2 + \sqrt{17}$ is the first and $\chi_2 = x^- = 2 - \sqrt{17}$ is the second root of $Q(T,1)$. Since $\lfloor 2 + \sqrt{17} \rfloor = 6$, we deduce that $\partial_{\chi_1} Q(X,Y) = \langle -1, 8, 1 \rangle_{X,Y}$ and since $\lfloor 2 - \sqrt{17} \rfloor = -3$, we deduce that $\partial_{\chi_2} Q(X,Y) = \langle 8, -10, 1 \rangle_{X,Y}$.

Remark 7.3. *Since the definition of derived forms is quite similar to* GAUSS' *definition of left neighbors, we want to highlight the difference. For any quadratic form $Q(X,Y) = \langle a,b,c \rangle_{X,Y} \in Quad(D)$, the left neighbor of $Q(X,Y)$ (in the sense of* GAUSS*) is given in our notation by $\partial_x \left(Q(X,Y) \circ \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix} \right) \right)$ where $Q(X,Y) \circ \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix} \right) = \langle a, -b, c \rangle_{X,Y}$ and where x denotes the first root of $aT^2 - bT + c$.*

Since $Q(T,1)$ and $-Q(T,1)$ have the same roots, proposition 7.2 implies the following corollary.

Corollary 7.4. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ and let x be any root of $Q(T,1)$. Then $\partial_x(-Q(X,Y)) = -\partial_x Q(X,Y)$.*

Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ and let x be any root of $Q(T,1)$. By a direct computation, one deduces that $\partial x = \frac{1}{x-\lfloor x \rfloor}$ and $\frac{1}{x'-\lfloor x \rfloor}$ are the roots of $\partial_x Q(T,1)$.

Example. Consider $Q(X,Y) = \langle 1,-4,-13\rangle_{X,Y} \in Quad(68)$ and $\partial_{\chi_1} Q(X,Y) = \langle -1,8,1\rangle_{X,Y}$. Then the two roots of $\partial_{\chi_1} Q(T,1)$ are given by $\partial\chi_1 = 4+\sqrt{17}$ and $\frac{1}{2-\sqrt{17}-6} = 4-\sqrt{17}$.

In what follows, we will discover some interesting properties of derived forms. However some of those properties will not take place if we take the (first) derived form of a given quadratic form, but only if we take again the derived form of its derived form and so on.

Unfortunately, we have always to choose the root with whose respect we compute the derived form. Thus, it might be difficult to determine, prior to some explicit computations, with respect to which root we compute the derived form. Even worse, one could think that one cannot even say anything about those forms. Although we do not have much information about derived forms of derived forms, we know one very strong fact for sure.

By computing the derived form of a given quadratic form $Q(X,Y)$ with respect to the root χ of $Q(T,1)$, one knows beforehand that one root of $\partial_\chi Q(T,1)$ will be $\partial\chi$. Thus, we are able to define the derived form of $\partial_\chi Q(X,Y)$ with respect to the root $\partial\chi$. By using this argumentation recursively, we obtain the following definition.

Definition 7.5. Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ and let x be any root of $Q(T,1)$. Then, the *0-th derived form of $Q(X,Y)$ with respect to x* is $\partial_x^0 Q(X,Y) = Q(X,Y)$ and the *n-th derived form of $Q(X,Y)$ with respect to x* is

$$\partial_x^n Q(X,Y) = \partial_{\partial^{n-1}x}(\partial_x^{n-1} Q(X,Y)) \text{ for all } n \geq 1.$$

In other words, for all $n \geq 1$, the *n-th derived form of $Q(X,Y)$ with respect to x* is

$$\partial_x^n Q(X,Y) = Q(X,Y) \circ T(\lfloor x \rfloor) T(\lfloor \partial x \rfloor) ... T(\lfloor \partial^{n-1} x \rfloor), \quad \text{where} \quad T(r) = \begin{pmatrix} r & 1 \\ 1 & 0 \end{pmatrix}.$$

Using the notation developed in proposition 3.19, we deduce that $\partial_x^n Q(X,Y) = Q(X,Y) \circ D_n = Q(X,Y) \circ \left(\begin{smallmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{smallmatrix}\right)$ and so if $\partial_n^x Q(X,Y) = \langle a_n, b_n, c_n\rangle_{X,Y}$, then, for all $n \in \mathbb{N} \setminus \{0\}$, the following equations are satisfied:

$$\begin{cases} a_n &= Q(p_{n-1}, q_{n-1}) \\ b_n &= (p_{n-2}\ q_{n-2}) \left(\begin{smallmatrix} 2a & b \\ b & 2d \end{smallmatrix}\right) \left(\begin{smallmatrix} p_{n-1} \\ q_{n-1} \end{smallmatrix}\right) \\ c_n &= Q(p_{n-2}, q_{n-2}) = a_{n-1} \end{cases}$$

7.2 The derived form of a weakly reduced quadratic form

Lemma 7.6. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y}$ be a weakly reduced binary quadratic form and let χ_1 be the first root of $Q(T,1)$. Then:*

1. *$\partial_{\chi_1} Q(X,Y)$ is weakly reduced.*
2. *If $\chi_1 > 1$, $\partial_{\chi_1} Q(X,Y)$ is strongly reduced.*
3. *$\partial_{\chi_1}^2 Q(X,Y)$ is strongly reduced.*

Proof. Let χ_2 be the second root of $Q(T,1)$. Then, the first root of $\partial_{\chi_1}Q(T,1)$ is $\partial\chi_1 > 1$ and the second root is $\frac{1}{\chi_2 - \lfloor\chi_1\rfloor} = (\partial\chi_1)'$.

1. Since $\chi_2 < 0$ and since $\lfloor\chi_1\rfloor \geq 0$, we deduce that $\chi_2 - \lfloor\chi_1\rfloor < 0$, whereby we conclude that the second root of $\partial_{\chi_1}Q(T,1)$ is negative. Thus $\partial\chi_1$ is weakly reduced and so is then $\partial_{\chi_1}Q(X,Y)$.

2. If $\chi_1 > 1$, then $\lfloor\chi_1\rfloor \geq 1$, we deduce that $\chi_2 - \lfloor\chi_1\rfloor < -1$, whereby we conclude that the second root of $\partial_{\chi_1}Q(T,1)$ lies strictly between -1 and 0. Thus $\partial\chi_1$ is strongly reduced and so is then $\partial_{\chi_1}Q(X,Y)$.

3. Since $\partial_{\chi_1}Q(X,Y)$ is weakly reduced and since the the first root of $\partial_{\chi_1}Q(T,1)$ is $\partial\chi_1 > 1$, we deduce by the second part that $\partial^2_{\chi_1}Q(X,Y)$ is strongly reduced.

□

Example. Consider $Q(X,Y) = \langle 1,-4,-13\rangle_{X,Y} \in Quad(68)$. By previous examples, we know that $Q(X,Y)$ is weakly reduced, that $\chi_1 = 2+\sqrt{17} > 1$ and that $\partial_{\chi_1}Q(X,Y) = \langle -1,8,1\rangle_{X,Y}$. The last lemma implies that $\partial_{\chi_1}Q(X,Y)$ is strongly reduced. This conclusion can be easily verified with lemma 6.11.

7.3 The derived form of a strongly reduced from

Since any strongly reduced quadratic form is also weakly reduced and the first root of its corresponding polynomial is greater than 1, lemma 7.6 implies the following result.

Corollary 7.7. *Let $Q(X,Y) \in Quad(D)$ be strongly reduced and let χ_1 be the first root of $Q(T,1)$. Then $\partial_{\chi_1}Q(X,Y)$ is strongly reduced.*

Example. The quadratic form $Q(X,Y) = \langle -1,10,3\rangle_{X,Y} \in Quad(112)$ is strongly reduced. The first root of $Q(T,1)$ is $\chi_1 = 5+2\sqrt{7}$. The last corollary yields that $\partial_{\chi_1}Q(X,Y) = \langle 3,-10,-1\rangle_{X,Y}$ is strongly reduced.

Unfortunately, the general method to compute the derived form of $Q(X,Y)$ with respect to the first or the second root of $Q(T,1)$ consists in computing first the roots of $Q(T,1)$ explicitly, then the integral part of the considered root and finally the derived form of $Q(X,Y)$ with respect to the considered root. We would like to shorten the computations. Whereas this is not possible for quadratic forms in general, it is possible for strongly reduced forms. This is due to the fact that we can predict the value of the integral part of the roots of $Q(T,1)$ whenever $Q(X,Y)$ is strongly reduced.

Indeed, if $Q(X,Y)$ is strongly reduced, then its second root χ_2 satisfies $-1 < \chi_2 < 0$ and so $\lfloor\chi_2\rfloor = -1$. By proposition 7.2, we deduce the following proposition.

Proposition 7.8. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ be a strongly reduced quadratic form and let χ_2 be the second root of $Q(T,1)$. Then $\partial_{\chi_2}Q(X,Y) = \langle a-b+c, b-2a, a\rangle_{X,Y}$.*

Example. The quadratic form $Q(X,Y) = \langle -1,10,3\rangle_{X,Y} \in Quad(112)$ is strongly reduced. The second root of $Q(T,1)$ is $\chi_2 = 5-2\sqrt{7}$. Then the last proposition gives $\partial_{\chi_2}Q(X,Y) = \langle -1-10+3, 10+2, -1\rangle_{X,Y} = \langle -8,12,-1\rangle_{X,Y}$.

Remark 7.9. *Observe that if $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ is a strongly reduced quadratic form and χ_2 is the second root of $Q(T,1)$, then $\partial_{\chi_2}Q(X,Y)$ is not strongly, nor weakly reduced since the first and last coefficient of $\partial_{\chi_2}Q(X,Y)$ have equal sign.*

If $Q(X,Y)$ is strongly reduced, then the integral part of the first root χ_1 of $Q(T,1)$ depends on the quadratic form $Q(X,Y)$. Since $\chi_1+\chi_2 = -\frac{b}{a}$ and since $-1 < \chi_2 < 0$, one obtains easily that $\lfloor -\frac{b}{a} \rfloor \in \{\lfloor \chi_1 \rfloor - 1, \lfloor \chi_1 \rfloor\}$. Since $\chi_2 < 0 < 1 \leq \lfloor \chi_1 \rfloor < \chi_1$ and since $\lfloor \chi_1 \rfloor < \chi_1 < \lfloor \chi_1 \rfloor + 1$, we conclude that $aQ(\lfloor \chi_1 \rfloor, 1) < 0$ and $aQ(\lfloor \chi_1 \rfloor + 1, 1) > 0$. By those two observations, we deduce that $aQ(\lfloor -\frac{b}{a} \rfloor + 1, 1) < 0$ if and only if $\lfloor \chi_1 \rfloor = \lfloor -\frac{b}{a} \rfloor + 1$ and that $aQ(\lfloor -\frac{b}{a} \rfloor + 1, 1) > 0$ if and only if $\lfloor \chi_1 \rfloor = \lfloor -\frac{b}{a} \rfloor$. This gives an efficient way to compute the integral part of the first root χ_1 of $Q(T,1)$ and can be used to efficiently compute the derived form of $Q(X,Y)$ with respect to χ_1 (without explicitly computing χ_1).

Proposition 7.10. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ be a strongly reduced quadratic form and let χ_1 be the first root of $Q(T,1)$. If $aQ(\lfloor -\frac{b}{a} \rfloor + 1, 1) > 0$, set $\rho = \lfloor -\frac{b}{a} \rfloor$, else set $\rho = \lfloor -\frac{b}{a} \rfloor + 1$. Then $\partial_{\chi_1}Q(X,Y) = \langle a\rho^2 + b\rho + c, b + 2a\rho, a\rangle_{X,Y}$.*

Example. The quadratic form $Q(X,Y) = \langle -1, 10, 3\rangle_{X,Y} \in Quad(112)$ is strongly reduced and $\left\lfloor -\frac{10}{-1} \right\rfloor = 10$. Since $Q(11,1) = 8 > 0$, we set $\rho = \left\lfloor -\frac{10}{-1} \right\rfloor = 10$. Let χ_1 be the first root of $Q(T,1)$. Then by the last proposition $\partial_{\chi_1}Q(X,Y) = \langle a\rho^2 + b\rho + c, b + 2a\rho, a\rangle_{X,Y} = \langle 3, -10, -1\rangle_{X,Y}$.

Remark 7.11. *In the last proposition, we did not use the fact that if $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ is strongly reduced and if χ_1 is the first root of $Q(T,1)$, then $\partial_{\chi_1}Q(X,Y)$ is also strongly reduced. One can show that $\lfloor \chi_1 \rfloor$ is the unique integer ξ such that $\sqrt{D} - 2|a| < \sigma(b + 2a\xi) < \sqrt{D}$ where σ denotes the sign of a. Hence, if we know the value of $\sqrt{D}$, then this observation can also be used to find $\lfloor \chi_1 \rfloor$ and to compute $\partial_{\chi_1}Q(X,Y)$ in a fast way without computing χ_1 explicitly. However, one should observe that the method given in the last proposition is more effective.*

8 Reduction of indefinite forms and rough upper bounds for the class number of positive discriminants

Sources: [1, p.13-14,16], [8, p.216-219, 321-322,354-358], [3, p.160-161], [10, p.113-114,120-125]
Other approaches: [9, p.22-23], [12, p.359-364], [11, art. 183-185]

We are finally able to prove that there are only finitely many equivalence classes in $Quad(D)$ for any $D > 0$.

In section 8.1, we will prove that this holds for all $D \in \mathbb{N}^2 \setminus \{0\}$. This proof could have been given earlier since it uses some number theoretical facts and the action of $GL(2,\mathbb{Z})$ on $Quad(D)$ only. However this chapter seemed to be the best place to develop this result.

In section 8.2, we will prove the claim for non-square positive discriminants. We are going to use the concepts introduced in chapter 6 and chapter 7 to prove first that any quadratic form is equivalent to a weakly reduced one and then that any quadratic form is equivalent to a strongly reduced one. Thereafter the conclusion will follow by giving an upper bound for the number of strongly reduced quadratic forms of a given discriminant.

8.1 The class number $h(D)$ for $D \in \mathbb{N}^2 \setminus \{0\}$ is finite

Lemma 8.1. *Any quadratic form $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ with $D \in \mathbb{N}^2 \setminus \{0\}$ is equivalent to a quadratic form of the form $\langle a', \sqrt{D}, 0\rangle$ where $0 \leq a' < \sqrt{D}$.*

Proof. Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$. Let $D = n^2$ with $n \in \mathbb{N} \setminus \{0\}$. Then at least one of the roots of $Q(T,1)$ is non-zero (either $x^+ = \frac{-b+\sqrt{D}}{2a} = \frac{-b+n}{2a}$ or $x^- = \frac{-b-\sqrt{D}}{2a} = \frac{-b-n}{2a}$). Let $x = \frac{p}{q}$ (where $p,q \in \mathbb{Z}$ with $\gcd(p,q) = 1$ and with $q > 0$) be a non-zero root of $Q(T,1)$. By Bezout's theorem, there exist $u,v \in \mathbb{Z}$ such that $pu + qv = 1$. Let

$$Q'(X,Y) = Q(X,Y) \circ \begin{pmatrix} v & p \\ -u & q \end{pmatrix} = \langle Q(v,-u), (2avp - ubp + bqv - 2ucq), 2Q(p,q)\rangle_{X,Y}.$$

Observe first that $Q(p,q) = ap^2 + bpq + cq^2 = q^2(a(\frac{p}{q})^2 + b\frac{p}{q} + c) = 0$ since $\frac{p}{q}$ is a root of $Q(T,1)$. Observe secondly that since $\left(\begin{smallmatrix} v & p \\ -u & q \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$, the discriminant of $Q(X,Y) \circ \left(\begin{smallmatrix} v & p \\ -u & q \end{smallmatrix}\right)$ is the same as the discriminant of $Q(X,Y)$. Hence $(2qvp - ubp + bqr - 2ucq)^2 = D = n^2$ and thus $2qvp - ubp + bqr - 2ucq = \pm n$ and so $Q'(X,Y) = \langle Q(v,-u), \pm n, 0\rangle_{X,Y}$.

If $2qvp - ubp + bqr - 2ucq = n$, then set $Q''(X,Y) = Q'(X,Y)$.

If $2qvp - ubp + bqr - 2ucq = -n$, then set $Q''(X,Y) = Q'(X,Y) \circ \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = \langle Q(v,-u), n, 0\rangle_{X,Y}$.

Hence, in any case Q is equivalent to $Q''(X,Y) = \langle Q(v,-u), n, 0\rangle_{X,Y}$. Choose now $\psi \in \mathbb{Z}$ such that $0 \leq Q(v,-u) + n\psi < n$ (which clearly exists) and set

$$Q'''(X,Y) = Q''(X,Y) \circ \begin{pmatrix} 1 & 0 \\ \psi & 1 \end{pmatrix} = \frac{1}{2} V^T \begin{pmatrix} 2Q(v,-u) & n \\ n & 0 \end{pmatrix} V \circ \begin{pmatrix} 1 & 0 \\ \psi & 1 \end{pmatrix} = \langle Q(v,-u) + n\psi, n, 0\rangle_{X,Y}.$$

Since $\left(\begin{smallmatrix} 1 & 0 \\ \psi & 1 \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$, we finally conclude that $Q(X,Y)$ is equivalent to $Q'''(X,Y) = \langle Q(v,-u) + n\psi, n, 0\rangle_{X,Y}$ which satisfies the prescribed condition. □

Theorem 8.2. *For all $D \in \mathbb{N}^2 \setminus \{0\}$, the class number $h(D)$ is finite.*

Proof. Let $D \in \mathbb{N}^2 \setminus \{0\}$ be fixed. By lemma 8.1, any quadratic form in $Quad(D)$ is equivalent to a quadratic form of the form $\langle a, \sqrt{D}, 0\rangle_{X,Y}$ where $0 \leq a < \sqrt{D}$. Thus, there are at most $\sqrt{D}$ non-equivalent quadratic forms in $Quad(D)$ (one for all possible value of a). Hence $h(D) \leq \sqrt{D}$. □

Example. By lemma 8.1, the equivalence classes of $H(1)$ are of the form $\langle a, 1, 0\rangle \circ GL(2,\mathbb{Z})$ where $0 \leq a < 1$. Thus, there is at most one equivalence class and so

$$H(1) = \{\langle 0, 1, 0\rangle_{X,Y} \circ GL(2,\mathbb{Z})\}.$$

A general method to find the equivalent quadratic forms of the form $\langle a, \sqrt{D}, 0\rangle_{X,Y}$ with $0 \leq a < \sqrt{D}$ for all $D \in \mathbb{N}^2$ is developed in appendix B. This completes the classification of the quadratic forms with non-zero square discriminant and so the determination of their class number. Hence, in the rest of the document, we shall consider positive non-square discriminants only.

8.2 The class number $h(D)$ for $D \in \mathbb{N} \setminus \mathbb{N}^2$ is finite

Any quadratic form of discriminant $D \in \mathbb{N} \setminus \mathbb{N}^2$ is equivalent to a strongly reduced one

Lemma 8.3. *1. For any quadratic surd x, there exists $N \in \mathbb{N}$ such that $\partial^n x$ is weakly reduced for all $n > N$.*

2. Any quadratic form of discriminant $D \in \mathbb{N} \setminus \mathbb{N}^2$ is equivalent to a weakly reduced one.

Proof. Let x be a quadratic surd and let x be a root of $Q(T,1) \in \mathbb{Z}[T]$ with $Q(X,Y) = \langle a, b, c\rangle_{X,Y}$. Consider the quadratic forms $\partial_x^n Q(X,Y) = \langle a_n, b_n, c_n\rangle_{X,Y}$ for all $n \in \mathbb{N}$. Then we know by chapter 6 that $\partial^n x$ is a root of $\partial_x^n Q(T,1)$.

By lemma 6.9, it is sufficient to show that there exists $N \in \mathbb{N}$ such that $a_n c_n < 0$ for all $n > N$. Since $c_n = a_{n-1}$, it is sufficient to show that the sequence $(a_n)_{n\geq 1}$ has alternating signs for sufficiently large n.

Since x is a quadratic surd, and a root of $Q(T,1)$, we know that the other root of $Q(T,1)$ is given by the conjugate x' of x. In particular we have that $x + x' = -\frac{b}{a}$ and $xx' = \frac{c}{a}$. Thus

$$a_n = Q(p_{n-1}, q_{n-1}) = a(p_{n-1} - xq_{n-1})(p_{n-1} - x'q_{n-1}).$$

1. In the proof of proposition 3.21, we deduced that

$$x - \frac{p_n}{q_n} = \frac{(-1)^{n+2}}{q_n(x_n q_n + q_{n+1})}$$

where $x_n = \partial^n x$ denotes the n-th complete quotient of the continued fraction of x. By multiplying both sides by q_n, we get

$$q_n x - p_n = \frac{(-1)^{n+2}}{x_n q_n + q_{n+1}}.$$

Since the denominator on the right side is positive for all $n \geq 1$, the sign of the fraction, and thus the one of the expression on the left hand side, is given by the parity of n. Hence, the sign of the expression on the left hand side is alternating. Hence the sign of the sequence $(p_n - xq_n)_{n\geq 1}$ is alternating.

2. Since $\left|x - \frac{p_n}{q_n}\right| \to 0$ when $n \to +\infty$ and since $x' \neq x$, there exists $N \in \mathbb{N} \setminus \{0\}$ such that $\left|x - \frac{p_n}{q_n}\right| < |x - x'|$ for all $n > N$. Thus $\frac{p_n}{q_n} - x' = (\frac{p_n}{q_n} - x) + (x - x')$ has the same sign as $(x - x')$. Therefore $p_n - x'q_n = q_n\left(\frac{p_n}{q_n} - x'\right)$ has the same sign as $q_n(x - x')$ for all $n > N$. Since $q_n > 0$ for all $n \geq 1$, the sign of $q_n(x - x')$ is non-alternating. Hence the sign of the sequence $(p_n - x'q_n)_{n>N}$ is non-alternating.

By the points 1. and 2., we conclude that the sequence $(a_n)_{n\in\mathbb{N}}$ has alternating signs for all $n > N$ and hence $\partial^n x$ is weakly reduced for all $n > N$.

Furthermore this implies that $\partial_x^n Q(X, Y)$ is weakly reduced for all $n > N$ and since $Q(X, Y)$ is equivalent to $\partial_x^n Q(X, Y)$ for all $n \in \mathbb{N}$, this proves the second statement. □

Example. Let $x = \frac{37+\sqrt{17}}{13}$. Since $x' = \frac{37-\sqrt{17}}{13} > 3$, x is not weakly reduced. The first complete quotient of x is given by $\partial x = \frac{1}{\frac{37+\sqrt{17}}{13} - 3} = \frac{1}{\frac{-2+\sqrt{17}}{13}} = 2 + \sqrt{17} > 1$. The conjugate of ∂x is $2 - \sqrt{17}$ and since $2 - \sqrt{17} < -2 < 0$, we deduce that ∂x is weakly reduced.

By lemma 8.3 and lemma 7.6, we can conclude the following theorem.

Theorem 8.4. *1. For any quadratic surd x, there exists $N \in \mathbb{N}$ such that $\partial^n x$ is strongly reduced for all $n > N$.*

2. Any quadratic form of discriminant $D \in \mathbb{N} \setminus \mathbb{N}^2$ is equivalent to a strongly reduced one.

Example. The quadratic form $Q(X, Y) = \langle 13, -74, 104\rangle_{X,Y} \in Quad(68)$ is not weakly reduced and its first root is $\chi_1 = \frac{37+\sqrt{17}}{13}$. The form $\partial_{\chi_1} Q(X, Y) = \langle -1, 4, 13\rangle_{X,Y}$ is weakly reduced and has first root $\partial\chi_1 = 2 + \sqrt{17}$. Then the form $\partial_{\chi_1}^2 Q(X, Y) = \langle 1, -8, -1\rangle_{X,Y}$ is strongly reduced and has first root $\partial^2\chi_1 = 4 + \sqrt{17}$.

By part 2. of theorem 8.4 any quadratic form $Q(X, Y)$ with positive non-square discriminant D is equivalent to a strongly reduced one. This strongly reduced quadratic form can be obtained by choosing a root x of $Q(T, 1)$, computing the sequence $(\partial_x^n Q(X, Y))_{n\in\mathbb{N}}$ of n-th derived forms with respect to x of $Q(X, Y)$ and checking for each $n \in \mathbb{N}$ whether $\partial_x^n Q(X, Y)$ is strongly reduced or not. Unfortunately this construction does not yield an upper bound for n and so, one cannot guess how many terms of the sequence $(\partial_x^n Q(X, Y))_{n\in\mathbb{N}}$ have to be calculated to obtain a strongly reduced one.

However, this proves that $H(D)$ is completely determined by the strongly reduced quadratic forms of discriminant D and hence, we can now discuss the class number $h(D)$.

The class number $h(D)$ for $D \in \mathbb{N} \setminus \mathbb{N}^2$ is finite

Theorem 8.5. *Let $D \in \mathbb{N} \setminus \mathbb{N}^2$. There are only finitely many strongly reduced quadratic forms in $Quad(D)$.*

Proof. Let $Q(X, Y) = \langle a, b, c\rangle_{X,Y} \in Quad(D)$ be strongly reduced. Then in particular the Gauss conditions are satisfied. Thus $0 < |a|$, $0 < |b| < \sqrt{D}$, $0 < |c|$ and $\sqrt{D} - |b| < 2|a| < \sqrt{D} + |b|$ and $\sqrt{D} - |b| < 2|c| < \sqrt{D} + |b|$. By definition of the discriminant, we have $D = b^2 - 4ac = |b|^2 + 4|a||c|$. Since $0 < |b| < \sqrt{D}$ and $2|a| < \sqrt{D} + |b|$, we deduce that

$|a| < \sqrt{D}$. Since $0 < |b| < \sqrt{D}$ and $2|c| < \sqrt{D} + |b|$, we deduce that $|c| < \sqrt{D}$. Therefore $a, b, c \in \mathbb{Z}$ satisfy

$$1 \leq \min\{|a|, |b|, |c|\} \leq \max\{|a|, |b|, |c|\} \leq \sqrt{D}.$$

Since there are at most $(2\sqrt{D})^3 = 8D\sqrt{D}$ (a finite number) possibilities to choose the values of a, b and c (among the numbers $\pm 1, \pm 2, ..., \pm \left\lfloor \sqrt{D} \right\rfloor$), there is only a finite number of strongly reduced forms in $Quad(D)$. □

Since by the second part of theorem 8.4 every binary quadratic form in $Quad(D)$ is equivalent to a strongly reduced one and since by the proof of theorem 8.5 there are at most $8D\sqrt{D}$ different strongly reduced forms in $Quad(D)$, we conclude that $h(D) \leq 8D\sqrt{D}$. Hence we conclude the following theorem.

Theorem 8.6. *For all $D \in \mathbb{N} \setminus \mathbb{N}^2$, the class number $h(D)$ is finite.*

Example. By theorem 8.4 the equivalence classes of $H(D)$ are given by $Q(X,Y) \circ GL(2,\mathbb{Z})$ where $Q(X,Y) = \langle a, b, c \rangle_{X,Y}$ is a strongly reduced integral binary quadratic form. By the Gauss conditions, we have $D = |b|^2 + 4|a||c|$.

If $D = 5$, the equation $5 = |b|^2 + 4|a||c|$ implies that $|a| = |c| = |b| = 1$ and thus $a = 1$ and $b = c = -1$ or $a = -1$ and $b = c = 1$. Hence the only strongly reduced quadratic forms of discriminant 5 are $\langle 1, -1, -1 \rangle_{X,Y}$ and $\langle -1, 1, 1 \rangle_{X,Y}$. Moreover $\langle -1, 1, 1 \rangle_{X,Y} = \langle 1, -1, -1 \rangle_{X,Y} \circ \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Whence $H(5) = \{\langle 1, -1, -1 \rangle \circ GL(2,\mathbb{Z})\}$ and $h(5) = 1$.

8.3 Rough upper bounds for the class number of positive discriminants

From the proof of theorem 8.2, we deduce the following corollary.

Corollary 8.7. *For all $D \in \mathbb{N}^2 \setminus \{0\}$, $h(D) \leq \sqrt{D}$.*

From the proof of theorem 8.6, we deduce the following corollary.

Corollary 8.8. *For all $D \in \mathbb{N} \setminus \mathbb{N}^2$, $h(D) \leq 8D\sqrt{D}$.*

Remark 8.9. *The upper bound on the class number given by corollary 8.8 is not the best possible bound and can be reduced via different number theoretical arguments. For example, one can notice in the proof of theorem 8.6 that b has to present the same parity as D and that the signs of b and c are always the opposite of the sign of a. Moreover, one can reduce the upper bound by taking into account that $D - b^2$ is divisible by 4 and by giving an upper bound for the number of factorizations of $\frac{D-b^2}{4}$ into a and c.*

Since we have upper bounds for the class number of positive discriminants and since we know that for each positive discriminant D there is a set S_{et}^D of quadratic forms that includes at least one representative for each equivalence class of $Quad(D)$ (this set is given by $S_{et}^D = \{\langle a, \sqrt{D}, 0 \rangle_{X,Y} : 0 \leq a < \sqrt{D}\}$ if $D \in \mathbb{N}^2 \setminus \{0\}$ and by $S_{et}^D = \{Q(X,Y) \in Quad(D) : Q(X,Y)$ is strongly reduced$\}$ if $D \in \mathbb{N} \setminus \mathbb{N}^2$), we can compute $h(D)$ by brute force by computing all quadratic forms of this set and by determining its equivalence classes.

However, we do not have an efficient method to check if the quadratic forms in S_{et}^D are equivalent. Hence, already for small D, the brute force computation might be impossible to carry out. In this logic, the next step of our analysis will be to develop an efficient method to determine the equivalence relation on the quadratic forms in S_{et}^D for all $D > 0$.

Part III
The fusion of both theories

The third part is divided into chapter 9, 10, 11, 12 and 13. The final part will merge the first two parts, in order to deduce new results about continued fractions, as well as about integral binary quadratic forms. This will finally lead to a characterization of irrational numbers with purely periodic continued fractions and to an easy method to compute the class number for a given positive non-square discriminant. Unfortunately, these conclusions do not follow instantly from the previous work and so some more properties have to be developed.

In chapter 9 yet another special form of integral binary quadratic forms will be defined. These integral binary quadratic forms are called associate forms and will present some convenient properties. Furthermore, they will be very important to formulate which strongly reduced integral binary quadratic forms are equivalent and which are not.

Chapter 10 will use associate forms to prove that if one starts with a strongly reduced integral binary quadratic form and one computes recursively the derived form with respect to the first root of its corresponding polynomial, then one ends up at the initial strongly reduced quadratic form (c.f. theorem 10.1). This gives rise to so called cycles of strongly reduced integral binary quadratic forms.

In chapter 11, the results from chapter 9 and 10 will be combined to prove in theorem 11.4 that purely periodic continued fractions are identified with strongly reduced quadratic surds. Then the continued fraction of the first root of the corresponding polynomial of the associate form of a given integral binary quadratic form will be computed, which plays a crucial role in chapter 13. Next, the continued fraction of a square root of a positive non-square integer will be computed which can be used to construct an algorithm to solve the Pell equation.

Chapter 12 will then concentrate on integral binary quadratic forms and start again the study of the class number of positive non-square discriminants. In order to do so, it revisits the definition of derived forms and analyzes in a first place why it was chosen as it is. Then in theorem 12.7 this discussion finally pays out when the resulting lemmas and propositions are used to prove that two strongly reduced integral binary quadratic forms are equivalent if and only if they belong to the same cycle or associate cycles. This is one of the most important results of this Master thesis.

Finally, chapter 13 uses the conclusion of chapter 12 to discover how one can find the equivalence of strongly reduced integral binary quadratic forms by considering the first root of their corresponding polynomial only (new results). Then this will be used to show some new facts about integral binary quadratic forms and to construct a basic algorithm to compute the class number for positive non-square discriminants.

9 Associate forms

Sources: [1, p.17-18], [8, p.356]
Other approaches: [9, p.24]

In this chapter, we are going to define the associate form of a given quadratic form. Those are actually important to conclude the general shape of the continued fraction of a square-root of a positive non-square integer and to conclude which strongly reduced quadratic forms are equivalent.

9.1 Associate forms

Definition 9.1. Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y}$ be a quadratic form. Then we call $\widetilde{Q}(X,Y) = Q(X,Y) \circ \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = \langle c,-b,a\rangle_{X,Y}$ the *associate form* of $Q(X,Y)$.

Example. Let $Q(X,Y) = \langle 1,2,3\rangle_{X,Y}$. Then the associate of $Q(X,Y)$ is $\widetilde{Q}(X,Y) = \langle 3,-2,1\rangle_{X,Y}$.

Notice that the associate form of a quadratic form is properly equivalent to the quadratic form and that the map $\widetilde{(\ )} : Quad(D) \longrightarrow Quad(D)$ defined by $Q(X,Y) \mapsto \widetilde{Q}(X,Y)$ is an involution.

Using lemma 6.9 and lemma 6.11, one can easily confirm the following result.

Proposition 9.2. *The quadratic form $Q(X,Y)$ is strongly (weakly) reduced if and only if its associate form $\widetilde{Q}(X,Y)$ is so.*

Example. By the Gauss conditions $Q(X,Y) = \langle 1,-2,-2\rangle_{X,Y}$ is strongly reduced and so $\widetilde{Q}(X,Y) = \langle -2,2,1\rangle_{X,Y}$ is also strongly reduced.

Let x be a non-zero root of $Q(T,1)$. Then

$$Q(x,1) = 0 \Leftrightarrow a + b\frac{1}{x} + c\frac{1}{x^2} = 0 \Leftrightarrow a - b\left(-\frac{1}{x}\right) + c\left(-\frac{1}{x}\right)^2 = 0 \Leftrightarrow \widetilde{Q}(-\frac{1}{x},1) = 0.$$

Hence, we conclude the following proposition.

Proposition 9.3. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ with $D > 0$ be weakly reduced and let $\widetilde{Q}(X,Y) = \langle c,-b,a\rangle_{X,Y}$ be its associate. Assume χ_1 is the first and χ_2 the second root of $Q(T,1)$. Let $\widetilde{\chi_1}$ be the first and $\widetilde{\chi_2}$ be the second root of $\widetilde{Q}(T,1)$. Then*

$$\widetilde{\chi_1} = -\frac{1}{\chi_2} \quad \textit{and} \quad \widetilde{\chi_2} = -\frac{1}{\chi_1}.$$

Example. Let $Q(X,Y) = \langle 1,-1,-10\rangle_{X,Y} \in Quad(41)$, then the roots of $Q(T,1)$ are $\chi_1 = \frac{1+\sqrt{41}}{2}$ and $\chi_2 = \frac{1-\sqrt{41}}{2}$. The roots of $\widetilde{Q}(T,1)$ are $\widetilde{\chi_1} = \frac{1+\sqrt{41}}{20} = -\frac{1}{\chi_2}$ and $\widetilde{\chi_2} = \frac{1-\sqrt{41}}{20} = -\frac{1}{\chi_1}$.

9.2 The relation between the n-th derived form and the associate form

Lemma 9.4. *Let $Q(X,Y), Q^{\triangleright}(X,Y) \in Quad(D)$ for $D \in \mathbb{N} \setminus \mathbb{N}^2$ be two strongly reduced quadratic forms. Let χ_1 be the first root of $Q(T,1)$ and let $\chi_1^{\triangleright}$ be the first root of $\widetilde{Q^{\triangleright}}(T,1)$. Then*

$$Q^{\triangleright}(X,Y) = \partial_{\chi_1} Q(X,Y) \Leftrightarrow \widetilde{Q}(X,Y) = \partial_{\chi_1^{\triangleright}} \widetilde{Q^{\triangleright}}(X,Y).$$

Proof. Since the map $Q \mapsto \widetilde{Q}$ is an involution, it is sufficient to prove the result in the forward direction only.

Notice first that the quadratic forms $Q(X,Y)$, $Q^{\triangleright}(X,Y) = \partial_{\chi_1} Q(X,Y)$, $\widetilde{Q}(X,Y)$ and $\partial_{\chi_1^{\triangleright}}\widetilde{Q^{\triangleright}}(X,Y)$ have all the same discriminant and are all strongly reduced.

Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y}$. Then notice that the last coefficient of $\widetilde{Q}(X,Y)$ and the last coefficient of $\partial\widetilde{Q^{\triangleright}}(X,Y)$ is given by a. Since both quadratic forms are strongly reduced, the sign of their first coefficient is the opposite of the sign of a. In particular their first coefficients have the same sign.

Let χ_2 be the second root of $Q(T,1)$ and let $\chi_2^{\triangleright}$ be the second root of $\widetilde{Q^{\triangleright}}(T,1)$. By proposition 9.3, the first root of $\widetilde{Q}(T,1)$ is given by $-\frac{1}{\chi_2}$. The second root of $Q^{\triangleright}(T,1) = \partial_{\chi_1} Q(T,1)$ is given by $\frac{1}{\chi_2 - \lfloor\chi_1\rfloor}$. By proposition 9.3, the first root of $\widetilde{Q^{\triangleright}}(T,1)$ is given by $\chi_1^{\triangleright} = \lfloor\chi_1\rfloor - \chi_2$. Hence the first root of $\partial_{\chi_1^{\triangleright}}\widetilde{Q^{\triangleright}}(T,1)$ is given by $\partial\chi_1^{\triangleright} = \frac{1}{\lfloor\chi_1\rfloor - \chi_2 - \lfloor\lfloor\chi_1\rfloor - \chi_2\rfloor} = -\frac{1}{\chi_2}$ since $\lfloor\lfloor\chi_1\rfloor - \chi_2\rfloor = \lfloor\chi_1\rfloor$ (since $-1 < \chi_2 < 0$). Hence $\widetilde{Q}(T,1)$ and $\partial_{\chi_1^{\triangleright}}\widetilde{Q^{\triangleright}}(T,1)$ have the same first root.

By lemma 6.6, we conclude the statement.

□

Example. The quadratic form $Q(X,Y) = \langle 3,-8,-1\rangle_{X,Y} \in Quad(76)$ is strongly reduced and the first root of $Q(T,1)$ is given by $\chi_1 = \frac{4+\sqrt{19}}{3}$. Thus $Q^{\triangleright}(X,Y) = \partial_{\chi_1} Q(X,Y) = \langle 5,-4,-3\rangle_{X,Y}$ is strongly reduced. Moreover $\widetilde{Q}(X,Y) = \langle 1,-8,-3\rangle_{X,Y}$ and $\widetilde{Q^{\triangleright}}(X,Y) = \langle 3,-4,-5\rangle_{X,Y}$. The first root of $\widetilde{Q^{\triangleright}}(T,1)$ is $\chi_1^{\triangleright} = \frac{2+\sqrt{19}}{3}$ and so $\partial_{\chi_1^{\triangleright}}\widetilde{Q^{\triangleright}}(X,Y) = \langle 1,-8,-3\rangle_{X,Y} = \widetilde{Q}(X,Y)$.

The last result can be extended to the n-th derived form of a quadratic form.

Corollary 9.5. *Let $Q(X,Y), Q^{\triangleright}(X,Y) \in Quad(D)$ for $D \in \mathbb{N} \setminus \mathbb{N}^2$ be two strongly reduced quadratic forms. Let χ_1 be the first root of $Q(T,1)$ and let $\chi_1^{\triangleright}$ be the first root of $\widetilde{Q^{\triangleright}}(T,1)$. Then*

$$Q^{\triangleright}(X,Y) = \partial_{\chi_1}^n Q(X,Y) \Leftrightarrow \widetilde{Q}(X,Y) = \partial_{\chi_1^{\triangleright}}^n \widetilde{Q^{\triangleright}}(X,Y)$$

for any $n \in \mathbb{N}$.

Proof. We prove the claim by induction on n. Since the map $\widetilde{(\,)}$ is an involution, the statement holds for $n = 0$ and by lemma 9.4 the statement holds for $n = 1$. Thus assume the claim holds for some $n \geq 1$. Then

$$\begin{aligned}
&Q^{\triangleright}(X,Y) = \partial_{\chi_1}^{n+1} Q(X,Y)\\
\Leftrightarrow &\widetilde{\partial_{\chi_1}^n Q}(X,Y) = \partial_{\chi_1^{\triangleright}}\widetilde{Q^{\triangleright}}(X,Y)\\
\Leftrightarrow &\widetilde{\partial_{\chi_1^{\triangleright}}\widetilde{Q^{\triangleright}}}(X,Y) = \partial_{\chi_1}^n Q(X,Y) = \partial_{\chi_1}^n \widetilde{\widetilde{Q}}(X,Y)\\
\Leftrightarrow &\widetilde{Q}(X,Y) = \partial_{\partial_{\chi_1^{\triangleright}}}^n \widetilde{\partial_{\chi_1^{\triangleright}} Q^{\triangleright}}(X,Y) = \partial_{\chi_1^{\triangleright}}^{n+1}\widetilde{Q^{\triangleright}}(X,Y)
\end{aligned}$$

where we get the first equivalence by lemma 9.4, the second equivalence since $\widetilde{(\,)}$ is an involution, the third equivalence by induction and the last equality since $\widetilde{(\,)}$ is an involution. □

10 Cycles of strongly reduced forms

Sources: [1, p.18-19], [8, p.354-358], [3, p.160-161], [10, p.127-131]
Other approaches: [9, p.23-24], [12, p.356-364] (with an innovative notation for cycles), [11, art. 186]

Chapter note: Assume throughout this chapter that $D \in \mathbb{N} \setminus \mathbb{N}^2$.
Note: Throughout the rest of the document, we shall denote the n-th derived form of a quadratic form $Q(X,Y)$ with respect to the first root χ_1 of $Q(T,1)$ by $\partial^n Q(X,Y) = \partial^n_{\chi_1} Q(X,Y)$. Whenever the second root is used, we will mention this explicitly.

In chapter 8 theorem 8.5, we proved that the number of strongly reduced forms with discriminant D is finite. Furthermore, we proved in corollary 7.7 that the derived form with respect to the first root of a strongly reduced quadratic form is also strongly reduced. Hence, if we start with a strongly reduced quadratic form, compute its derived form with respect to the first root and then compute recursively the derived form with respect to the first root of the obtained quadratic form, we necessarily end up with a cycle of strongly reduced forms.

This chapter describes those cycles and proves some of their basic properties as for example that they partition the set of strongly reduced quadratic forms of a given discriminant, which helps us later on to study the equivalence relations of strongly reduced quadratic forms.

10.1 The definition of cycles of strongly reduced quadratic forms

Let us first formally prove the above stated observation.

Theorem 10.1. *Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ be strongly reduced. Then there exists $m \in \mathbb{N} \setminus \{0\}$ such that $\partial^m Q(X,Y) = Q(X,Y)$.*

Proof. Consider the sequence $(Q_n(X,Y))_{n\geq 0}$ where $Q_n(X,Y) = \partial^n Q(X,Y)$ is the n-th derived form with respect to the first root of $Q(X,Y)$. Since $Q(X,Y)$ is strongly reduced, by corollary 7.7 $\partial^n Q(X,Y)$ is strongly reduced for all $n \geq 0$. By theorem 8.5, there are only finitely many different strongly reduced quadratic forms. Thus there exists $m, n \in \mathbb{N}$ where $m > 0$ such that $Q_{m+n}(X,Y) = Q_n(X,Y)$ which rewrites $\partial^n Q_m(X,Y) = Q_n(X,Y)$. By corollary 9.5 we have $\partial^n \widetilde{Q_n}(X,Y) = \widetilde{Q_m}(X,Y)$. Since $\partial^n Q_0(X,Y) = Q_n(X,Y)$, corollary 9.5 implies $\partial^n \widetilde{Q_n}(X,Y) = \widetilde{Q_0}(X,Y)$. Thus $\widetilde{Q_m}(X,Y) = \widetilde{Q_0}(X,Y)$.
Since the map $Q(X,Y) \mapsto \widetilde{Q}(X,Y)$ is an involution, this implies $Q_m(X,Y) = Q_0(X,Y)$. Thus $Q(X,Y) = \partial^m Q(X,Y)$ for some $m \in \mathbb{N} \setminus \{0\}$. □

Example. Consider the strongly reduced quadratic form $Q(X,Y) = \langle 15, -18, -5\rangle_{X,Y} \in Quad(156)$. Then, we have the following

$$\begin{aligned}
Q(X,Y) &= \langle 15, -18, -5\rangle_{X,Y};\\
\partial Q(X,Y) &= \langle -8, 12, 15\rangle_{X,Y};\\
\partial^2 Q(X,Y) &= \langle 7, -20, -8\rangle_{X,Y};\\
\partial^3 Q(X,Y) &= \langle -5, 22, 7\rangle_{X,Y};\\
\partial^4 Q(X,Y) &= \langle 15, -18, -5\rangle_{X,Y}.
\end{aligned}$$

Hence $\partial^4 Q(X,Y) = Q(X,Y)$.

This gives rise to the following definition.

Definition 10.2. Let $Q(X,Y) \in Quad(D)$ be a strongly reduced quadratic form and let $M \in \mathbb{N} \setminus \{0\}$ be the smallest integer such that $\partial^M Q(X,Y) = Q(X,Y)$. Then we call $\{Q(X,Y), \partial Q(X,Y), ..., \partial^{M-1} Q(X,Y)\}$ *the cycle of* $Q(X,Y)$ and we call M the *length* of the cycle of $Q(X,Y)$.

Example. By the last example, we conclude that the cycle of $Q(X,Y) = \langle 15, -18, -5\rangle_{X,Y} \in Quad(156)$ is given by

$$\{\langle 15, -18, -5\rangle_{X,Y}, \langle -8, 12, 15\rangle_{X,Y}, \langle 7, -20, -8\rangle_{X,Y}, \langle -5, 22, 7\rangle_{X,Y}\}$$

and the length of the cycle of $Q(X,Y)$ is 4.

Remark 10.3. *The cycles of strongly reduced forms will play later on in this document the same role as* Gauss' *cycles of right neighbors in [11]. However, using the cycles of strongly reduced forms has two advantages. Using proposition 7.10, we can compute the cycles in a fast manner and we will see later on that the cycles can be used to find result about continued fractions. However, cycles of strongly reduced forms might include more quadratic forms than right neighbor cycles.*

10.2 Basic properties of cycles of strongly reduced quadratic forms

General properties

Since the derived form of a quadratic form is unique, one can easily prove the following lemma.

Lemma 10.4. *If the strongly reduced quadratic form $Q'(X,Y)$ belongs to the cycle of $Q(X,Y)$. Then the cycle of $Q'(X,Y)$ is equal to the cycle of $Q(X,Y)$.*

Hence, cycles of strongly reduced quadratic forms partition the set of strongly reduced forms with discriminant D. In this logic, they can be used to classify (strongly reduced) quadratic forms.

Lemma 10.5. *Let $Q(X,Y)$ be a strongly reduced quadratic form. Then the length of the cycle of $Q(X,Y)$ is even.*

Proof. Let $Q(X,Y) = \langle a, b, c\rangle_{X,Y}$, then the sign of the first coefficient of $\partial^k Q(X,Y)$ is the same than the sign of a if and only if k is even. □

Cycles that can be deduced from another cycle

Let us assume that we know the cycle of some strongly reduced quadratic form $Q(X,Y)$. Then, we would like to know if we can deduce another cycle from the cycle of $Q(X,Y)$. Actually, this is possible.

By corollary 7.4, we know that $\partial(-Q(X,Y)) = -\partial Q(X,Y)$. Hence, we conclude the following lemma.

Lemma 10.6. *Let $Q(X,Y) \in Quad(D)$ be a strongly reduced quadratic form and let*

$$\{Q(X,Y), \partial Q(X,Y), ..., \partial^{M-1} Q(X,Y)\}$$

be the cycle of $Q(X,Y)$. Then the cycle of $-Q(X,Y)$ is given by

$$\{-Q(X,Y), -\partial Q(X,Y), ..., -\partial^{M-1} Q(X,Y)\}.$$

Example. Since the cycle of $Q(X,Y) = \langle 15, -18, -5\rangle_{X,Y} \in Quad(156)$ is given by

$$\{\langle 15, -18, -5\rangle_{X,Y}, \langle -8, 12, 15\rangle_{X,Y}, \langle 7, -20, -8\rangle_{X,Y}, \langle -5, 22, 7\rangle_{X,Y}\}$$

we deduce that the cycle of $-Q(X,Y) = \langle -15, 18, 5\rangle_{X,Y}$, is given by

$$\{\langle -15, 18, 5\rangle_{X,Y}, \langle 8, -12, -15\rangle_{X,Y}, \langle -7, 20, 8\rangle_{X,Y}, \langle 5, -22, -7\rangle_{X,Y}\}.$$

Furthermore, we can deduce the cycle of the associate form of a quadratic form from its cycle.

Proposition 10.7. *Let $Q(X,Y) \in Quad(D)$ be a strongly reduced quadratic form and let*

$$\{Q(X,Y), \partial Q(X,Y), ..., \partial^{M-1}Q(X,Y)\}$$

be the cycle of $Q(X,Y)$. Then the cycle of its associate form $\widetilde{Q}(X,Y)$ (also called associate cycle of $Q(X,Y)$) is given by

$$\{\widetilde{Q}(X,Y), \widetilde{\partial^{M-1}Q}(X,Y), ..., \widetilde{\partial Q}(X,Y)\}.$$

Proof. By corollary 9.5, we have that $Q(X,Y) = \partial^M Q(X,Y)$ if and only if $\widetilde{Q}(X,Y) = \partial^M \widetilde{Q}(X,Y)$. Hence, we deduce that the length of the cycle of $\widetilde{Q}$ is equal to the length of the cycle of $Q(X,Y)$. Thus the length of the cycle of $\widetilde{Q}$ is equal to M. Hence the cycle of $\widetilde{Q}(X,Y)$ is given by

$$\{\widetilde{Q}(X,Y), \partial\widetilde{Q}(X,Y), ..., \partial^{M-1}\widetilde{Q}(X,Y)\}.$$

Since $Q(X,Y) = \partial^m Q(X,Y) = \partial^n(\widetilde{\widetilde{\partial^{m-n}Q}}(X,Y))$ for all $n \in \{0, 1, ..., m\}$, corollary 9.5 implies that $\widetilde{\partial^{m-n}Q}(X,Y) = \partial^n \widetilde{Q}(X,Y)$. Hence the cycle of $\widetilde{Q}(X,Y)$ is given by

$$\{\widetilde{Q}(X,Y), \widetilde{\partial^{M-1}Q}(X,Y), ..., \widetilde{\partial Q}(X,Y)\}.$$

□

Example. Since the cycle of $Q(X,Y) = \langle 15, -18, -5\rangle_{X,Y} \in Quad(156)$ is given by

$$\{\langle 15, -18, -5\rangle_{X,Y}, \langle -8, 12, 15\rangle_{X,Y}, \langle 7, -20, -8\rangle_{X,Y}, \langle -5, 22, 7\rangle_{X,Y}\}$$

the cycle of $\widetilde{Q}(X,Y) = \langle -5, 18, 15\rangle_{X,Y} \in Quad(156)$, the cycle of $\langle -8, 12, 15\rangle_{X,Y}$ is given by

$$\{\langle -5, 18, 15\rangle_{X,Y}, \langle 7, -22, -5\rangle_{X,Y}, \langle -8, 20, 7\rangle_{X,Y}, \langle 15, -12, -8\rangle_{X,Y}\}.$$

11 Ultimately periodic continued fractions and their identification with quadratic surds

Sources: [1, p.16-20], [2, p.47-50], [3, p.344-350], [4, p.88-96,104-113], [13]
Other approaches: [6, p.101-107], [5, p.8-10]

In part II, we developed several results about quadratic forms. The constructions that we used seemed independent of continued fractions. However, at different places, we used implicitly the familiar concepts of continued fractions to obtain results about quadratic forms.

In particular, for any $Q(X,Y) \in Quad(D)$ for $D \in \mathbb{N} \setminus \mathbb{N}^2$ and any root x of $Q(T,1)$, the definition of the n-th derived form of $Q(X,Y)$ with respect to x hides the continued fraction structure of x. Indeed, if the continued fraction of x is known, say $x = [r_0, r_1, r_2, ...]$, then one can easily compute $\partial_x^n Q(X,Y) = Q(X,Y) \circ T(r_0)T(r_1)...T(r_{n-1})$ with $T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$ for any $n \in \mathbb{N}$. Since derived forms were used on a regular basis to develop the previous theory, it is not surprising that the continued fraction structure is deeply involved in the theory about quadratic forms. However, up to now, we did not have the opportunity, nor the necessity to highlight the connection between those two fields.

In this chapter we will begin to fusion both theories and use both theories explicitly to obtain the wished results. Indeed, we are going to show in this chapter that a strongly reduced quadratic surd (which is the root of a polynomial which corresponds to a strongly reduced quadratic form) has a so called *purely periodic* continued fraction. This will be one of the key results to analyze the set $H(D)$ for $D \in \mathbb{N} \setminus \mathbb{N}^2$. Several other sources use either the theory of continued fractions or the one of quadratic forms to prove this result. However, both theories are always implicitly used. We are not going to disguise the used material and we will present an elegant modern proof.

After proving this identification, we are going to develop the continued fraction of some particular quadratic surds.

11.1 The definition of ultimately periodic continued fractions

Definition 11.1. An infinite continued fraction $[r_0, r_1, r_2, ...]$ is called *ultimately periodic* if there exist $n, m \in \mathbb{N}$ such that $r_{k+m} = r_k$ for all $k \geq n$. In other words, the continued fraction of some irrational number x is *ultimately periodic* if there exist $n, m \in \mathbb{N}$ such that $\partial^{k+m} x = \partial^k x$ for all $k \geq n$.

Example. The continued fraction of $\sqrt{23}$ is $[4, 1, 3, 1, 8, 1, 3, 1, 8, 1, 3, 1, 8, 1, ...]$ and thus $\sqrt{23}$ has an ultimately periodic continued fraction.

If the continued fraction of $x \in \mathbb{R}$ is ultimately periodic, then its continued fraction is given by

$$x = [s_0, s_1, ..., s_{n-1}, r_0, r_1, ..., r_{m-1}, r_0, r_1, ..., r_{m-1}, r_0, ...].$$

To shorten the presentation of those infinite repetitions of the same block of partial quotients, we are going to use the following notation

$$x = [s_0, s_1, ..., s_{n-1}, \overline{r_0, r_1, ..., r_{m-1}}].$$

Example. The continued fraction of $\sqrt{23}$ is $[4, \overline{1, 3, 1, 8}] = [4, \overline{1, 3, 1, 8, 1, 3, 1, 8}]$.

Definition 11.2. Let $x \in \mathbb{R}$ have an ultimately periodic continued fraction. The minimal positive integer m such that $\partial^{k+m}x = \partial^k x$ holds for all $k \geq n$ is called the *caliber* of x. The ordered m-tuple of partial quotients $(r_0, ..., r_{m-1})$ that satisfies $x = [s_0, ..., s_n, \overline{r_0, ..., r_{m-1}}]$ and minimizes $n \in \mathbb{N}$ is called the *period* of x.

Example. The caliber of $\sqrt{23}$ is 4 and its period is $(1, 3, 1, 8)$.

Definition 11.3. An infinite continued fraction $[r_0, r_1, r_2, ...]$ is called *purely periodic* if there exists $m \in \mathbb{N}$ such that $r_{k+m} = r_k$ for all $k \geq 0$. In other words, the continued fraction of some irrational number x is *purely periodic* if there exists $m \in \mathbb{N}$ such that $\partial^{k+m}x = \partial^k x$ for all $k \geq 0$.

Example. The continued fraction of $1 + \sqrt{2}$ is $[\overline{2}]$ and so it is purely periodic with caliber 1.

Although the computation of the continued fraction of an irrational number may be an infinite process, we observe that the computation of an ultimately periodic continued fraction can be performed in a finite number of steps. Indeed, if one computes the continued fraction of an irrational number x whose continued fraction is known to be ultimately periodic, then one can stop the computation after obtaining the period once which is exactly when for some minimal $n \in \mathbb{N}$, we have $\partial^n x = \partial^k x$ for some $k \in \{0, 1, ..., n-1\}$. Hence, we would like to know prior to any computation if the considered number has an ultimately continued fraction.

11.2 The continued fraction of $x \in \mathbb{R}$ is purely periodic if and only if x is a strongly reduced quadratic surd

The next theorem is one of if not the most important results of this document.

Theorem 11.4. *The continued fraction of $x \in \mathbb{R}$ is purely periodic if and only if x is a strongly reduced quadratic surd.*

Proof. 1. Let $x \in \mathbb{R}$ be a strongly reduced quadratic surd and let x be the first (positive) root of $Q(T, 1)$ where $Q(X, Y)$ denotes a strongly reduced quadratic form. Then, by theorem 10.1, there exists $M \in \mathbb{N} \setminus \{0\}$ such that $\partial^M Q(X, Y) = Q(X, Y)$. Since $\partial^m x$ is the first root of $\partial^M Q(T, 1)$, we deduce that $\partial^M x = x$. Hence the continued fraction of x is purely periodic.

2. Let $x \in \mathbb{R}$ have a purely periodic continued fraction. Then

$$x = [\overline{r_0, ..., r_{m-1}}] = [r_0, ..., r_{m-1}, x]$$

for some $r_0, .., r_{m-1} \in \mathbb{N}\setminus\{0\}$. (Notice that $[r_0, ..., r_{m-1}, x]$ is a partial continued fraction since $x = [\overline{r_0, ..., r_{m-1}}] > 1$ since $r_0 \geq 1$). By the first part of proposition 3.19, we deduce that $(x : 1) = D_m \star (x : 1)$ where $D_m = \left(\begin{smallmatrix} r_0 & 1 \\ 1 & 0 \end{smallmatrix}\right) ... \left(\begin{smallmatrix} r_{m-1} & 1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} p_{m-1} & p_{m-2} \\ q_{m-1} & q_{m-2} \end{smallmatrix}\right)$.
Therefore $(x : 1) = \left(\frac{p_{m-1}x + p_{m-2}}{q_{m-1}x + q_{m-2}} : 1\right)$. Hence $x = \frac{p_{m-1}x + p_{m-2}}{q_{m-1}x + q_{m-2}}$ and so $q_{m-1}x^2 + q_{m-2}x = p_{m-1}x + p_{m-2}$ which implies that $q_{m-1}x^2 + (q_{m-2} - p_{m-1})x - p_{m-2} = 0$. In other words, x is a root of $f(T) = q_{m-1}T^2 + (q_{m-2} - p_{m-1})T - p_{m-2}$.
Since x has an infinite continued fraction, x is irrational. Since x is irrational and the root of the quadratic polynomial $f(T)$ which has integral coefficients, x is a quadratic surd. By the properties of the sequences $(p_n)_{n\geq -1}$ and $(q_n)_{n\geq -1}$ developed in proposition 3.19 and proposition 3.21, we deduce that $f(0) = -p_{m-2} < 0$ and $f(-1) = q_{m-1} - q_{m-2} +$

$p_{m-1} - p_{m-2} > 0$. Thus the second root x' of $f(T)$ satisfies $-1 < x' < 0$. Hence x is strongly reduced.

□

As stated previously, the above theorem can be used to compute efficiently the continued fraction of strongly reduced quadratic surds. Actually, if x is a quadratic surd, the above theorem yields that the continued fraction of x is purely periodic. Thus, there is a minimal $M \in \mathbb{N}$ such that $\partial^M x = x$. Hence, we can optimize the algorithm to calculate the continued fraction of x by including the condition that the algorithm computes $r_k = \lfloor \partial^k x \rfloor$ of x as long as $\partial^k x \neq x$ and if $\partial^M x = x$, then the algorithm concludes that $x = [\overline{r_0, r_1, ..., r_{M-1}}]$. In particular the algorithm will end after finitely many steps.

Let us consider an example of such a computation and let us consider at the same time the corresponding derived forms.

Example. Consider the strongly reduced quadratic surd $x = 4+\sqrt{19} > 1$. Then x is the first root of the polynomial T^2-8T+3. By the previous theorem, we know that the continued fraction of $4+\sqrt{19}$ is purely periodic. Let us verify this by computing the continued fraction. Since

$$\begin{aligned}
x_0 &= 4+\sqrt{19} & r_0 &= 8 & Q_0(X,Y) &= \langle 1,-8,-3\rangle_{X,Y};\\
x_1 &= \frac{4+\sqrt{19}}{3} & r_1 &= 2 & Q_1(X,Y) &= \langle -3,8,1\rangle_{X,Y};\\
x_2 &= \frac{2+\sqrt{19}}{5} & r_2 &= 1 & Q_2(X,Y) &= \langle 5,-4,-3\rangle_{X,Y};\\
x_3 &= \frac{3+\sqrt{19}}{2} & r_3 &= 3 & Q_3(X,Y) &= \langle -2,6,5\rangle_{X,Y};\\
x_4 &= \frac{3+\sqrt{19}}{5} & r_4 &= 1 & Q_4(X,Y) &= \langle 5,-6,-2\rangle_{X,Y};\\
x_5 &= \frac{2+\sqrt{19}}{3} & r_5 &= 2 & Q_5(X,Y) &= \langle -3,4,5\rangle_{X,Y};\\
x_6 &= 4+\sqrt{19} & r_6 &= 8 & Q_6(X,Y) &= \langle 1,-8,-3\rangle_{X,Y};
\end{aligned}$$

Thus we conclude that $4+\sqrt{19} = [\overline{8,2,1,3,1,2}]$.

Observe that in the last example the quadratic forms $Q_0(X,Y), Q_1(X,Y), ..., Q_5(X,Y)$ are all strongly reduced and are all equivalent. Furthermore, by considering the last example, one could think that the number of equivalent strongly reduced forms obtained by derivation is equal to the caliber of the first root of one such form. However, this is not always the case as the next example shows. A formal proof of these observations will be given in the next chapter.

Example. Consider the strongly reduced quadratic surd $x = \frac{11+\sqrt{197}}{19} > 1$. Then x is the first root of the polynomial $19T^2-22T-4$. By the previous theorem, we know that the continuous

fraction of $\frac{11+\sqrt{197}}{19}$ is purely periodic. We compute:

$$\begin{aligned}
x_0 &= \frac{11+\sqrt{197}}{19} & r_0 &= 1 & Q_0(X,Y) &= \langle 19,-22,-4\rangle_{X,Y};\\
x_1 &= \frac{8+\sqrt{197}}{7} & r_1 &= 3 & Q_1(X,Y) &= \langle -7,16,19\rangle_{X,Y};\\
x_2 &= \frac{13+\sqrt{197}}{4} & r_2 &= 6 & Q_2(X,Y) &= \langle 4,-26,-7\rangle_{X,Y};\\
x_3 &= \frac{11+\sqrt{197}}{19} & r_3 &= 1 & Q_3(X,Y) &= \langle -19,22,4\rangle_{X,Y};\\
x_4 &= \frac{8+\sqrt{197}}{7} & r_4 &= 3 & Q_4(X,Y) &= \langle 7,-16,-19\rangle_{X,Y};\\
x_5 &= \frac{13+\sqrt{197}}{4} & r_5 &= 6 & Q_5(X,Y) &= \langle -4,26,7\rangle_{X,Y};\\
x_6 &= \frac{11+\sqrt{197}}{19} & r_6 &= 1 & Q_6(X,Y) &= \langle 19,-22,-7\rangle_{X,Y};
\end{aligned}$$

Thus we conclude $\frac{11+\sqrt{197}}{19} = [\overline{1,3,6}]$, although we count six equivalent strongly reduced equivalent forms obtained by derivation.

Theorem 11.5. *The continued fraction of $x \in \mathbb{R}$ is ultimately periodic if and only if x is a quadratic surd.*

Proof. Let x be a quadratic surd. By theorem 8.4, there exists a sufficiently large $N \in \mathbb{N}$ such that $x_n = \partial^n x$ is strongly reduced for all $n \geq N$. Thus by theorem 11.4 we know that x_n has a purely periodic continued fraction and so $x_n = [\overline{r_0,...,r_{m-1}}]$ for some $m \in \mathbb{N}$ and $r_0,...,r_{m-1} \in \mathbb{N} \setminus \{0\}$. Thus

$$x = [s_0,...,s_{n-1},x_n] = [s_0,...,s_{n-1},\overline{r_0,...,r_{m-1}}]$$

for some $n \in \mathbb{N}$ and $s_0,...,s_{n-1} \in \mathbb{N} \setminus \{0\}$.

Assume x has an ultimately periodic continued fraction. Then $x = [s_0,...,s_{n-1},y]$ where y has a purely periodic continued fraction. Thus by theorem 11.4 y is a strongly reduced quadratic surd. Since $(x:1) = D_n \star (y:1)$ where $D_n = T(s_0)...T(s_{n-1})$, corollary 6.3 implies that x is a quadratic surd. □

Example. $1+\sqrt{19} = 4+\sqrt{19}-3 = [\overline{8,2,1,3,1,2}]-3 = [5,\overline{2,1,3,1,2,8}]$.

By applying a similar condition on our algorithm as in the case of strongly reduced quadratic surds, we can define an efficient method to compute the continued fraction of quadratic surds.

11.3 Some particular ultimately periodic continued fractions

We finish this chapter by developing the continued fraction of some particular quadratic surds.

The continued fraction of the first root of the associate form of a strongly reduced form

Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ be a strongly reduced quadratic form and let χ_1 be the first root of $Q(T,1)$. GALOIS discovered that one can express the continued fraction of the first root of $\widetilde{Q}(T,1)$ by the continued fraction of χ_1.

Corollary 11.6. *Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ be a strongly reduced quadratic form. Let $\chi_1 = [\overline{r_0, ..., r_{m-1}}]$ be the first root of $Q(T,1)$. Then the first root $\widetilde{\chi}_1$ of $\widetilde{Q}(T,1)$ is given by $\widetilde{\chi}_1 = [\overline{r_{m-1}, ..., r_0}]$.*

Proof. By the calculus of conjugates, we deduce that $\chi_1' = [r_0, ..., r_{m-1}, \chi_1]' = [r_0, ..., r_{m-1}, \chi_1']$ where $[r_0, ..., r_{m-1}, \chi_1']$ denotes a formal partial continued fraction since $\chi_1' < 0$. Let us rewrite the equation as follows

$$\chi_1' = [r_0, r_1, r_2, ..., r_{m-1}, \chi_1'] \Leftrightarrow [0, \chi_1' - r_0] = [r_1, r_2, ..., r_{m-1}, \chi_1']$$

with formal partial continued fractions on both sides of the last equation (since $\chi_1' - r_0 < 0$). Repeating the above computation for all r_i ($i \in \{1, 2, ..., m-1\}$) yields that

$$\begin{aligned}
&\chi_1' = [r_0, r_1, r_2, ..., r_{m-1}, \chi_1'] \\
\Leftrightarrow &[0, \chi_1' - r_0] = [r_1, r_2, ..., r_{m-1}, \chi_1'] \\
\Leftrightarrow &[0, -r_1, \chi_1' - r_0] = [r_2, ..., r_{m-1}, \chi_1'] \\
\Leftrightarrow &[0, -r_2, -r_1, \chi_1' - r_0] = [r_3, ..., r_{m-1}, \chi_1'] \\
&\vdots \qquad\quad \vdots \qquad\quad \vdots \qquad\quad \vdots \\
\Leftrightarrow &[0, -r_{m-1}, ..., -r_2, -r_1, \chi_1' - r_0] = [\chi_1'] = \chi_1'.
\end{aligned}$$

Thus we conclude that

$$-\chi_1' = -[0, -r_{m-1}, ..., -r_2, -r_1, \chi_1' - r_0] = [0, r_{m-1}, ..., r_2, r_1, r_0 - \chi_1']$$

where the rightmost expression denotes a partial continued fraction since $r_0 - \chi_1' > 1$ because $r_0 \geq 1$ and $\chi_1' < 0$. Furthermore, we can replace recursively $-\chi_1'$ into the expression on the right hand side and deduce that

$$-\chi_1' = [0, \overline{r_{m-1}, ..., r_2, r_1, r_0}].$$

Thus, since the first root of $\widetilde{Q}(T,1)$ is given by $\widetilde{\chi}_1 = -\frac{1}{\chi_1'}$, we finally conclude that

$$\widetilde{\chi}_1 = -\frac{1}{\chi_1'} = [\overline{r_{m-1}, ..., r_2, r_1, r_0}].$$

□

Corollary 11.7. *Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ be a strongly reduced quadratic form. Let $\chi_1 = [\overline{r_0, ..., r_{m-1}}]$ be the first and χ_2 be the second root of $Q(T,1)$. Then there exists $n \in \mathbb{N}$ such that $\partial^n \chi_2 = [\overline{r_{m-1}, ..., r_0}]$.*

Proof. By corollary 11.6, the first root $\widetilde{\chi}_1$ of $\widetilde{Q}(T,1)$ is given by $\widetilde{\chi}_1 = [\overline{r_{m-1},...,r_0}]$. Furthermore, we know that $\widetilde{\chi}_1 = -\frac{1}{\chi_2}$. In other words, $(\widetilde{\chi}_1 : 1) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \star (\chi_2 : 1)$. Hence $(\widetilde{\chi}_1 : 1)$ and $(\chi_2 : 1)$ are equivalent in $P^1(\mathbb{R})$ and so proposition 4.8 implies that there exist $p, q \in \mathbb{N}$ such that $\partial^p \widetilde{\chi}_1 = \partial^q \chi_2$. Since $\widetilde{\chi}_1 = [\overline{r_{m-1},...,r_0}]$, there exists $k \in \mathbb{N}$ such that $\partial^{q+k}\chi_2 = \partial^{p+k}\widetilde{\chi}_1 = [\overline{r_{m-1},...,r_0}]$. □

Example. Since $4 + \sqrt{19} = [\overline{8,2,1,3,1,2}]$ is the first root of $T^2 - 8T - 3$, corollary 11.6 yields that $[\overline{2,1,3,1,2,8}]$ is the first root of $-3T^2 + 8T + 1$. Indeed $[\overline{2,1,3,1,2,8}] = \frac{4+\sqrt{19}}{3}$ and $-3\left(\frac{4+\sqrt{19}}{3}\right)^2 + 8\frac{4+\sqrt{19}}{3} + 1 = 0$. Moreover, corollary 11.7 is verified since $4 - \sqrt{19} = [-1,1,1,\overline{1,3,1,2,8,2}]$ is the first root of $\langle 1,-8,-3\rangle_{T,1}$ and $\partial^8(4-\sqrt{19}) = [\overline{2,1,3,1,2,8}]$.

Corollary 11.8. *Let $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ and let χ_1 be the first root of $Q(T,1)$. Then $Q(X,Y) = \langle a,-b,-a\rangle_{X,Y}$ for some $a,b \in \mathbb{Z}$ with $ab > 0$ if and only if χ_1 has a purely periodic symmetric continued fraction (i.e. if $\chi_1 = [\overline{r_0,...,r_{m-1}}]$ then $r_i = r_{m-1-i}$ for all $i \in \{0,1,...,m-1\}$).*

Proof. Let $Q(X,Y) = \langle a,-b,-a\rangle_{X,Y}$ for some $a,b \in \mathbb{Z}$. Then since $ab > 0$, the Gauss conditions imply that $Q(X,Y)$ is strongly reduced. By theorem 11.4, χ_1 is purely periodic. Let $\chi_1 = [\overline{r_0,...,r_{m-1}}]$. By the last corollary we deduce that the first (positive) root $\widetilde{\chi}_1$ of $\widetilde{Q}(X,Y)$ is given by $\widetilde{\chi}_1 = [\overline{r_{m-1},...,r_0}]$. Since $\widetilde{Q}(X,Y) = -Q(X,Y)$ and since $Q(T,1)$ and $-Q(T,1)$ have the same first root, we deduce that $\chi_1 = \widetilde{\chi}_1$. Hence the continued fraction of χ_1 is symmetric.

Conversely, let χ_1 have a purely periodic symmetric continued fraction. Then theorem 11.4 implies that χ_1 is a strongly reduced quadratic surd. Let $\chi_1 = a+b\sqrt{D}$ with $a,b \in \mathbb{Q}$. Then χ_1 is the first (positive) root of $(T-\chi_1)(T-\chi_1') = T^2 - 2aT + (a^2 - b^2D)$ where $\chi_1' = a - b\sqrt{D}$ denotes the conjugate of χ_1. Since χ_1 is symmetric $\chi_1 = \widetilde{\chi}_1$. Thus $\chi_1' = -\frac{1}{\widetilde{\chi}_1} = -\frac{1}{\chi_1}$. Hence $a - b\sqrt{D} = -\frac{1}{a+b\sqrt{D}} = -\frac{a-b\sqrt{D}}{a^2-b^2D}$. Hence $a^2 - b^2D = -1$. Thus χ_1 is a root of $x^2 - 2ax - 1$. (Since χ_1 is strongly reduced, we deduce that $0 < \chi_1 + \chi_1' = 2a$.) Since $a \in \mathbb{Q}$, there exists $k \in \mathbb{N}$ such that $2ak = a'$ with $a' \in \mathbb{N}$ and $\gcd(a',k) = 1$. Thus χ_1 and χ' are the roots of $k(T-\chi_1)(T-\chi_1') = kT^2 - a'T - k$ and the roots of $Q(T,1)$. Hence $Q(X,Y) = l\langle k,-a',-k\rangle_{X,Y} = \langle lk,-la',-lk\rangle_{X,Y}$ for some $l \in \mathbb{Z}$. □

Example. The quadratic form $Q(X,Y) = \langle 2,-5,-2\rangle_{X,Y}$ is strongly reduced and its first (positive) root is $\chi_1 = x^+ = \frac{5+\sqrt{41}}{4} = [\overline{2,1,5,1,2}]$.

Continued fractions of square-roots

At last, we would like to highlight the form of the continued fractions of square-roots of non-square positive integers.

Corollary 11.9. *For any $r \in \mathbb{N} \setminus \mathbb{N}^2$, $\sqrt{r} = [r_0, \overline{r_1, r_2, r_3, ..., r_3, r_2, r_1, 2r_0}]$.*

Proof. Let $\sqrt{r} = [r_0, r_1, r_2, ...]$. Then $\sqrt{r}$ is the first root of $Q(T,1)$ where $Q(X,Y) = \langle 1,0,-r\rangle_{X,Y}$ (which is weakly reduced). Consider $x = \partial\sqrt{r} = \frac{1}{\sqrt{r}-\lfloor\sqrt{r}\rfloor} = \frac{\sqrt{r}+\lfloor\sqrt{r}\rfloor}{r-\lfloor\sqrt{r}\rfloor^2} > 1$, then its conjugate is $x' = \frac{-\sqrt{r}+\lfloor\sqrt{r}\rfloor}{r-\lfloor\sqrt{r}\rfloor^2} = -\frac{1}{\sqrt{r}+\lfloor\sqrt{r}\rfloor} \in]-1,0[$. Thus $\partial\sqrt{r}$ is strongly reduced and has

by theorem 11.4 a purely periodic continued fraction. Let $x = [\overline{r_1, ..., r_k}]$ for some $k \in \mathbb{N} \setminus \{0\}$. Thus

$$\sqrt{r} = [r_0, x] = [r_0, \overline{r_1, ..., r_k}].$$

From an other point of view, we can observe that x and x' are the roots of $\partial Q(T, 1)$. Observe that $\partial Q(X, Y) = \partial\langle 1, 0, -r\rangle_{X,Y} = \langle \lfloor\sqrt{r}\rfloor^2 - r, 2\lfloor\sqrt{r}\rfloor, 1\rangle_{X,Y}$ is strongly reduced. Hence corollary 11.6 states that the first root of $\widetilde{Q}(T, 1)$ is given by $[\overline{r_k, ..., r_1}] = -\frac{1}{x'} = \sqrt{r} + \lfloor\sqrt{r}\rfloor$. Since $\lfloor\sqrt{r} + \lfloor\sqrt{r}\rfloor\rfloor = 2\lfloor\sqrt{r}\rfloor$, we deduce that $r_k = 2r_0$. Since $\partial(\sqrt{r} + \lfloor\sqrt{r}\rfloor) = \frac{1}{\sqrt{r} - \lfloor\sqrt{r}\rfloor} = \partial\sqrt{r}$, we conclude that $r_{k-1-i} = r_{1+i}$ for all $i \in \{0, ..., k-2\}$. Thus

$$\sqrt{r} = [r_0, \overline{r_1, ...r_{k-1}, r_k}] = [r_0, \overline{r_1, r_2, ..., r_2, r_1, 2r_0}].$$

□

Example. $\sqrt{19} = 4 + \sqrt{19} - 4 = [\overline{8, 2, 1, 3, 1, 2}] - 4 = [4, \overline{2, 1, 3, 1, 2, 8}]$.

Using the special form of the continued fractions as well as the properties of strongly reduced quadratic forms that have been developed so far, one can construct a basic algorithm to solve a particular Diophantine equation, namely the *Pell equation*. An easy step by step construction of this algorithm is given in appendix C.

12 Equivalence of strongly reduced forms

Sources: [8, p.353-359]
Other approaches:[1, p.21-23], [9, p.23-26], [10, p.120-132], [12, p.359-364], [11, art. 186-189,193,223-227]

Chapter note: Assume throughout this chapter that $D \in \mathbb{N} \setminus \mathbb{N}^2$.

By combining the theory of continued fractions and the theory of quadratic forms, we were able to prove some strong properties of continued fractions. Next, we would like to do the same for quadratic forms. This approach will be developed in chapter 13.

However, before we start with this development, we would like to know more about quadratic forms and their equivalence classes.

In chapter 8, we concluded that the equivalence classes of $Quad(D)$ are completely determined by the strongly reduced forms. More precisely, we deduced that

$$H(D) = Quad(D)/GL(2,\mathbb{Z}) = \bigcup_{Q(X,Y)\in S_{et}^D} Q(X,Y) \circ GL(2,\mathbb{Z})$$

where $S_{et}^D = \{Q(X,Y) \in Quad(D) : Q(X,Y) \text{ is strongly reduced}\}$ for all $D \in \mathbb{N} \setminus \mathbb{N}^2$. Hence, in order to study the set of equivalence classes $H(D)$, it is sufficient to study the equivalence relations inside the set S_{et}^D.

In this chapter, we are going to develop general results that indicate which strongly reduced quadratic forms are equivalent and which are not. Indeed, theorem 12.8 states that two strongly reduced forms are equivalent if and only if they belong to the same cycle or associate cycles.

In order to obtain this result, we go back to the very beginning and analyze finally the definition of derived forms in more detail. At a first glance, the results of section 12.1 might not seem to bring along new information about strongly reduced quadratic forms, however they will be the key results needed to prove theorem 12.7. This particular development comes from the third appendix of Weil's book [8] and seemed to the author of this Master thesis to be one of the best approaches to obtain theorem 12.8. Furthermore, as Weil describes himself, this method goes back to Lagrange and can be seen historically as being the first such approach. We are going to discover how well this fits into our construction.

12.1 Various facts about weakly and strongly reduced quadratic forms

Up to now, we used derived forms in order to obtain most results about quadratic forms. However, we did not yet try to look behind the definition of derived form. This will be done in this section which gives eventually a good reason why the definition of derived forms is stated as given in definition 7.1.

Let us start to analyze the action of similar matrices onto weakly reduced quadratic forms.

Lemma 12.1. *Let $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ be weakly reduced and let χ_1 be the first root of $Q(T,1)$. Let $Q'(X,Y) = Q(X,Y) \circ T(r)$ for some $r \in \mathbb{N} \setminus \{0\}$ where $T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Then*

1. *$Q'(X,Y)$ is weakly reduced if and only if $r < \chi_1$.*

2. *$Q'(X,Y)$ is strongly reduced if and only if $r = \lfloor \chi_1 \rfloor$.*

Proof. Let χ_2 be the second root of $Q(T,1)$. By proposition 5.7, we know that $Q'(X,Y) = \langle ar^2 + br + c, b + 2ar, a \rangle_{X,Y}$. Since $\chi_1 + \chi_2 = -\frac{b}{a}$ and since $\chi_1\chi_2 = \frac{c}{a}$, we conclude that

$$\begin{aligned} Q'(X,Y) &= (ar^2 + br + c)X^2 + (b + 2ar)XY + aY^2 \\ &= (ar^2 - a(\chi_1 + \chi_2)r + a\chi_1\chi_2)X^2 + (-a(\chi_1 + \chi_2) + 2ar)XY + aY^2 \\ &= a(Y - (\chi_2 - r)X)(Y - (\chi_1 - r)X) \end{aligned}$$

Thus the roots of $Q'(T,1)$ are $\frac{1}{\chi_1 - r}$ and $\frac{1}{\chi_2 - r}$. Since $\chi_2 < 0$ and $r \geq 1$, we deduce that $-1 < \frac{1}{\chi_2 - r} < 0$.

1. Hence $r < \chi_1$ if and only if $\chi_1 - r > 0$ if and only if $\frac{1}{\chi_1 - r} > 0$. This holds if and only if $Q'(X,Y)$ is weakly reduced.

2. Hence $r = \lfloor \chi_1 \rfloor$ if and only if $1 > \chi_1 - r > 0$ if and only if $\frac{1}{\chi_1 - r} > 1$. This holds if and only if $Q'(X,Y)$ is strongly reduced.

□

We observe that our definition of derived forms is the only definition that works out fine to obtain strongly reduced forms when we consider derived forms of weakly reduced forms whose first root of their corresponding polynomial is greater than 1 (c.f. lemma 7.6).

The next lemma will only be used later on but since its proof is quite similar to the previous one, we are going to state it already at this point.

Lemma 12.2. *Let $Q(X,Y) = \langle a,b,c \rangle_{X,Y} \in Quad(D)$ be weakly reduced. Let $Q'(X,Y) = Q(X,Y) \circ \left(\begin{smallmatrix} 1 & r \\ 0 & 1 \end{smallmatrix}\right)$ for some $r \in \mathbb{N}$. Then $Q'(X,Y)$ is not strongly reduced.*

Proof. Let χ_2 be the second root of $Q(T,1)$. By proposition 5.7, we know that $Q'(X,Y) = \langle a, b + 2ar, ar^2 + br + c \rangle_{X,Y}$. Since $\chi_1 + \chi_2 = -\frac{b}{a}$ and since $\chi_1\chi_2 = \frac{c}{a}$, we conclude in a similar way than in the last proof that

$$Q'(X,Y) = a(X - (\chi_2 - r)Y)(X - (\chi_1 - r)Y)$$

Thus the roots of $Q'(T,1)$ are $\chi_1 - r$ and $\chi_2 - r$. Since $\chi_2 < 0$ and $r \geq 1$, we deduce that $\chi_2 - r < -1$. Hence $Q'(X,Y)$ cannot be strongly reduced. □

Since by definition two quadratic forms $Q(X,Y)$ and $Q'(X,Y)$ are equivalent if and only if there exists $S \in GL(2,\mathbb{Z})$ such that $Q'(X,Y) = Q(X,Y) \circ S$, we would like to reduce the equivalence to easier matrices. Fortunately, we proved in lemma 4.4, proposition 4.5 and corollary 4.7 that any matrix $S \in GL(2,\mathbb{Z})$ has a product decomposition into easier matrices $Z = \left(\begin{smallmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{smallmatrix}\right)$, $W = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $U^r = \left(\begin{smallmatrix} 1 & r \\ 0 & 1 \end{smallmatrix}\right)$ where $U^r W = T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Hence, we restrict ourselves in a first instance to those easy matrices.

Proposition 12.3. *Let $S' = \left(\begin{smallmatrix} b & \rho \\ d & \sigma \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$ such that $b, \rho, d, \sigma \geq 0$. Let $S = S'T(r) = S\left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ for some $r \in \mathbb{N} \setminus \{0\}$ where $a = rb + \rho \geq 0$ and $c = rd + \sigma \geq 0$. Let $Q_0(X,Y) = \langle a_0, b_0, c_0 \rangle_{X,Y}$ be weakly reduced. Let $Q'(X,Y) = Q_0(X,Y) \circ S' = \langle A', B', C' \rangle_{X,Y}$ and $Q(X,Y) = Q_0(X,Y) \circ S = Q'(X,Y) \circ T(r) = \langle A, B, C \rangle_{X,Y}$. If $Q(X,Y)$ is weakly reduced, then so is $Q'(X,Y)$.*

Proof. Let χ_1 be the first and χ_2 be the second root of $Q_0(T,1)$. Then, since $\chi_1+\chi_2=-\frac{b_0}{a_0}$ and since $\chi_1\chi_2=\frac{c_0}{a_0}$, we conclude that $Q_0(X,Y)=a_0(X-\chi_1Y)(X-\chi_2Y)$.

By proposition 5.7, we know that :

$$\begin{aligned} A=Q_0(a,c)&=a_0(a-\chi_1c)(a-\chi_2c),\\ C=A'=Q_0(b,d)&=a_0(b-\chi_1d)(b-\chi_2d),\\ C'=Q_0(\rho,\sigma)&=a_0(\rho-\chi_1\sigma)(\rho-\chi_2\sigma). \end{aligned}$$

Since $Q_0(X,Y)$ is weakly reduced, we deduce that $\chi_2<0$. Since $\chi_2<0$ and since neither a and c, nor b and d, nor ρ and σ can be equal to zero at the same time (since the matrices $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right),\left(\begin{smallmatrix}b&\rho\\d&\sigma\end{smallmatrix}\right)\in GL(2,\mathbb{Z})$), we deduce that $(a-\chi_2c)>0$, $(b-\chi_2d)>0$ and $(\rho-\chi_2\sigma)>0$. Since $Q(X,Y)$ is weakly reduced, lemma 6.9 yields that $AC<0$ hence $(a-\chi_1c)(b-\chi_1d)<0$. Since

$$\rho-\chi_1\sigma=(a-rb)-\chi_1(c-rd)=(a-\chi_1c)-r(b-\chi_1d),$$

we conclude that

$$\begin{aligned} A'C'&=a_0^2(b-\chi_1d)(b-\chi_2d)(\rho-\chi_1\sigma)(\rho-\chi_2\sigma),\\ &=a_0^2(b-\chi_1d)(b-\chi_2d)((a-\chi_1c)-r(b-\chi_1d))(\rho-\chi_2\sigma),\\ &=a_0^2(b-\chi_1d)(b-\chi_2d)(a-\chi_1c)(\rho-\chi_2\sigma)-ra_0^2(b-\chi_1d)^2(b-\chi_2d)(\rho-\chi_2\sigma). \end{aligned}$$

Since $a_0^2>0$, $r>0$, $(b-\chi_2d)>0$ and $(\rho-\chi_2\sigma)>0$, we deduce that $-ra_0^2(b-\chi_1d)^2(b-\chi_2d)(\rho-\chi_2\sigma)<0$. For the same reasons and since $(b-\chi_1d)(a-\chi_1c)<0$, we deduce that $a_0^2(b-\chi_1d)(b-\chi_2d)(a-\chi_1c)(\rho-\chi_2\sigma)<0$. Hence $A'C'<0$ and so $Q'(X,Y)$ is weakly reduced by lemma 6.9. □

By considering some particular matrices, we can even obtain the strongly reduced property of quadratic forms.

Proposition 12.4. *Let $Q(X,Y)$ be weakly reduced and let χ_1 be the first root of $Q(T,1)$. Let $r_1,r_2\in\mathbb{N}\setminus\{0\}$. Let $Q_1(X,Y)=Q(X,Y)\circ T(r_1)$ and $Q_2(X,Y)=Q_1(X,Y)\circ T(r_2)$ with $T(r)=\left(\begin{smallmatrix}r&1\\1&0\end{smallmatrix}\right)$. If $Q_2(X,Y)$ is weakly reduced, then $Q_1(X,Y)$ is strongly reduced, $\chi_1>1$ and $r_1=\lfloor\chi_1\rfloor$.*

Proof. Set $S'=T(r_1)$, $S=S'T(r_2)$ and $Q_0(X,Y)=Q(X,Y)$. Since $Q_2(X,Y)$ is weakly reduced, proposition 12.3 implies that $Q_1(X,Y)$ is weakly reduced.

Furthermore, if χ_2 denotes the second root of $Q(T,1)$, we know that $Q_1(X,Y)=a(Y-(\chi_1-r)X)(Y-(\chi_2-r)X)$ and that the roots of $Q_1(T,1)$ are given by $\frac{1}{\chi_1-r_1}$ and $\frac{1}{\chi_2-r_1}$. Since $\frac{1}{\chi_2-r_1}<0$ and since $Q_1(X,Y)$ is weakly reduced, we deduce that $\frac{1}{\chi_1-r_1}>0$ is the first root of $Q_1(T,1)$ and so $r_1<\chi_1$.

Since $Q_1(X,Y)$ is weakly reduced, since the first root of $Q_1(T,1)$ is $\frac{1}{\chi_1-r_1}$ and since $Q_2(X,Y)=Q_1(X,Y)\circ T(r_2)$ is weakly reduced, lemma 12.1 implies that $r_2<\frac{1}{\chi_1-r}$.

Since $r_2\geq 1$, we deduce that $1\leq\frac{1}{\chi_1-r_1}$ and so $0<\chi_1-r_1\leq 1$. Since $\chi_1\notin\mathbb{N}$, we deduce that $0<\chi_1-r_1<1$.

Hence $r_1<\chi_1<r_1+1$ which implies that $r_1=\lfloor\chi_1\rfloor$. Since $Q(X,Y)$ is weakly reduced, since χ_1 is the first root of $Q(T,1)$, since $Q_1(X,Y)=Q(X,Y)\circ T(r_1)$ and since $r_1=\lfloor\chi_1\rfloor$, lemma 12.1 implies that $Q_1(X,Y)$ is strongly reduced. □

The last proposition can be generalized to several quadratic forms which are one obtained from each other by similar matrix actions.

Proposition 12.5. *Let $m \in \mathbb{N}\setminus\{0\}$ and $r_1, r_2, r_3, ..., r_m \in \mathbb{N}\setminus\{0\}$. Let $S_i = T(r_1)T(r_2)...T(r_i)$ where $T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Let $Q_0(X,Y)$ be weakly reduced. For all $i \in \{1, 2, ..., m\}$, let $Q_i(X,Y) = Q_0(X,Y) \circ S_i$. Denote the first root of $Q_i(T,1)$ by χ_1^i for all $i \in \{0, 1, 2, ..., m\}$.*

If $Q_m(X,Y)$ is weakly reduced, then $Q_i(X,Y)$ is strongly reduced for all $i \in \{1, 2, ..., m-1\}$ and $r_i = \lfloor \chi_1^{i-1} \rfloor$ for all $i \in \{1, 2, ..., m-1\}$.

Proof. Since all the entries of S_i are non-negative for all $i \in \{1, 2, ..., m\}$ and since $S_{i+1} = S_i T(r_{i+1})$ for all $i \in \{1, 2, ..., m-1\}$ and since $Q_m(X,Y)$ is weakly reduced, proposition 12.3 implies recursively that $Q_{m-1}(X,Y), Q_{m-2}(X,Y), ..., Q_2(X,Y)$ and $Q_1(X,Y)$ are weakly reduced. Since for any $i \in \{0, 1, 2, ..., m-2\}$, the quadratic forms $Q_i(X,Y)$, $Q_{i+1}(X,Y) = Q_i(X,Y) \circ T(r_{i+1})$ and $Q_{i+2}(X,Y) = Q_{i+1}(X,Y) \circ T(r_{i+2})$ are weakly reduced, proposition 12.4 implies that $Q_{i+1}(X,Y)$ is strongly reduced and that $r_{i+1} = \lfloor r_i \rfloor$ for all $i \in \{0, 1, 2, ..., m-2\}$. □

At this point, we should notice that in view of the preceding results, we could not have defined derived forms in another way if we would like to conclude the results about strongly reduced forms and continued fractions of the last two chapters. Indeed, we can state the last proposition differently in more familiar terms as follows.

Corollary 12.6. *Let $m \in \mathbb{N} \setminus \{0\}$ and $r_1, r_2, r_3, ..., r_m \in \mathbb{N} \setminus \{0\}$. Let $S_i = T(r_1)T(r_2)...T(r_i)$ where $T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Let $Q_0(X,Y)$ be weakly reduced. For all $i \in \{1, 2, ..., m\}$, let $Q_i(X,Y) = Q_0(X,Y) \circ S_i$. Denote the first root of $Q_i(T,1)$ by χ_1^i for all $i \in \{0, 1, 2, ..., m\}$.*

If $Q_m(X,Y)$ is weakly reduced, then $Q_i(X,Y)$ is strongly reduced for all $i \in \{1, 2, ..., m-1\}$, $Q_i(X,Y) = \partial^i Q_0(X,Y)$ for all $i \in \{1, 2, ..., m-1\}$, $r_i = \lfloor \chi_1^{i-1} \rfloor$ for all $i \in \{1, 2, ..., m-1\}$ and $\chi_1^i = \partial^i \chi_1^0$ for all $i \in \{1, 2, ..., m-1\}$.

Furthermore, in this case $\chi_1^0 = [r_1, r_2, ...r_{m-1}, \chi_1^m]$.

If $Q_m(X,Y)$ is strongly reduced, then, beside the conclusions that follow since $Q_m(X,Y)$ is weakly reduced, $Q_m(X,Y) = \partial^m Q_0(X,Y)$, $r_m = \lfloor r_{m-1} \rfloor$ and $\chi_1^m = \partial^m \chi_1^0$.

Now, we are prepared to get back to the equivalence of strongly reduced forms in general. Since there is a certain loss of information by doing so, we need to include the associate form of a quadratic form into our statements of the next section.

12.2 Equivalence of strongly reduced quadratic forms

Let us prove the main result of this chapter.

Theorem 12.7. *Let $Q_0(X,Y) = \langle a_0, b_0, c_0 \rangle_{X,Y}$ and $Q_1(X,Y) = \langle a_1, b_1, c_1 \rangle_{X,Y}$ be two strongly reduced quadratic forms. If $Q_0(X,Y)$ and $Q_1(X,Y)$ are equivalent, then $Q_1(X,Y)$ or $\widetilde{Q_1}(X,Y)$ is in the cycle of $Q_0(X,Y)$.*

Proof. By definition, $Q_0(X,Y)$ and $Q_1(X,Y)$ are equivalent if and only if there exists $S \in GL(2,\mathbb{Z})$ such that $Q_1(X,Y) = Q_0(X,Y) \circ S$. We divide our proof into two parts depending on the signs of the entries of S.

1. Let the entries of S be all non-negative. Then, by proposition 4.5, we have

$$S = WU^{r_0}WU^{r_1}W...U^{r_{m-1}}WU^{r_m}W = WT(r_0)T(r_1)...T(r_m)$$

for some $m \in \mathbb{N}$, $r_0, r_m \geq 0$ and if $m \geq 2$ then $r_i \geq 1$ for all $i = 1, ..., m-1$. Let us distinguish four sub-cases depending on the values of r_0 and r_m.

(a) Assume $r_0 = r_m = 0$ and $m = 0$. Then

$$S = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and so $Q_1(X,Y) = Q_0(X,Y)$.

(b) Assume $r_0 = 0$ and $r_m > 0$ ($m > 0$). Then

$$S = U^{r_1}W...U^{r_{m-1}}WU^{r_m}W = T(r_1)...T(r_m).$$

Since $Q_1(X,Y)$ is strongly reduced, corollary 12.6 implies that $Q_1(X,Y) = \partial^m Q_0(X,Y)$.

(c) Assume by contradiction that $r_0 = r_m = 0$ and $m > 0$. Then

$$S = U_1^r W...U^{r_{m-1}} = T(r_1)...T(r_{m-2})U^{r_{m-1}}.$$

Hence,

$$SW = T(r_1)...T(r_{m-2})T(r_{m-1}) = S'T(r_{m-1}),$$

where $S' = T(r_1)...T(r_{m-2})$. Since $Q_1(X,Y) = Q_0(X,Y) \circ S$, we deduce that $Q_1(X,Y) \circ W = Q_0(X,Y) \circ S'T(r_{m-1})$. Since $Q_1(X,Y)$ is weakly reduced, so is $Q_1(X,Y) \circ W$ (which can be easily concluded by lemma 6.9). Since $Q_0(X,Y)$ and $Q_1(X,Y) \circ W$ are weakly reduced, since S' has non-negative entries only and since $Q_1(X,Y) \circ W = Q_0(X,Y) \circ S'T(r_{m-1})$, lemma 12.3 implies that $Q_0(X,Y) \circ S'$ is weakly reduced. By lemma 12.2 and since $\left(\begin{smallmatrix} 1 & r \\ 0 & 1 \end{smallmatrix}\right) = T(r)W$, we deduce that

$$Q_1(X,Y) = (Q_0(X,Y) \circ S') \circ T(r_{m-1})W$$

is not strongly reduced, which contradicts our assumption.

(d) Assume by contradiction that $r_0 > 0$. Then, set $S_0 = WS$ such that

$$Q_1(X,Y) = (Q_0(X,Y) \circ W) \circ S_0 = (Q_0(X,Y) \circ W) \circ T(r_0)T(r_1)...T(r_m).$$

If $m = 0$, by assumption $Q_1(X,Y) = (Q_0(X,Y) \circ W) \circ T(r_0)$ is strongly reduced. If $m > 1$, then proposition 12.5 implies that $(Q_0(X,Y) \circ W) \circ T(r_0)$ is strongly reduced. By lemma 12.1, we deduce that if the first root of $Q_0(T,1) \circ W$ is given by ψ, then $1 \leq r_0 = \lfloor \psi \rfloor$. Let the first root of $Q_0(T,1)$ be given by χ_1, then, the first root of $Q_0(T,1) \circ W$ is equal to $\psi = \frac{1}{\chi_1}$ (which can be easily verified by a direct computation). Since $Q_0(X,Y)$ is strongly reduced $\chi_1 > 1$. Hence $1 < \psi \leq 1$ which clearly is a contradiction.

2. Let $S \in GL(2,\mathbb{Z})$. Then, by lemma 4.4, there exist $Z_1, Z_2 \in \Delta$ such that $S = Z_1 S' Z_2$ with all the entries of S' being non-negative.

   Notice that $\Delta = \{\pm \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \pm \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\} = \{\pm \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \pm \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) W\}$ and that $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) W = -W \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$. Hence, we can rewrite S as one of the following

$$\pm S'', \qquad \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} S'', \qquad \pm S'' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \pm \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} S'' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

   where S'' has non-negative entries only.

   Since $Q_1(X,Y) = Q_0(X,Y) \circ S$ and since the action of a matrix and its negative are the same, $Q_1(X,Y)$ is given by one of the forms

$$Q_0(X,Y) \circ S'', \qquad Q_0(X,Y) \circ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} S'',$$

$$Q_0(X,Y) \circ S'' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad Q_0(X,Y) \circ \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} S'' \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

   In the first case, we conclude by the first part of the proof that $Q_1(X,Y)$ is in the cycle of $Q_0(X,Y)$. In the second case, we conclude by the first part of the proof that $Q_1(X,Y)$ is in the cycle of $Q_0(X,Y) \circ \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = \widetilde{Q_0}(X,Y)$. In the latter two cases, we observe that then $\widetilde{Q_1}(X,Y) = Q_1(X,Y) \circ \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = Q_0(X,Y) \circ S''$ or $\widetilde{Q_1}(X,Y) = Q_0(X,Y) \circ \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) S''$. Furthermore, since $Q_1(X,Y)$ is strongly reduced, $\widetilde{Q_1}(X,Y)$ is strongly reduced. Hence, in the third case, we conclude by the first part of the proof that $\widetilde{Q_1}(X,Y)$ is in the cycle of $Q_0(X,Y)$. Finally, in the fourth case, we conclude by the first part of the proof that $\widetilde{Q_1}(X,Y)$ is in the cycle of $Q_0(X,Y) \circ \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = \widetilde{Q_0}(X,Y)$.

   Since $\widetilde{Q_1}(X,Y)$ is in the cycle of $Q_0(X,Y)$ if and only if $Q_1(X,Y)$ is in the cycle of $\widetilde{Q_0}(X,Y)$ and since $\widetilde{Q_1}(X,Y)$ is in the cycle of $\widetilde{Q_0}(X,Y)$ if and only if $Q_1(X,Y)$ is in the cycle of $Q_0(X,Y)$, we conclude the theorem.

□

Thus, we conclude that any strongly reduced quadratic form that is equivalent to the strongly reduced quadratic form $Q(X,Y)$ belongs necessarily to the cycle of $Q(X,Y)$ or its associate cycle. Since furthermore any quadratic form in the cycle of $Q(X,Y)$ or its associate cycle is strongly reduced and equivalent to $Q(X,Y)$, we deduce the following theorem.

Theorem 12.8. *Let $Q_1(X,Y), Q_2(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ be two strongly reduced quadratic forms. Then $Q_2(X,Y)$ is equivalent to $Q_1(X,Y)$ if and only if $Q_2(X,Y) = \partial^n Q_1(X,Y)$ for some $n \in \mathbb{N}$ or $Q_2(X,Y) = \partial^n \widetilde{Q_1}(X,Y)$ for some $n \in \mathbb{N}$.*

Example. Consider the strongly reduced quadratic form $Q(X,Y) = \langle 19, -22, -4 \rangle_{X,Y}$. By the last theorem the strongly reduced forms that are equivalent to $Q(X,Y)$ are exactly the

following ones :

$$\begin{aligned}
Q_0(X,Y) &= \langle 19,-22,-4\rangle_{X,Y}, & \widetilde{Q_0}(X,Y) &= \langle -4,22,19\rangle_{X,Y};\\
Q_1(X,Y) &= \langle -7,16,19\rangle_{X,Y}, & \widetilde{Q_1}(X,Y) &= \langle 19,-16,-7\rangle_{X,Y};\\
Q_2(X,Y) &= \langle 4,-26,-7\rangle_{X,Y}, & \widetilde{Q_2}(X,Y) &= \langle -7,26,4\rangle_{X,Y};\\
Q_3(X,Y) &= \langle -19,22,4\rangle_{X,Y}, & \widetilde{Q_3}(X,Y) &= \langle 4,-22,-19\rangle_{X,Y};\\
Q_4(X,Y) &= \langle 7,-16,-19\rangle_{X,Y}, & \widetilde{Q_4}(X,Y) &= \langle -19,16,7\rangle_{X,Y};\\
Q_5(X,Y) &= \langle -4,26,7\rangle_{X,Y}, & \widetilde{Q_5}(X,Y) &= \langle 7,-26,-4\rangle_{X,Y}
\end{aligned}$$

where $Q_i(X,Y) = \partial^i Q(X,Y)$ for all $i \in \{0,1,2,3,4,5\}$. Recall that we deduced in a previous example that $\partial^6 Q(X,Y) = Q(X,Y)$.

We have now an easy tool to verify equivalence of strongly reduced forms. Thus, we could now try to find all the equivalent quadratic forms in S_{et}^D and obtain in this manner the class number. However, this method needs still a lot of computational work and so one might try to optimize it first.

13 The equivalence of strongly reduced quadratic forms through the comparison of the continued fraction of the first root of their corresponding polynomials and a method to compute the class number

```
Sources:  Personal work
Other approaches:  [10, p.118-119]
```

Chapter note: Assume throughout this chapter that $D \in \mathbb{N} \setminus \mathbb{N}^2$.

Chapter 12 concluded that the only strongly reduced quadratic forms that are equivalent to a given strongly reduced form are the quadratic form in its cycle and in its associate cycle. Since $H(D) = Quad(D)/GL(2,\mathbb{Z}) = \bigcup_{Q(X,Y)\in S_{et}^D} Q(X,Y) \circ GL(2,\mathbb{Z})$ where $S_{et}^D = \{Q(X,Y) \in Quad(D) : Q(X,Y) \text{ is strongly reduced}\}$ for all $D \in \mathbb{N} \setminus \mathbb{N}^2$, we conclude that $H(D)$ is now completely determined and we could start by developing an algorithm to obtain the class number. This is exactly GAUSS's approach and solution to the problem. However, by pushing our analysis a bit further, we will obtain some more results which will simplify the computations.

Indeed, we discovered in chapter 10 that if the cycle of $Q(X,Y)$ is known, then the cycle of $\widetilde{Q}(X,Y)$ and the cycle of $-Q(X,Y)$ are known. The results of section 13.1 present some surprising and innovative results about when those cycles are identified and so, when the cycle of $Q(X,Y)$ and of $-Q(X,Y)$ denote the same or different equivalence classes.

The whole analysis will make clear which computations are necessary and which are not to compute the class number. This will lead finally to a method to compute the class number $h(D)$.

13.1 On the number of cycles and equivalence classes that can be deduced from one cycle

Recall that in chapter 10 in section 10.2, we discovered that if one knows the cycle of a given quadratic form, then one knows also its associate cycle and the cycle of its negative form and therefore also the associate cycle of its negative form.

Hence, one could think that from a single cycle, one obtains four distinct cycles and four distinct equivalence classes. However this is wrong. Indeed, the associate form of a given quadratic form and this form are equivalent and so are then their cycles.

Hence, from the four cycles above, we get at most two distinct equivalence classes. Then, because of the first counter intuitive result, one might think that a given quadratic form and its negative are always equivalent and hence define a single equivalence class, however this does not hold in general. An easy counterexample to this fact which does not rely on our development and a late motivation for our non-standard approach of this document is given in appendix D.

Because of these confusing possibilities, we would like to determine the number of distinct cycles as well as the number of equivalence classes that can be obtained from the cycle of $Q(X,Y)$ depending on the forms that belong to this cycle. To achieve this task, we will take a small deviation and develop results that depend on the continued fraction of the first root of $Q(T,1)$.

To facilitate this analysis, we are first going to study the relation between the length of the cycle of $Q(X,Y)$ and the caliber of the first root of $Q(T,1)$. Then, we are going to develop some general results about which cycles are identical and finally, we will have confirmation that a quadratic form and its negative are not always equivalent.

The relation between the length of the cycle of $Q(X,Y)$ and the caliber of the first root of $Q(T,1)$

To begin with, we would like to find the relation between the length of the cycle of $Q(X,Y)$ and the caliber of the first root of $Q(T,1)$.

Lemma 13.1. *Let $Q(X,Y)$ be a strongly reduced quadratic form and let χ_1 be the first root of $Q(T,1)$. Let m denote the caliber of χ_1 and let M denote the length of the cycle of $Q(X,Y)$. Then M is a multiple of m.*

Proof. By definition of the caliber of χ_1, we know that m is the smallest positive integer such that $\partial^m \chi_1 = \chi_1$. By definition of the length of the cycle of $Q(X,Y)$, we know that M is the smallest positive integer such that $\partial^M Q(X,Y) = Q(X,Y)$.

Since the first root of $\partial^M Q(T,1)$ is given by $\partial^M \chi_1$ and since the first root of $Q(T,1)$ is given by χ_1 and since $\partial^M Q(T,1) = Q(T,1)$, we deduce that $\partial^M \chi_1 = \chi_1$. Thus, there exist $k, r \in \mathbb{N}$ such that $M = km + r$ and $0 \leq r < m$. Hence $\chi_1 = \partial^M \chi_1 = \partial^{km+r} \chi_1 = \partial^r(\partial^{km}\chi_1) = \partial^r \chi_1$. Since $\partial^r \chi$ and since $0 \leq r < m$, the minimality of m implies that $r = 0$. Furthermore, since $M, m > 0$, we conclude that $M = km$ for some positive integer k. □

Proposition 13.2. *Let $Q(X,Y)$ be a strongly reduced quadratic form and let χ_1 be the first root of $Q(T,1)$. Let m denote the caliber of χ_1 and let M denote the length of the cycle of $Q(X,Y)$.*

1. *Firstly, the following are equivalent:*
   - *(a) m is even.*
   - *(b) $\partial^m Q(X,Y) = Q(X,Y)$.*
   - *(c) $M = m$*
2. *Secondly, the following are equivalent:*
   - *(a) m is odd.*
   - *(b) $\partial^m Q(X,Y) = -Q(X,Y)$.*
   - *(c) $M = 2m$*

Proof. First of all, we observe that since $\partial^m \chi_1 = \chi_1$, the polynomials $Q(T,1)$ and $\partial^m Q(T,1)$ have the same roots. Moreover, since they are equivalent, the quadratic forms $Q(X,Y)$ and $\partial^m Q(X,Y)$ have the same discriminant. Hence, lemma 6.6 implies that $\partial^m Q(X,Y) = \pm Q(X,Y)$.

1. Let m be even. Since the signs of the first coefficient of $\partial^m Q(X,Y)$ and of the first coefficient of $Q(X,Y)$ are equal, this implies that $\partial^m Q(X,Y) = Q(X,Y)$.

   Let $\partial^m Q(X,Y) = Q(X,Y)$. Since by lemma 13.1 $M = km$ for some $k \in \mathbb{N} \setminus \{0\}$, we deduce by minimality of M that $M = m$.

   Let $M = m$. Then, since by lemma 10.5, M is even, so is then m.

2. Let m be odd. Since the signs of the first coefficient of $\partial^m Q(X,Y)$ and of the first coefficient of $Q(X,Y)$ are equal, this implies that $\partial^m Q(X,Y) = -Q(X,Y)$.

   Let $\partial^m Q(X,Y) = -Q(X,Y)$. Then $\partial^{2m} Q(X,Y) = \partial^m(\partial^m Q(X,Y)) = \partial^m(-Q(X,Y)) = -\partial^m Q(X,Y) = -(-Q(X,Y)) = Q(X,Y)$. Since by lemma 13.1 $M = km$ for some $k \in \mathbb{N} \setminus \{0\}$ and since $k \neq 1$ by assumption, we deduce by minimality of M that $M = 2m$.

   Let $M = 2m$. Assume by contradiction that m is even. Then by part 1., $M = m$ which contradicts our assumption. Hence m is odd.

□

Example. Let $Q(X,Y) = \langle 1, -8, -3\rangle_{X,Y} \in Quad(76)$. Then the first root of $Q(T,1)$ is given by $4+\sqrt{19}$. By a previous example from chapter 11 section 10.2, we know that the caliber of χ_1 is 6 and that indeed the length of the cycle of $Q(X,Y)$ is also 6 as is predicted by the last proposition.

Example. Let $Q(X,Y) = \langle 19, -22, -4\rangle_{X,Y} \in Quad(788)$. Then the first root of $Q(T,1)$ is given by $\frac{11+\sqrt{197}}{19}$. By a previous example from chapter 11 section 10.2, we know that the caliber of χ_1 is 3 and that indeed $\partial^3 Q(X,Y) = -Q(X,Y)$ and that the length of the cycle of $Q(X,Y)$ is 6 as is predicted by the last proposition.

Some general observations

Let $Q(X,Y) \in Quad(D)$. Then, we would like to compute the cycle of $Q(X,Y)$ and deduce other cycles from this cycle.

While computing the cycle of $Q(X,Y)$, one can easily check if $-Q(X,Y)$, $\widetilde{Q}$ or $-\widetilde{Q}$ belong to the cycle of $Q(X,Y)$ by comparing them with each of the quadratic forms that belong to the cycle of $Q(X,Y)$.

Then, by the properties of associate forms and derived forms, we can deduce the following statements:

1. $-Q(X,Y)$ belongs to the cycle of $Q(X,Y)$ if and only if $-\widetilde{Q}(X,Y)$ belongs to the cycle of $\widetilde{Q}$.
2. $\widetilde{Q}(X,Y)$ belongs to the cycle of $Q(X,Y)$ if and only if $-\widetilde{Q}(X,Y)$ belongs to the cycle of $-Q(X,Y)$.
3. $-\widetilde{Q}(X,Y)$ belongs to the cycle of $Q(X,Y)$ if and only if $\widetilde{Q}$ belongs to the cycle of $-Q(X,Y)$.

Hence, if two of the three quadratic forms $-Q(X,Y)$, $\widetilde{Q}$ and $-\widetilde{Q}$ belong to the cycle of $Q(X,Y)$ so does then also the third. At this point, we should ask if it is even possible that the cycle of $Q(X,Y)$ includes one of those forms or all of them.

The three following subsections will answer this question and give direct methods to conclude the inclusion of one of those forms in the cycle of $Q(X,Y)$ by analyzing the general shape of the continued fraction of the first root of $Q(T,1)$. Then in a fourth subsection, we will conclude when a strongly reduced quadratic form and its negative are equivalent.

Condition on when $-Q(X,Y)$ belongs to the cycle of $Q(X,Y)$

Using arguments similar to those in the proofs of lemma 13.1 and of proposition 13.2, one can prove the following corollary.

Corollary 13.3. *Let $Q(X,Y)$ be a strongly reduced quadratic form, let χ_1 be the first root of $Q(T,1)$, let m be the caliber of χ_1 and let M be the length of the cycle of $Q(X,Y)$. Then the following are equivalent:*

1. *$-Q(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{0,1,2,...,M-1\}$.*
2. *$k = m$ and m is odd.*

Example. Using again the two examples above, we observe that the cycle of $Q(X,Y) = \langle 1,-8,-3\rangle_{X,Y} \in Quad(76)$ does not include $-Q(X,Y)$ and that the caliber of the first root of $Q(T,1)$ is even, whereas the cycle of $Q'(X,Y) = \langle 19,-22,-4\rangle_{X,Y} \in Quad(788)$ includes $-Q'(X,Y)$ and the caliber of the first root of $Q'(T,1)$ is odd.

Condition on when $\widetilde{Q}(X,Y)$ belongs to the cycle of $Q(X,Y)$

Proposition 13.4. *Let $Q(X,Y)$ be a strongly reduced quadratic form, let χ_1 be the first root of $Q(T,1)$, let m be the caliber of χ_1 and let M denote the length of the cycle of $Q(X,Y)$.*

1. *If m is even, then the following are equivalent:*
   (a) *$\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{0,1,2,...,M-1\}$.*
   (b) *For some $j,n \in \mathbb{N}$,*
   $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_{j-1}, ..., r_1, r_0, \quad r_{j+1}, ..., r_{j+n-1}, r_{j+n}, r_{j+n-1}, ..., r_{j+1}}]$.
2. *If m is odd, then the following are equivalent:*
   (a) *$\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{0,1,2,...,m-1\}$.*
   (b) *For some $j,n \in \mathbb{N}$,*
   $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_{j-1}, ..., r_1, r_0, \quad r_{j+1}, ..., r_{j+n-1}, r_{j+n}, r_{j+n}, r_{j+n-1}, ..., r_{j+1}}]$.
3. *If m is odd, then the following are equivalent:*
   (a) *$\widetilde{Q}(X,Y) = \partial^m Q(X,Y)$.*
   (b) *For some $j \in \mathbb{N}$,*
   $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_{j-1}, ..., r_1, r_0}]$.
4. *If m is odd, then the following are equivalent:*
   (a) *$\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{m+1, m+2, ..., M-1\}$.*
   (b) *For some $j,n \in \mathbb{N}$,*
   $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_j, r_{j-1}, ..., r_1, r_0, \quad r_{j+1}, ..., r_{j+n-1}, r_{j+n}, r_{j+n-1}, ..., r_{j+1}}]$.

Proof. Because of the similarity of the arguments, we are going to prove first that if $\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{0,1,2,...,M-1\}$, then the continued fraction of χ_1 is given by one of the above forms depending on the parity of m and the value of k. Then we are going to show that the respective converses also hold.

1. Observe that if $\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{0,1,2,...,M-1\}$, then the sign of the first coefficient of $\widetilde{Q}(X,Y)$ and of $\partial^k Q(X,Y)$ are equal and thus $k = 2j+1$ for some $j \in \mathbb{N}$. Since $Q(X,Y)$ is strongly reduced, so is its first root. Let $\chi_1 = [\overline{r_0,...,r_{m-1}}]$. We observe that $\widetilde{\chi_1} = \partial^k \chi_1 = \partial^{2j+1}\chi_1$ since both real numbers are the first root of $\widetilde{Q}(T,1) = \partial^k Q(T,1)$.

   (a) Assume m is even. Then since $M = m$, we deduce that $k < m$. Let $m = 2m_0$ for some $m_0 \in \mathbb{N} \setminus \{0\}$. We deduce that

   $$[\overline{r_{2m_0-1},...,r_{2j+1},\quad r_{2j},...,r_0}] = \widetilde{\chi_1} = \partial^{2j+1}\chi_1 = [\overline{r_{2j+1},...,r_{2m_0-1},\quad r_0,...,r_{2j}}].$$

   Thus, we conclude that the continued fraction of χ_1 is given by

   $$[\overline{r_0,...,r_{j-1},r_j,r_{j-1},...,r_0,\quad r_{2j+1},...,r_{m_0+j-1},r_{m_0+j},r_{m_0+j-1},...,r_{2j+1}}].$$

   (b) Assume m is odd. Let $m = 2m_0+1$ for some $m_0 \in \mathbb{N}$.

      i. If $2j+1 = k < m = 2m_0+1$, then we deduce that

      $$[\overline{r_{2m_0},...,r_{2j+1},\quad r_{2j},...,r_0}] = \widetilde{\chi_1} = \partial^{2j+1}\chi_1 = [\overline{r_{2j+1},...,r_{2m_0},\quad r_0,...,r_{2j}}].$$

      Thus, we conclude that the continued fraction of χ_1 is given by

      $$[\overline{r_0,...,r_{j-1},r_j,r_{j-1},...,r_0,\quad r_{2j+1},...,r_{m_0+j-1},r_{m_0+j},r_{m_0+j},r_{m_0+j-1},...,r_{2j+1}}].$$

      ii. If $2j+1 = k = m = 2m_0+1$, then we deduce that

      $$[\overline{r_{2m_0},...,r_0}] = \widetilde{\chi_1} = \partial^{2j+1}\chi_1 = \chi_1 = [\overline{r_0,...,r_{2m_0}}].$$

      Thus, we conclude that the continued fraction of χ_1 is given by

      $$[\overline{r_0,...,r_{m_0-1},r_{m_0},r_{m_0-1},...,r_0}].$$

      iii. If $2j+1 = k > m = 2m_0+1$, then $k = m+2l$ for some $l \in \mathbb{N} \setminus \{0\}$ (and $2l < m$) and we deduce that $\widetilde{\chi_1} = \partial^k \chi_1 = \partial^{2l+m}\chi_1 = \partial^{2l}\chi_1$ and so

      $$[\overline{r_{2m_0},...,r_{2l},\quad r_{2l-1},...,r_0}] = \widetilde{\chi_1} = \partial^{2l}\chi_1 = [\overline{r_{2l},...,r_{2m_0},\quad r_0,...,r_{2l-1}}].$$

      Thus, we conclude that the continued fraction of χ_1 is given by

      $$[\overline{r_0,...r_{l-2},r_{l-1},r_{l-1},r_{l-2},...,r_0,\quad r_{2l},...,r_{m_0+l-1},r_{m_0+l},r_{m_0+l-1},...,r_{2l}}].$$

2. Let $\widetilde{\chi_1}$ denote the first root of $\widetilde{Q}(T,1)$ and assume that χ_1 is of one of the given forms.

   (a) If $\chi_1 = [\overline{r_0,r_1,...,r_{j-1},r_j,r_{j-1},...,r_1,r_0,\quad r_{j+1},...,r_{j+n-1},r_{j+n},r_{j+n-1},...,r_{j+1}}]$, then $\widetilde{\chi_1} = \partial^{2j+1}\chi_1$. Since $2j+1$ is odd, the first coefficient of $Q(X,Y)$ and $\partial^{2j+1}Q(X,Y)$ have opposite signs as do the first coefficients of $Q(X,Y)$ and $\widetilde{Q}(X,Y)$. Since the signs of the first coefficients of $\partial^{2j+1}Q(X,Y)$ and $\widetilde{Q}(X,Y)$ are equal and since $\partial^{2j+1}Q(T,1)$ and $\widetilde{Q}(T,1)$ have the same first root and since both quadratic forms have the same discriminant (as does $Q(X,Y)$), we conclude by lemma 6.6 that $\widetilde{Q}(X,Y) = \partial^{2j+1}Q(X,Y)$.

(b) If $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_{j-1}, ..., r_1, r_0, \quad r_{j+1}, ..., r_{j+n-1}, r_{j+n}, r_{j+n}, r_{j+n-1}, ..., r_{j+1}}]$, then the same arguments as in the first case show that $\widetilde{Q}(X,Y) = \partial^{2j+1}Q(X,Y)$.

(c) If $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_{j-1}, ..., r_1, r_0}]$, then the same arguments as in the first case show that $\widetilde{Q}(X,Y) = \partial^{2j+1}Q(X,Y)$.

(d) If $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_j, r_{j-1}, ..., r_1, r_0, \quad r_{j+1}, ..., r_{j+n-1}, r_{j+n}, r_{j+n-1}, ..., r_{j+1}}]$, then we observe that the caliber m of χ_1 is odd (and given by $m = 2j+2+2n-1 = 2j+2n+1$) and that $\widetilde{\chi_1} = \partial^{2j+2}\chi_1 = \partial^{2j+2+m}\chi_1$. Since $2j+2+m$ is odd, the first coefficient of $Q(X,Y)$ and $\partial^{2j+2+m}Q(X,Y)$ have opposite signs as do the first coefficients of $Q(X,Y)$ and $\widetilde{Q}(X,Y)$. Since the signs of the first coefficients of $\partial^{2j+2+m}Q(X,Y)$ and $\widetilde{Q}(X,Y)$ are equal and since $\partial^{2j+2+m}Q(T,1)$ and $\widetilde{Q}(T,1)$ have the same first root and since both quadratic forms have the same discriminant (as does $Q(X,Y)$), we conclude by lemma 6.6 that $\widetilde{Q}(X,Y) = \partial^{2j+2+m}Q(X,Y)$.

□

Example. Consider again the example of $Q(X,Y) = \langle 1, -8, -3\rangle_{X,Y} \in Quad(76)$. Then, we observed that $\partial Q(X,Y) = \langle -3, 8, 1\rangle_{X,Y} = \widetilde{Q}(X,Y)$. Furthermore, the first root of $Q(T,1)$ is given by $4 + \sqrt{19} = [\overline{8, \quad 2, 1, 3, 1, 2}]$.

Condition on when $-\widetilde{Q}(X,Y)$ belongs to the cycle of $Q(X,Y)$

Proposition 13.5. *Let $Q(X,Y)$ be a strongly reduced quadratic form, let χ_1 be the first root of $Q(T,1)$, let m be the caliber of χ_1 and let M denote the length of the cycle of $Q(X,Y)$.*

1. *If m is even, then the following are equivalent:*
   (a) $-\widetilde{Q}(X,Y) = Q(X,Y)$.
   (b) *For some $j \in \mathbb{N}$,* $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_j, r_{j-1}, ..., r_0}]$.
2. *If m is even, then the following are equivalent:*
   (a) $-\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ *for some* $k \in \{1, 2, ..., M-1\}$.
   (b) *For some $j, n \in \mathbb{N}$,* $\chi_1 = [\overline{r_0, ..., r_{j-1}, r_j, r_j, r_{j-1}, ..., r_0, \quad r_{j+1}, ..., r_{j+n}, r_{j+n}, ..., r_{j+1}}]$.
3. *If m is odd, then the following are equivalent:*
   (a) $-\widetilde{Q}(X,Y) = Q(X,Y)$.
   (b) *For some $j \in \mathbb{N}$,* $\chi_1 = [\overline{r_0, ..., r_{j-1}, r_j, r_{j-1}, ..., r_0}]$.
4. *If m is odd, then the following are equivalent:*
   (a) $-\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ *for some* $k \in \{1, 2, ..., m-1\}$.
   (b) *For some $j, n \in \mathbb{N}$,* $\chi_1 = [\overline{r_0, ..., r_j, r_j, ..., r_0, \quad r_{j+1}, ..., r_{n+j-1}, r_{n+j}, r_{n+j-1}, ..., r_{j+1}}]$. .
5. *If m is odd, then the following are equivalent:*

(a) $-\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ *for some* $k \in \{m+1, m+2, ..., M-1\}$.

(b) For some $j, n \in \mathbb{N}$,
$\chi_1 = [\overline{r_0, ...r_{j-1}, r_j, r_{j-1}, ..., r_0, \quad r_{j+1}, ..., r_{j+n-1}, r_{j+n}, r_{j+n}, r_{j+n-1}, ..., r_{j+1}}]$.

Proof. Because of the similarity of the arguments, we are going to prove first that if $-\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{0,1,2,...,M-1\}$, then the continued fraction of χ_1 is given by one of the above forms depending on the parity of m and the value of k. Then we are going to show that the respective converses also hold.

1. Observe that if $-\widetilde{Q}(X,Y) = \partial^k Q(X,Y)$ for some $k \in \{0,1,2,...,M-1\}$, then the sign of the first coefficient of $-\widetilde{Q}(X,Y)$ and of $\partial^k Q(X,Y)$ are equal and thus $k = 2j$ for some $j \in \mathbb{N}$. Since $Q(X,Y)$ is strongly reduced, so is its first root. Let $\chi_1 = [\overline{r_0, ..., r_{m-1}}]$. We observe that $-\widetilde{\chi_1} = \partial^k \chi_1 = \partial^{2j} \chi_1$ since both real numbers are the first root of $-\widetilde{Q}(T,1) = \partial^k Q(T,1)$.

   (a) Assume m is even. Then since $M = m$, we deduce that $k < m$. Let $m = 2m_0$ for some $m_0 \in \mathbb{N} \setminus \{0\}$.

      i. If $k = 0$, we deduce that

      $$[\overline{r_{2m_0-1}, r_{2m_0-2}, ..., r_0}] = \widetilde{\chi_1} = \chi_1 = [\overline{r_0, r_1, ..., r_{2m_0-1}}].$$

      Thus, we conclude that the continued fraction of χ_1 is given by

      $$[\overline{r_0, r_1, ..., r_{m_0-2}, r_{m_0-1}, r_{m_0-1}, r_{m_0-2}, ..., r_0}].$$

      ii. If $k > 0$, we deduce that

      $$[\overline{r_{2m_0-1}, ..., r_{2j}, \quad r_{2j-1}, ..., r_0}] = \widetilde{\chi_1} = \partial^{2j}\chi_1 = [\overline{r_{2j}, ..., r_{2m_0-1}, \quad r_0, ..., r_{2j-1}}].$$

      Thus, we conclude that the continued fraction of χ_1 is given by

      $$[\overline{r_0, ..., r_{j-2}, r_{j-1}, r_{j-1}, r_{j-2}, ..., r_0, \quad r_{2j}, ..., r_{m_0+j-1}, r_{m_0+j-1}, ..., r_{2j}}].$$

   (b) Assume m is odd. Let $m = 2m_0 + 1$ for some $m_0 \in \mathbb{N}$.

      i. If $2j = k = 0$, then we deduce that

      $$[\overline{r_{2m_0}, ..., r_0}] = \widetilde{\chi_1} = \chi_1 = [\overline{r_0, ..., r_{2m_0}}].$$

      Thus, we conclude that the continued fraction of χ_1 is given by

      $$[\overline{r_0, ..., r_{m_0-1}, r_{m_0}, r_{m_0-1}, ..., r_0}].$$

      ii. If $0 < 2j = k < m = 2m_0 + 1$, then we deduce that

      $$[\overline{r_{2m_0}, ..., r_{2j}, \quad r_{2j-1}, ..., r_0}] = \widetilde{\chi_1} = \partial^{2j}\chi_1 = [\overline{r_{2j}, ..., r_{2m_0}, \quad r_0, ..., r_{2j-1}}].$$

      Thus, we conclude that the continued fraction of χ_1 is given by

      $$[\overline{r_0, ..., r_{j-1}, r_{j-1}, ..., r_0, \quad r_{2j}, ..., r_{m_0+j-1}, r_{m_0+j}, r_{m_0+j-1}, ..., r_{2j}}].$$

iii. If $2j = k > m = 2m_0 + 1$, then $k = m + 2l + 1$ for some $l \in \mathbb{N}$ (and $2l + 1 < m$) and we deduce that $\widetilde{\chi_1} = \partial^k \chi_1 = \partial^{2l+1+m}\chi_1 = \partial^{2l+1}\chi_1$ and so

$$[\overline{r_{2m_0}, ..., r_{2l+1}, \quad r_{2l}, ..., r_0}] = \widetilde{\chi_1} = \partial^{2l+1}\chi_1 = [\overline{r_{2l+1}, ..., r_{2m_0}, \quad r_0, ..., r_{2l}}].$$

Thus, we conclude that the continued fraction of χ_1 is given by

$$[\overline{r_0, ...r_{l-1}, r_l, r_{l-1}, ..., r_0, \quad r_{2l+1}, ..., r_{m_0+l-1}, r_{m_0+l}, r_{m_0+l}, r_{m_0+l-1}, ..., r_{2l+1}}].$$

2. Let $\widetilde{\chi_1}$ denote the first root of $\widetilde{Q}(T, 1)$ and assume that χ_1 is of one of the given forms.

   (a) If $\chi_1 = [\overline{r_0, r_1, ..., r_{j-1}, r_j, r_j, r_{j-1}, ..., r_0}]$.,
   then $\widetilde{\chi_1} = \chi_1$. Since the first coefficient of $Q(X, Y)$ and $-\widetilde{Q}(X, Y)$ have the same sign and since $Q(T, 1)$ and $-\widetilde{Q}(T, 1)$ have the same first root and since both quadratic forms have the same discriminant (as does $Q(X, Y)$), we conclude by lemma 6.6 that $-\widetilde{Q}(X, Y) = Q(X, Y)$.

   (b) If $\chi_1 = [\overline{r_0, ..., r_{j-1}, r_j, r_j, r_{j-1}, ..., r_0, \quad r_{j+1}, ..., r_{j+n}, r_{j+n}, ..., r_{j+1}}]$,
   then $\widetilde{\chi_1} = \partial^{2j+2}\chi_1$. Since $2j + 2$ is even, the first coefficient of $Q(X, Y)$ and $\partial^{2j+2}Q(X, Y)$ have equal signs as do the first coefficients of $Q(X, Y)$ and $-\widetilde{Q}(X, Y)$. Since the signs of the first coefficients of $\partial^{2j+2}Q(X, Y)$ and $-\widetilde{Q}(X, Y)$ are equal and since $\partial^{2j+2}Q(T, 1)$ and $-\widetilde{Q}(T, 1)$ have the same first root and since both quadratic forms have the same discriminant (as does $Q(X, Y)$), we conclude by lemma 6.6 that $-\widetilde{Q}(X, Y) = \partial^{2j+2}Q(X, Y)$.

   (c) If $\chi_1 = [\overline{r_0, ..., r_{j-1}, r_j, r_{j-1}, ..., r_0}]$, then the same arguments as in the first case show that $-\widetilde{Q}(X, Y) = Q(X, Y)$.

   (d) If $\chi_1 = [\overline{r_0, ..., r_j, r_j, ..., r_0, \quad r_{j+1}, ..., r_{n+j-1}, r_{n+j}, r_{n+j-1}, ..., r_{j+1}}]$, then the same arguments as in the second case show that $-\widetilde{Q}(X, Y) = \partial^{2j+2}Q(X, Y)$.

   (e) If $\chi_1 = [\overline{r_0, ...r_{j-1}, r_j, r_{j-1}, ..., r_0, \quad r_{j+1}, ..., r_{j+n-1}, r_{j+n}, r_{j+n}, r_{j+n-1}, ..., r_{j+1}}]$,
   then we observe that the caliber m of χ_1 is odd (and given by $m = 2j + 1 + 2n = 2j + 2n + 1$) and that $\widetilde{\chi_1} = \partial^{2j+1}\chi_1 = \partial^{2j+1+m}\chi_1$. Since $2j + 1 + m$ is even, the first coefficient of $Q(X, Y)$ and $\partial^{2j+1+m}Q(X, Y)$ have equal signs as do the first coefficients of $Q(X, Y)$ and $-\widetilde{Q}(X, Y)$. Since the signs of the first coefficients of $\partial^{2j+1+m}Q(X, Y)$ and $-\widetilde{Q}(X, Y)$ are equal and since $\partial^{2j+1+m}Q(T, 1)$ and $-\widetilde{Q}(T, 1)$ have the same first root and since both quadratic forms have the same discriminant (as does $Q(X, Y)$), we conclude by lemma 6.6 that $-\widetilde{Q}(X, Y) = \partial^{2j+1+m}Q(X, Y)$.

□

Example. Consider the strongly reduced quadratic surd $x = 9 + \sqrt{221} > 1$. Then x is the first root of the polynomial $7T^2 - 9T - 5$. The continued fraction of x is given by $x = [\overline{1, 1, 2, 2}]$ and the cycle of $Q(X, Y) = \langle 7, -9, -5\rangle_{X,Y}$ is given by

$$\{\langle 7, -9, -5\rangle_{X,Y}, \langle -7, 5, 7\rangle_{X,Y}, \langle 5, -9, -7\rangle_{X,Y}, \langle -5, 11, 5\rangle_{X,Y}\}.$$

We observe that $-\widetilde{Q}(X, Y) = \langle 5, -9, -7\rangle_{X,Y}$ belongs to the cycle of $Q(X, Y)$ which verifies the proposition.

An interesting coincidence is given in the next corollary.

Corollary 13.6. *Let $Q(X,Y)$ be a strongly reduced quadratic form. Then $-\widetilde{Q}(X,Y)$ belongs to the cycle of $Q(X,Y)$ if and only if there exists $d \in \mathbb{N}$ such that $\partial^d Q(X,Y) = \langle a, -b, -a\rangle_{X,Y}$ for some $a, b \in \mathbb{Z}$ such that $ab > 0$.*

Proof. Let χ_1 be the first root of $Q(T,1)$ and let m be the caliber of χ_1.

Let $-\widetilde{Q}(X,Y)$ belong to the cycle of $Q(X,Y)$. Then the continued fraction of χ_1 is of one of the forms given in proposition 13.5. One notices that for any such form, there exists $d \in \mathbb{N}$ such that $\partial^d \chi_1$ is symmetric. We can assume without loss of generality that $d < M$ where M denotes the length of the cycle of $Q(X,Y)$. Since $\partial^d \chi_1$ is the first root of $\partial^d Q(T,1)$, corollary 11.8 implies that $\partial^d Q(X,Y) = \langle a, -b, -a\rangle_{X,Y}$ for some $a, b \in \mathbb{Z}$ such that $ab > 0$.

Let there be $d \in \mathbb{N}$ such that $\partial^d Q(X,Y) = \langle a, -b, a\rangle_{X,Y}$ for some $a, b \in \mathbb{Z}$ such that $ab > 0$. Then by corollary 11.8, the first root of $\partial^d Q(T,1)$ is symmetric.
If $m = 2n+2$ is even, then we notice that $\partial^d \chi_1 = [\overline{r_0, r_1, ..., r_n, r_n, ..., r_1, r_0}]$. Hence

$$\chi_1 = [r_d, ..., r_1, r_0, r_0, r_1, ..., r_d, r_{d+1}, ..., r_n, r_n, ..., r_{d+1}].$$

If $m = 2n+1$ is odd, then we notice that $\partial^d \chi_1 = [\overline{r_0, r_1, ..., r_n, ..., r_1, r_0}]$. Hence, if $0 < d < m$, then

$$\chi_1 = [r_d, ..., r_1, r_0, r_0, r_1, ..., r_d, r_{d+1}, ..., r_n, ..., r_{d+1}],$$

if $d = 0$ or $d = m$, then

$$\chi_1 = [\overline{r_0, r_1, ..., r_n, ..., r_1, r_0}],$$

and if $d > m$, then $d = m + d'$ and so

$$\chi_1 = [r_{d'}, ..., r_1, r_0, r_0, r_1, ..., r_{d'}, r_{d'+1}, ..., r_n, ..., r_{d'+1}].$$

We observe that in each case, we can conclude that $-\widetilde{Q}(X,Y)$ belongs to the cycle of $Q(X,Y)$ by proposition 13.5. □

Condition on when $-Q(X,Y)$ is equivalent to $Q(X,Y)$

By theorem 12.8, we know that $-Q(X,Y)$ is equivalent to $Q(X,Y)$ if and only if either $-Q(X,Y)$ is in the cycle of $Q(X,Y)$ or $-Q(X,Y)$ is in the cycle of $\widetilde{Q}(X,Y)$. The first condition is easy to verify by corollary 13.3. The second condition is equivalent to $-\widetilde{Q}(X,Y)$ belonging to the cycle of $Q(X,Y)$. By corollary 13.6, this latter condition holds if and only if the cycle of $Q(X,Y)$ includes a quadratic form of the form $\langle a, -b, -a\rangle_{X,Y}$ where $ab > 0$.

Hence, we conclude the following theorem.

Theorem 13.7. *Let $Q(X,Y)$ be a strongly reduced quadratic form and let χ_1 be the first root of $Q(T,1)$. Then the following are equivalent:*

1. *$-Q(X,Y)$ is equivalent to $Q(X,Y)$.*
2. *The caliber of χ_1 is odd or there exists $a, b \in \mathbb{Z}$ such that $ab > 0$ and such that $\langle a, -b, -a\rangle_{X,Y}$ belongs to the cycle of $Q(X,Y)$.*

This solves one of the motivating questions for this Master thesis as is described in appendix D. Furthermore, it confirms that we have to pay attention to not identify the equivalence classes of $Q(X,Y)$ and $-Q(X,Y)$ without testing the above conditions or similar ones.

Conclusions

In the last four subsections, we discovered conditions on when $-Q(X,Y)$, $\widetilde{Q}(X,Y)$ or $-\widetilde{Q}(X,Y)$ belong to the cycle of $Q(X,Y)$. By inspecting these conditions, we observe that it is possible that either none or only one or all three of those quadratic forms belong to the cycle of $Q(X,Y)$. Furthermore, the conditions give us easy methods to construct all examples and counter-examples that we want for each possible case.

The last subsection gives us an easy method to verify if a given strongly reduced quadratic form is equivalent to its negative or not. In particular, we concluded that not every quadratic form is equivalent to its negatives.

Because of all the possibilities, we must indeed pay attention when concluding on the number of cycles and equivalence classes that can be deduced from a given cycle and check all the possibilities when we develop the algorithm which computes the class number for a given positive non-square discriminant.

Unfortunately the above results rely on the continued fraction of the first root of the corresponding polynomial of a given strongly reduced quadratic form $Q(X,Y)$ which might not be given. Instead of checking if the first root is of one of the forms, it certainly is more effective to check if one of the quadratic forms $-Q(X,Y)$, $\widetilde{Q}(X,Y)$ and $-\widetilde{Q}(X,Y)$ belongs to the cycle of $Q(X,Y)$. In this regard, the previous results have to be seen as a confirmation that all these verifications have to be done in order to conclude the number of cycles that are obtained from one cycle, as well as to conclude the number of equivalence classes that are obtained from one cycle.

13.2 A basic algorithm to compute the class number for positive non-square discriminants

In order to compute the class number for a given positive non-square discriminant D, one of the most basic approaches is to list all the elements of

$$S_{et}^{D} = \{Q(X,Y) \in Quad(D) : Q(X,Y) \text{ is strongly reduced}\},$$

find all equivalence relations between those elements, conclude the number of equivalence classes in S_{et}^{D} and deduce so the class number $H(D)$.

Using proposition 7.10, we can effectively compute the cycle of a given strongly reduced form $Q(X,Y)$. Then, using the conclusions of section 13.1, we can effectively compute the number of equivalence classes that can be deduced from its cycle and deduce by theorem 12.8 all the strongly reduced equivalent forms of $Q(X,Y)$. Hence, the only part that is missing is to describe how one can list all the strongly reduced quadratic forms. This is actually still an unsolved problem since no effective method to do so has been found yet.

The algorithm which will be described below includes all possible optimizations (as far as the author knows) that can be obtained by our development. The algorithm divides into two parts which will be developed separately.

Listing all strongly reduced quadratic forms of a given positive non-square discriminant

We know that if $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ is strongly reduced, so is then $-Q(X,Y)$, $\widetilde{Q}(X,Y)$ and $-\widetilde{Q}(X,Y)$. Thus, in order to find all strongly reduced quadratic forms of dis-

criminant D, it is sufficient to find all strongly reduced quadratic forms $\langle a,b,c\rangle_{X,Y}$ such that $0 < a \leq |c|$.

Let us consider a strongly reduced quadratic form $Q(X,Y) = \langle a,b,c\rangle_{X,Y} \in Quad(D)$ where $0 < a \leq |c|$. Then, from the Gauss conditions and the definition of the discriminant, we can deduce the following properties for b:

1. D and b have the same parity.
2. $-\sqrt{D} < b < 0$.
3. $-\frac{D-b^2}{4} = ac$.
4. $b + \sqrt{D} < 2a < -b + \sqrt{D}$.

Thus, in order to find all strongly reduced quadratic forms $\langle a,b,c\rangle_{X,Y}$ such that $0 < a \leq |c|$, we list all positive b with the same parity as D which are smaller than or equal to $\left\lfloor \sqrt{D} \right\rfloor$. Then, for each b, we set $R_b = \frac{D-b^2}{4}$. Next, for each b, we list all $a_b > 0$ which satisfy $\left\lfloor \frac{\lfloor \sqrt{D} \rfloor - b}{2} \right\rfloor < a_b$ (which is deduced from the Gauss conditions) and which satisfy $a_b \geq \lfloor \sqrt{R_b} \rfloor$ (which is deduced from the assumption that $0 < a < |c|$ and since $|ac| = \frac{D-b^2}{4}$). Finally, for each a_b, we compute $\frac{R_b}{a_b} = c_{a_b}$. Now for each b and each a_b, if c_{a_b} is an integer, then we write $\langle a_b, -b, -c_{a_b}\rangle_{X,Y}$ into our list of strongly reduced quadratic forms, (else we restart with other values).

Remark 13.8. *Notice that the integral part of a square root of a given positive integer N can be efficiently computed (for example by Newton's iterative method). Furthermore, subtraction and division by 4 are efficient computations. Thus the quantities $\left\lfloor \sqrt{D} \right\rfloor$, R_b, $\left\lfloor \sqrt{D} \right\rfloor$ and $\left\lfloor \frac{\sqrt{D}-b}{2} \right\rfloor$ can be efficiently computed. However, the global approach to list all the strongly reduced forms is not effective since it relies partially on the factorization of each quantity R_b. Furthermore, as one notices fast, this approach does a lot of unnecessary computations as to find a_b and c_{a_b}. However, up to now it is one of the only approaches that seem to work out fine.*

Let us now formalize the given approach.

Algorithm (list(D)).
Input: A positive non-square integer D.
Output: A list L of strongly reduced quadratic forms whose first coefficient is smaller than or equal to their last coefficient in absolute value which is arranged such that if $l,k \in \{1,2,...,length(L)\}$ with $k < l$ then the two elements $L[k] = (a_k,b_k,c_k)$ and $L[l] = (a_l,b_l,c_l)$ satisfy $b_k < b_l$ or $b_k = b_l$ and $a_k < a_l$.
Functioning:
$D_0 \leftarrow \left\lfloor \sqrt{D} \right\rfloor$
$i \leftarrow 1$
If $D \equiv 0 \pmod 2$, then $b \leftarrow 2$, else $b \leftarrow 1$
While $b \leq D_0$ do:
$R \leftarrow \frac{D-b^2}{4}$
$R_0 \leftarrow \left\lfloor \sqrt{R} \right\rfloor$

$a \leftarrow \left\lfloor \frac{D_0 - b}{2} \right\rfloor + 1$
While $a \leq R_0$ do:
If $R \equiv 0 \pmod{a}$, then
$L[i] \leftarrow (a, -b, -\frac{R}{a})$
$i \leftarrow i + 1$
else nothing
$a \leftarrow a + 1$
$b \leftarrow b + 2$
Return L

Let us give a minimal example on how the algorithm works.

Example. Consider $D = 41$. Then list(41) computes $D_0 = 6$ and sets $i = 1$. Since $29 \not\equiv 0 \pmod{2}$, it sets $b = 1$. Then it starts the first iteration of the first *While* loop (since $b = 1 \leq 5 = D_0$).

The first iteration of the first *While* loop works as follows. It starts by setting $R = 10$, $R_0 = 3$ and $a = 3$. Since $a \leq R_0$, the condition of the second *While* loop is fulfilled. The first iteration of the second *While* loop works as follows. Since $10 \not\equiv 0 \pmod{3}$, the if loop is skipped. Then $a = 4$ and since the condition of the second *While* loop is no longer fulfilled, it ends. Finally, the first *While* loop sets $b = 3$.

Since $b = 3 \leq 5$, the condition of the first *While* loop is still fulfilled. The second iteration of the first *While* loop works as follows. It starts by setting $R = 8$, $R_0 = 2$ and $a = 2$. Since $a \leq R_0$, the condition of the second *While* loop is fulfilled. The first iteration of the second *While* loop works as follows. Since $8 \equiv 0 \pmod{2}$, the first entry of L is given by $L[1] = (2, -3, -4)$ and then i is set to be $i = 2$. Then $a = 3$ and since the condition of the second *While* loop is no longer fulfilled, it ends. Finally, the first *While* loop sets $b = 5$.

Since $b = 5 \leq 5$, the condition of the first *While* loop is still fulfilled. The third iteration of the first *While* loop works as follows. It starts by setting $R = 4$, $R_0 = 2$ and $a = 1$. Since $a \leq R_0$, the condition of the second *While* loop is fulfilled. The first iteration of the second *While* loop works as follows. Since $4 \equiv 0 \pmod{1}$, the second entry of L is given by $L[2] = (1, -5, -4)$ and then i is set to be $i = 3$. Then $a = 2$ and since the condition of the second *While* loop is still fulfilled, we enter a second time the second *While* loop. The second iteration of the second *While* loop works as follows. Since $4 \equiv 0 \pmod{2}$, the third entry of L is given by $L[3] = (2, -5, -2)$ and then i is set to be $i = 4$. Then $a = 3$ and since the condition of the second *While* loop is no longer fulfilled, it ends. Finally, the first *While* loop sets $b = 7$.

Since the condition of the first *While* loop is no longer fulfilled, it ends.

Finally, the algorithm returns $L = [(2, -3, -4), (1, -5, -4), (2, -5, -2)]$.

A real optimization for this algorithm could be obtained by finding a sequence of strongly reduced quadratic forms that cannot be equivalent one to each other. Even if this sequence would not contain one strongly reduced quadratic form of each equivalence class (which would make this development unnecessary since it would give directly the class number), it would still reduce the computations above. Unfortunately no such sequence has been found yet, nor a general applicable criterion on when quadratic forms cannot be equivalent which does not rely on theorem 12.8 or an equivalent statement.

Finding all equivalence relations of strongly reduced quadratic forms

In order to deduce all equivalence relations of strongly reduced quadratic forms, it is sufficient by theorem 12.8 to construct all the cycles of strongly reduced quadratic forms, to identify their associate cycles and then conclude the class number.

By the first algorithm explained above, we obtain a list of strongly reduced quadratic forms whose first coefficient is smaller than or equal to their last coefficient in absolute value.

Let us choose one such quadratic form $Q(X,Y) = \langle a,b,c\rangle_{X,Y}$. Proposition 7.10 can be used to efficiently compute the cycle of $Q(X,Y)$. Instead of doing so and then check to conclude the number of cycles that can be deduced from this cycle, as well as the number of equivalence classes, we are going to do those verifications after each application of the proposition. This has the advantage that one does not necessarily have to compute the whole cycle.

Indeed, if the first coefficient of some quadratic form $Q'(X,Y)$ in the cycle of $Q(X,Y)$ is equal in absolute value to a or to $|c|$ and if at the same time its second coefficient is equal in absolute value to $|b|$, then $Q'(X,Y)$ is necessarily one of the forms $Q(X,Y)$, $\widetilde{Q}(X,Y)$, $-Q(X,Y)$ or $-\widetilde{Q}(X,Y)$. If $Q'(X,Y) = -Q(X,Y)$, then we can stop the computation of the cycle, deduce that the second half of the cycle is given by the negative of the first part and that the cycle denotes exactly one equivalence class. If $Q'(X,Y)$ is one of the other two forms, then we cannot shorten the computation. However, if $Q'(X,Y) = -\widetilde{Q}(X,Y)$, then we can conclude that the cycle of $Q(X,Y)$ comprises exactly one equivalence class.

The last part to analyze is how we can use our initial list inside this computation. Actually, when we compute the cycle, as soon as a quadratic form $Q'(X,Y) = \langle a',b',c'\rangle_{X,Y}$ appears inside the resulting cycle, exactly one of the forms $Q'(X,Y)$, $\widetilde{Q'}(X,Y)$, $-Q'(X,Y)$ or $-\widetilde{Q'}(X,Y)$ is inside our list (except if some of them are identified), we can delete this form then from our list since by our conclusions about the cycle and the equivalence classes, we won't need it later on.

All the observations above lead to the following algorithm.

Algorithm (relations(L))**.**
Input: A list L of strongly reduced quadratic forms whose first coefficient is smaller than or equal to their last coefficient in absolute value which is arranged such that if $l,k \in \{1,2,...,length(L)\}$ such that $k < l$ then the two elements $L[k] = (a_k,b_k,c_k)$ and $L[l] = (a_l,b_l,c_l)$ satisfy $b_k < b_l$ or $b_k = b_l$ and $a_k < a_l$.
Output: The positive integer N which denotes the number of equivalence classes represented by the quadratic forms in L.
Functioning:
$N \leftarrow 0$
$C \leftarrow 0$
For $j \leftarrow 1$ to $length(L)$ with step 1 do:
    $L'[j] \leftarrow 1$
For $k \leftarrow 1$ to $length(L)$ with step 1 do:
    If $L'[k] = 1$, then
        $L'[k] \leftarrow 0$
        $a \leftarrow L[k][1]$
        $b \leftarrow L[k][2]$
        $c \leftarrow L[k][3]$
        If $a\left(a(\lfloor -\frac{b}{a}\rfloor + 1)^2 + b(\lfloor -\frac{b}{a}\rfloor + 1) + c\right) > 0$, then $\rho = \lfloor -\frac{b}{a}\rfloor$ else $\rho = \lfloor -\frac{b}{a}\rfloor + 1$

$Z' \leftarrow (a\rho^2 + b\rho + c, b + 2a\rho, a)$
While $|Z'[1]| \neq a$ or $|Z'[2]| \neq -b$ do:
$a_0 \leftarrow \min\{|Z'[1]|, |Z'[3]|\}$
$b_0 \leftarrow -|Z'[2]|$
For $p \leftarrow k$ to $length(L)$ with step 1 do:
If $L[p][1] = a_0$ and $L[p][2] = b_0$, then
$L'[p] \leftarrow 0$
break
else if $|L[p][2]| > |b_0|$, then break
else nothing
If $Z'[1] = -c$ and $Z'[2] = b$, then $C = 1$, else nothing
If $Z'[1]\left(Z'[1](\left\lfloor -\frac{Z'[2]}{Z'[1]} \right\rfloor + 1)^2 + Z'[2](\left\lfloor -\frac{Z'[2]}{Z'[1]} \right\rfloor + 1) + Z'[3]\right) > 0$, then
$\rho = \left\lfloor -\frac{Z'[2]}{Z'[1]} \right\rfloor$
else $\rho = \left\lfloor -\frac{Z'[2]}{Z'[1]} \right\rfloor + 1$
$Z' \leftarrow (Z'[1]\rho^2 + Z'[2]\rho + Z'[3], Z'[2] + 2Z'[1]\rho, Z'[1])$
If $Z'[1] = -a$ or $C = 1$ then $N \leftarrow N + 1$, else $N = N + 2$
else nothing
Return N

Remark 13.9. *In the above algorithm* $|\quad|$ *denotes the absolute value,* $\lfloor\quad\rfloor$ *denotes the integral part of a real number,* $L[k][i]$ *denotes the* i*-th coefficient of the* k*-th quadratic form inside* L *and similarly* $Z'[i]$ *denotes the* i*-th coefficient of the quadratic form* Z'*.*

Example. Consider the list $L = [(2, -3, -4), (1, -5, -4), (2, -5, -2)]$. Then the algorithm relations(L) will first set $N = 0$ and $C = 0$. Then, the first *For* loop produces the list $L'[1, 1, 1]$. Then the second *For* loop begins which we will detail a little bit more than the previous part.

First $k = 1$.

Then, the algorithm changes L' into $L' = [0, 1, 1]$. Then it sets $a = 2$, $b = -3$ and $c = -4$. This step could be skipped but is given to visualize the process. Next the algorithm computes the first derived form of the form $\langle 2, -3, -4\rangle_{X,Y}$. Since $a\left(a(\left\lfloor -\frac{b}{a} \right\rfloor + 1)^2 + b(\left\lfloor -\frac{b}{a} \right\rfloor + 1) + c\right) = -4 < 0$, the algorithm sets $\rho = \left\lfloor -\frac{b}{a} \right\rfloor + 1 = 2$. Then it computes the derived form $Z' = (-2, 5, 2)$.

Since the second coefficient of Z' is not equal to $3 = |b|$, the conditions of the *While* loop are fulfilled and so the algorithm starts the first iteration.

In a first step, it will identify the quadratic form inside the list L that corresponds to Z'. To do so, it sets $a_0 = 2$ and $b_0 = -5$. Now it verifies for all $p = 2, 3$ if first coefficient of the p-th quadratic form is given by a_0 and the second by b_0, in which case it sets $L'[p] = 0$. Here, it will find this for $p = 3$ and set $L' = [0, 1, 0]$. Notice here that the quadratic forms of L are ordered lexicographically via their second and then first coefficients, it would not make sense to let the algorithm verify all the possibilities. Thus it ends as soon as it passes the values of the given form Z'. Let us return to our computation.

Since the quadratic form has been identified and tagged in our list L', the algorithm verifies if maybe the quadratic form Z' is the negative of the associate form of the given form $\langle 2, -3, -4\rangle_{X,Y}$ in which case we know, the negative of $\langle 2, -3, -4\rangle_{X,Y}$ and $\langle 2, -3, -4\rangle_{X,Y}$. would be equivalent and the cycle of $\langle 2, -3, -4\rangle_{X,Y}$ would give us only one equivalence class. Since this is not the case here, nothing happens.

Finally, the algorithm computes the derived form of Z' in the same manner as Z' has been computed in the first place and then it replaces Z' by its derived form. Here, it obtains $Z' = (4, -3, -2)$.

Since the conditions of the *While* loop are still fulfilled, the algorithm repeats the procedure. In order to not repeat ourselves, we state the important observations only. This time, the identification step will we be interrupted by the fact that $|L[k]| = |L[2]| = |-5| > |b_0| = 3$. Furthermore, since $Z'[1] = -c$ and $Z'[2] = b$, the algorithm sets $C = 1$. Then the algorithm repeats the *While* loop three more times to obtain at each iteration respectively $Z' = (-1, 5, 4)$ and $L' = [0, 0, 0]$, $Z' = (4, -5, -1)$ and finally $Z' = (-2, 3, 4)$. Then the conditions of the *While* loop are no longer satisfied since $|-2| = 2 = a$ and $|3| = -(-3) = -b$.

Since $Z'[1] = -2 = -a$ and also $C = 1$ (although one of both conditions is sufficient), the algorithm sets $N = N + 1 = 1$.

Next it restarts the *For* loop where it will find for $k = 2, 3$ that $L'[k] = 0$ and thus nothing is changed.

Finally the algorithm returns $N = 1$.

Finding all equivalence relations of strongly reduced quadratic forms

After analyzing each part of the algorithm, the final algorithm can be given in the following easy terms.

Algorithm (h(D)).
Input: A positive non-square integer D.
Output: A positive integer N which corresponds to the class number $h(D)$.
Functioning:
Return relations(list(D))

Example. Consider $D = 41$. By the two previous examples, the algorithm $h(D)$ returns $N = 1$.

Remark 13.10. *As stated before, the algorithm $h(D)$ is not effective and some strong theoretical results would be necessary to obtain an effective and polynomial algorithm to compute the class number. Any such result would indirectly also be a result to factorize a given class of numbers which is still one of the most important questions in number theory.*

13.3 Perspectives

Today, quadratic forms are still used in several algebraic domains and their connections to quadratic fields have been proved by several mathematicians. Because of this relation, any result or optimization of results of quadratic forms consists at the same time in an advancement in other algebraic fields. This document shows that also continued fractions can be used to discover such results. Since neither of those directions have been studied yet by the author of this Master thesis, this document should not end with a final conclusion but rather with ideas describing where one could be looking for some new results.

A natural way to generalize the development of the first part of this document would be to study continued fractions in all generality and to connect the resulting theory with different algebraic fields. The theory about integral continued fractions allowed to see the real number field from a different perspective and so should also other algebraic fields be observable from other points of views by considering continued fractions defined by elements from those fields.

We discovered another method (instead of the traditional gaussian method) to study quadratic forms. It might be useful to find a third independent and complementing method to do so. Actually, our approach showed quite naturally the connection between quadratic forms and continued fractions and solved the question which quadratic forms are equivalent to their negatives. In this regard, one might asks what can be deduced by another third method and what could be the gain by combining all those methods.

Since some important results have been obtained by discussing the shape of continued fractions, a detailed investigation about the n-th convergent and tails of continued fractions might bring along some unexpected results. Since we discovered the general shape of the continued fractions of square-roots and quadratic surds in general, one might be tempted to look out for the general shape of the continued fraction of other algebraic numbers. This could help to study eventually n-ary quadratic forms and solve some number theoretical problems.

Because of a lack of time, this document did not outline the use of the developed theory in the study of lattices, quadratic fields, ideal theory and class groups. This has been developed by many of the cited sources and should not be neglected.

At last, let us mention that quadratic forms, as well as continued fractions have already been used to factorize big numbers (method of SHANKS or LEHMANN). In particular, as we highlighted above, a polynomial algorithm to compute the class number would also consist in a polynomial algorithm to factorize a special class of numbers. The unconventional factorization methods which use the developed theory could in this logic become breakthroughs in number theory. Further, this application shows once more the wide spectrum of this theory and which interesting results can be achieved by some supplementary analysis of the results presented in this Master thesis.

Part IV
Appendix

A A natural way to construct continued fractions of rational numbers

Sources: [1, p.1-3], [3, p.325-327], [4, p.8-13]
Other approach: [5, p.1-5]

The Euclidean division algorithm

Continued fractions arise quite naturally from some number theoretical results. Let us start with some reminders.

Proposition A.1. *For any two integers a and $b \neq 0$, there exist two unique integers r and c such that $a = br + c$ and $0 \leq c < |b|$.*

The expression of the last proposition is well known and has a precise definition.

Definition A.2. The function $f : \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \longrightarrow \mathbb{Z} \times \mathbb{N}$ defined by $(a, b) \mapsto (r, c)$ where $a = br + c$ and $0 \leq c < |b|$ is called *Euclidean division.*
For $a, b \in \mathbb{Z}$ and $b \neq 0$, we call $f(a, b)$ the *Euclidean division of a by b.* Furthermore, if $f(a, b) = (r, c)$, then we call a the *dividend*, b the *divisor*, r the *quotient* and c the *remainder* of the Euclidean division of a by b.

Example. The euclidean division of $a = 21$ by $b = 15$ is $f(21, 15) = (1, 6)$ since $21 = 21.1+6$.

Consider the Euclidean division of $a \in \mathbb{Z}$ by $b \in \mathbb{Z} \setminus \{0\}$. Let $f(a, b) = (r, c)$. If $c \neq 0$, then one can compute the Euclidean division of b by c. Thus, as long as the remainder of a Euclidean division does not vanish, we can compute recursively the Euclidean division of the divisor by the remainder of the previous Euclidean division. Since the remainder of an Euclidean division is strictly smaller than the divisor, Fermat's principle of descent yields that after a finite number of steps the remainder will vanish and the process will end.

Definition A.3. Let f be the Euclidean division. Then, the *Euclidean division algorithm* is the function $F : \mathbb{Z} \times (\mathbb{Z} \setminus \{0\}) \longrightarrow \bigcup_{n=1}^{\infty}(\mathbb{Z} \times \mathbb{N})$ defined by $(a, b) \mapsto \{(a_i, b_i)\}_{i \in \mathbb{N}}$ where $(a_0, b_0) = (a, b)$ and for all $i \in \mathbb{N} \setminus \{0\}$ the pair (a_i, b_i) is given by $(a_i, b_i) = f(a_{i-1}, b_{i-1})$ if $b_{i-1} \neq 0$ and by $(a_i, b_i) = (0, 0)$ if $b_{i-1} = 0$.

Let $a, b \in \mathbb{Z}$ and $b \neq 0$ and let

$$F(a, b) = \{(a_0, a_1), (r_0, a_2), (r_1, a_3), ..., (r_{n-2}, a_n), (r_{n-1}, 0), (0, 0), (0, 0), (0, 0), ...\}.$$

Then, by definition, the pairs satisfy the following relations to which we will refer as (Relation

system 1):

$$\begin{array}{rlrll}
f(a_0,a_1)=(r_0,a_2) & \text{where} & a_0=a_1r_0+a_2 & \text{and} & (0<a_2<|a_1|);\\
f(a_1,a_2)=(r_1,a_3) & \text{where} & a_1=a_2r_1+a_3 & \text{and} & (0<a_3<a_2);\\
\vdots & \vdots & \vdots & \vdots & \vdots\ ;\\
f(a_{n-2},a_{n-1})=(r_{n-2},a_n) & \text{where} & a_{n-2}=a_{n-1}r_{n-2}+a_n & \text{and} & (0<a_n<a_{n-1});\\
f(a_{n-1},a_n)=(r_{n-1},0) & \text{where} & a_{n-1}=a_nr_{n-1} & &
\end{array}$$

Example. The Euclidean algorithm of 21 by 15 is $F(21,15)=\{(21,15),(1,6),(2,3),(2,0),(0,0),(0,0),...\}$ since

$$\begin{aligned}
21&=15\cdot 1+6 && (0<6<15);\\
15&=6\cdot 2+3 && (0<3<6);\\
6&=3\cdot 2
\end{aligned}$$

The Euclidean division algorithm was first (from the historical point of view) used to determine the greatest common divisor of two non-zero integers a_0 and a_1 which is denoted by $\gcd(a_0,a_1)$. Indeed, one can easily verify that if $F(a_0,a_1)=\{(a_0,a_1),(r_0,a_2),...,(r_{n-2},a_n),(r_{n-1},0),(0,0),...\}$, then $\gcd(a_0,a_1)=a_n$.

Example. By the previous example $\gcd(21,15)=3$.

Moreover, since $a_i-a_{i+1}r_i=a_{i+2}$ for all $i\in\{0,1,...,n-2\}$ and since $\gcd(a_0,a_1)=a_n$, one can express $\gcd(a_0,a_1)$ as a linear combination of a_0 and a_1. This result is known as *Bezout's theorem.*

Theorem A.4 (Bezout's theorem)**.** *Let $a,b\neq 0$ be integers. Then there exist $x,y\in\mathbb{Z}$ such that* $\gcd(a,b)=ax+by$.

A natural way to construct continued fractions of rational numbers

The purpose of the Euclidean division algorithm in the following analysis will differ from its historical use. Before we can go on, we want to highlight some of its maybe unnoticed aspects.

Remark A.5. *1. The remainder of an Euclidean division is always non-negative.*

2. *By construction, the Euclidean division algorithm gives a unique sequence of integers $r_0,r_1,...,r_{n-1}$ which satisfy the equations in (Relation system 1).*

3. *Each divisor of the Euclidean divisions computed by the Euclidean division algorithm, except maybe a_1, is positive.*

4. *If the divisor of the Euclidean division algorithm (a_1) is positive, then each r_i, except maybe r_0, is positive.*

Let $a,b\in\mathbb{Z}$ and $b>0$ and let

$$F(a,b)=\{(a_0,a_1),(r_0,a_2),(r_1,a_3),...,(r_{n-2},a_n),(r_{n-1},0),(0,0),(0,0),(0,0),...\}.$$

Then, by definition, the pairs satisfy the following relations to which we will refer as (Relation system 2):

$$\begin{array}{rcccc}
f(a_0,a_1)=(r_0,a_2) & \text{where} & \frac{a_0}{a_1}=r_0+\frac{a_2}{a_1} & \text{and} & (0<\frac{a_2}{a_1}<1);\\
f(a_1,a_2)=(r_1,a_3) & \text{where} & \frac{a_1}{a_2}=r_1+\frac{a_3}{a_2} & \text{and} & (0<\frac{a_3}{a_2}<1);\\
\vdots & \vdots & \vdots & \vdots & \vdots\ ;\\
f(a_{n-2},a_{n-1})=(r_{n-2},a_n) & \text{where} & \frac{a_{n-2}}{a_{n-1}}=r_{n-2}+\frac{a_n}{a_{n-1}} & \text{and} & (0<\frac{a_n}{a_{n-1}}<1);\\
f(a_{n-1},a_n)=(r_{n-1},0) & \text{where} & \frac{a_{n-1}}{a_n}=r_{n-1}. & &
\end{array}$$

Observe that $r_0, r_1, ..., r_{n-1}$ is the unique sequence of integers which satisfy the equations in (Relation system 2).

Definition A.6. The *integral part* of $x \in \mathbb{R}$ is $\lfloor x \rfloor \in \mathbb{Z}$ where $x - 1 < \lfloor x \rfloor \leq x$.

Example. $\lfloor \frac{5}{2} \rfloor = 2$, $\lfloor -\frac{16}{5} \rfloor = -4$, $\lfloor \frac{1}{2} \rfloor = 0$ and $\lfloor 1 \rfloor = 1$.

Proposition A.7. *The integral part of $x \in \mathbb{R}$ is unique.*

Proof. Let r_1 and r_2 be integral parts of x, then $x-1 < r_1, r_2 \leq x$ implies that $-1 < r_1 - r_2 < 1$. Since $r_1, r_2 \in \mathbb{Z}$, $r_1 = r_2$. □

Since in (Relation system 2) of the Euclidean division algorithm, we have $0 < \frac{a_{i+2}}{a_{i+1}} < 1$ for all $i \in \{0, ..., n-2\}$, we deduce that $\frac{a_i}{a_{i+1}} - 1 < r_i \leq \frac{a_i}{a_{i+1}}$ for all $i \in \{0, ..., n-1\}$ and so $r_i = \left\lfloor \frac{a_i}{a_{i+1}} \right\rfloor$.

Definition A.8. The *derived number* of $x \in \mathbb{R} \setminus \mathbb{Z}$ is the unique number ∂x that satisfies $x = \lfloor x \rfloor + \frac{1}{\partial x}$ or equivalently $\partial x = \frac{1}{x - \lfloor x \rfloor}$ (thus $\partial x > 1$).
The *derived number* of $x \in \mathbb{Z} \cup \{\infty\}$ is by convention $\partial x = \infty$.

We use the notation $\partial^0 x = x$, $\partial^1 x = \partial x$ and $\partial^k x = \partial(\partial^{k-1} x)$ for all $k \in \mathbb{N}$.

Example. $\partial\left(\frac{5}{2}\right) = 2$, $\partial^2\left(\frac{5}{2}\right) = \partial 2 = \infty$ and $\partial\left(-\frac{16}{5}\right) = \frac{5}{4}$.

In (Relation system 2) of the Euclidean division algorithm, we can observe that $\partial\left(\frac{a_i}{a_{i+1}}\right) = \frac{a_{i+1}}{a_{i+2}}$ for all $i \in \{0, 1, ..., n-2\}$.

Using definition A.7 and A.8, we can completely rewrite the (Relation system 2) of the Euclidean division algorithm. Indeed, let $a, b \in \mathbb{Z}$ and $b > 0$ and let

$$F(a,b) = \{(a_0,a_1), (r_0,a_2), (r_1,a_3), ..., (r_{n-2},a_n), (r_{n-1},0), (0,0), (0,0), (0,0), ...\}.$$

Then the pairs satisfy the following relations to which we refer as (Relation system 3):

$$\begin{array}{llll}
f(a_0,a_1)=(r_0,a_2) & \text{where} & \dfrac{a_0}{a_1}=\left\lfloor\dfrac{a_0}{a_1}\right\rfloor+\dfrac{1}{\partial\frac{a_0}{a_1}} & \text{and} \quad (1<\partial\dfrac{a_0}{a_1});\\
f(a_1,a_2)=(r_1,a_3) & \text{where} & \partial\dfrac{a_0}{a_1}=\left\lfloor\partial\dfrac{a_0}{a_1}\right\rfloor+\dfrac{1}{\partial^2\frac{a_0}{a_1}} & \text{and} \quad (1<\partial^2\dfrac{a_0}{a_1});\\
\vdots & \vdots & \vdots & \vdots \quad \vdots \quad ;\\
f(a_{n-2},a_{n-1})=(r_{n-2},a_n) & \text{where} & \partial^{n-2}\dfrac{a_0}{a_1}=\left\lfloor\partial^{n-2}\dfrac{a_0}{a_1}\right\rfloor+\dfrac{1}{\partial^{n-1}\frac{a_0}{a_1}} & \text{and} \quad (1<\partial^{n-1}\dfrac{a_0}{a_1});\\
f(a_{n-1},a_n)=(r_{n-1},0) & \text{where} & \partial^{n-1}\dfrac{a_0}{a_1}=\left\lfloor\partial^{n-1}\dfrac{a_0}{a_1}\right\rfloor &
\end{array}$$

Example. The Euclidean algorithm of $a=21$ by $b=15$ is $F(21,15)=\{(21,15),(1,6),(2,3),(2,0),(0,0),(0,0),...\}$ and we have

$$\begin{aligned}
\frac{21}{15}&=1+\frac{1}{\frac{15}{6}};\\
\frac{15}{6}&=2+\frac{1}{\frac{6}{3}};\\
\frac{6}{3}&=2.
\end{aligned}$$

with $\left\lfloor\frac{21}{15}\right\rfloor=1$, $\partial\frac{21}{15}=\frac{15}{6}$, $\left\lfloor\partial\frac{21}{15}\right\rfloor=2$, $\partial^2\frac{21}{15}=2$.

From (Relation system 3) of the Euclidean division algorithm we deduce that $\frac{a_0}{a_1}$ is given by

$$\frac{a_0}{a_1}=\left\lfloor\frac{a_0}{a_1}\right\rfloor+\cfrac{1}{\left\lfloor\partial\frac{a_0}{a_1}\right\rfloor+\cfrac{1}{\ddots\quad+\cfrac{1}{\left\lfloor\partial^{n-2}\frac{a_0}{a_1}\right\rfloor+\cfrac{1}{\left\lfloor\partial^{n-1}\frac{a_0}{a_1}\right\rfloor}}}}. \tag{4}$$

Example. From the previous examples, we deduce that

$$\frac{21}{15}=1+\frac{1}{2+\frac{1}{2}}.$$

Remark A.9. *Since a_0 and $a_1>0$ were arbitrary integers in this whole construction, we showed that any rational number can be written in a form as in (4).*

An object of the form (4) will be called a *finite continued fraction*. Unfortunately, the above construction allows to rewrite only rational numbers under such a form. A method to construct continued fractions for any real number and more general continued fractions are presented in chapter 2.

B A set of representatives of the equivalence classes of $Quad(D)$ for $D \in \mathbb{N}^2 \setminus \{0\}$

Sources: [11, art. 206-211]

Chapter note: Assume throughout this appendix that $D = n^2$ for some $n \in \mathbb{N} \setminus \{0\}$.

Lemma 8.1 states that any quadratic form $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N}^2 \setminus \{0\}$ is equivalent to $\langle a, \sqrt{D}, 0\rangle_{X,Y}$ for some $0 \leq a < \sqrt{D}$. This gives rise to theorem 8.2 which states that $h(D)$ is finite and to corollary 8.7 which gives a rough upper bound ($h(D) \leq \sqrt{D}$) for the class number $h(D)$.

We would like to determine a set of quadratic forms which includes exactly one representative of every equivalence class of $Quad(D)$, in other words a set of quadratic forms with discriminant D with $h(D)$ elements that are non-equivalent one-by-one.

To achieve this, we can use the previous work and find a general condition to determine the equivalence of two quadratic forms $\langle a, \sqrt{D}, 0\rangle_{X,Y}$ and $\langle b, \sqrt{D}, 0\rangle_{X,Y}$ with $0 \leq a, b < \sqrt{D}$.

Let us first consider proper equivalence.

Lemma B.1. *Let* $\langle b, n, 0\rangle_{X,Y} = \langle a, n, 0\rangle_{X,Y} \circ \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$ *for some* $0 \leq a, b < n$ *and* $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$ *such that* $\det \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = 1$. *Then* $a = b$.

Proof. By proposition 5.7 and since $\det \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = 1$, we deduce the following system of equations.

$$\begin{cases} b = ap^2 + npr & (5) \\ n = 2apq + n(ps + qr) & (6) \\ 0 = aq^2 + nqs & (7) \\ 1 = ps - qr & (8) \end{cases}$$

Assume $q = 0$, then equation (6) becomes $n = nps$ which implies that $p = s = \pm 1$ (since $n, p, s \in \mathbb{Z}$). Then equation (5) becomes $b = a \pm nr$ which implies that $b - a \equiv 0(\mod n)$. Since $-n < b - a < n$, we deduce that $b - a = 0$ and hence $b = a$.

Assume $q \neq 0$, then by dividing equation (7) by q, we obtain $0 = aq + ns$ and so $aq = -ns$. Substituting this into equation (6) gives $n = -2nps + n(ps + qr) = -nps + nqr$. Hence, we deduce that $1 = -ps + qr$ and so $-1 = ps - qr$ which contradicts equation (8). □

Hence, if two distinct quadratic forms of the set $\{\langle a, \sqrt{D}, 0\rangle_{X,Y} : 0 \leq a < \sqrt{D}\}$ are equivalent then they are necessarily improperly equivalent. This implies the following corollary.

Corollary B.2. *The equivalence classes of the set* $\{\langle a, n, 0\rangle_{X,Y} : 0 \leq a < n\}$ *include at most two elements.*

Proof. Assume by contradiction that there are three distinct quadratic forms $\langle a_1, n, 0\rangle_{X,Y}$, $\langle a_2, n, 0\rangle_{X,Y}$, $\langle a_3, n, 0\rangle_{X,Y}$ with $0 \leq a_1, a_2, a_3 < n$ which are equivalent. Let $\langle a_2, n, 0\rangle_{X,Y} = \langle a_1, n, 0\rangle_{X,Y} \circ S$ and $\langle a_3, n, 0\rangle_{X,Y} = \langle a_2, n, 0\rangle_{X,Y} \circ S'$ for some $S, S' \in GL(2,\mathbb{Z})$. Since $a_1 \neq a_2 \neq a_3$, lemma B.1 implies that $\det(S) = \det(S') = -1$. Then $\langle a_3, n, 0\rangle_{X,Y} = (\langle a_1, n, 0\rangle_{X,Y} \circ S) \circ S' = \langle a_1, n, 0\rangle_{X,Y} \circ SS'$. Since $\det(SS') = 1$, lemma B.1 implies that $a_3 = a_1$ which is a contradiction. □

Hence, we deduce that for $D = n^2 \in \mathbb{N}^2 \setminus \{0\}$, the class number $h(D)$ satisfies $\frac{n}{2} \leq h(D) \leq n$. Let us now consider improper equivalence.

Proposition B.3. *Two quadratic forms* $\langle b,n,0\rangle_{X,Y}$ *and* $\langle a,n,0\rangle_{X,Y}$ *with* $0 \leq a,b < n$ *are improperly equivalent if and only if* $\gcd(a,n) = m = \gcd(b,n)$ *and* $m^2 \equiv ab \pmod{mn}$.

Proof. Assume first that $\langle b,n,0\rangle_{X,Y}$ and $\langle a,n,0\rangle_{X,Y}$ are improperly equivalent. Let $\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$ such that $\det\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = -1$ and $\langle b,n,0\rangle_{X,Y} = \langle a,n,0\rangle_{X,Y} \circ \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right)$. By proposition 5.7 and since $\det\left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) = -1$, we deduce the following system of equations.

$$\begin{cases} b = ap^2 + npr & (9) \\ n = 2apq + n(ps+qr) & (10) \\ 0 = aq^2 + nqs & (11) \\ -1 = ps - qr & (12) \end{cases}$$

Then $(10) - n(12)$ gives $2n = 2apq + n(ps+qr) - nps + nqr = 2apq + 2nqr$ and hence:

$$n = q(ap + nr) \tag{13}$$

Similarly $rs(10) - 2r^2(11) - (2a + 2aqr + nrs)(12)$ gives after simplifications the following:

$$a = -s(ap + nr) \tag{14}$$

Consider now $a(9)$ which gives $ab = ap(ap+nr)$. Substituting this into $(ap+nr)^2 - ab$ gives

$$(ap+nr)^2 - ab = (ap+nr)^2 - ap(ap+nr) = nr(ap+nr) \tag{15}$$

By equation (13), we know that $(ap+nr)$ divides n and by equation (14), we know that $(ap+nr)$ divides a. Hence $(ap+nr)$ divides $\gcd(a,n) = m$. Conversely, since m divides a and since m divides n, we deduce that m divides $(ap+nr)$. Thus $(ap+nr) = \pm m$. Hence by equation (15), we conclude that $m^2 \equiv ab \pmod{mn}$.

By equation (9), we deduce that $b = \pm pm$. Since $m^2 \equiv ab \pmod{mn}$, we deduce that $m^2 = ab + kmn$ for some $k \in \mathbb{Z}$. Let $a = k_a m$, $b = k_b m$ and $n = k_n m$ for some $k_a, k_b, k_n \in \{0,1,...,n\}$. Then $m^2 = ab + kmn$ implies $m^2 = k_a k_b m^2 + k k_n m^2$. Since $m > 0$, we deduce that $1 = k_a k_b + k k_n$. Hence $\gcd(k_a k_b, k k_n) = 1$ and so $\gcd(k_b, k_n) = 1$. Hence $\gcd(b,n) = \gcd(k_b m, k_n m) = m$.

Assume that $\gcd(a,n) = m = \gcd(b,n)$ and $m^2 \equiv ab \pmod{mn}$. Then $-\frac{b}{m}, -\frac{n}{m}, \frac{ab-m^2}{mn}, \frac{a}{m} \in \mathbb{Z}$. A direct computation shows that

$$\langle b,n,0\rangle_{X,Y} = \langle a,n,0\rangle_{X,Y} \circ \begin{pmatrix} -\frac{b}{m} & -\frac{n}{m} \\ \frac{ab-m^2}{mn} & \frac{a}{m} \end{pmatrix}$$

and that $\det\begin{pmatrix} -\frac{b}{m} & -\frac{n}{m} \\ \frac{ab-m^2}{mn} & \frac{a}{m} \end{pmatrix} = -1$. □

Remark B.4. *One can prove that if, for fixed* a,n *(and* $\gcd(a,n) = m$*), the equivalence* $m^2 \equiv ab \pmod{mn}$ *has a solution* b*, then there is a unique* B *satisfying* $0 \leq B < n$ *and* $m^2 \equiv aB \pmod{mn}$ *which verifies corollary B.2.*

Using proposition B.3, we can easily compute the class number $h(D)$ for any positive square discriminant D.

Example. Consider $D = 36$. Then $\sqrt{D} = n = 6$. Hence, consider the set $S_{et}^{36} = \{\langle a, 6, 0\rangle_{X,Y} : 0 \leq a < 6\}$. In order to determine $h(6)$, we only have to determine the number of equivalence classes in S_{et}^{36}.

First of all, we can distribute the numbers $0, 1, ..., 5$ into sets depending on their greatest common divisor with 6 and obtain $\{0\}_6, \{1, 5\}_1, \{2, 4\}_2, \{3\}_3$ (where the indices of the sets denote the greatest common divisor of the elements of the set and 6).

By the first condition of proposition B.3 ($\gcd(a, n) = \gcd(b, n)$), we conclude that $\langle 0, 6, 0\rangle_{X,Y}$ and $\langle 3, 6, 0\rangle_{X,Y}$ form their own equivalence classes.

Since $1^2 \not\equiv 5 \,.\, 1 \pmod{1 \,.\, 6}$, we conclude that $\langle 1, 6, 0\rangle_{X,Y}$ and $\langle 5, 6, 0\rangle_{X,Y}$ form their own equivalence classes.

Since $2^2 \not\equiv 2 \,.\, 4 \pmod{2 \,.\, 6}$, we conclude that $\langle 2, 6, 0\rangle_{X,Y}$ and $\langle 4, 6, 0\rangle_{X,Y}$ form their own equivalence classes.

Hence $h(36) = 6$.

Example. Consider $D = 49$. Then $\sqrt{D} = n = 7$. Hence, consider the set $S_{et}^{49} = \{\langle a, 7, 0\rangle_{X,Y} : 0 \leq a < 7\}$. In order to determine $h(6)$, we only have to determine the number of equivalence classes in S_{et}^{49}.

First of all, we can distribute the numbers $0, 1, ..., 5, 6$ into sets depending on their greatest common divisor with 6 and obtain $\{0\}_7$, $\{1, 2, 3, 4, 5, 6\}_1$ (where the indices of the sets denotes the greatest common divisor of the elements of the set and 7).

By the first condition of proposition B.3 ($\gcd(a, n) = \gcd(b, n)$), we conclude that $\langle 0, 7, 0\rangle_{X,Y}$ forms its own equivalence class.

Since $1^2 \equiv 1 \,.\, 1 \pmod{7}$, we conclude that $\langle 1, 7, 0\rangle_{X,Y}$ forms its own equivalence class.

Since $1^2 \equiv 2 \,.\, 4 \pmod{7}$, we conclude that $\langle 2, 7, 0\rangle_{X,Y}$ and $\langle 4, 7, 0\rangle_{X,Y}$form an equivalence class.

Since $1^2 \equiv 3 \,.\, 5 \pmod{7}$, we conclude that $\langle 3, 7, 0\rangle_{X,Y}$ and $\langle 5, 7, 0\rangle_{X,Y}$form an equivalence class.

Since $1^2 \equiv 6 \,.\, 6 \pmod{7}$, we conclude that $\langle 6, 7, 0\rangle_{X,Y}$ forms its own equivalence class.

Hence $h(49) = 5$.

C A basic algorithm to solve the Pell equations

Sources: [1, p.23-28], [5, p.10-11], [3, p.351-357], [3, p.351-357], [4, p.113-121]
Other approaches: [9, p.31-34,42-43], [6, p.413-418], [10, p.133-136]

The Pell equations are special Diophantine equations that were already studied by Brahmagupta and Fermat. The developed theory about quadratic forms and continued fractions offers an easy way to develop an algorithm to solve these equations. This algorithm is the topic of this appendix.

Definition C.1. The *(positive) Pell equation* is the equation $X^2-NY^2=1$ where $N \in \mathbb{N}\backslash\{0\}$. The *negative Pell equation* is the equation $X^2 - NY^2 = -1$ where $N \in \mathbb{N} \setminus \{0\}$. Any pair of integers (x, y) such that $x^2 - Ny^2 = \pm 1$ is called a *solution* to the (positive or negative) Pell equation.

Since (x, y) is a solution of a Pell equation if and only if $(-x, y)$, $(x, -y)$ and $(-x, -y)$ are so, it is sufficient to study the solutions (x, y) of the Pell equations where x, y are non-negative integers. We will call such a solution a non-negative solution and we will ignore any other solutions in the preceding.

Solutions to the Pell equations with N being a non-zero square integer

Proposition C.2. *1. The Pell equation $X^2 - NY^2 = 1$ with $N \in \mathbb{N}^2 \setminus \{0\}$ has exactly one non-negative solution, namely $(x, y) = (1, 0)$.*

2. The negative Pell equation $X^2 - Y^2 = -1$ (i.e. $N = 1$) has exactly one non-negative solution, namely $(x, y) = (0, 1)$.

3. The negative Pell equation $X^2 - NY^2 = -1$ with $1 < N \in \mathbb{N}^2 \setminus \{0\}$ does not have solutions.

Proof. Let $N \in \mathbb{N}^2 \setminus \{0\}$, then $N = m^2$ for some $m \in \mathbb{N} \setminus \{0\}$. Let (x, y) be a non-negative solution to the Pell equation $X^2 - NY^2 = \delta$ where $\delta \in \{-1, 1\}$. Then

$$x^2 - Ny^2 = x^2 - (my)^2 = x^2 - z^2 = \delta$$

where $z = my \in \mathbb{N}$. We observe that $x \neq z$ since else $0 = \delta$ (which would be a contradiction). Hence $z = x + k$ for some $k \in \mathbb{Z} \setminus \{0\}$. Thus

$$\delta = x^2 - z^2 = x^2 - (x + k)^2 = x^2 - (x^2 + 2xk + k^2) = -k(2x + k).$$

Thus, $\delta = -k(2x + k)$ and so $-k = \pm 1$ and $2x + k = \pm 1$. Let us distinguish the different cases and deduce so all the possible solutions.

1. If $\delta = 1$, then $1 = -k(2x + k)$ and so $-1 = k(2x + k)$.

   If $k = -1$, then $2x - 1 = 1$ and so $x = 1$. Thus, $z = x + k = 0$. Since $z = my$ and $m \neq 0$, this implies that $y = 0$ and hence $(x, y) = (1, 0)$.

   If $k = 1$, then $2x+1 = -1$ and so $x = -1$ which is a contradiction since x is non-negative by assumption.

2. If $\delta = -1$, then $-1 = -k(2x+k)$ and so $1 = k(2x+k)$.

   If $k = -1$, then $2x - 1 = -1$ and so $x = 0$. Thus, $z = x + k = -1$. Since $z = my$ and $m > 0$, this implies that $y < 0$ which is a contradiction since y is non-negative by assumption.

   If $k = 1$, then $2x + 1 = 1$ and so $x = 0$. Thus, $z = x + k = 1$. Since $z = my$ and $m \neq 0$, this implies that $m = 1$ and $y = 1$ and hence $(x, y) = (0, 1)$.

□

Solutions to the Pell equations with N being a positive non-square integer

For the sake of simplicity, we are going to deduce one property at a time. Hence the approach below might not be the most effective one but surely one of the easiest.

Before we start solving the Pell equations, we can notice that the polynomial $X^2 - NY^2$ is actually a particular binary quadratic form. Indeed, we can rewrite this as $Q_0(X,Y) = Q(X,Y) = \langle 1, 0, -N\rangle_{X,Y} \in Quad(4N)$. By noticing that the roots of $Q(T,1)$ are $\sqrt{N} > 1 > 0$ and $-\sqrt{N} < 0$, one deduces that $Q(X,Y)$ is weakly reduced. By lemma 7.6 and corollary 7.7, $Q_n(X,Y) = \partial^n Q(X,Y)$ is strongly reduced for all $n \in \mathbb{N} \setminus \{0\}$.

Since $\sqrt{N}$ is irrational, proposition 3.10 implies that $\sqrt{N}$ has an infinite continued fraction. Let $\sqrt{D} = [r_0, r_1, r_2, r_3, ...]$. Since $\partial^n\sqrt{N}$ is the first root of $\partial^n Q(T,1)$, we deduce that $Q_n(X,Y) = \langle a_n, b_n, c_n\rangle_{X,Y}$ is given by

$$\partial^n Q(X,Y) = Q(X,Y) \circ T(r_0)T(r_1)...T(r_{n-1}) = Q(X,Y) \circ D_n$$

where $T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $D_n = T(r_0)T(r_1)...T(r_{n-1}) = \left(\begin{smallmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{smallmatrix}\right)$ and $r_n = \lfloor \partial^n \rfloor \sqrt{N}$. Hence

$$\begin{cases} a_n &= & Q(p_{n-1}, q_{n-1}) & = & a_{n-1}r_{n-1}^2 + b_{n-1}r_{n-1} + c_{n-1} \\ b_n &= & (p_{n-2}\ q_{n-2}) \left(\begin{smallmatrix} 2a & b \\ b & 2d \end{smallmatrix}\right) \left(\begin{smallmatrix} p_{n-1} \\ q_{n-1} \end{smallmatrix}\right) & = & b_{n-1} + 2a_{n-1}r_{n-1} \\ c_n &= & a_{n-1} & = & a_{n-1} \end{cases}$$

Furthermore, we know by corollary 11.9 that $\partial\sqrt{N}$ is strongly reduced and so is then $\partial Q(X,Y)$. Hence, theorem 10.1 implies that there exists minimal $M \in 2\mathbb{N} \setminus \{0\}$ such that $\partial^{M+1}Q(X,Y) = \partial Q(X,Y)$.

Since $\partial\sqrt{N}$ is strongly reduced, its continued fraction is purely periodic and the general form of this continued fraction is given in corollary 11.9. Let m be the caliber of $\partial\sqrt{N}$.

One can easily prove that if m is odd, then $M = 2m$ and $Q_{m+1}(X,Y) = -Q_1(X,Y)$ and if m is even then $M = m$ and $Q_{m+1}(X,Y) = Q_1(X,Y)$ (c.f. proposition 13.2). Hence, we deduce that $Q_{km+1+n}(X,Y) = \pm Q_{1+n}(X,Y)$ for all $k \geq 0$ and $n \geq 0$.

Let us find in a first step some particular non-negative solutions of the Pell equations.

The Pell equation $X^2 - NY^2 = 1$ has the trivial non-negative solution $(1,0)$.The negative Pell equation does not have trivial solutions. Furthermore, the next proposition gives a lot of non-trivial solutions.

Proposition C.3. *For all $n \geq 1$, the following holds*

$$p_{n-1}^2 - Nq_{n-1}^2 = a_n.$$

In particular

$$p_{km-1}^2 - Nq_{km-1}^2 = (-1)^{km}.$$

and so all the couples $(x,y)=(p_{km-1},q_{km-1})\in\mathbb{N}^2$ *with* $k\geq 1$ *are non-negative solutions of the Pell equations.*

Proof. We observe that for all $n\geq 1$, we have $Q_n(1,0)=Q(1,0)\circ D_n=Q_0(p_{n-1},q_{n-1})=p_{n-1}^2-Nq_{n-1}^2$. At the same time, we have $Q_n(1,0)=a_n$. Hence $p_{n-1}^2-Nq_{n-1}^2=a_n$.

Assume first that m is odd. Then, $Q_{m+1}(X,Y)=-Q_1(X,Y)$ and so $a_m=c_{m+1}=-c_1=-a_0=-1=(-1)^m$. Similarly, since $Q_{km+1}(X,Y)=(-1)^mQ_1(X,Y)$, we deduce that $a_{km}=(-1)^{km}$ for all $k\geq 1$.
Assume next that m is even. Then, $Q_{m+1}(X,Y)=Q_1(X,Y)$ and so $a_m=c_{m+1}=c_1=a_0=1=(-1)^m$. Similarly, since $Q_{km+1}(X,Y)=(-1)^mQ_1(X,Y)$, we deduce that $a_{km}=(-1)^{km}$ for all $k\geq 1$.
Using now the first part of the proposition, we deduce the second part. □

The previous proposition gives us infinitely many solutions to the Pell equations. Next, we should find out if there are any other solutions. Actually this is not the case as proposition C.5 will show. In order to prove this fact, we need the following lemma.

Lemma C.4. *Let* $Q(X,Y)=\langle a,b,c\rangle_{X,Y}$ *be a strongly reduced quadratic form. Let* (X_0,Y_0) *be any integer solution of the equation* $Q(X,Y)=\delta$ *where* $\delta\in\{-1,1\}$ *such that* $0<Y_0<X_0$. *Let* χ_1 *be the first root of* $Q(T,1)$ *and let* $r=\lfloor\chi_1\rfloor$. *Then*

1. $(X',Y')=(X_0,Y_0)T(r)^{-1}=(X_0,Y_0)\left(\begin{smallmatrix}0&1\\1&-r\end{smallmatrix}\right)=(Y_0,X_0-rY_0)$ *is a solution of* $\partial Q(X,Y)=\delta$.

2. *If* $Y_0>1$, *then* $0<Y'<X'$, $X'<X_0$ *and* $Y'<Y_0$.

3. *If* $Y_0=1$, *then* $X'=1$. *Furthermore, in this case, if* $a\delta<0$, *then* $Y'=0$ *and if* $a\delta>0$, *then* $Y'=1$.

Proof. Since $(X',Y')T(r)=(X,Y)$, we conclude that $\partial Q(X',Y')=Q(X',Y')\circ T(r)=Q(X,Y)=\delta$. Hence (X',Y') is a solution of $\partial Q(X,Y)=\delta$.

Set $f(T)=Q(T,1)$, denote the second root of $Q(T,1)$ by χ_2 and set $\phi(T)=f(T)-\delta\frac{1}{Y_0^2}$.

The function $f(T)=Q(T,1)$ satisfies $af(t)>0$ for all $t\in]-\infty,\chi_2[\cup]\chi_1,+\infty[$, $af(t)=0$ for $t=\chi_1,\chi_2$ and $af(t)<0$ for all $t\in]\chi_2,\chi_1[$.

Since $-1<\chi_2<0<1\leq r<\chi_1<r+1$, we deduce that

$$af(-1)>0,\quad af(0)<0,\quad af(r)<0,\quad af(r+1)>0.$$

1. Assume $Y_0>1$.

   Then $-1<\delta\frac{1}{Y^2}<1$. The above observation yields that

   $$a\phi(-1)>0,\quad a\phi(0)<0,\quad a\phi(r)<0,\quad a\phi(r+1)>0.$$

   Since $Q(X_0,Y_0)=\delta$, we have that $f(\frac{X_0}{Y_0})=Q(\frac{X_0}{Y_0},1)=\frac{1}{Y_0^2}Q(X_0,Y_0)=\delta\frac{1}{Y_0^2}$. Thus $\frac{X_0}{Y_0}$ is a root of $\phi(T)$. Since by construction $X'=Y_0$ and since by hypothesis $0<Y_0<X_0$, we deduce that $0<X'<X_0$. Since $\frac{X_0}{Y_0}>0$, we deduce by the above that $r<\frac{X_0}{Y_0}<r+1$. This can be rewritten as $rY_0<X_0<(r+1)Y_0$. Finally, we deduce that $0<X_0-rY_0<Y_0$ and so $0<Y'<Y_0=X'$.

2. Assume $Y_0 = 1$. Then, by construction $X' = 1$ and it suffices to show that if $a\delta < 0$, then $X_0 = r$ and if $a\delta > 0$, then $X_0 - 1 = r$.

   (a) Assume $af(X_0) = a(aX_0^2 + bX_0 + c) = a\delta < 0$. Since $X_0 > 0$, we have also $a^2X_0 + ab + \frac{ac}{X_0} = \frac{a\delta}{X_0}$. Thus

$$\begin{aligned} af(X_0+1) &= a(a(X_0+1)^2 + b(X_0+1) + c) \\ &= a^2X_0^2 + 2a^2X_0 + a^2 + abX_0 + ab + ac \\ &= a\delta + 2a^2X_0 + a^2 + ab \\ &= a\delta + a^2X_0 + a^2 + (a^2X_0 + ab) \\ &= a\delta + a^2X_0 + a^2 + \frac{a\delta - ac}{X_0} \\ &= a^2X_0 + a(a+\delta) + \frac{a\delta - ac}{X_0} \end{aligned}$$

   Hence $af(X_0 + 1) = a^2X_0 + a(a + \delta) + a\frac{\delta - c}{X_0}$.
   Since $Q(X, Y)$ is strongly reduced, the Gauss conditions imply that $ac < 0$. Since $X_0 > 1$ and $ac < 0$, we deduce that $a^2X_0 > 0$, $a(a + \delta) \geq 0$ and $a\frac{\delta - c}{X_0} \geq 0$. Thus $af(X_0 + 1) > 0$.
   Since $af(X_0) < 0$ and $af(X_0 + 1) > 0$ and $X_0 > Y_0 = 1$, we deduce by the above observation that $X_0 < \chi_1 < X_0 + 1$ which implies that $X_0 = r$.
   Therefore $X' = Y_0 = 1$ and $Y' = X_0 - r = 0$.

   (b) Assume $af(X_0) = a(aX_0^2 + bX_0 + c) = a\delta > 0$. Since $X_0 > 0$, we have also $a^2X_0 + ab + \frac{ac}{X_0} = \frac{a\delta}{X_0}$. Thus

$$\begin{aligned} af(X_0-1) &= a(a(X_0-1)^2 + b(X_0-1) + c) \\ &= a^2X_0^2 - 2a^2X_0 + a^2 + abX_0 - ab + ac \\ &= a\delta - 2a^2X_0 + a^2 - ab \\ &= a\delta - a^2X_0 + a^2 - (a^2X_0 + ab) \\ &= a\delta - a^2X_0 + a^2 - \frac{a\delta - ac}{X_0} \\ &= -a^2\left(X_0 - 1 - \frac{\delta}{a}\right) + \frac{ac - a\delta}{X_0} \end{aligned}$$

   Hence $af(X_0 - 1) = -a^2\left(X_0 - 1 - \frac{\delta}{a}\right) + \frac{ac - a\delta}{X_0}$.
   Since $Q(X, Y)$ is strongly reduced, the Gauss conditions imply that $ac < 0$. Since $X_0 > 1$, $0 < \frac{\delta}{a} \leq 1$ and $ac < 0$, we deduce that $-a^2\left(X_0 - 1 - \frac{\delta}{a}\right) \leq 0$ and $a\frac{c - \delta}{X_0} \leq 0$. Since $f(T, 1)$ does not have integer roots, we deduce that $af(X_0 - 1) \neq 0$. Thus $af(X_0 - 1) < 0$.
   Since $af(X_0) > 0$ and $af(X_0 - 1) < 0$ and $X_0 - 1 > Y_0 - 1 = 1 - 1 = 0$, we deduce by the above observation that $X_0 - 1 < \chi_1 < X_0$ which implies that $X_0 - 1 = r$.
   Therefore $X' = Y_0 = 1$ and $Y' = X_0 - r = 1$.

□

Example. Consider the strongly reduced quadratic form $Q(X,Y) = \langle 5,-7,-3\rangle_{X,Y} \in Quad(109)$. Then observe that the equation $Q(X,Y) = 1$ has the solution $(X_0,Y_0) = (7,4)$. The first root of $Q(T,1)$ is $\frac{7+\sqrt{109}}{10}$ and set $r = \left\lfloor \frac{7+\sqrt{109}}{10} \right\rfloor = 1$. Set $(X',Y') = (X_0,Y_0)\left(\begin{smallmatrix} 0 & 1 \\ 1 & -r \end{smallmatrix}\right) = (7,4)\left(\begin{smallmatrix} 0 & 1 \\ 1 & -1 \end{smallmatrix}\right) = (4,3)$. Now it is easy to verify that $(4,3)$ is a solution of $1 = \partial Q(X,Y) = \langle -5,3,5\rangle_{X,Y}$.

We can repeat this to obtain the solution $(3,1)$ of $1 = \partial^2 Q(X,Y) = \langle 3,-7,-5\rangle_{X,Y}$.

Then the first root of $\partial^2 Q(T,1)$ is $\frac{7+\sqrt{109}}{6}$ and so $\left\lfloor \frac{7+\sqrt{109}}{6} \right\rfloor = 2$. One easily verifies that $(3,1)\left(\begin{smallmatrix} 0 & 1 \\ 1 & -2 \end{smallmatrix}\right) = (1,1)$ is a solution of $1 = \partial^3 Q(T,1) = \langle -7,5,3\rangle_{X,Y}$.

At last, we can repeat the process to obtain the solution $(1,0)$ of $1 = \partial^4 Q(T,1) = \langle 1,-9,-7\rangle_{X,Y}$.

Lemma C.4 gives a method to reduce a solution (X_0,Y_0) of an equation $Q(X,Y) = \delta$ where $\delta \in \{-1,1\}$ and $Q(X,Y)$ denotes a strongly reduced quadratic form to a solution (X',Y') of $\partial Q(X,Y) = \delta$ where $0 \le Y' \le X'$, $X' < X_0$ and $Y' \le Y_0$. Furthermore, by applying the lemma recursively at the resulting solutions, one will end up at a solution $(1,0)$ as indicates already the last example.

In the proof of proposition C.5, we are going to see that any solution of the Pell equation $Q(X,Y) = X^2 - NY^2 = \pm 1$ gives rise to a solution of the kind of lemma C.4 of the equation $\partial Q(X,Y) = \pm 1$ where $\partial Q(X,Y)$ is strongly reduced. Thus, we can use lemma C.4 to obtain recursively easier solutions for the derived forms of $Q(X,Y)$. This will then be the key ingredient to prove that the only solutions of the Pell equation are the one given in proposition C.3, namely the pairs (p_{km-1}, q_{km-1}) where $k \in \mathbb{N} \setminus \{0\}$, m is the caliber of $\sqrt{N}$ and $\left(\begin{smallmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{smallmatrix}\right) = T(r_0)T(r_1)...T(r_{n-1})$ with $r_n = \left\lfloor \partial^n \sqrt{N} \right\rfloor$.

Proposition C.5. *The non-trivial non-negative solutions (x,y) of the Pell equations are exactly the couples*

$$(x,y) = (p_{km-1}, q_{km-1}) \qquad \text{with } k \ge 1.$$

In particular the solution $(x_0,y_0) = (p_{m-1}, q_{m-1})$ is the minimal non-trivial non-negative solution, i.e. $x > x_0$ and $y > y_0$ for every other non-trivial non-negative solution (x,y) of the Pell equations.

Proof. By proposition C.3, we know already that the couples (p_{km-1}, q_{km-1}) are solutions of the Pell equations. Thus, it suffices to show that there are no other solutions.

Assume (u,v) is a non-negative solution of the Pell equation $X^2 - NY^2 = \delta$ with $\delta \in \{-1,1\}$. Assume by contradiction that $u = 0$.

with $u > 0$ and $v > 0$ and $u^2 + Nv^2 = \delta$ with $\delta \in \{-1,1\}$.

In order to show that (u,v) has to be one of the couples (p_{mk-1}, q_{mk-1}), we are going to show, after dealing with some special cases, that in general the solution (u,v) of $Q(X,Y) = \delta$ where $Q(X,Y) = X^2 - NY^2$ and $\delta \in \{-1,1\}$ gives rise to a solution (x_1,y_1) of the equation $\partial Q(X,Y)$ with $0 < y_1 < x_1$. Then re can apply lemma C.4 recursively to obtain a solution (x_k,y_k) of $\partial^k Q(X,Y) = \delta$ that satisfies $0 < y_k < x_k$ for all $k \in \{1,...,n\}$ for some n and $y_n = 1$. Then lemma C.4 will be used again once or twice to obtain the solution of $\partial^\psi Q(X,Y) = \delta$ for some $\psi > n$. A short analysis will show that actually $\partial^\psi Q(X,Y) = \pm \partial^m Q(X,Y)$ which leads to the conclusion that $\psi = k'm$ for some $k' \in \mathbb{N} \setminus \{0\}$ and $(u,v) = (p_{k'm-1}, q_{k'm-1})$.

Now, that the outline of the proof is clear, let us start with the formal proof.

Assume by contradiction that $v > u > 0$. Then $v \geq u + 1 > 1$ and so $v^2 \geq u^2 + 2u + 1 > u^2 + 3$. This implies that $Nv^2 > v^2 > u^2 + 3$ and so $-Nv^2 < -u^2 - 3$. Hence $u^2 - Nv^2 < u^2 - u^2 - 3 = -3$. This contradicts the fact that (u, v) is a solution of a Pell equation.

Assume that $v = u > 0$. Then $u^2 - Nv^2 = (1 - N)u^2 < 0$. Since (u, v) is a solution of a Pell equation we deduce that $u^2 - Nv^2 = \pm 1$. Since $u^2 - Nv^2 = (1-N)u^2 < 0$, $u^2 - Nv^2 \neq 1$. Furthermore, $u^2 - Nv^2 = -1$ holds if and only if $N = 2$ and $u = v = 1$. One can easily observe that in this case (u, v) is one of the considered couples, namely it is the minimal non-trivial solution of $X^2 - 2Y^2 = -1$.

Assume $0 < v < u$. Set $r_0 = \left\lfloor \sqrt{N} \right\rfloor \geq 1$ and consider

$$(x_1, y_1) = (u, v)T(r_0)^{-1} = (u, v) \begin{pmatrix} 0 & 1 \\ 1 & -r_0 \end{pmatrix} = (v, u - r_0 v).$$

Since N is not a square $0 < r_0 < \sqrt{N} < r_0 + 1$ and since (u, v) satisfies the Pell equation $X^2 - NY^2 = \delta$, we deduce that $u^2 = Nv^2 + \delta$. Thus

$$\begin{aligned} r_0^2 < N < (r_0 + 1)^2 \\ r_0^2 + 1 \leq N \leq (r_0 + 1)^2 - 1 \\ (r_0^2 + 1)v^2 \leq Nv^2 \leq ((r_0 + 1)^2 - 1)v^2 \\ (r_0^2 + 1)v^2 - 1 \leq Nv^2 + \delta \leq ((r_0 + 1)^2 - 1)v^2 + 1 \\ r_0^2 v^2 + (v^2 - 1) \leq u^2 \leq r_0^2 v^2 + 2r_0 v^2 + 1 \end{aligned}$$

If $v > 1$, then $v^2 - 1 > 0$ and $r_0^2 v^2 + 2r_0 v^2 + 1 < r_0^2 v^2 + 2r_0 v^2 + v^2 = (r_0 + 1)^2 v^2$ and so the last line implies that $r_0^2 v^2 < u^2 < (r_0 + 1)^2 v^2$. Since $r_0, u, v > 0$, we obtain

$$r_0 v < u < (r_0 + 1)v.$$

Since $v < u$ and since $x = v$, we deduce that $x < u$. By the right inequality, we have $u - r_0 v < v$ and so $y < v$. Furthermore, by the same arguments as in the proof of lemma C.4, we deduce that (x_1, y_1) is a solution of $\partial Q(X, Y) = \delta$.

Consider now the initial quadratic form $Q_0(X, Y) = \langle 1, 0, -N \rangle_{X,Y}$. Then we showed that $Q_1(X, Y) = \partial Q(X, Y) = \langle r_0^2 - N, 2r_0, 1 \rangle_{X,Y}$ is strongly reduced. By lemma C.4, we deduce that if we start from $Q_1(X, Y)$ and the solution (x_1, y_1) of $Q_1(X, Y) = \delta$, then we can obtain a solution (x_2, y_2) of the equation $Q_2(X, Y) = \partial Q_1(X, Y) = \delta$ such that $0 < y_2 < y_1$ and $0 < x_2 < x_1$. We can reiterate this process as long as the resulting second entry is strictly greater than 1. By Fermat's principle of descent, there is one $n \in \mathbb{N}$ such that $y_{n-1} > 1$ and $y_n = 1$. Thus the reiterating process ends. At this point we obtained two decreasing sequences of integers $1 = y_n < y_{n-1} < ... < y_2 < y_1$ and $x_n < x_{n-1} < ... < x_2 < x_1$ that are defined by

$$(x_i, y_i) = (u, v)T(r_0)^{-1}T(r_1)^{-1}...T(r_{i-1})^{-1}$$

for all $i \in \{1, ..., n\}$ where since r_i is the integral part of the first root of $Q_i(T, 1) = \partial^i Q_1(T, 1)$ since $\sqrt{N} = [r_0, r_1, r_2, ...]$ (i.e. $\left\lfloor \partial^i \sqrt{N} \right\rfloor$). Furthermore, we know that $x_i = y_{i-1}$ and $Q_i(x_i, y_i) = \delta$ for all $i = 1, ..., n$.

Let us use lemma C.4 again on $(x_n, y_n) = (x_n, 1)$ and $Q_n(X, Y) = \langle a_n, b_n, c_n \rangle_{X,Y}$. If $a_n \delta < 0$, then we obtain $x_n > x_{n+1} = 1$, $1 = y_n > y_{n+1} = 0$ and $Q_{n+1}(x_{n+1}, y_{n+1}) = \delta$.

If $a_n\delta > 0$, then we obtain $x_n > x_{n+1} = 1$, $1 = y_n = y_{n+1}$ and $Q_{n+1}(x_{n+1}, y_{n+1}) = \delta$. In the latter case one uses lemma C.4 again on $(x_{n+1}, y_{n+1}) = (x_{n+1}, 1)$ and $Q_{n+1}(X,Y) = \langle a_{n+1}, b_{n+1}, c_{n+1}\rangle_{X,Y}$. Since the signs of a_{n+1} and a_n are opposite, we then have $a_{n+1}\delta < 0$ and so we obtain $y_{n+2} = 0$, $x_{n+2} = 1$ and $Q_{n+2}(x_{n+2}, y_{n+2}) = \delta$.

Either way, we end up at the final pair $(x_\psi, y_\psi) = (1,0)$ for some $\psi \in \mathbb{N} \setminus \{0\}$. Since $Q_\psi(x_\psi, y_\psi) = \delta$, we conclude that $Q_\psi(1,0) = a_\psi = \delta$. Thus $Q_\psi(X,Y)$ is a strongly reduced quadratic form with discriminant $4N$ which satisfies $Q_\psi = \langle \delta, b, c\rangle$.

Since $Q_m(X,Y)$ is strongly reduced, we know that $|b| < \sqrt{4N} = 2\sqrt{N}$. Moreover by the Gauss conditions, $2\sqrt{N} - |b| < 2$ and so $2\sqrt{N} - 2 < |b|$. Since $r_0 = \left\lfloor \sqrt{N} \right\rfloor < \sqrt{N} < r_0 + 1$, we obtain

$$2r_0 - 2 < 2\sqrt{N} - 2 < |b| < 2\sqrt{N} < 2r_0 + 2.$$

Hence $|b| \in \{2r_0 - 1, 2r_0, 2r_0 + 1\}$. Since $b^2 - 4\delta c = 4N$, 4 divides b^2 and so 2 divides b. Thus we conclude that $|b| = 2r_0$. Since $Q_m(X,Y)$ is strongly reduced, the sign of b is the opposite sign of a, we finally obtain $b = -\delta 2r_0$. From the relation $b^2 - 4\delta c = 4N$ we deduce that $c = \delta(r_0^2 - N)$. Hence $Q_\psi(X,Y) = \delta\langle 1, -2r_0, r_0^2 - N\rangle_{X,Y} = \delta\widetilde{Q_1}(X,Y)$.

By corollary 11.9, we know that $\sqrt{N} = [r_0, \overline{r_1, r_2, ..., r_2, r_1, 2r_0}]$ which is the first positive root of $Q(T,1)$. Thus the first root of $Q_1(T,1)$ is $[\overline{r_1, r_2, ..., r_2, r_1, 2r_0}]$. By corollary 11.6 the first root of $\widetilde{Q_1}(T,1)$ is $[\overline{2r_0, r_1, r_2, ..., r_2, r_1}]$, which is also the first root of $Q_m(T,1)$ (since the caliber of $\sqrt{N}$ is m). Since $Q_1(X,Y)$ is strongly reduced, proposition 9.2 yields that $\widetilde{Q_1}(X,Y)$ is strongly reduced. Since $\widetilde{Q_1}(X,Y)$ and $Q_m(X,Y)$ are strongly reduced, have the same first root and have the same discriminant, lemma 6.6 yields that $\widetilde{Q_1}(X,Y) = \pm Q_m(X,Y)$.

Since $Q_\psi(X,Y) = \delta\widetilde{Q_1}(X,Y) = \pm\delta Q_m(X,Y)$, we conclude that $\psi = km$ for some $k \geq 1$. Thus

$$(u,v) = (1,0)T(r_0)T(r_1)...T(r_{km-1}) = (1,0)D_{km} = (p_{km-1}, q_{km-1}).$$

□

Example. Consider the equation $X^2 - 14Y^2 = \pm 1$. Then since $\sqrt{14} = [3, \overline{1,2,1,6}]$ (and so $m = 4$), the last two propositions yield that the positive integers solutions of $X^2 - 14Y^2 = \pm 1$ are given by $(p_{4k-1}, q_{4k-1}) = D_{4k}^k(1,0)$ and moreover, since m is even, we have that $p_{4k-1}^2 - 14q_{4k-1}^2 = 1$. Indeed $D_4 = T(3)T(1)T(2)T(1) = \left(\begin{smallmatrix} 15 & 11 \\ 4 & 3 \end{smallmatrix}\right)$ and the smallest positive integer solution of $X^2 - 14Y^2 = \pm 1$ is $(15,4)$.

Corollary C.6. *There exist integer solutions to the equation*

$$x^2 - Ny^2 = -1$$

if and only if the caliber m of $\sqrt{N}$ is odd.

Proof. By proposition C.5 the solutions (x,y) to the Pell equation with $x > 0$ and $y > 0$ are exactly the couples $(x,y) = (p_{km-1}, q_{km-1})$ with $k \geq 1$. By proposition C.3 one has $p_{km-1} - Nq_{km-1} = (-1)^{km}$. Thus if m is odd, then (p_{m-1}, q_{m-1}) is a solution of $x^2 - Ny^2 = -1$. However if m is even, then $p_{km-1} - Nq_{km-1} = 1$ for all $k \geq 1$ and so $x^2 - Ny^2 = -1$ does not have solutions. □

Example. Consider the equation $X^2 - 13Y^2 = \pm 1$. Using the same method as in the last example, we deduce from the continued fraction $\sqrt{13} = [3, \overline{1,1,1,1,6}]$ (and so $m = 5$) that (18,5) is a solution to $X^2 - 13Y^2 = -1$.

Let us summarize the previous results in a single theorem.

Theorem C.7. *Let $N \in \mathbb{N} \setminus \mathbb{N}^2$ and let m be the caliber of N.*

1. *If m is odd, then:*
   (a) *The non-negative solutions of the Pell equation $X^2 - NY^2 = 1$ are exactly the pairs (p_{2km-1}, q_{2km-1}) with $k \in \mathbb{N} \setminus \{0\}$ and the trivial pair $(1, 0)$.*
   (b) *The non-negative solutions of the Pell equation $X^2 - NY^2 = -1$ are exactly the pairs $(p_{(2k+1)m-1}, q_{(2k+1)m-1})$ with $k \in \mathbb{N}$.*
2. *If m is even, then:*
   (a) *The non-negative solutions of the Pell equation $X^2 - NY^2 = 1$ are exactly the pairs (p_{km-1}, q_{km-1}) with $k \in \mathbb{N} \setminus \{0\}$ and the trivial pair $(1, 0)$.*
   (b) *The Pell equation $X^2 - NY^2 = -1$ does not have solutions.*

D Another proof that in general a quadratic form and its negative are not equivalent and a late motivation for the non-standard definitions used in this document

In corollary 13.7, we developed a general condition on when a strongly reduced quadratic form $Q(X,Y) \in Quad(D)$ with $D \in \mathbb{N} \setminus \mathbb{N}^2$ is equivalent to $-Q(X,Y)$. Using corollary 7.4, we deduce that any general quadratic form $Q'(X,Y)$ with positive square discriminant is equivalent to $-Q'(X,Y)$ if and only if the strongly reduced forms that are equivalent to $Q'(X,Y)$ satisfy this same condition.

We are going to exhibit a general counterexample which shows that some strongly reduced forms are not equivalent to their negative forms. This counterexample will be developed independently from the theory that we developed in this document.

Counterexample

Proposition D.1. *The quadratic form $Q(X,Y) = \langle 2,-2,-1\rangle_{X,Y} \in Quad(12)$ is not equivalent to $-Q(X,Y) = \langle -2,2,1\rangle_{X,Y}$.*

By comparing all the sources in the bibliography, one can convince oneself that the quadratic form $Q(X,Y) = \langle 2,-2,-1\rangle_{X,Y}$ is strongly reduced in almost every source (and for the ones where it is not, $-Q(X,Y)$ is strongly reduced). Furthermore, one of the only common definitions in all of the given sources is the definition of equivalence of quadratic forms which is formulated as in definition 5.8 and can be computed using proposition 5.7. Since those definitions are the only theoretical points on which most classical texts agree, we are going to use no more theory to prove the next proposition.

Proof. Assume by contradiction that $-Q(X,Y)$ is equivalent to $Q(X,Y)$. Then there exists $S = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in GL(2,\mathbb{Z})$ such that $-Q(X,Y) = Q(X,Y) \circ S$. Then, by proposition 5.7, we deduce that

$$\begin{cases} -2 &= & 2p^2 - 2pr - r^2 \\ 2 &= & 4pq - 2qr - 2ps - 2rs \left(\begin{smallmatrix} p \\ r \end{smallmatrix}\right) \\ 1 &= & 2q^2 - 2qs - s^2 \end{cases}$$

By the last equation, we deduce that

$$0 = s^2 + 2qs + (1 - 2q^2).$$

We solve for s and deduce that

$$s = \frac{-2q \pm \sqrt{12q^2 - 4}}{2} = -q \pm \sqrt{3q^2 - 1}. \tag{16}$$

Since $S \in GL(2,\mathbb{Z})$, we deduce that $q, s \in \mathbb{Z}$. Since $q, s \in \mathbb{Z}$, this implies that $s + q \in \mathbb{Z}$ and hence (16) implies that $\sqrt{3q^2 - 1} \in \mathbb{Z}$. Hence $3q^2 - 1 = y^2$ for some $y \in \mathbb{Z}$. Then $-1 \equiv y^2 \pmod 3$ or in other words $2 \equiv y^2 \pmod 3$. But this contradicts the fact that the only squares modulo 3 are 0 and 1. □

Because of the universality of this counterexample, we deduce that a general approach to quadratic forms cannot rely on the equivalence of quadratic forms and their negatives.

A late motivation for the non-standard definitions used in this document

Throughout the whole document, we compared regularly our approach to quadratic forms with the method used by GAUSS. This does not only have historical reasons but is also due to the fact that, after a detailed analysis, one can get to the conclusion that his method is well developed and without apparent mistakes. However, GAUSS' approach is not well suited to solve the problem of the equivalence of a quadratic form and its negative, nor can it be applied directly to continued fractions which are two points in which we were especially interested in.

Thus, one can think of an easier method to do so. That's what we developed in this document. Indeed, we gave in chapter 7 section 7.3 an easy computation method to compute the derived form of a quadratic form. Furthermore, the n-th derived form of a given quadratic form with respect to the first or second root of the corresponding polynomial is completely determined by the continued fraction of this root. Hence, the computation of derived forms is an effective task.

The main idea of our approach is due to the document [1] and was strengthened partially by the book [10]. Actually both sources end up at the same conclusions as did GAUSS. Furthermore, both sources conclude some more sophisticated results. Because of the simplicity of this approach and the supplementary results which can be obtained through it, this approach seemed perfect to study quadratic forms and to use the obtained results to discover a relation between quadratic forms and lattices of quadratic fields. Unfortunately both cited sources present minor inconsistencies throughout their development whereby one could not have concluded some results, neither the wished supplementary results.

Example. Let us set ourself in the context and notation of [1]. In this document, the definition of a quadratic form Q, the equivalence between quadratic forms, the definition of the two roots x^+, x^- of the polynomial $Q(T,1)$ and the definition of strongly quadratic surds are the same as in our development. However a quadratic form $Q(X,Y) = aX^2 + bXY + cY^2$ is strongly reduced if $a > 0$ and x^+ is strongly reduced. (We have generalized this to accept also $a < 0$ but then we need to consider the x^-instead of x^+, which explains the definition of the first root.)

At p.15, the derived form of a weakly quadratic form Q is defined to be $-Q \circ T(r)$ where $T(r) = \left(\begin{smallmatrix} r & 1 \\ 1 & 0 \end{smallmatrix}\right)$ with $r = \lfloor x^+ \rfloor$. At p.21, the document concludes that the only reduced forms equivalent to a reduced form Q are its derived forms. Furthermore a development that appears at the end of this same page yields that all of them are equivalent to Q. This development goes as follows: *"Consider the whole set $K(D)$ of reduced quadratic forms with discriminant D. [...]. Take $Q \in K(D)$; [...]. C denotes the wide equivalence class of Q; [...], we see that the set $K(D) \cap C$ is the orbit of Q under the operator ∂"*. Unfortunately, we see that if Q is equivalent to $-Q \circ T(r)$, then since $-Q \circ T(r)$ is equivalent to $-Q$, this yields that Q is equivalent to $-Q$ which by our counterexample is not true in general.

At p.17, the associate form of a quadratic form Q is defined to be $-Q \circ \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$. Furthermore, immediately after the definition the document concludes the equivalence of Q and its associate. This leads to the same contradiction as the definition and equivalence of derived forms.

Since most of the results rely on the derived and associate forms, the conclusions become insecure.

Example. Let us set ourself in the context and notation of [10]. In this document, the definition of a quadratic form f, the equivalence between quadratic forms are the same as in our development. However a quadratic form $f(X,Y) = aX^2 + bXY + cY^2$ is (strongly) reduced

if $|\sqrt{D} - 2|a|| < b < \sqrt{D}$. This differs slightly from our definition which is strengthened by the fact that if $(\theta_1(f), \theta_2(f)) = (\frac{b+\sqrt{D}}{2a}, \frac{b-\sqrt{D}}{2a})$, then f is reduced if and only if $|\theta_1(f)| > 1$, $|\theta_2(f)| < 1$ and $\theta_1(f)\theta_2(f) < 0$. (For example $-2X^2 + 2XY + Y^2$ is reduced). The whole reduction algorithm which finds an equivalent reduced form for any given quadratic form presented from p.112 to p.126 is really effective and has a low complexity.

At p.127, the book defines in formula (6.27) for given $f = (a, b, c)$, the transformation $\tau(f) = f\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = (-a, b, -c)$. However, by the computation rules, one has $f\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = (a, -b, c)$. Hence the second equality is wrong. Furthermore, immediately after the definition, the authors claim that $\tau(f)$ is equivalent to f which holds certainly for $\tau(f) = f\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, but does not hold in general for $\tau(f) = (-a, b, -c)$ as our counterexample shows.

Although, the subsequent development seems to be correct, it should be verified using different methods since most results rely on the definition of this transformation.

Since the initial sources presented some minor inconsistencies, another reference had to be found. One such reference is the book [8]. Unfortunately, the theory described in this book was not detailed and did not present all the wanted results. Furthermore, the definitions used in this book left space for generalizations, which we used to combine quadratic forms with continued fractions in an elegant manner. However, the first hint that the strongly reduced forms equivalent to a of a given strongly reduced form $Q(X, Y)$ are split into the cycle of $Q(X, Y)$ and the cycle of its associate form is due to the last observation in the last appendix of this book.

Finally, several other books like [11], [12], [9], which follow the approach of Gauss confirmed our guesses and pushed us to prove the results formally.

At last, let us highlight the fact that, although most of the results were first obtained by the computation of multiple examples, each non-standard definition was obtained naturally from the previously developed theory in such a way that it bypasses the problem of the equivalence of a quadratic form and its negative.

References

[1] Gilles Lachaud. Continued fractions, binary quadratic forms, quadratic fields, and zeta functions. In *Algebra and topology 1988 (Taejŏn, 1988)*, pages 1–56. Korea Inst. Tech., Taejŏn, 1988.

[2] A. Ya. Khinchin. *Continued fractions*. The University of Chicago Press, Chicago, Ill.-London, 1964.

[3] Ivan Niven, Herbert S. Zuckerman, and Hugh L. Montgomery. *An introduction to the theory of numbers*. John Wiley & Sons, Inc., New York, fifth edition, 1991.

[4] C. D. Olds. *Continued fractions*. Random House, New York, 1963.

[5] Doug Hensley. *Continued fractions*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2006.

[6] Lisa Lorentzen and Haakon Waadeland. *Continued fractions with applications*, volume 3 of *Studies in Computational Mathematics*. North-Holland Publishing Co., Amsterdam, 1992.

[7] G. Serret, J.-A. et Darboux. *Oeuvres de LAGRANGE*. Gauthier-Villars,Paris, 1867-1892. 14 volumes.

[8] André Weil. *Number theory*. Birkhäuser Boston, Inc., Boston, MA, 1984. An approach through history, From Hammurapi to Legendre.

[9] Duncan A. Buell. *Binary quadratic forms*. Springer-Verlag, New York, 1989. Classical theory and modern computations.

[10] Johannes Buchmann and Ulrich Vollmer. *Binary quadratic forms*, volume 20 of *Algorithms and Computation in Mathematics*. Springer, Berlin, 2007. An algorithmic approach.

[11] Carl Friedrich Gauss. *Untersuchungen über höhere Arithmetik*. Deutsch herausgegeben von H. Maser. Chelsea Publishing Co., New York, 1965. Disquisitiones Arithmeticae.

[12] J. H. Conway and N. J. A. Sloane. *Sphere packings, lattices and groups*, volume 290 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, third edition, 1999. With additional contributions by E. Bannai, R. E. Borcherds, J. Leech, S. P. Norton, A. M. Odlyzko, R. A. Parker, L. Queen and B. B. Venkov.

[13] Évariste Galois. Analyse algébrique. Démonstration d'un théorème sur les fractions continues périodiques. *Ann. Math. Pures Appl. [Ann. Gergonne]*, 19:294–301, 1828/29.