SCUOLA
NORMALE
SUPERIORE

Tesi di Perfezionamento in Matematica

Reduction and specialization of hyperelliptic continued fractions

Olaf Merkert

Anno accademico 2015-2016

Relatore: Prof. Umberto Zannier

To my friends, for all the time spent together.

Contents

1. Introduction **9**

1.1. Acknowledgements . . . 14

1.2. Notation reference . . . 15

2. Pell equation **17**

2.1. Multiplication law and unit-norm equation . . . 17

2.2. Units of hyperelliptic function fields . . . 18

2.3. Laurent series and valuation . . . 19

2.4. Group structure of Pell solutions . . . 21

3. Rational approximations **23**

3.1. Convergents . . . 23

3.2. Pell solutions are convergents . . . 25

3.3. The universal property of best-approximation . . . 26

3.4. A linear system for computing convergents . . . 28

4. A (hyper)elliptic curve **33**

4.1. Defining the (hyper)elliptic curve . . . 33

4.2. Divisors and the Jacobian variety . . . 34

4.2.1. Divisors . . . 35

4.2.2. Order functions . . . 35

4.2.3. Embedding the curve in the Jacobian . . . 36

4.3. Divisors of convergents . . . 37

5. Continued fractions **41**

5.1. Finite continued fractions . . . 41

5.2. Continued fractions and matrix products . . . 42

5.3. Infinite continued fractions . . . 43

5.4. Continued fraction process . . . 45

5.5. Canonical convergents and classification of best-approximations . . . 46

5.6. Multiplication of a continued fraction by a constant . . . 48

5.7. Periodic continued fractions . . . 49

5.7.1. Periods . . . 49

5.7.2. Quasi-periods . . . 50

5.8. Reduced complete quotients . . . 51

6. Computation of hyperelliptic continued fractions **55**

6.1. Representing complete quotients with polynomials . . . 56

6.2. Complete quotients are eventually σ-reduced 58
6.3. Periodicity and Pell equation 60
6.4. Torsion order and period length 61
6.5. Period lengths over finite fields 61
6.6. Divisors of complete quotients 62

7. Specialization of continued fractions 65
7.1. Specialization of Laurent series 65
7.1.1. Discrete valuation rings 65
7.1.2. Gauss norms 66
7.1.3. Criterion for bounded square roots 69
7.2. Specialization of polynomial continued fractions 69
7.2.1. Good reduction 70
7.2.2. Bad reduction 72
7.2.3. Reduction and normalisation of continued fractions 73
7.3. Normalisation and reduction of convergents 73
7.4. Calculating valuations 77
7.4.1. Relating complete quotients with convergents 77
7.4.2. Fibre analysis of λ 79

8. Specialization of hyperelliptic continued fractions 83
8.1. Reduction of periodic quadratic continued fractions 83
8.1.1. Reduction for $\deg D = 2$ 84
8.2. Reduction of non-periodic quadratic continued fractions 85
8.2.1. Reduction to square 86
8.2.2. Reduction to periodic and denominators 86
8.2.3. Primes occurring in infinitely many denominators 87
8.3. Genus 1 valuation patterns 90
8.4. Reduction of abelian varieties 95
8.4.1. Reduction of curve and its Jacobian 95
8.4.2. Reduction of torsion points and periodicity test 97
8.4.3. Shortening of quasi-period 98
8.4.4. Reduction of convergent divisors 99
8.5. Good reduction at infinitely many primes 101
8.6. Complex functions case 103
8.6.1. Results of Masser and Zannier 103
8.6.2. Repeated occurrences of $t - t_0$ 105

9. Heights 107
9.1. Heights 107
9.1.1. Places and Product formula 107
9.1.2. Height on projective and affine space 108
9.1.3. Height of polynomials 109
9.1.4. Weil's Height Machine and Néron-Tate height 110

9.1.5. Heights on the Jacobian . . . 111
9.2. Height of convergents . . . 112
9.2.1. Height bounds for series coefficients of $\sqrt{D}$ . . . 112
9.2.2. Lower bound . . . 114
9.2.3. Upper bound . . . 115
9.3. Connecting heights and valuations . . . 117

10. Examples **119**
10.1. Reduction to a square . . . 119
10.2. Reduction periodic to periodic . . . 120
10.3. Reduction non-periodic to periodic . . . 122
10.3.1. Genus 1 . . . 122
10.3.2. Genus 2 . . . 125
10.4. Non-constant degrees . . . 131
10.5. Recurring partial quotients . . . 133

A. Appendix **135**
A.1. Polynomial Pell equation in characteristic 2 . . . 135
A.2. Valuations in Laurent series quotients . . . 136
A.3. A lemma for a quadratic form . . . 138

Bibliography **139**

1. Introduction

This thesis investigates how prime factors arise in denominators of polynomial continued fractions, with a focus on continued fractions of the square root of a polynomial. This is strongly related to the problem of reducing polynomial continued fractions modulo a prime.

Continued fractions have a very long history – those of rational numbers express the Euclidean Algorithm which was already known in ancient Greece. In modern times, mathematicians such as Lagrange and Galois studied continued fractions of irrational numbers, in particular quadratics (for example square roots). Even today, continued fractions of real numbers remain an important research topic in number theory and other branches of mathematics.

We write a continued fraction as

$$\alpha = [a_0, a_1, a_2, \dots] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}.$$

For the classical continued fractions with $\alpha \in \mathbb{R}$, the *partial quotients* a_n are integers, positive for $n \geq 1$. Instead, one may also take the $a_n \in \mathbb{Q}[X]$ to be polynomials, non-constant for $n \geq 1$, to build the continued fraction of a Laurent series in X^{-1}, i.e. $\alpha \in \mathbb{Q}((X^{-1}))$. The role of the nearest integer is then played by the polynomial part of the Laurent series.

We are interested for which n a given prime number $\mathfrak{p}$ divides the denominator of the coefficients of the a_n (for brevity, we say the "prime $\mathfrak{p}$ appears in the denominator of a_n"). We are especially interested when it first appears and whether it can disappear again.

Of particular interest is the continued fraction of $\sqrt{D}$, where $D \in \mathbb{Q}[X]$ is a monic non-square polynomial of even degree $2d$. It was first considered by Abel in 1826 [Abe26], who used it to study the integration in elementary terms of certain algebraic functions. Abel showed that periodicity of this continued fraction is equivalent to the existence of a non-trivial solution $p, q \in \mathbb{Q}[X]$, $q \neq 0$ of the polynomial Pell equation $p^2 - D\,q^2 = 1$ (see Chapter 2 and Theorem 6.3). We say that D is *Pellian* if such a solution exists. Later, Chebyshev expanded upon these results [Che57].

We call continued fractions of this type *hyperelliptic* because they encode information about the (hyper)elliptic curve $Y^2 = D(X)$, given that $d \geq 1$ and D is also squarefree. For example, if $O_\pm$ are the two points at infinity in a smooth model, the class of $(O_+) - (O_-)$ is torsion in the Jacobian of the curve if and only if D is Pellian, i.e. the continued fraction is periodic (see Theorem 4.1).

Note that the polynomials of degree $2d$, after some normalisation, form an affine variety of dimension $2d - 2$. The Pellian polynomials are then contained in a denumerable union of subvarieties of dimension at most $d-1$ (see [Zan14], a survey focusing on the geometric aspects of the polynomial Pell equation). This implies that unlike positive square-free integers which are always "Pellian", most polynomials D are not Pellian, and usually we do not expect a periodic continued fraction. But other results for the classical continued fractions have direct analogues for polynomial continued fractions, see for example [Sch00].

Let us also introduce the *canonical convergents* which are defined via the recurrence relations

$$p_n = a_n\, p_{n-1} + p_{n-2}, \qquad\qquad q_n = a_n\, q_{n-1} + q_{n-2}$$

and $p_0 = a_0,\ p_{-1} = q_0 = 1,\ q_{-1} = 0$. These imply that $p_n, q_n \in \mathbb{Q}[X]$ are coprime for any integer $n \geq 0$, via the identity $p_n\, q_{n-1} - q_n\, p_{n-1} = (-1)^{n+1}$. The canonical convergents arise by calculating the numerator and denominator of the finite continued fraction

$$\frac{p_n}{q_n} = [a_0, a_1, \dots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \cfrac{1}{a_n}}}.$$

Note that they are usually not monic nor have content 1. This is related to prime numbers suddenly appearing in the denominators of the coefficients of the a_n, something van der Poorten was already aware of (see [vdP01]).

This follows from the fact that $D_\mathfrak{p}$, the reduction of D modulo $\mathfrak{p}$, is Pellian unless it is a square (the Jacobian over $\mathbb{F}_\mathfrak{p}$ is finite, all points on it are torsion), so the continued fraction of $\sqrt{D_\mathfrak{p}}$ is automatically periodic. This leads to one of the main results of this thesis:

Theorem 1.1. *Let $\sqrt{D} = [a_0, a_1, a_2, \dots]$. If $D \in \mathbb{Q}[X]$ is not Pellian, then for all prime numbers $\mathfrak{p}$ except finitely many, $\mathfrak{p}$ appears in infinitely many polynomials a_n in a denominator (of the coefficients).*

We prove this in Theorem 8.1 more generally for arbitrary number fields. Note that the formula for multiplying polynomial continued fractions with a constant,

$$\mathfrak{p}^e\, [b_0, b_1, b_2, \dots] = [\mathfrak{p}^e\, b_0, \mathfrak{p}^{-e}\, b_1, \mathfrak{p}^e\, b_2, \dots], \quad (e \in \mathbb{Z}), \tag{1.1}$$

raises the question if – at least for a fixed prime $\mathfrak{p}$ – the infinite occurrences in the denominators of the a_n arise in this rather trivial way. Indeed, this is not the case; we can show that for any $e \in \mathbb{Z}$, the continued fraction of $\sqrt{\mathfrak{p}^{-2e} D} = [b_0, b_1, b_2, \dots]$ enjoys the property that $\mathfrak{p}$ appears in infinitely many b_n as a denominator.

The primes which are excluded in Theorem 1.1 are the prime 2, any primes appearing already in a denominator of D and those with $D_\mathfrak{p}$ square. For technical reasons, we may also need to exclude further primes, depending on the first occurrence of an a_n with minimal degree. These primes can be determined effectively, too (see Remark 8.15). The prime 2 is of course excluded because we are taking square roots.

Remark 1.1. This result is true only for $\sqrt{D}$, and does not apply to other elements of the hyperelliptic function field $\mathbb{Q}(X, \sqrt{D})$. With an analogue of the fact that there are infinitely many primes $\mathfrak{p}$ such that $2^n \not\equiv 5 \pmod{\mathfrak{p}}$ for all n, we construct an example of type $\alpha = \left(r + \sqrt{D}\right)/X$ where there are infinitely many primes $\mathfrak{p}$ that never appear in the denominators of the a_n (see Theorem 8.4 in Section 8.5). The proof relies on the Čebotarev density theorem, and represents a variant of the results of Schinzel [Sch60] and Corrales-Rodrigáñez-Schoof [CRS97].

For $\deg D = 4$, another more explicit approach avoids the issue of excluding additional primes. This is described in the rather technical Theorem 8.2 and Corollary 8.20. The former has another consequence for the Gauss norm of the convergents.

Recall that, given some valuation on a field K, we may extend the valuation to polynomials. Define the valuation of a polynomial in $K[X]$ as the minimum of the valuation on the coefficients (see Section 7.1.2 for details). The corresponding absolute value is usually called a *Gauss norm.* For $D \in \mathbb{Q}[X]$, we naturally use the $\mathfrak{p}$-adic valuation $\nu_{\mathfrak{p}}$. A negative $\nu_{\mathfrak{p}}(f)$ then indicates that $\mathfrak{p}$ appears in at least one denominator of the coefficients of the polynomial f.

As a special case of Corollary 8.18, we obtain:

Theorem 1.2. *Let D be a non-Pellian polynomial of degree 4, and let $\mathfrak{p}$ an odd prime with $D_{\mathfrak{p}}$ square-free and the class of $(O_+) - (O_-)$ of* even *torsion order m in the (finite) Jacobian of the elliptic curve $Y^2 = D_{\mathfrak{p}}(X)$. Then*

$$\begin{aligned}(-1)^n \nu_{\mathfrak{p}}(a_n) &\geq 2\left\lfloor (n-1)/m \right\rfloor_{\mathbb{Z}} + 2\left\lfloor (n+1)/m \right\rfloor_{\mathbb{Z}}, \\ (-1)^n \nu_{\mathfrak{p}}(q_n) &\geq 2\left\lfloor (n+1)/m \right\rfloor_{\mathbb{Z}},\end{aligned}$$

where $\lfloor \cdot \rfloor_{\mathbb{Z}}$ denotes the floor function. In particular, the Gauss norms of the partial quotients and the convergents are unbounded both from above and below.

In the case of *odd* torsion order m, the negative valuations are possibly cancelled out by positive valuations coming from phenomena as in (1.1); this currently prevents any similar prediction (see Example 5, in particular table 10.2). Moreover, the precise growth of these Gauss norms is not understood at all right now. This is an even bigger issue for $\deg D > 4$, where we have to keep track of further unknowns. This makes an exact estimation of the valuations for higher degrees much more difficult.

The Gauss norms are also related to the height of polynomials. However, we have no information on the archimedean place and the 2-adic valuations, so we have to be careful if we want to compare with known results about the height of the convergents (see Section 9.3).

Indeed, the convergents (p_n, q_n) are also Padé approximations of $\sqrt{D}$, i.e. they satisfy

$$\operatorname{ord}_\infty(p_n - \sqrt{D}\, q_n) > \deg q_n, \tag{1.2}$$

where $\operatorname{ord}_\infty$ is the non-archimedean valuation with $\operatorname{ord}_\infty X = -1$ and which makes $\mathbb{Q}((X^{-1}))$ the completion of $\mathbb{Q}(X)$. In other words $p_n - \sqrt{D}\, q_n$ has a zero of high order at infinity.

Then by a general result of Bombieri and Paula Cohen [BC97] on the height of Padé approximations, it follows in the non-periodic case that the logarithmic projective height of the convergents grows quadratically in n. In this thesis, we have worked out the details of a simpler proof for the hyperelliptic case suggested by Zannier, see Theorem 9.3 and Theorem 9.5 for lower respectively upper bounds. This leads to upper bounds for the projective height of the partial quotients as well (see Corollary 9.12). The corresponding lower bounds for the height of the partial quotients require different arguments, see [Zan16].

The main approach to prove results like Theorem 1.1 and 1.2 is to study reduction of continued fractions modulo primes. This is interesting in itself, as it gives an example of a map between two "spaces" of continued fractions. Chapter 7 contains a general exposition of reduction of continued fractions, using the theory of discrete valuation rings.

The idea is to compare the continued fractions of $\sqrt{D}$ and $\sqrt{D_{\mathfrak{p}}}$. Their partial quotients are contained in $\mathbb{Q}[X]$ respectively in $\mathbb{F}_p[X]$. A naive approach would be to try to reduce the partial quotients, but this does not capture the structure of the continued fraction sufficiently. Instead we have to try to reduce the complete quotients $\alpha_n = [a_n, a_{n+1}, \dots]$ of $\sqrt{D}$ which are Laurent series in X^{-1} over $\mathbb{Q}$.

We say that a continued fraction has *good reduction in* $\mathfrak{p}$ if we can reduce the complete quotients of $\sqrt{D}$ and obtain exactly the complete quotients of $\sqrt{D_{\mathfrak{p}}}$. If this fails, we speak of *bad reduction of the continued fraction.* The latter is the usual situation for non-Pellian D over $\mathbb{Q}$ – and this is a key ingredient for the proof of Theorem 1.1. Other equivalent characterisations for good reduction of the continued fraction are given in Theorem 7.1. Note that this notion of good or bad reduction for the continued fraction of $\sqrt{D}$ is very different from the good or bad reduction of the corresponding (hyper)elliptic curve.

If the continued fraction of $\sqrt{D}$ is periodic, it trivially has good reduction at almost all primes $\mathfrak{p}$. This implies that the period length of the continued fraction of $\sqrt{D_{\mathfrak{p}}}$ is essentially independent of $\mathfrak{p}$. This can also be stated and deduced directly in terms of reducing minimal solutions of the polynomial Pell equation, and has recently been used by Platonov [Pla14], also together with Benyash-Krivets [BKP07] and Petrunin [PP12], to construct hyperelliptic curves over $\mathbb{Q}$ of genus 2, where the Jacobian contains a torsion point of a specific order. These examples are relevant for the uniform boundedness conjecture for torsion points of abelian varieties.

Van der Poorten's approach to reduction of continued fractions deals primarily with reduction of the convergents: the inequality (1.2) essentially characterises the convergents up to a common factor of small degree, constant if p and q are coprime (see Corollary 5.20). If we normalise p_n and q_n correctly, their reduction modulo $\mathfrak{p}$ remains a convergent of $\sqrt{D_{\mathfrak{p}}}$. Moreover, the following theorem holds (both for Pellian and non-Pellian D):

Theorem 1.3 (van der Poorten)**.** *If the prime $\mathfrak{p}$ does not appear in a denominator in D, then the reductions modulo $\mathfrak{p}$ of the normalised convergents $(\widetilde{p_n}, \widetilde{q_n})$ of $\sqrt{D}$ yield* all *the convergents of $\sqrt{D_{\mathfrak{p}}}$.*

Unfortunately, the proofs given by van der Poorten (there are slightly different versions in [vdP98], [vdP99] and [vdP01]) do not appear to be complete. So one of the main goals of Chapter 7 is to give a more precise statement and a rigorous proof of van der Poorten's result (as in Theorem 7.2).

As might be expected, the reduction of the convergents is strongly related to the reduction of the continued fraction. For example, the bad reduction of the continued fraction is caused by two (or more) convergents of $\sqrt{D}$ reducing to the same convergent modulo $\mathfrak{p}$ – see Proposition 7.32 and example 6, in particular table 10.3.

Finally, we remark that periodicity of the continued fraction of $\sqrt{D}$ is equivalent to $\deg a_n = d$ for at least one $n \geq 1$, where $2d = \deg D$ (see Corollary 6.1). Bad reduction of the continued fraction is also determined by how these degrees increase under reduction (see the discussion in Section 7.4.2) which connects periodicity of the continued fraction of $\sqrt{D_\mathfrak{p}}$ and occurrences of $\mathfrak{p}$ in the denominators. The interplay with the normalisation factors of the canonical convergents then allows us to exclude issues related to (1.1), and leads to a proof of Theorem 1.1.

On specialization

The reduction theory for continued fractions of Chapter 7 applies also to specialization. Instead of reducing $D \in \mathbb{Q}[X]$ modulo a prime, we take for example $D \in \mathbb{C}(t)[X]$, and try to specialize t to some $t_0 \in \mathbb{C}$. Searching for the values t_0 of t that specialize to a periodic continued fraction of $\sqrt{D_{t=t_0}}$ corresponds to a special case of the relative Manin-Mumford conjecture, which in turn is a consequence of Pink's conjecture. Recall that periodicity is equivalent to the class of $(O_+) - (O_-)$ being torsion in the Jacobian of the curve $Y^2 = D(X)$.

The periodicity of the reduction of the continued fraction was a crucial ingredient for the proof of Theorem 1.1. It is therefore natural to ask for specialization analogues of this theorem. The answer depends on the geometry:

For example Masser and Zannier showed that for $D = X^6 + X + t$, the continued fraction of $\sqrt{D}$ is non-periodic, the Jacobian of the curve $Y^2 = D(X)$ is simple and there are only finitely many $t_0 \in \mathbb{C}$ such that $\sqrt{D_{t=t_0}}$ has a periodic continued fraction (see [MZ15], here we have reformulated the results in the language of continued fractions). For these t_0, all of them algebraic numbers, we can reuse the arguments from Theorem 1.1 and show that $t - t_0$ appears in infinitely many a_n of the generic continued fraction as a denominator of a coefficient.

However, from the results of Masser and Zannier follows also that there are infinitely many $t_1 \in \overline{\mathbb{Q}}$ for which $t - t_1$ appears at least once as a denominator of a coefficient of some a_n. They might appear infinitely often, but we will show that this can happen only for the trivial reason that we excluded in Theorem 1.1. More precisely we can find $e \in \mathbb{Z}$ (perhaps not effectively), such that in

$$(t - t_1)^e \sqrt{D} = [b_0, b_1, b_2, \dots], \qquad b_n \in \mathbb{C}(t)[X]$$

the "prime" $(t - t_1)$ appears only in finitely many b_n as a denominator. We will discuss this in more detail in Section 8.6.2.

1.1. Acknowledgements

First and foremost, I would like to thank my supervisor Prof. Umberto Zannier, for pointing me to interesting mathematical problems and sharing his mathematical insight. You have helped me to see number theory in a new light, and improved my understanding of various problems. This thesis would not exist without his input and support. Thank you for answering my many questions and teaching me not to give up and to be independent. I am indebted to you and Scuola Normale Superiore for offering me the chance to pursue a Perfezionamento.

I also would like to dearly thank Prof. David Masser for introducing me to the polynomial Pell equation, and sending me towards Pisa in the first place. I thank Prof. Vistoli for teaching me some algebraic geometry.

A very big “thank you” goes to Lars, for many discussions about mathematics and other more trivial topics, putting up with me as a flatmate, and actually reading a draft of this thesis.

Big thanks also to Laura, Fabrizio, Michele and Soli, for countless lunches, game nights and for working together. Thanks for all your help, and for listening to me, even if I made rather less sense. Special thanks to Laura for helping me from my first day in Italy, and to Michele for participating in many sometimes crazy activities.

I would like to thank Francesca for working together, and being a very diligent mathematician.

I wish to thank all the wonderful and interesting people I met at Scuola Normale Superiore, for silly and serious conversations and reminding me that there are people in this world. Many of you I consider now my friends.

Let me thank in particular Josefine for showing me Florence and the beach, Alex for early morning runs and literally talking to everybody, Sara for teaching me about real friendship, Alexey for extraordinary observations and highly entertaining discussions, Clélie for not being afraid to talk of anything, Błażej for making me a better table tennis player, and Giacomo for his delicious chinese cooking and strange questions.

Thanks to Mario and Simone for explaining Italy, and Michele (the other one) for explaining biology with a passion. Thanks to Ilir, Marcello and Renata for being loyal hikers, to Umesh for playing table-tennis, to Adam for trying to take silly things seriously,to Elisa and Henry for chatting about fotography and to François and Max for reminding me that I am german.

Thanks also to all the people I spent time with at conferences, for interesting discussions and experiences from other places. Harry and Jung-Kyu, thanks for inviting me to visit the math department of Basel every once in a while.

To Aki, even if we have never met in real life, thank you for the countless hours in the skys of Georgia, Nevada and elsewhere, and in the woods of Chernarus, and for sharing your knowledge of aviation.

Finally, I want to thank my parents, my brother Sven and my sister Heike, for your support (logistical and otherwise) and for *always* believing that I could complete my PhD. It looks like you were right in the end.

1.2. Notation reference

Symbol	Description
$\mathbb{N}$	natural numbers: $\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	natural numbers with 0: $\{0, 1, 2, 3, \dots\}$
$\mathbb{K}$	field of characteristic $\neq 2$
$\mathsf{Fr}(R)$	fraction field of integral domain R
$\mathbb{K}(\!(X^{-1})\!)$	Laurent series in X^{-1} with coefficients in $\mathbb{K}$
$\operatorname{ord}(f)$	zero-order at $X = \infty$, sometimes denoted $\operatorname{ord}_\infty$
$\ell c(f)$	leading coefficient of polynomial or Laurent series
$\mathcal{Q}(\mathbb{K})$	$\{(p, q) \in \mathbb{K}[X]^2 \mid q \neq 0\}$
$\mathcal{C}_\alpha(\mathbb{K})$	set of convergents of α
$\mathcal{B}_\alpha(\mathbb{K})$	set of best-approximations of α
D, d	polynomial, with $\deg D = 2d$ and $\ell c(D)$ a square
$\mathcal{P}(D)$	solutions of polynomial Pell equation (2.1)
$\mathcal{P}^\times(D)$	solutions of unit-norm equation (2.2)
σ	involution $\sqrt{D} \to -\sqrt{D}$
O, $\mathfrak{m} = (\pi)$	discrete valuation ring and maximal ideal with uniformiser
K, k	fraction field and residue field of O, of characteristic $\neq 2$
ν	valuation, usually of O
$K(\!(X^{-1})\!)_\nu$	Laurent series with coefficient valuations bounded from below
$\widetilde{x}$	normalisation of $x \in K(\!(X^{-1})\!)_\nu$ to valuation $\nu(\widetilde{x}) = 0$
$\overline{x} = \rho\,(x)$	reduction/specialization of $x \in O(\!(X^{-1})\!)$
$\widehat{x} = \rho\,(\widetilde{x})$	reduction of normalisation
$\mathfrak{p}$	prime *number* $\mathfrak{p}$ (positive integer)
$\mathfrak{P}$	prime *ideal* $\mathfrak{P}$ (usually over $\mathfrak{p}$)
$\mathbf{CF}(\alpha)$	continued fraction of α
a_n	partial quotient of α
α_n	complete quotient of α
(p_n, q_n)	canonical convergent of α
g_n	normalisation factor of canonical convergent, $\nu(g_n) = \nu(q_n)$.
ϑ_n	$p_n - \alpha\, q_n$ normalised to $\nu(\vartheta_n) = 0$
c_n	partial quotient of $\gamma = \overline{\alpha}$
γ_n	complete quotient of γ
(u_n, v_n)	canonical convergent of γ
h_n	correction factor (in $k[X]$) for reduced convergents
$\lambda : \mathbb{N}_0 \to \mathbb{N}_0$	$(\widehat{p_n}, \widehat{q_n}) = h_n \cdot (u_{\lambda(n)}, v_{\lambda(n)})$, see Corollary 7.27
(P)	point as divisor
$[P]$	divisor class of point
$\mathbf{D}$	divisor (bold)
$[\mathbf{D}]$	divisor class
$\mathcal{C}_{\text{aff}}, \mathcal{C}$	smooth affine and projective models of $Y^2 = D(X)$
$O_\pm$, $\mathbf{O}$	the two points of $\mathcal{C}$ at infinity; $\mathbf{O} = (O_+) - (O_-)$
σ	conjugation of points, $Y \to -Y$

2. Pell equation

We begin by exploring some well-known basic properties of the Pell equation over polynomials, usually called the *polynomial Pell equation.* We also explain how to write square roots of polynomials in X as Laurent series in X^{-1}, and use this to show that the group of solutions of the polynomial Pell equation has rank at most 1.

Given a base field $\mathbb{K}$ with $\operatorname{char}\mathbb{K} \neq 2$, let $D \in \mathbb{K}[X]$ a *non-constant* polynomial consider the *polynomial Pell equation*

$$p^2 - D\,q^2 = 1. \tag{2.1}$$

Clearly, there *always* exist the trivial solutions $(p, q) = (\pm 1, 0)$, so naturally we ask if there exist other solutions $(p, q) \in \mathbb{K}[X]^2$ with $q \neq 0$, which we call the *non-trivial solutions.* If this is the case, we say D is *Pellian.* If $\mathbb{K}$ is finite, one may show as for the classical Pell equation over $\mathbb{Z}$ that D is always Pellian. If $\mathbb{K}$ is infinite, it is unlikely that D is Pellian – because D Pellian is equivalent to a torsion condition on a point in the Jacobian of a (hyper)elliptic curve, see Chapter 4 for details.

Proposition 2.1. *Suppose D is Pellian. Then $\deg D$ must be even, and the leading coefficient $\ell c(D)$ is a square in $\mathbb{K}$. However D cannot be a square in $\mathbb{K}[X]$.*

Proof. By the hypotheses D non-constant and $q \neq 0$, we have $\deg(D\,q^2) > 0$. Then p^2 must cancel out the non-constant terms, hence $\deg p^2 = \deg(D\,q^2)$ which implies $\deg D = 2(\deg p - \deg q)$ and that $\ell c(D) = \ell c(p)^2/\ell c(q)^2$ is a square.

Finally, we show that D is not a square in $\mathbb{K}[X]$: It is obvious that for $D = 1$, i.e. $p^2 - q^2 = (p - q)(p + q) = 1$ there are only constant solutions because $\mathbb{K}[X]^\times = \mathbb{K}^\times$. So if $D = E^2$ with $E \in \mathbb{K}[X] \setminus \mathbb{K}$, then for any solution (p, q) we must have $p, E\,q$ constant which implies $q = 0$. □

So these three conditions are necessary (but not sufficient) for the existence of non-trivial solutions. The situation in characteristic 2 is however completely different, see Section A.1 in the Appendix.

2.1. Multiplication law and unit-norm equation

We assume from now on that D has even degree $2d$, is not a square, but $\ell c(D)$ is a square in K (for example 1 if D is monic).

The set of solutions $\mathcal{P}(D)$ (including trivial solutions) of (2.1) carries an abelian group structure[1] via the multiplication

$$(p,q) * (p',q') = (p\,p' + D\,q\,q', p\,q' + p'\,q)$$

which comes from the map

$$\mathcal{P}(D) \longrightarrow \mathbb{K}[X,\sqrt{D}]^{\times}, \qquad (p,q) \mapsto p + q\sqrt{D}$$

which is an (injective) group homomorphism (see Section 2.2 below).

Note that $(p,q) * (p,-q) = (p^2 - D\,q^2, 0) = (1,0)$ for any Pell solution, so $(1,0)$ is the neutral element, and $(p,-q)$ is the inverse of (p,q).

Actually, we will not really work with (2.1). To study the structure of the solution set, it is far more convenient to relax to the unit-norm equation (see [HMPLR87] for a general treatment)

$$p^2 - D\,q^2 = \omega \in \mathbb{K}^{\times} \tag{2.2}$$

where ω is an arbitrary unit of $\mathbb{K}$. Clearly, any Pell solution satisfies also this equation. The converse does of course not hold, but from a non-trivial solution of (2.2) we can recover a non-trivial solutions of (2.1):

Proposition 2.2. *Suppose* (2.2) *has a non-trivial solution* $(p,q) \in \mathbb{K}[X]^2$ *(with* $q \neq 0$*). Then* D *is Pellian.*

Proof. The multiplication law from above generalises to (2.2), with $(p,q) * (p,-q) = (\omega, 0)$, hence

$$(p,q) * (p,q) * (p,-q) * (p,-q) = (\omega^2, 0).$$

Set

$$(p',q') = (\omega^{-1}, 0) * (p,q) * (p,q) = \omega^{-1} \cdot (p^2 + D\,q^2, 2\,p\,q),$$

so that (p',q') remains in $\mathbb{K}[X]$ and is clearly a solution of (2.1). As observed in the proof of Proposition 2.1, $q \neq 0$ implies $p \neq 0$, hence $\omega^{-1}\,2\,p\,q \neq 0$, so (p',q') is a non-trivial Pell solution. □

From now on, we refer also to (2.2) as the *Pell equation*, and mostly forget about (2.1). We denote by $\mathcal{P}^{\times}(D)$ the set of all solutions of (2.2). We will see that for the purposes of this thesis, it is more natural to work with the unit-norm equation.

2.2. Units of hyperelliptic function fields

The quadratic field extension $\mathbb{K}(X,\sqrt{D})$ of $\mathbb{K}(X)$ is called a hyperelliptic function field – specifically it is the function field of the hyperelliptic curve $\mathcal{C}_{\text{aff}} : Y^2 = D(X)$ which

[1] This group is a twisted $\mathbb{G}_m$. We can see $D(X)\,Q^2 = P^2 - 1$ as a twist of $Q^2 = P^2 - 1$ by the (hyper)elliptic curve $Y^2 = D(X)$, via $(P,Q) \mapsto (P, Y\,Q)$. Of course $Q^2 = P^2 - 1$ written as $P^2 - Q^2 = 1$ is isomorphic to $\mathbb{G}_m$. See [Haz97] for more details.

we will study in more detail in Chapter 4. The subring $\mathbb{K}[X, \sqrt{D}]$ of $\mathbb{K}(X, \sqrt{D})$ is the integral closure of $\mathbb{K}[X]$, describing the regular functions on the affine curve. For now, we show that the units of $\mathbb{K}[X, \sqrt{D}]$ correspond to solutions of the Pell equation (2.2). See also [HMPLR87] for generalisations to other algebraic functions.

Theorem 2.1. *The map*

$$\pi : \mathcal{P}^\times(D) \longrightarrow \mathbb{K}[X, \sqrt{D}]^\times, \quad (p, q) \mapsto p + q\sqrt{D}$$

is bijective, and via the multiplication $*$ on $\mathcal{P}^\times(D)$ gives an isomorphism of abelian groups.

Observe that there is a single non-trivial $\mathbb{K}(X)$-automorphism σ of $\mathbb{K}(X, \sqrt{D})$, defined by $\sigma(\sqrt{D}) = -\sqrt{D}$.

Proof. Actually, we defined $*$ as the pullback under π of the multiplication on $\mathbb{K}[X, \sqrt{D}]$, so clearly

$$\pi(\phi * \psi) = \pi(\phi) * \pi(\psi) \text{ for all } \phi, \psi \in \mathcal{P}^\times(D).$$

And by the identity

$$(p + q\sqrt{D})(p - q\sqrt{D}) = p^2 - D\,q^2 = \omega \in \mathbb{K}^\times$$

it follows that $\operatorname{im}\pi \subset \mathbb{K}[X, \sqrt{D}]^\times$, so π is well defined.

Recall that we assume that D is not a square, so the ring $\mathbb{K}[X, \sqrt{D}]$ is a free rank 2 module over $\mathbb{K}[X]$ with basis $(1, \sqrt{D})$: this implies that π is injective.

It remains to check that π is also surjective: Let $\phi = p + q\sqrt{D} \in \mathbb{K}[X, \sqrt{D}]^\times$ with $p, q \in \mathbb{K}[X]$. Then we have

$$\phi \cdot \sigma(\phi) = (p + q\sqrt{D})(p - q\sqrt{D}) = p^2 - D\,q^2 \in \mathbb{K}[X]$$

Applying the same argument to the inverse $1/\phi$, we find $p^2 - D\,q^2 \in \mathbb{K}[X]^\times = \mathbb{K}^\times$, so (p, q) is a solution of (2.2). This proves that π is surjective. □

Remark 2.3. Observe that the trivial solutions of (2.2) correspond precisely to the elements of $\mathbb{K}^\times$.

2.3. Laurent series and valuation

Define the field of Laurent series over $\mathbb{K}$

$$\mathbb{K}(\!(X^{-1})\!) = \left\{ \sum_{n=-\infty}^{N} t_n X^n \;\middle|\; N \in \mathbb{Z}, t_n \in \mathbb{K} \right\}.$$

It contains $\mathbb{K}[X]$ and its fraction field $\mathbb{K}(X)$. Note that $\mathbb{K}(\!(X^{-1})\!)$ is the completion of $\mathbb{K}(X)$ with respect to the discrete valuation $\operatorname{ord} = \operatorname{ord}_\infty$ (the zero-order at infinity), defined by

$$\operatorname{ord}(f) = \operatorname{ord}_\infty(f) = -N \text{ where } f = \sum_{n=-\infty}^{N} t_n X^n,\; f_N \neq 0.$$

Remark 2.4. For example if $f \in \mathbb{K}[X]$, then $\operatorname{ord}(f) = -\deg f$. Moreover,

$$\operatorname{ord}(f) > 0, \quad f \in \mathbb{K}[X] \text{ implies } f = 0. \tag{2.3}$$

There is a truncation operation which takes a Laurent series and returns a polynomial, essential for the continued fraction process:

Definition 2.5. For $\alpha = \sum_{n=-\infty}^{N} t_n X^n \in \mathbb{K}(\!(X^{-1})\!)$, we define the *truncation* (or *principal part*)

$$\lfloor \alpha \rfloor = \begin{cases} 0 & \text{if } \operatorname{ord}(\alpha) > 0, \text{ i.e. } N < 0 \\ t_N X^N + \cdots + t_0 & \text{if } \operatorname{ord}(\alpha) \leq 0, \text{ i.e. } N \geq 0 \end{cases}$$

as the *polynomial part* of α.

Remark 2.6. We could also define $\lfloor \alpha \rfloor$ as the *unique* $a \in \mathbb{K}[X]$ satisfying $\operatorname{ord}(\alpha - a) > 0$ – unicity is a consequence of Remark 2.4.

Remark 2.7. The preceding remark implies for $\alpha, \beta \in \mathbb{K}(\!(X^{-1})\!)$ that $\lfloor \alpha + \beta \rfloor = \lfloor \alpha \rfloor + \lfloor \beta \rfloor$.

Remark 2.8. Recall that $\mathbb{K}[X]$ is Euclidean with respect to deg. So for $p, q \in \mathbb{K}[X]$ with $q \neq 0$ there exist $a, r \in \mathbb{K}[X]$ satisfying $p = a\,q + r$ and $\deg r < \deg q$. Then

$$\frac{p}{q} - a = \frac{r}{q} \text{ with } \operatorname{ord}(r/q) > 0$$

implies $\lfloor p/q \rfloor = a$, and moreover a, r are uniquely determined, again by Remark 2.6.

We now explain how to compute $\sqrt{D}$ as a Laurent series in X^{-1}:

Proposition 2.9. *Let $D \in \mathbb{K}[X]$ with $\deg D = 2d$ and $\ell c(D) \in \mathbb{K}$ a square. Then $\sqrt{D} \in \mathbb{K}(\!(X^{-1})\!)$, so D is a square in $\mathbb{K}(\!(X^{-1})\!)$.*

Proof. Let $D = d_{2d} X^{2d} + \cdots + d_0$, where d_{2d} is a square in $\mathbb{K}$. Hence we may reduce to the case $d_{2d} = 1$, and write

$$D = X^{2d}\,(1 + f(X)) \text{ where } f(X) = d_{2d-1} X^{-1} + \cdots + d_0 X^{-2d}.$$

Of course X^{2d} is a square in $\mathbb{K}(\!(X^{-1})\!)$, and because $\operatorname{ord}(f(X)) > 0$, we find that

$$\sqrt{1 + f(X)} = \sum_{n=0}^{\infty} \binom{1/2}{n} f(X)^n$$

converges in $\mathbb{K}(\!(X^{-1})\!)$, so also $(1 + f(X))$ is a square. □

Definition 2.10. We choose once and for all one square root of D, and denote it by $\sqrt{D}$. We also set $A = \left\lfloor \sqrt{D} \right\rfloor$. For example, if D is monic of degree $2d$, then we choose $\sqrt{D} = X^d + \ldots$.

Proposition 2.11. *We have $\deg A = \frac{1}{2} \deg D$, and $\deg(D - A^2) < \deg A$.*

Proof. As ord is a valuation, clearly $-\deg D = \operatorname{ord} D = 2 \operatorname{ord} \sqrt{D} < 0$, hence $-\deg A = \operatorname{ord} A = \operatorname{ord} \left\lfloor \sqrt{D} \right\rfloor = \operatorname{ord} \sqrt{D}$ which implies the first claim.

Moreover, we can write

$$\sqrt{D} = A + \varepsilon \text{ with } \varepsilon \in \mathbb{K}(\!(X^{-1})\!) \text{ and } \operatorname{ord}(\varepsilon) > 0. \tag{2.4}$$

So

$$D = A^2 + 2\, A\, \varepsilon + \varepsilon^2$$

where of course

$$\operatorname{ord}\bigl(2\, A\, \varepsilon + \varepsilon^2\bigr) = \min(\operatorname{ord}(A)\, , \operatorname{ord}(\varepsilon)) + \operatorname{ord}(\varepsilon) = \operatorname{ord}(A) + \operatorname{ord}(\varepsilon) > \operatorname{ord}(A)$$

implies the second claim. □

We can rephrase this as

Lemma 2.12 (Completion of the square)**.** *There exist* $A, \Omega \in \mathbb{K}[X]$ *with* $\deg \Omega < \deg A = \frac{1}{2} \deg D$ *satisfying*

$$D = A^2 + \Omega$$

where A is unique up to a factor -1.

Remark 2.13. Note that the lemma also holds if D is a square.

Remark 2.14. If $\deg \Omega = 0$, then clearly $(A, 1)$ is a solution of the Pell equation (2.2).

2.4. Group structure of Pell solutions

We apply the definitions of the previous section directly to study the structure of the Pell solutions. The group of solutions of (2.2) is essentially cyclic:

Proposition 2.15. *If D is not Pellian, then $\mathcal{P}(D) = \{\pm 1\}$ and $\mathcal{P}^{\times}(D) = K^{\times}$. But if D is Pellian, then*

$$\mathcal{P}(D) \simeq \{\pm 1\} \times \mathbb{Z} \qquad \textit{and} \qquad \mathcal{P}^{\times}(D) \simeq K^{\times} \times \mathbb{Z}.$$

Proof. We use that $\mathcal{P}^{\times}(D) \simeq \mathbb{K}[X, \sqrt{D}]^{\times}$. By Proposition 2.9, we can embed $\mathbb{K}[X, \sqrt{D}]$ into $\mathbb{K}(\!(X^{-1})\!)$, and define

$$o(p, q) = \operatorname{ord}\Bigl(p + \sqrt{D}\, q\Bigr) \text{ for } (p, q) \in \mathcal{P}^{\times}(D).$$

This defines a group homomorphism $o : \mathcal{P}^{\times}(D) \to \mathbb{Z}$. The kernel is made precisely of the trivial solutions:

$$\operatorname{ord}(p) = \operatorname{ord}\Bigl(p + \sqrt{D} + p - \sqrt{D}\Bigr) \geq \min\Bigl(\operatorname{ord}\Bigl(p + \sqrt{D}\, q\Bigr)\, , \operatorname{ord}\Bigl(p - \sqrt{D}\Bigr)\Bigr)$$

and

$$\operatorname{ord}\Bigl(p+\sqrt{D}\,q\Bigr)+\operatorname{ord}\Bigl(p-\sqrt{D}\Bigr)=0$$

so $\operatorname{ord}\Bigl(p+\sqrt{D}\Bigr)=0$ implies $\deg p=-\operatorname{ord}(p)\le 0$, hence $q=0$.

If D is not Pellian, then the image of o is 0. But if D is Pellian, then the image of o is isomorphic to $\mathbb{Z}$.

We can of course restrict o to $\mathcal{P}(D)$, and then the kernel becomes $\{(\pm 1,0)\}\simeq\{\pm 1\}$.

The structure of $\mathcal{P}(D)$ and $\mathcal{P}^{\times}(D)$ now follows from standard theorems about group homomorphisms. □

We conclude our discussion of the polynomial Pell equation with the following observation:

Corollary 2.16. *If* $\deg D=2$ *and the leading coefficient* $\ell c(D)$ *is a square, then* D *is* always *Pellian (unless it is square).*

Proof. By Lemma 2.12, in this case $\deg\Omega<\deg A=1$ so forcefully $\deg\Omega=0$, and Remark 2.14 says that $(A,1)$ is a Pell solution. □

3. Rational approximations

As mentioned before, for the polynomial Pell equation the existence of non-trivial solutions is not guaranteed. But one observes that the Pell solutions produce very good rational approximations for $\sqrt{D}$ (as in the numerical case). This chapter introduces two notions of rational approximation: convergents and best-approximations. We will study in this chapter how they are related to each other and to the non-trivial Pell solutions. Their complete classification is however best understood with the help of continued fractions, to be discussed later in Section 5.5.

For our purposes, it is convenient to keep track of common factors in the numerator and denominator of the rational approximation. *Instead* of $\mathbb{K}(X)$, our candidate set for rational approximations is the set of tuples representing quotients

$$\mathcal{Q}(\mathbb{K}) = \{(p, q) \in \mathbb{K}[X]^2 \mid q \neq 0\}.$$

We loosely refer to p as the *numerator* and to q as the *denominator*, in spirit of the obvious map $\mathcal{Q}(\mathbb{K}) \longrightarrow \mathbb{K}(X),\ (p, q) \mapsto p/q$.

For $r, p, q \in \mathbb{K}[X]$ with $r, q \neq 0$ we also write

$$r \cdot (p, q) = (r\,p, r\,q).$$

With this terminology established, we can begin the study of different types of approximations. Of course, we are using the valuation ord introduced in Section 2.3 to measure how well we can approximate any Laurent series in $\mathbb{K}(\!(X^{-1})\!)$.

3.1. Convergents

A classical type of rational approximation is given by the convergents. They arise very naturally from the continued fraction expansion – we will see details later in Chapter 5. For now, we give a different characterisation in the spirit of the famous Dirichlet Lemma. This definition also shows immediately that the convergents are a special case of Padé approximations.

Definition 3.1. Let $\alpha \in \mathbb{K}(\!(X^{-1})\!)$. A tuple $(p, q) \in \mathcal{Q}(\mathbb{K})$ is called a *convergent* of α over $\mathbb{K}[X]$ if it satisfies

$$\operatorname{ord}(p - \alpha\,q) > \deg q. \tag{3.1}$$

We denote the set of all convergents by $\mathcal{C}_\alpha(\mathbb{K})$.

Remark 3.2. It can easily be seen that convergents exist: The condition (3.1) is a linear condition on the coefficients of p and q. Clearly p removes the coefficients of X^n for $n \geq 0$ in $\alpha\, q$; then only the coefficients of $X^{-1}, \ldots, X^{-\deg q}$ need to vanish, which can be accomplished by choosing the $1 + \deg q$ coefficients of q appropriately. See Section 3.4 for more details.

Remark 3.3. Suppose $r, p, q \in \mathbb{K}[X]$. Then

$$r \cdot (p, q) \in \mathcal{C}_\alpha(\mathbb{K}) \implies (p, q) \in \mathcal{C}_\alpha(\mathbb{K})$$

because $0 \geq \operatorname{ord}(r)$ implies

$$\operatorname{ord}(p - \alpha\, q) \geq \operatorname{ord}(r\, p - \alpha\, r\, q) > \deg(r\, q) \geq \deg q.$$

Note that the implication in the converse direction does not hold in general because multiplication with r decreases ord and increases deg.

In principle, one could for any convergent (p, q) assume that p and q are coprime, and identify it with the fraction. This might improve the approximation quality, however it turns out that the common factors help to understand the reduction of convergents modulo a prime better (to be discussed in Chapter 7).

Anyway the common factor usually has a small and controllable degree:

Proposition 3.4. *Let $(p, q) \in \mathcal{Q}(\mathbb{K})$ and $r \in \mathbb{K}[X] \setminus \{0\}$. Suppose*

$$\operatorname{ord}(p - \alpha\, q) = \xi + \deg q.$$

Then $r \cdot (p, q) \in \mathcal{C}_\alpha(\mathbb{K})$ is a convergent if and only if $\deg r < \xi/2$.

In particular, suppose $r' \in \mathbb{K}[X] \setminus \{0\}$ with $\deg r \leq \deg r'$. Then $r' \cdot (p, q) \in \mathcal{C}_\alpha(\mathbb{K})$ implies $r \cdot (p, q) \in \mathcal{C}_\alpha(\mathbb{K})$.

Remark 3.5. Note that the Proposition holds also when $\xi = \infty$ – but that happens only for $\alpha \in \mathbb{K}(X)$.

Proof. In order for $r \cdot (p, q)$ to be a convergent, the following expression must be positive:

$$\operatorname{ord}(r\, p - \alpha\, r\, q) - \deg\,(r\, q) = \operatorname{ord}(r) + \operatorname{ord}(p - \alpha\, q) - \deg r - \deg q = \xi - 2\, \deg r. \quad (3.2)$$

The second part of the Proposition follows immediately. □

Remark 3.6. The above (3.2) also suggests that for (p, q) coprime we have the optimal relative approximation quality: higher is better.

Remark 3.7. Let $\alpha \in \mathbb{K}((X^{-1}))$, set $a = \lfloor \alpha \rfloor$. Then $(a, 1) \in \mathcal{C}_\alpha(\mathbb{K})$ because $\operatorname{ord}(a - \alpha) > 0 = \deg 1$.

Proposition 3.8. *If $(p, q) \in \mathcal{C}_\alpha(\mathbb{K})$ is a convergent, then p is uniquely determined by q via $p = \lfloor \alpha\, q \rfloor$.*

Proof. This follows immediately from $\operatorname{ord}(p - \alpha\, q) > \deg q \geq 0$, and Remark 2.6 characterising $\lfloor \cdot \rfloor$. □

3.2. Pell solutions are convergents

Let us for a moment return to the polynomial Pell equation, and show that the non-trivial Pell solutions (up to conjugate) are convergents of $\sqrt{D}$. Obviously, not all convergents of $\sqrt{D}$ need to be Pell solutions.

Proposition 3.9. *Let $(p,q) \in \mathcal{Q}(\mathbb{K})$ and $p^2 - D\,q^2 = \Omega$. Then the inequality*

$$\deg \Omega < \tfrac{1}{2} \deg D \tag{3.3}$$

holds if and only if either $(p,q) \in \mathcal{C}_{\sqrt{D}}(\mathbb{K})$ or $(p,-q) \in \mathcal{C}_{\sqrt{D}}(\mathbb{K})$ is a convergent of $\sqrt{D}$.

In particular, if $(p,q) \in \mathcal{P}^{\times}(D)$ is a Pell solution with $q \neq 0$, then one of $(p,q), (p,-q)$ is a convergent of $\sqrt{D}$.

Proof. Let us begin with some observation useful to both directions of implication. Note that

$$\operatorname{ord}(\Omega) = \operatorname{ord}\bigl(p^2 - D\,q^2\bigr) = \operatorname{ord}\Bigl(p + \sqrt{D}\,q\Bigr) + \operatorname{ord}\Bigl(p - \sqrt{D}\,q\Bigr)\,. \tag{3.4}$$

And if $\operatorname{ord}\Bigl(p - \sqrt{D}\,q\Bigr) > 0$, the ultrametric inequality and $\operatorname{ord}\Bigl(\sqrt{D}\,q\Bigr) \leq 0$ imply

$$\operatorname{ord}\Bigl(p + \sqrt{D}\,q\Bigr) = \min\Bigl(\operatorname{ord}\Bigl(2\sqrt{D}\,q\Bigr), \operatorname{ord}\Bigl(p - \sqrt{D}\,q\Bigr)\Bigr) = \operatorname{ord}\Bigl(\sqrt{D}\,q\Bigr) < 0. \tag{3.5}$$

Now assume that $(p,q) \in \mathcal{C}_{\sqrt{D}}(\mathbb{K})$ is a convergent, hence $\operatorname{ord}\Bigl(p - \sqrt{D}\,q\Bigr) > \deg q \geq 0$. Then (3.4) and (3.5) yield

$$\operatorname{ord}(\Omega) > \deg q + \operatorname{ord}\Bigl(\sqrt{D}\,q\Bigr) = \operatorname{ord}\Bigl(\sqrt{D}\Bigr)$$

which implies $\deg \Omega < \frac{1}{2} \deg D$.

For the other direction, assume that (p,q) satisfies (3.3), hence $\operatorname{ord}(\Omega) > \operatorname{ord}\Bigl(\sqrt{D}\Bigr) \geq \operatorname{ord}\Bigl(\sqrt{D}\,q\Bigr)$. Without loss of generality, we may further assume $\operatorname{ord}\Bigl(p - \sqrt{D}\,q\Bigr) \geq \operatorname{ord}\Bigl(p + \sqrt{D}\,q\Bigr)$. It follows

$$\operatorname{ord}(\Omega) > \operatorname{ord}\Bigl(2\sqrt{D}\,q\Bigr) = \operatorname{ord}\Bigl(p + \sqrt{D}\,q - (p - \sqrt{D}\,q)\Bigr) \geq \operatorname{ord}\Bigl(p + \sqrt{D}\,q\Bigr)$$

so by (3.4) $\operatorname{ord}\Bigl(p - \sqrt{D}\,q\Bigr) > 0$, which in turn implies (3.5). Using (3.4) again, we arrive at

$$\begin{aligned}\operatorname{ord}\Bigl(p - \sqrt{D}\,q\Bigr) &= \operatorname{ord}(\Omega) - \operatorname{ord}\Bigl(p + \sqrt{D}\,q\Bigr)\\ &= \operatorname{ord}(\Omega) - \operatorname{ord}\Bigl(\sqrt{D}\,q\Bigr) > -\operatorname{ord}(q) = \deg q\end{aligned}$$

as desired. □

3.3. The universal property of best-approximation

The convergents have a useful universal property: they are in some sense the optimal approximations that we can find. For a discussion on where this particular universal property comes from, see [Khi56]. See also [Cas57] where the continued fraction process for real numbers is defined using best-approximations.[1]

As we did with the convergents, we modify our definition so that it allows common factors; and we prefer a category theoretic style of universal property.

Definition 3.10. Let $\alpha \in \mathbb{K}(\!(X^{-1})\!)$. A tuple $(p,q) \in \mathcal{Q}(\mathbb{K})$ is called a *best-approximation* (of second type) in $\mathbb{K}[X]$, if for every other tuple $(p',q') \in \mathcal{Q}(\mathbb{K})$ satisfying

$$\operatorname{ord}\bigl(p' - \alpha\, q'\bigr) \geq \operatorname{ord}(p - \alpha\, q) \text{ and } \deg q' \leq \deg q \tag{3.6}$$

we have $p'/q' = p/q$.

We denote by $\mathcal{B}_\alpha(\mathbb{K})$ the set of all best-approximations of α.

Remark 3.11. If $(p,q) \in \mathcal{Q}(\mathbb{K})$ and $r, r' \in \mathbb{K}[X] \setminus \{0\}$ with $\deg r' \leq \deg r$ (for example $r' = 1$), then

$$r \cdot (p,q) \in \mathcal{B}_\alpha(\mathbb{K}) \implies r' \cdot (p,q) \in \mathcal{B}_\alpha(\mathbb{K}).$$

because

$$\operatorname{ord}\bigl(r'\, p - \alpha\, r'\, q\bigr) \geq \operatorname{ord}(r\, p - \alpha\, r\, q) \text{ and } \deg\bigl(r'\, q\bigr) \leq \deg\,(r\, q)\,.$$

So without loss of generality, one *could* assume that for a best-approximation (p,q), we have p and q coprime. This could also be enforced by changing the phrasing of the definition slightly, as is in fact usually done in the literature. However, in that case, (3.6) becomes harder to satisfy because removing a common (non-constant) factor decreases $\deg q$ and increases $\operatorname{ord}(p - \alpha\, q)$.

As alluded to before, when studying the reduction of convergents modulo a prime, it is useful to allow common factors. The notion of best-approximation presented here gives even more freedom for such common factors than our notion of convergent. We can indeed find best-approximations (p,q) for arbitrary $\deg q$, which may not be possible with convergents (see Section 5.5). This simplifies their classification, and hence the classification of convergents.

Before we investigate the relation between convergents and best-approximations, let us show that there are not so many best-approximations:

Proposition 3.12. *Let $(p,q) \in \mathcal{Q}(\mathbb{K})$ coprime and $r \in \mathbb{K}[X] \setminus \{0\}$. Suppose $r \cdot (p,q) \in \mathcal{B}_\alpha(\mathbb{K})$ is a best-approximation.*

Then any (other) best-approximation $(p',q') \in \mathcal{B}_\alpha(\mathbb{K})$ with $\deg q' = \deg\,(r\, q)$ has the shape

$$(p',q') = r' \cdot (p,q) \text{ where } r' \in \mathbb{K}[X], \deg r = \deg r'.$$

[1] The polynomial case is even simpler than the integer case treated there: because the absolute value (corresponding to the valuation ord) is non-archimedean, there are no intermediate fractions to worry about.

Proof. Because $(p', q'), r \cdot (p, q) \in \mathcal{B}_\alpha(\mathbb{K})$ with $\deg q' = \deg (r\, q)$, at least one of

$$\operatorname{ord}\bigl(p' - \alpha\, q'\bigr) \geq \operatorname{ord}(r\, p - \alpha\, r\, q) \ \text{ or } \ \operatorname{ord}\bigl(p' - \alpha\, q'\bigr) \leq \operatorname{ord}(r\, p - \alpha\, r\, q)$$

must be satisfied. Together with $\deg q' = \deg (r\, q)$ this implies $\frac{p'}{q'} = \frac{r\,p}{r\,q} = \frac{p}{q}$ by the best-approximation property of either $r \cdot (p, q)$ or (p', q').

Finally because we assume p, q are coprime, there exists $r' \in \mathbb{K}[X]$ with $q' = r'\, q$ and $p' = r'\, p$. □

This proposition has two important consequences:

Corollary 3.13. *For any best-approximation $(p, q) \in \mathcal{B}_\alpha(\mathbb{K})$, the numerator p is uniquely determined by the denominator q.*

Corollary 3.14. *Given an integer $n \geq 0$, there exists up to a constant factor at most one best-approximation (p, q) with $\deg q = n$ and p, q coprime.*

We proceed to show that best-approximations generalise the convergents.

Proposition 3.15. *Let $(p, q) \in \mathcal{Q}(\mathbb{K})$ and $r \in \mathbb{K}[X] \setminus \{0\}$. Suppose*

$$\operatorname{ord}(p - \alpha\, q) = \xi + \deg q.$$

Then $\deg r < \xi$ implies $r \cdot (p, q) \in \mathcal{B}_\alpha(\mathbb{K})$ is a best-approximation.

Putting $r = 1$ with $\deg r = 0 < \xi$ by definition of convergents, we get:

Corollary 3.16. *Every convergent is a best-approximation: $\mathcal{C}_\alpha(\mathbb{K}) \subset \mathcal{B}_\alpha(\mathbb{K})$.*

With Corollary 3.9 this implies also:

Corollary 3.17. *For every non-trivial solution (p, q) of the Pell equation (2.2), either (p, q) or $(p, -q)$ is a best-approximation of $\sqrt{D}$.*

Proof of Proposition 3.15. Let $(p', q') \in \mathcal{Q}(\mathbb{K})$ satisfy

$$\deg q' \leq \deg (r\, q) \ \text{ and } \ \operatorname{ord}\bigl(p' - \alpha\, q'\bigr) \geq \operatorname{ord}(r) + \operatorname{ord}(p - \alpha\, q)\,.$$

Now

$$\det \begin{pmatrix} p & p' \\ q & q' \end{pmatrix} = \det \begin{pmatrix} 1 & -\alpha \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p & p' \\ q & q' \end{pmatrix} = \det \begin{pmatrix} p - \alpha\, q & p' - \alpha\, q' \\ q & q' \end{pmatrix}$$

and taking the valuation ord we get

$$\begin{aligned} \operatorname{ord}\bigl(p\, q' - p'\, q\bigr) &\geq \min\Bigl(\operatorname{ord}\bigl(q'\bigr) + \operatorname{ord}(p - \alpha\, q)\,, \operatorname{ord}(q) + \operatorname{ord}\bigl(p' - \alpha\, q'\bigr)\Bigr) \\ &\geq \operatorname{ord}(r) + \operatorname{ord}(q) + \operatorname{ord}(p - \alpha\, q) = \xi - \deg r > 0. \end{aligned}$$

But $p\, q' - p'\, q \in \mathbb{K}[X]$, so it must be 0. This implies $p'/q' = p/q = r\, p/r\, q$ as desired. □

Remark 3.18. Note that unlike Proposition 3.4, this is only a sufficient condition. It is not necessary: if we start with (p, q) with $\xi > 1$ (for example $\xi = 2$), then multiplying with r of maximal degree (for example $\deg r = 1$), we obtain a best-approximation $(p', q') = r \cdot (p, q)$ with $\xi' \le 0$ (in the example $\xi' = 0$). Then $r' = 1$ does not satisfy $\deg r' < \xi'$, even though (p', q') is a best-approximation.

We conclude our study of best-approximations by investigating their ordering. Indeed we expect that increasing the "height" of the convergent (i.e. $\deg q$) should also increase the approximation quality:

Proposition 3.19. *Let $(p, q), (p', q') \in \mathcal{B}_\alpha(\mathbb{K})$ different best-approximations, i.e. $p/q \ne p'/q'$. Then*

$$\deg q < \deg q' \iff \operatorname{ord}(p - \alpha\, q) < \operatorname{ord}\bigl(p' - \alpha\, q'\bigr)\,.$$

Proof. By the universal property, the statement

$$\operatorname{ord}(p - \alpha\, q) \ge \operatorname{ord}\bigl(p' - \alpha\, q'\bigr) \text{ and } \deg q \le \deg q' \tag{3.7}$$

is false under the hypothesis of the fractions being different.

So if $\deg q < \deg q'$, necessarily the first inequality must not hold, giving the ⇒ part.

Conversely, if $\operatorname{ord}(p - \alpha\, q) > \operatorname{ord}(p' - \alpha\, q')$, then the second inequality is false, i.e. $\deg q > \deg q'$. But this is clearly the ⇐ part, with the roles of (p, q) and (p', q') swapped. □

If we restrict to coprime approximations, we don't even need strict inequalities:

Proposition 3.20. *Let $(p, q), (p', q') \in \mathcal{B}_\alpha(\mathbb{K})$ where p, q and p', q' respectively are coprime. Then*

$$\deg q \le \deg q' \iff \operatorname{ord}(p - \alpha\, q) \le \operatorname{ord}\bigl(p' - \alpha\, q'\bigr)\,.$$

Proof. This is also covered by Proposition 3.19, unless $p/q = p'/q'$. But in this case, the best-approximations differ only by a constant factor, so both inequalities actually become equalities. □

3.4. A linear system for computing convergents

This thesis contains three different proofs for the existence of convergents of arbitrary approximation quality. There is a geometric argument to be explained in Chapter 4. The most elegant approach uses the continued fraction expansion, and yields a complete classification of convergents and best-approximations at the same time; it is one of the main goals of Chapter 5. But here, we give an elementary proof which uses only some linear algebra and other results from this chapter.

We describe a linear system which allows to compute the convergents, alluded to already in Remark 3.2. This already demonstrates the existence of convergents. We will also use these results in Chapter 9 to produce estimates for the projective height of the convergents.

See also [Pla14], where a version of this linear system with additional conditions/rows is used to determine the existence of Pell solutions.

From 3.8 we know $p = \lfloor \alpha\, q \rfloor$ which gives a linear condition on the coefficients of p. Moreover, from the Cauchy product formula, it is clear that every coefficient of $\alpha\, q$ is a linear expression in the coefficients of q. And (3.1) requires just finitely many coefficients of $p - \alpha\, q$ to vanish, so this produces a linear condition on the coefficients of q as well.

We make this more precise now, and start by fixing notation:

Write $\alpha = \sum_{j=-\infty}^{N} A_j\, X^j$ (with $A_N \neq 0$, so $\operatorname{ord}(\alpha) = -N$), and $q = Q_n\, X^n + \dots + Q_0$, $p = P_{n+N}\, X^{n+N} + \dots + P_0$. The A_n are given, and we are solving for $P_{n+N}, \dots, P_0, Q_n, \dots, Q_0$, a total of $N + 2n + 2$ unknowns. For simplicity, we assume $N \geq 0$, but the argument actually works for negative N as well. We get

$$\begin{array}{rllllll}
\alpha\, q - p = & X^{n+N} & (-P_{n+N} & +A_N\, Q_n) & & \\
& +X^{n+N-1} & (-P_{n+N-1} & +A_{N-1}\, Q_n & +A_N\, Q_{n-1}) & \\
& & \quad\vdots & & & \\
& +X^n & (-P_n & +A_0\, Q_n & \dots & +A_n\, Q_0) \\
& +X^0 & (-P_0 & +A_{-n}\, Q_n & \dots & +A_0\, Q_0) \\
& & \quad\vdots & & & \\
& +X^{-n} & (& +A_{-2n}\, Q_n & \dots & +A_{-n}\, Q_0) \\
& +\dots & & & &
\end{array}$$

and the condition $\operatorname{ord}(\alpha\, q - p) > \deg q = n$ means that at the very least the coefficients of $X^{n+N}, \dots, X^{-n}$ vanish. We count a total of $N + 2n + 1$ conditions linear in the P_i and Q_i.

So the matrix describing the linear system has $N + 2n + 2$ columns and $N + 2n + 1$ rows; the right part (and also the left) on its own has the shape of a Toeplitz matrix:[2]

$$\mathcal{M}_n = \left(\begin{array}{cccc|ccc}
-1 & & & 0 & A_N & & 0 \\
 & \ddots & & & A_{N-1} & \ddots & \\
 & & \ddots & & \vdots & \ddots & A_N \\
 & & & \ddots & \vdots & \ddots & \vdots \\
0 & & & -1 & A_{-n} & \dots & A_0 \\ \hline
 & & & & A_{-n-1} & \dots & A_{-1} \\
 & & 0 & & \vdots & \ddots & \vdots \\
 & & & & A_{-2n} & \dots & A_{-n}
\end{array}\right) \tag{3.8}$$

Proposition 3.21. *Every non-zero element of* $\ker \mathcal{M}_n$ *yields a convergent* (p, q)*. As always* $\ker \mathcal{M}_n \neq 0$*, this implies that for any* $\alpha \notin \mathbb{K}(X)$ *there exist convergents with arbitrarily high* $\operatorname{ord}(p - \alpha\, q)$*.*

[2] Or the shape of a Hankel matrix if we reverse the ordering of the columns.

Proof. From the discussion above, it is evident that an element of the kernel gives polynomials (p, q) which are a convergent of α as soon as $q \neq 0$. But if an element of $\ker \mathcal{M}_n$ has all $Q_i = 0$, then clearly it follows that also all $P_i = 0$. So we only need to avoid the zero element. And elementary linear algebra tells as that $\ker \mathcal{M}_n \neq 0$ because there are more columns than rows.

If $\alpha \in \mathbb{K}(X)$, then of course at some point $\operatorname{ord}(p - \alpha\, q) = \infty$, so the approximation quality can no longer be improved. □

Note that for a single $\mathcal{M}_n$, we do not get different convergents:

Proposition 3.22. *If (p, q) and (p', q') correspond to non-zero kernel elements, then $p/q = p'/q'$.*

Proof. Let (p_i, q_i) for $i = 1, \ldots, r$ correspond to a basis of $\ker \mathcal{M}_n$. Then for any (p, q) corresponding to a solution, we get

$$(p, q) = \sum_{i=1}^{r} \eta_i \cdot (p_i, q_i) \text{ where } \eta_i \in \mathbb{K}$$

and hence

$$\operatorname{ord}(p - \alpha\, q) = \operatorname{ord}\left(\sum_{i=1}^{r} \eta_i\, (p_i - \alpha\, q_i)\right) \geq \min_{i=1,\ldots,r} (\operatorname{ord}(p_i - \alpha\, q_i))$$

so there exists (p, q) in the kernel with $\operatorname{ord}(p - \alpha\, q)$ minimal. Write $\operatorname{ord}(p - \alpha\, q) = \xi + \deg q > n$. By Proposition 3.15 also $X^{\xi-1} \cdot (p, q)$ is a best-approximation.[3] And by minimality of $\operatorname{ord}(p - \alpha\, q)$, we have for every (p', q') in the kernel

$$\operatorname{ord}\big(p' - \alpha\, q'\big) \geq \operatorname{ord}(p - \alpha\, q) \geq \operatorname{ord}\Big(X^{\xi-1}\, (p - \alpha\, q)\Big)$$

and moreover $\deg q' \leq n \leq \deg\left(X^{\xi-1}\, q\right)$ which implies $p'/q' = p/q$. □

We can also compute the dimension of the kernel (i.e. the rank of $\mathcal{M}_n$):

Proposition 3.23. *There exists (p, q) in the kernel with p and q coprime.*

If $\operatorname{ord}(p - \alpha\, q) = \xi + \deg q$, then

$$\dim \ker \mathcal{M}_n = \min(1 + \lfloor(\xi - 1)/2\rfloor_{\mathbb{Z}}, 1 + n - \deg q, \xi + \deg q - n)$$

where $\lfloor\cdot\rfloor_{\mathbb{Z}}$ denotes the next lowest integer. So if $\xi \leq 2$ or $n = \deg q$, the matrix $\mathcal{M}_n$ has full rank.

Proof. Removing a common factor decreases $\deg q$ and increases $\operatorname{ord}(p - \alpha\, q)$, so the existence of any solutions implies the existence of a coprime solution. Of course, by the previous Proposition, we can produce all other solutions by adding back a common factor r, with has to satisfy $\deg r \leq n - \deg q$, $\deg r < \xi/2$, and also

$$\operatorname{ord}(r) + \operatorname{ord}(p - \alpha\, q) = \operatorname{ord}(r) + \xi + \deg q > n$$

which is equivalent to $\deg r < \xi + \deg q - n$. □

[3] Here we profit already from allowing common factors for best-approximations.

These results hold for any $\alpha \in \mathbb{K}(\!(X^{-1})\!)$, even if $\alpha \in \mathbb{K}(X)$.

With Cramer's rule we can compute an element of the kernel:

Remark 3.24. Denote by $\det \mathcal{M}_n(i)$ the ith minor obtained by striking the ith column. Then

$$\begin{multline}(P_{n+N}, \dots, P_0, Q_n, \dots, Q_0) = \\ (\det \mathcal{M}_n(1), -\det \mathcal{M}_n(2), \det \mathcal{M}_n(3), \dots \\ \dots, (-1)^{N+2n} \det \mathcal{M}_n(N+2n+1), \\ (-1)^{N+2n+1} \det \mathcal{M}_n(N+2n+2)\big) \quad (3.9)\end{multline}$$

is an element of the kernel. If $\mathcal{M}_n$ has full rank, then it is clearly non-zero.

These formulas present an alternative to computing convergents via the continued fraction, and we will later show that the convergents obtained in this way are actually optimally normalised (see Proposition 7.33).

4. A (hyper)elliptic curve

In this chapter, we describe the (hyper)elliptic curve corresponding to a given polynomial Pell equation. We additionally assume that D is square-free, to avoid complications and so that we may work with the Jacobian of the curve.[1]

We also explain how the convergents give rise to principal divisors of particular shape (Lemma 4.7), and this gives rise to the torsion condition for D being Pellian (Theorem 4.1).

Most of the results of this chapter have long been known, probably already to Abel [Abe26] and Chebyshev [Che57], albeit not in our modern mathematical language. More recent publications are [AR80] for elliptic curves, or [Ber90] for arbitrary genus.

As in the previous chapters, we assume that $\mathbb{K}$ is a field of characteristic not 2.

4.1. Defining the (hyper)elliptic curve

Let $D \in \mathbb{K}[X] \setminus \mathbb{K}$ *square-free* with even degree $2(g+1)$ and $\ell c(D)$ a square in $\mathbb{K}$. Then

$$\mathcal{C}_{\mathrm{aff}} : Y^2 = D(X)$$

defines an affine (plane) curve over $\mathbb{K}$ of genus g.

Proposition 4.1. *The curve $\mathcal{C}_{\mathrm{aff}}$ is smooth and normal in $\mathbb{A}^2_{\mathbb{K}}$.*

Proof. The curve is defined by the equation

$$F = Y^2 - D(X, Z).$$

Applying the Jacobian criterion we calculate

$$\frac{\partial}{\partial X} F = -\partial_X D(X) = D'(X) \qquad \frac{\partial}{\partial Y} F = 2\,Y$$

which are never simultaneously 0 because D square-free implies that D and D' are coprime.

For normality, we need to show that if $p + Y\,q \in \mathsf{Fr}(\mathbb{K}[X,Y]/\big(Y^2 - D(X)\big)) = \mathbb{K}(X)[Y]/\big(Y^2 - D(X)\big)$ is integral, it is already contained in $\mathbb{K}[X,Y]/\big(Y^2 - D(X)\big)$, i.e. $p, q \in \mathbb{K}[X]$ are polynomials. Recall that the integral closure is a subring of the fraction field, and $p + Y\,q$ integral implies that the conjugate $p - Y\,q$ is integral as well. It follows that $2\,p$ and $p^2 - D\,q^2$ are integral. As we assume $\operatorname{char}\mathbb{K} \neq 2$, this implies p and also $D\,q^2$ are integral over $\mathbb{K}[X,Y]/\big(Y^2 - D(X)\big)$, so in particular over the subring $\mathbb{K}[X]$. As D is square-free, it follows that $p, q \in \mathbb{K}[X]$ as desired, and $\mathcal{C}_{\mathrm{aff}}$ is normal. □

[1] If D is not square-free, we have to use generalised Jacobians instead. See [Zan16] on how this relates to the Pell equation and continued fractions, and [Ser88] for an introduction to generalised Jacobians.

Remark 4.2. If $\deg D = 2$, then $\overline{\mathcal{C}_{\text{aff}}} \subset \mathbb{P}^2_{\mathbb{K}}$ remains smooth at infinity, so it is isomorphic to $\mathbb{P}^1$ (see Proposition 7.4.1 in [Liu02]).

But if $\deg D > 2$, then $\overline{\mathcal{C}_{\text{aff}}} \subset \mathbb{P}^2_{\mathbb{K}}$ has a singularity at infinity (easily verified with the Jacobian criterion).

We build a smooth projective model for $\mathcal{C}_{\text{aff}}$, as in Lemma III.1.7 of [Mir95]:

Define the curve

$$\mathcal{C}_\infty : V^2 = D^\flat(U) = U^{2(g+1)} D(1/U)$$

where $D^\flat(U)$ is a polynomial of degree at most $2(g+1)$ – its coefficients are those of D in reverse order. Note that $D^\flat(0) \neq 0$ because $\deg D = 2(g+1)$, and by Proposition 4.1 the curve $\mathcal{C}_\infty$ is smooth in $\mathbb{A}^2_{\mathbb{K}}$.

The relations $X\,U = 1$ and $U^{g+1}\,Y = V$ (respectively $X^{g+1}\,V = Y$) describe a birational map between $\mathcal{C}_{\text{aff}}$ and $\mathcal{C}_\infty$ which is an isomorphism outside of $U = 0$ and $X = 0$. So we may glue $\mathcal{C}_{\text{aff}}$ and $\mathcal{C}_\infty$ together to obtain a curve $\mathcal{C}$. This simply adds two points $O_\pm$ with $U = 0$ to $\mathcal{C}_{\text{aff}}$, the points at infinity.

Proposition 4.3. *The curve $\mathcal{C}$, glued together from $\mathcal{C}_{\text{aff}}$ and $\mathcal{C}_\infty$ is a normal smooth projective curve over $\mathbb{K}$.*

Proof. Normality and smoothness of $\mathcal{C}$ are local conditions, hence they follow from Proposition 4.1 applied to $\mathcal{C}_{\text{aff}}$ and $\mathcal{C}_\infty$.

We get a finite morphism $\mathcal{C} \to \overline{\mathcal{C}_{\text{aff}}} \subset \mathbb{P}^2_{\mathbb{K}}$, hence $\mathcal{C}$ is proper over $\mathbb{K}$. As $\mathcal{C}$ is an algebraic variety, this implies by Remark 3.3.33 (1) in [Liu02] that it is projective. □

There is an involution σ defined by $X \mapsto X,\ Y \mapsto -Y$, or $U \mapsto Y,\ V \mapsto -V$. By abuse of notation, we also consider it as an automorphism of the function field $\mathbb{K}(X, Y)$. If we quotient $\mathcal{C}$ by the group $\{1, \sigma\}$, we find that $\mathcal{C}$ is (hyper)elliptic (we use Definition 7.4.7 from [Liu02] which is essentially the content of the following proposition):

Proposition 4.4. *There is finite morphism $\pi : \mathcal{C} \to \mathbb{P}^1$ of degree 2 defined by $(x, y) \mapsto (x : 1)$ on $\mathcal{C}_{\text{aff}}$ and $\pi(O_\pm) = (1 : 0)$. For $g = 1$, the curve $\mathcal{C}$ is elliptic, and for $g \geq 2$ it is hyperelliptic.*

Proof. The map π is defined on $\mathcal{C}_{\text{aff}}$ via $(x, y) \mapsto (x : 1)$, and on $\mathcal{C}_\infty$ via $(u, v) \mapsto (1 : u)$. Clearly the definitions are compatible on the intersection (because there we have $x\,u = 1$). It is also clear that π is a finite morphism of degree 2 which means that $\mathcal{C}$ is elliptic for $g = 1$ and hyperelliptic for $g \geq 2$. □

4.2. Divisors and the Jacobian variety

We recall some basic notions about divisors and the Jacobian variety now. For more details, consult your favourite algebraic geometry book, for instance [Har77], [GW10] or [Liu02]. For the rest of the chapter, we work over the algebraic closure $\overline{\mathbb{K}}$ to avoid complications.

4.2.1. Divisors

For any $P \in \mathcal{C}(\overline{\mathbb{K}})$, there is a discrete valuation

$$\mathrm{ord}_P : \overline{\mathbb{K}}(X,Y)^{\times} \to \mathbb{Z},$$

the zero-order of P of a function on $\mathcal{C}$. In fact, all non-trivial discrete $\overline{\mathbb{K}}$-valuations (up to equivalence) on $\overline{\mathbb{K}}(X,Y)$ arise in this way.

By the *group of divisors* $\mathrm{Div}(\mathcal{C})$ we understand the free abelian group generated by all points of $\mathcal{C}(\overline{\mathbb{K}})$ (we mark divisors in **bold**). For every divisor

$$\mathbf{D} = \sum_{P \in \mathcal{C}(\overline{\mathbb{K}})} n_P \; (P)\,, \text{ where } n_P \in \mathbb{Z}$$

we define the *degree*

$$\deg \mathbf{D} = \sum_{P \in \mathcal{C}(\overline{\mathbb{K}})} n_P.$$

A divisor is called *effective* if $n_P \geq 0$ for all P.

For every element $f \in \overline{\mathbb{K}}(X,Y)^{\times}$, only finitely many $\mathrm{ord}_P\, f$ are non-zero, so we can define the *divisor of f* as

$$\mathrm{div}\, f = \sum_{P \in \mathcal{C}(\overline{\mathbb{K}})} (\mathrm{ord}_P\, f) \; (P)\,.$$

The divisors arising in this way are called *principal divisors*, and they all have degree 0. So there is group homomorphism $\mathrm{div} : \overline{\mathbb{K}}(X,Y)^{\times} \to \mathrm{Div}^0(\mathcal{C})$ where $\mathrm{Div}^0(\mathcal{C})$ denotes the divisors of degree 0.

Recall that the Jacobian $\mathcal{J}$ of $\mathcal{C}$ is an abelian variety of dimension g. If $g = 1$ (i.e. $\deg D = 4$), the curve $\mathcal{C}$ is an elliptic curve (Corollary 7.4.5 in [Liu02]), and it is isomorphic to its Jacobian.

The $\overline{\mathbb{K}}$-rational points of the *Jacobian* can be seen as the cokernel of the divisor map, more precisely the quotient

$$\mathcal{J} = \mathcal{J}(\mathcal{C}) = \mathrm{Div}^0(\mathcal{C}) / \operatorname{im} \mathrm{div},$$

with the projection $\mathrm{Div}^0(\mathcal{C}) \to \mathcal{J}$. By abuse of language, we call both the algebraic variety and its set of $\overline{\mathbb{K}}$-rational points “Jacobian”.

We write a divisor class in the Jacobian as

$$[\mathbf{D}] = \sum_{P \in \mathcal{C}(\overline{\mathbb{K}})} n_P \; [P]\,.$$

4.2.2. Order functions

If we restrict $\mathrm{ord}_{O_\pm}$ to $\overline{\mathbb{K}}(X)$, it becomes exactly $\mathrm{ord}_\infty = \mathrm{ord}$ from Section 2.3. As mentioned before, there are precisely two embeddings of $\overline{\mathbb{K}}(X,Y)$ into the completion $\overline{\mathbb{K}}(\!(X^{-1})\!)$. To distinguish them properly, we set for $p, q \in \overline{\mathbb{K}}(X)$

$$\mathrm{ord}_{O_+}(p + Y\,q) = \mathrm{ord}\Big(p + \sqrt{D}\,q\Big)\,, \quad \mathrm{ord}_{O_-}(p + Y\,q) = \mathrm{ord}\Big(p - \sqrt{D}\,q\Big)$$

for a fixed choice of $\sqrt{D}$ (see Definition 2.10). Apart from this, the roles of O_+ and O_- are essentially interchangeable (by the involution σ).

In a similar way, one may compute order functions for a finite point $P = (x, y) \in \mathcal{C}_{\text{aff}}$ by choosing an uniformiser. Sending X to $T + x$ gives a homomorphism $\overline{\mathbb{K}}(X) \to \overline{\mathbb{K}}((T))$, and if $y \neq 0$, one may compute $\sqrt{D(T + x)}$ in $\overline{\mathbb{K}}((T))$ with the constant coefficient y determining the choice of square root. Sending Y to $\sqrt{D(T + x)}$ then establishes a homomorphism $\mathbb{K}(X, \sqrt{D}) \to \overline{\mathbb{K}}((T))$ with ord_P corresponding to $\text{ord}_{T=0}$.

If $y = 0$, one sends instead X to $T^2 + x$, to ensure $\sqrt{D(T^2 + x)} \in \overline{\mathbb{K}}((T))$. Because the latter has odd $\text{ord}_{T=0}$, the choice of root does not matter, and one obtains as before the correspondence between ord_P and $\text{ord}_{T=0}$. Note that in this case for $f \in \overline{\mathbb{K}}(X)$ the zero-order $\text{ord}_P(f)$ is always *even*.

4.2.3. Embedding the curve in the Jacobian

Choosing the base point O_+ (a natural choice here, but any other point on $\mathcal{C}$ would do as well), define the map $j : \mathcal{C} \to \mathcal{J}$ via $P \mapsto [P] - [O_+]$ which is an embedding for $g \geq 1$ (see Theorem A8.1.1 in [HS00]). Actually, for $g = 1$, when $\mathcal{C}$ is an elliptic curve, it is an isomorphism of curves, determined uniquely by the choice of the base point.

Of course, we can extend $j : \text{Div}(\mathcal{C}) \to \mathcal{J}$ as a homomorphism of groups (using that $\text{Div}(\mathcal{C})$ is a free group on $\mathcal{C}$).

For each $r \geq 0$, we may also define a subvariety of $\mathcal{J}$

$$W_r = j(\mathcal{C}) + \cdots + j(\mathcal{C}) \quad (r \text{ copies})$$

remarking $W_g = \mathcal{J}$ (see again Theorem A8.1.1 in [HS00]), while the Theta divisor $\Theta = W_{g-1}$ forms a proper subvariety which depends on the embedding j. We will use this divisor with the Weil height machine later, and likewise the canonical divisor.

Proposition 4.5. *The canonical divisor* $\mathbf{K}_{\mathcal{C}}$ *on* $\mathcal{C}$ *is represented by*

$$\text{div}\left(\frac{\mathrm{d}X}{Y}\right) = (g - 1)\left((O_+) + (O_-)\right).$$

Proof. From Riemann-Roch one deduces easily that $\deg \mathbf{K}_{\mathcal{C}} = 2(g - 1)$. It is also clear that we obtain the canonical divisor class by computing the divisor of any differential on $\mathcal{C}$.

Now outside of infinity, the sheaf of differentials is clearly generated by $\mathrm{d}X$ and $\mathrm{d}Y$ which enjoy the relation

$$2\,Y\,\mathrm{d}Y = D'(X)\,\mathrm{d}X$$

obtained by differentiating the equation of the curve. This tells us that outside of $D(X) = 0$, the sheaf of differentials is generated by $\mathrm{d}X$, while outside of $D'(X) = 0$ it is generated by $\mathrm{d}Y$. Moreover we see that $\mathrm{d}X$ vanishes only on $Y = 0$ (i.e. $D(X) = 0$), while $\mathrm{d}Y$ vanishes only on $D'(X) = 0$.

It follows that $\frac{\mathrm{d}X}{Y}$ has poles and zeroes only at infinity. As the divisor of this differential is invariant under the involution σ (which changes the differential only by a factor -1), and has to have degree $2(g - 1)$, we obtain the above formula. □

4.3. Divisors of convergents

Given a rational approximation (p, q), it is very natural to build the function $p - Y\,q$ and study its divisor. For the convergents, we will see that this divisor describes how the multiples of the divisor at infinity

$$\mathbf{O} = (O_+) - (O_-) \in \mathrm{Div}^0(\mathcal{C}) \tag{4.1}$$

are represented as sums of g points, i.e. as elements of $W_g = \mathcal{J}$. Note that $\mathbf{O}$ is actually a $\mathbb{K}$-rational divisor, so $[\mathbf{O}]$ is a $\mathbb{K}$-rational point of $\mathcal{J}$.

Proposition 4.6. *Let $p, q \in \overline{\mathbb{K}}(X)$ and $\phi_\pm = p \pm Y\,q \neq 0$. The following are equivalent:*

1. *$p, q \in \overline{\mathbb{K}}[X]$*
2. *For all $P \neq O_\pm$ holds $\mathrm{ord}_P\,\phi_+ \geq 0$*
3. *For all $P \neq O_\pm$ holds $\mathrm{ord}_P\,\phi_- \geq 0$*

Proof. 1. and 3. are clearly equivalent because $\mathrm{ord}_P\,\phi_+ = \mathrm{ord}_{\sigma(P)}\,\phi_-$ for all P.

Together, they imply 1.:

$$\begin{aligned} \mathrm{ord}_P(p) = \mathrm{ord}_P(2\,p) = \mathrm{ord}_P(\phi_+ + \phi_-) \geq \min(\mathrm{ord}_P(\phi_+), \mathrm{ord}_P(\phi_-)) \geq 0 \\ \mathrm{ord}_P(Y\,q) = \mathrm{ord}_P(2\,Y\,q) = \mathrm{ord}_P(\phi_+ - \phi_-) \geq \min(\mathrm{ord}_P(\phi_+), \mathrm{ord}_P(\phi_-)) \geq 0 \end{aligned}$$

so clearly p has no poles outside infinity, hence it is a polynomial. If $\mathrm{ord}_P(Y) \neq 0$, then $\mathrm{ord}_P(Y) = 1$ because D is square-free. But at the same time $\mathrm{ord}_P(q)$ must be even (see Section 4.2.2). This shows that $\mathrm{ord}_P(q) \geq 0$, and that q has no poles outside infinity which means it is a polynomial.

Conversely 1. implies 2.: if $p, q \in \overline{\mathbb{K}}[X]$, then $\mathrm{ord}_P(p) \geq 0$, $\mathrm{ord}_P(q) \geq 0$ and of course $\mathrm{ord}_P(Y) \geq 0$ for all $P \neq O_\pm$. Hence

$$\mathrm{ord}_P(\phi_\pm) = \mathrm{ord}_P(p \pm Y\,q) \geq \min(\mathrm{ord}_P(p), \mathrm{ord}_P(Y) + \mathrm{ord}_P(q)) \geq 0$$

as desired. □

Lemma 4.7. *Let $p, q \in \overline{\mathbb{K}}(X)$, and $\phi_\pm = p \pm Y\,q \neq 0$. Set $m = \deg p$. Then $(p, q) \in \mathcal{C}_{\sqrt{D}}(\mathbb{K})$ (it is a convergent of $\sqrt{D}$) if and only if $m > 0$ and there exists $0 \leq r \leq \min(g, m)$ and $P_1, \dots, P_r \in \mathcal{C}_{\mathrm{aff}}$ such that*

$$\mathrm{div}\,\phi_- = -m\ (O_-) + (m - r)\ (O_+) + (P_1) + \cdots + (P_r)\,. \tag{4.2}$$

We call $\mathrm{div}\,\phi_-$ *a* convergent divisor.

Proof. By Proposition 4.6 we can clearly restrict to the case $p, q \in \overline{\mathbb{K}}[X]$ as the divisor in (4.2) allows only poles at infinity, and convergents are always made of polynomials. The rest of the proof boils down to distinguishing the points at infinity and calculating r.

Now $\operatorname{ord}_{O_+} \phi_- = \operatorname{ord}\left(p - \sqrt{D}\, q\right) \geq 0$ holds for both conditions and implies

$$\begin{aligned}\operatorname{ord}_{O_-} \phi_- = \operatorname{ord}_{O_+} \phi_+ &= \operatorname{ord}\left(p + \sqrt{D}\, q\right) \\ &= \min(\operatorname{ord}(p)\,, \operatorname{ord}\left(p - \sqrt{D}\, q\right)) = \operatorname{ord}(p) = -m.\end{aligned}$$

Similarly, $m = \deg q + g + 1$ (see also the proof of Proposition 3.9).

Now $P_1, \ldots, P_r$ are the finite zeroes of ϕ_-, accounted for with multiplicities. Of course $\operatorname{div} \phi_-$ has degree 0, hence

$$\operatorname{ord}\left(p - \sqrt{D}\, q\right) = \operatorname{ord}_{O_+} \phi_- = (m - r) = \deg q + g + 1 - r.$$

This is $> \deg q$ (i.e. $(p, q) \in \mathcal{C}_{\sqrt{D}}(\overline{\mathbb{K}})$) if and only if $r \leq g$, so we have the desired equivalence. □

We will give a slight generalisation (extending to other elements of the function field) later in Section 6.6, to illustrate the connection with the continued fraction.

Remark 4.8. In the Jacobian, we can write this divisor relation as

$$m \cdot j(O_-) = j(P_1) + \cdots + j(P_r).$$

Remark 4.9. With the notation from Proposition 3.9, we get $\deg \Omega = r$ because

$$\operatorname{ord}(\Omega) = \operatorname{ord}_{O_+}(\phi_+ \cdot \phi_-) = \operatorname{ord}_{O_+}(\phi_+) + \operatorname{ord}_{O_+}(\phi_-) = -m + (m - r) = -r.$$

In the same proposition, the condition to obtain a convergent (up to sign of q) was $r = \deg \Omega < \frac{1}{2} \deg D = g + 1$ which matches the above lemma.

Proposition 4.10. *For every $n \in \mathbb{N}$ there exists $\phi_n \in \overline{\mathbb{K}}(X, Y) \setminus \{0\}$ such that*

$$\operatorname{div} \phi_n = -m\ (O_-) + (m - r)\ (O_+) + (P_1) + \cdots + (P_r)$$

with $m \geq n$, $r \leq \min(g, m)$ and $P_1, \ldots, P_r \in \mathcal{C}_{\text{aff}}$.

Proof. For $n \in \mathbb{N}$ define the divisor

$$\mathbf{D}_n = (n + g)\ (O_-) - n\ (O_+)\,.$$

which has degree $\deg \mathbf{D}_n = g$. Then the Riemann-Roch theorem (see Theorem IV.1.3 in [Har77]) implies

$$\dim\{\phi \in \overline{\mathbb{K}}(X, Y) \setminus \{0\} \mid \operatorname{div} \phi + \mathbf{D}_n \geq 0\} \geq \deg \mathbf{D}_n - g + 1 = 1$$

so there exists ϕ_n with $\operatorname{div} \phi_n \geq -\mathbf{D}_n$. More precisely, we get

$$\operatorname{div} \phi_n = -(n + g)\ (O_-) + n\,(O_+) + (P_1) + \cdots + (P_g)$$

where $P_i \in \mathcal{C}$ (possibly $O_\pm$). We can write this as

$$\operatorname{div} \phi_n = -m\,(O_-) + (m - r)\,(O_+) + (P_{i_1}) + \cdots + (P_{i_r})\,,$$

cancelling out any O_- among the P_i (hence $m \geq n + g - g = n$) and absorbing any O_+ from the P_i (hence $m - r \geq n + g \geq 0$). And of course $r \leq g$. □

Via the above lemma, we now have another proof for the existence of convergents:

Corollary 4.11. *$\sqrt{D}$ has convergents (p, q) of arbitrarily high $\deg p$ (or $\deg q$).*

Theorem 4.1. *The Pell equation (2.2) has a non-trivial solution if and only if $[\mathbf{O}]$ is a torsion point in the Jacobian $\mathcal{J}$ of $\mathcal{C}$.*

Proof. From Proposition 3.9 we know that the Pell solutions (up to conjugation) form a subset of the convergents. By Remark 4.9 it is precisely the non-trivial Pell solutions for which we have $r = 0$ in Lemma 4.7.

By Remark 4.8, this implies $m\,[\mathbf{O}] = m\,j(O_-) = 0$ with $m > 0$. In other words, $[\mathbf{O}]$ is a torsion point in the Jacobian $\mathcal{J}$.

Conversely, if $[\mathbf{O}]$ is torsion, then there exists some function ϕ with divisor $\operatorname{div}\phi = m\,((O_+) - (O_-))$ and $m > 0$. By Lemma 4.7, we have $\phi = p - Y\,q$ where (p, q) is a convergent (actually a Pell solution because $r = 0$). □

Remark 4.12. Recall that the genus g corresponds to the dimension of the Jacobian. For $g = 0$, the Jacobian is the trivial group, hence $[\mathbf{O}]$ is trivially torsion. Hence D is always Pellian as observed before in Corollary 2.16.

Remark 4.13. If the base field $\mathbb{K}$ is finite, i.e. $\mathbb{K} = \mathbb{F}_q$ with q some prime power, the $\mathbb{K}$-rational points of the Jacobian form a finite group. The Hasse-Weil interval (conjectured by E. Artin in his thesis, then proved by Hasse for elliptic curves [Has36a, Has36b, Has36c], and generalised by Weil to higher genus curves in [Wei49]) then provides the following bounds for the number of elements of the Jacobian:

$$\operatorname{ord}([\mathbf{O}]) \in [(\sqrt{q} - 1)^{2g}, (\sqrt{q} + 1)^{2g}].$$

Note that $\mathcal{J}(\mathbb{F}_q)$ can be cyclic (for elliptic curves, see for example [GM90]), so we cannot hope to improve this bound for the point $[\mathbf{O}]$.

Remark 4.14. If $\mathbb{K}$ is not finite, then as mentioned in the introduction, this torsion condition allows to demonstrate the scarcity of Pellian polynomials. The polynomials of degree $2d$, after some normalisation, form an affine variety of dimension $2d - 2$. The Pellian polynomials are then contained in a denumerable union of subvarieties of dimension at most $d - 1$, corresponding to the possible torsion orders. See Section 12.2.2 in [Zan14] for details.

5. Continued fractions

In this chapter, we develop the theory of polynomial continued fractions, to build a solid foundation for the specialization questions that form the main results of this thesis. Beginning with formal continued fractions, moving on to convergence questions in $\mathbb{K}(\!(X^{-1})\!)$ and a classification of the best-approximations, we conclude with a discussion of periodic continued fractions and reducedness which is relevant mostly for hyperelliptic continued fractions.

The first sections reiterate well-known facts about continued fractions in modern language. Already Abel [Abe26] and Chebyshev [Che57] worked with this type of polynomial continued fractions which they adapted from the numerical continued fraction expansion for square roots. Indeed there are not many differences with the theory of continued fractions for real numbers.

Most results in this chapter may already be found the literature, albeit presented differently. The formal definitions of continued fractions can be found in classical books on continued fractions, for example [Per54], [Per57], [Khi56] and others. For polynomial continued fractions, see [Abe26], [Ber90], [Sch00] or the survey paper [vdPT00].

As before $\mathbb{K}$ is a field characteristic not 2 throughout the chapter

5.1. Finite continued fractions

For our formal continued fractions, we begin by using a double index notation, as this should make some calculations much clearer and precise. We will drop the first index once we no longer need it.

Definition 5.1. Let $m, n \in \mathbb{Z}, n \geq 0$. The expression

$$\alpha_{m,n} = [a_m, a_{m+1}, \dots, a_{m+n}] = a_m + \cfrac{1}{a_{m+1} + \cfrac{1}{\ddots + \cfrac{1}{a_{m+n}}}}$$

where we consider the a_i as free variables is called a *finite continued fraction.* We define it recursively by

$$\alpha_{m,0} = a_m \quad \text{and} \quad \alpha_{m,n} = a_m + \frac{1}{\alpha_{m+1,n-1}} \text{ for } n \geq 1$$

respectively

$$[a_m] = a_m \quad \text{and} \quad [a_m, a_{m+1}, \dots, a_{m+n}] = a_m + \frac{1}{[a_{m+1}, \dots, a_{m+n}]} \text{ for } n \geq 1$$

in the square bracket notation.

Remark 5.2. By induction one obtains also for $l \geq 1$

$$\alpha_{m,n} = [a_m, a_{m+1}, \dots, a_{m+l-1}, \alpha_{m+l,n-l}],$$

so the concatenation of $[a_m, a_{m+1}, \dots, a_{m+l-1}]$ and $[a_{m+l}, \dots, a_{m+n}]$ is the same as inserting the second at the end of the first finite continued fraction:

$$[a_m, a_{m+1}, \dots, a_{m+n}] = [a_m, a_{m+1}, \dots, a_{m+l-1}, [a_{m+l}, \dots, a_{m+n}]].$$

5.2. Continued fractions and matrix products

Clearly a continued fraction $\alpha_{m,n}$ can be seen as an element of $\mathbb{P}^1(\mathbb{Z}[a_m, \dots, a_{m+n}])$, where the empty continued fraction corresponds to $[\,] = \frac{1}{0} \in \mathbb{P}^1$. This motivates the following viewpoint:

We can think of a finite continued fractions as a map on $\mathbb{P}^1$, via

$$x \in \mathbb{P}^1 \mapsto [a_m, \dots, a_{m+n}, x] \in \mathbb{P}^1.$$

We can relate such a map to the natural (left) action of $\mathrm{GL}_2(\mathbb{Z}[a_i \mid i \in \mathbb{Z}])$ on $\mathbb{P}^1$ via Moebius transformations:

$$x \mapsto \frac{a\,x + b}{c\,x + d} \longleftrightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Then clearly

$$x \mapsto [a_m, x] = a_m + \frac{1}{x} \longleftrightarrow \begin{pmatrix} a_m & 1 \\ 1 & \end{pmatrix}.$$

As concatenation is the same as composition, this extends to

$$x \mapsto [a_m, \dots, a_{m+n}, x] \longleftrightarrow \begin{pmatrix} a_m & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} a_{m+n} & 1 \\ 1 & \end{pmatrix}.$$

By multiplying out these matrices, we can canonically compute the numerator and denominator of the fraction represented by a finite continued fraction.

Proposition 5.3. *For every $m, n \in \mathbb{Z}, n \geq -1$, there exist polynomials $p_{m,n}, q_{m,n} \in \mathbb{Z}[a_m, \dots, a_{m+n}]$ such that*

$$\begin{pmatrix} a_m & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} a_{m+n} & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} p_{m,n} & p_{m,n-1} \\ q_{m,n} & q_{m,n-1} \end{pmatrix}, \tag{5.1}$$

satisfying $p_{m,n}/q_{m,n} = \alpha_{m,n}$.

Proof. Take $x = [\,] = \frac{1}{0}$ the empty continued fraction, then we define

$$\frac{p_{m,n}}{q_{m,n}} := \begin{pmatrix} a_m & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} a_{m+n} & 1 \\ 1 & \end{pmatrix} \frac{1}{0} = [a_m, \dots, a_{m+n}, [\,]] = \alpha_{m,n}$$

where clearly $p_{m,n}, q_{m,n} \in \mathbb{Z}[a_m, \dots, a_{m+n}]$ because the matrix entries are in that ring. Also note that $\begin{pmatrix} a_{m+n} & 1 \\ 1 & \end{pmatrix} \frac{0}{1} = \frac{1}{0}$, so

$$\begin{pmatrix} a_m & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} a_{m+n} & 1 \\ 1 & \end{pmatrix} \frac{0}{1} = \begin{pmatrix} a_m & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} a_{m+n-1} & 1 \\ 1 & \end{pmatrix} \frac{1}{0} = \frac{p_{m,n-1}}{q_{m,n-1}}.$$

□

Remark 5.4. Transposing the matrix $\begin{pmatrix} p_{m,n} & p_{m,n-1} \\ q_{m,n} & q_{m,n-1} \end{pmatrix}$ corresponds to reversing the ordering of the variables $a_m, \dots, a_{m+n}$; and $p_{m,n-1}$ depends only on $a_m, \dots, a_{m+n-1}$, so one easily deduces that $q_{m,n}$ is independent of a_m: it follows $q_{m,n} \in \mathbb{Z}[a_{m+1}, \dots, a_{m+n}]$.

By taking the determinants of the matrix product, we get

Corollary 5.5. *For fixed m and n, we have the relation*

$$p_{m,n}\, q_{m,n-1} - q_{m,n}\, p_{m,n-1} = (-1)^{n+1}. \tag{5.2}$$

Consequently, the $p_{m,n}$ and $q_{m,n}$ are coprime. This holds even if we assign values to the a_i.

The sequences in n of the $p_{m,n}$ and $q_{m,n}$ may also be computed independently:

Corollary 5.6. *The $p_{m,n}$ and $q_{m,n}$ satisfy the recursion relations*

$$\begin{aligned} p_{m,n} &= a_{m+n}\, p_{m,n-1} + p_{m,n-2} \text{ for } n \geq 0, \quad p_{m,-1} = 1, \quad p_{m,-2} = 0, \\ q_{m,n} &= a_{m+n}\, q_{m,n-1} + q_{m,n-2} \text{ for } n \geq 1, \quad\quad q_{m,0} = 1, \quad q_{m,-1} = 0. \end{aligned} \tag{5.3}$$

5.3. Infinite continued fractions

To give sense to infinite continued fraction, we need some topology. In our case, we use $\mathbb{K}[X]$ with the previously defined (non-archimedean) absolute valuation $\operatorname{ord} = \operatorname{ord}_\infty$ (see Section 2.3). We assume that all $a_n \in \mathbb{K}[X]$. Then the $\alpha_{m,n}$ are contained in $\mathbb{K}(X)$, and we can hope to find a limit in the completion $\mathbb{K}(\!(X^{-1})\!)$.

Definition 5.7. We define the *infinite continued fraction*

$$\alpha_m = \alpha_{m,\infty} = [a_m, a_{m+1}, \dots] = \lim_{n \to \infty} \alpha_{m,n}$$

if the limit exists.

From now on, we assume that all $a_n \in \mathbb{K}[X]$, and search for a sufficient condition for the convergence of $(\alpha_{m,n})_{n\in\mathbb{N}}$.

Proposition 5.8. *If* $\deg a_n \geq 1$ *holds for all* $n \geq m+1$, *then*

$$\deg p_{m,n} = \sum_{j=0}^{n} \deg a_{m+j}, \qquad \deg q_{m,n} = \sum_{j=1}^{n} \deg a_{m+j} \tag{5.4}$$

and

$$\ell c(p_{m,n}) = \prod_{j=0}^{n} \ell c(a_{m+j}), \qquad \ell c(q_{m,n}) = \prod_{j=1}^{n} \ell c(a_{m+j}). \tag{5.5}$$

The proposition is a consequence of the following lemma:

Lemma 5.9. *Let* $(a_n)_{n\in\mathbb{N}}$ *a sequence in* $\mathbb{K}[X]$, *with* $\deg a_n \geq 1$ *for all* $n \geq 1$. *Define a sequence* $(b_n)_{n\geq -1}$ *via*

$$b_{-1} = 0, \quad b_0 = 1, \quad b_n = a_n\, b_{n-1} + b_{n-2} \text{ for } n \geq 1 \tag{5.6}$$

Then $\deg b_n$ *is strictly increasing and in fact*

$$\deg b_n = \sum_{j=1}^{n} \deg a_j \quad \text{for } n \geq 0.$$

Moreover, one can easily compute the leading coefficient:

$$\ell c(b_n) = \prod_{j=1}^{n} \ell c(a_j) \quad \text{for } n \geq 0.$$

Proof. We prove the statement by induction on n. For $n = 0$ we clearly have $\deg b_{-1} < \deg b_0 = 0$ and $\ell c(b_0) = 1$. For the induction step, note that by hypothesis $\deg b_{n-2} < \deg b_{n-1} < \deg(a_n\, b_{n-1})$, so (5.6) implies

$$\deg b_n = \deg(a_n\, b_{n-1} + b_{n-2}) = \deg a_n + \deg b_{n-1} = \deg a_n + \sum_{j=1}^{n-1} \deg a_j$$

as desired, and clearly $\deg b_{n-1} < \deg b_n$. It follows

$$\ell c(b_n) = \ell c(a_n\, b_{n-1}) = \ell c(a_n) \prod_{j=1}^{n-1} \ell c(a_j).$$

□

We can now answer the question about the convergence of infinite continued fractions:

Proposition 5.10. *Suppose* $\deg a_n \geq 1$ *holds for* $n \geq m+1$. *Then* $(\alpha_{m,n})_{n\in\mathbb{N}}$ *is a Cauchy sequence and converges in* $\mathbb{K}(\!(X^{-1})\!)$. *We denote the limit by* $\alpha_m = \alpha_{m,\infty} = [a_m, a_{m+1}, a_{m+2}, \dots]$.

Proof. Dividing (5.2) by $q_{m,n-1} \cdot q_{m,n}$ implies

$$\alpha_{m,n} - \alpha_{m,n-1} = \frac{p_{m,n}}{q_{m,n}} - \frac{p_{m,n-1}}{q_{m,n-1}} = \frac{(-1)^{n+1}}{q_{m,n-1} \cdot q_{m,n}},$$

hence

$$\operatorname{ord}(\alpha_{m,n} - \alpha_{m,n-1}) = \deg q_{m,n} + \deg q_{m,n-1} \geq 2n - 1 \tag{5.7}$$

by Proposition 5.8. This means the “distance” between $\alpha_{m,n}$ and $\alpha_{m,n-1}$ converges to 0 as $n \to \infty$. Because we are working with a non-archimedean valuation, this already implies that $(\alpha_{m,n})_{n\in\mathbb{N}}$ is a Cauchy sequence. □

So a continued fraction (with non-constant coefficients $a_n \in \mathbb{K}[X]$) produces an element of $\mathbb{K}(\!(X^{-1})\!)$ (actually a sequence of elements of $\mathbb{K}(\!(X^{-1})\!)$). In the next section, we will reverse the process and produce a continued fraction for every element of $\mathbb{K}(\!(X^{-1})\!)$, thus establishing a bijection between $\mathbb{K}(\!(X^{-1})\!)$ and (a subset of) continued fractions over $\mathbb{K}[X]$.

5.4. Continued fraction process

We now define a process which produces a continued fraction for elements of $\mathbb{K}(\!(X^{-1})\!)$, using the truncation $\lfloor\cdot\rfloor$ from Definition 2.5. This is mostly analogous to classical continued fractions over $\mathbb{Z}$, but slightly nicer because here we have a unique truncation operation, and we avoid ambiguity as for example with $[2] = 2 = 1 + \frac{1}{1} = [1,1]$ in the integer case.

Definition 5.11. Let $\alpha \in \mathbb{K}(\!(X^{-1})\!)$. We define the *complete quotients* of α as the (possibly finite) sequence

$$\alpha_0 = \alpha, \quad \alpha_{n+1} = \frac{1}{\alpha_n - \lfloor\alpha_n\rfloor} \qquad \text{for } n \geq 0 \text{ and } \alpha_n \notin \mathbb{K}[X]. \tag{5.8}$$

One defines also the *partial quotients* $a_n = \lfloor\alpha_n\rfloor$ whenever the corresponding complete quotient is defined. As $\alpha_n = a_n + \alpha_{n+1}^{-1}$, this clearly gives rise to a (finite or infinite) continued fraction

$$\mathbf{CF}(\alpha) = [a_0, a_1, \dots].$$

Remark 5.12. By definition of $\lfloor\cdot\rfloor$ we have always $\operatorname{ord}(\alpha_n - \lfloor\alpha_n\rfloor) > 0$ which implies $\operatorname{ord}(\alpha_{n+1}) < 0$ whenever α_{n+1} is defined. Then $\operatorname{ord}(a_{n+1}) = \operatorname{ord}(\alpha_{n+1}) < 0$ which means $\deg a_{n+1} \geq 1$. So if $\mathbf{CF}(\alpha)$ is an infinite continued fraction, it converges by Proposition 5.10.

The Euclidean algorithm works also in the ring $\mathbb{K}[X]$, establishing a complete correspondence between finite continued fraction and rational functions.

Proposition 5.13. *The continued fraction* $\mathbf{CF}(\alpha)$ *is finite if and only if* $\alpha \in \mathbb{K}(X)$*.*

Proof. If $\mathbf{CF}(\alpha)$ is finite, it produces an element of $\mathbb{K}(X)$, and obviously $\alpha = \mathbf{CF}(\alpha)$.

Conversely, assume $\alpha \in \mathbb{K}(X)$. Write $\alpha = \frac{r_0}{r_1}$ with $r_0, r_1 \in \mathbb{K}[X]$ and $r_1 \neq 0$. In fact, we can write $\alpha_n = \frac{r_n}{r_{n+1}}$ whenever defined, with $r_n, r_{n+1} \in \mathbb{K}[X]$.

Indeed, by Remark 2.8 we write $r_n = a_n\, r_{n+1} + r_{n+2}$ where $\deg r_{n+2} < \deg r_{n+1}$ because $a_n = \lfloor r_n / r_{n+1} \rfloor$. Hence

$$\alpha_n = \frac{r_n}{r_{n+1}} = a_n + \frac{r_{n+2}}{r_{n+1}} = a_n + \frac{1}{\alpha_{n+1}}.$$

So in this case, the continued fraction process corresponds to the Euclidean algorithm which is well known to terminate in a finite number of steps; so eventually $r_{n+1} = 0$ for some n which means that $\alpha_n \in \mathbb{K}[X]$ and that consequently $\mathbf{CF}(\alpha)$ is finite. □

5.5. Canonical convergents and classification of best-approximations

Definition 5.14. The sequence of *canonical convergents* of α is defined by

$$(p_n, q_n) = (p_{0,n}, q_{0,n}) \in \mathcal{Q}(\mathbb{K}) \quad \text{for } n \geq -1$$

where we plug the partial quotients into the formulas from Section 5.1. If $\mathbf{CF}(\alpha)$ is finite, this sequence is also finite.

Note that Corollary 5.5 implies that p_n and q_n are coprime for a given n. And the canonical convergents are in fact convergents, and we have precise information about their approximation quality:

Proposition 5.15. *Let* $n \geq 0$*. Then unless* $\alpha = p_n/q_n$*,*

$$\operatorname{ord}(p_n - \alpha\, q_n) = \deg q_{n+1} = \deg a_{n+1} + \deg q_n > \deg q_n,$$

so $(p_n, q_n) \in \mathcal{C}_\alpha(\mathbb{K})$*.*

Remark 5.16. If $\alpha = p_n/q_n$, then clearly $\operatorname{ord}(p_n - \alpha\, q_n) = \infty > \deg q_n$ and obviously $(p_n, q_n) \in \mathcal{C}_\alpha(\mathbb{K})$.

Proof. Unless $\mathbf{CF}(\alpha) = [a_0, \ldots, a_n]$ is finite of length exactly $n+1$ which directly implies $p_n - \alpha\, q_n = 0$ by the Proposition 5.13, we have

$$\alpha = [a_0, \ldots, a_n, \alpha_{n+1}] \quad \text{i.e. } \alpha = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \cdot \alpha_{n+1}.$$

Multiplying with the inverse matrix, we get the important formula

$$\alpha_{n+1} = (-1)^{n+1} \begin{pmatrix} q_{n-1} & -p_{n-1} \\ -q_n & p_n \end{pmatrix} \cdot \alpha = -\frac{p_{n-1} - \alpha\, q_{n-1}}{p_n - \alpha\, q_n}. \tag{5.9}$$

Recall that $p_{-1} = 1, q_{-1} = 0$, so a telescoping product yields

$$(-1)^{n+1} \prod_{j=0}^{n} \alpha_{j+1} = \prod_{j=0}^{n} \frac{p_{j-1} - \alpha\, q_{j-1}}{p_j - \alpha\, q_j} = \frac{1}{p_n - \alpha\, q_n}.$$

Taking valuations, note that $\operatorname{ord}(\alpha_j) = \operatorname{ord}(a_j) = -\deg a_j$ for $j \geq 1$, hence

$$\operatorname{ord}(p_n - \alpha\, q_n) = -\sum_{j=1}^{n+1} \operatorname{ord}(\alpha_j) = \sum_{j=1}^{n+1} \deg a_j = \deg a_{n+1} + \deg q_n = \deg q_{n+1},$$

the last two equalities being a consequence of Proposition 5.8. □

Proposition 5.17. *The continued fraction of α represents α as an element of $\mathbb{K}(\!(X^{-1})\!)$, i.e.*

$$\alpha = \mathbf{CF}(\alpha) \text{ in } \mathbb{K}(\!(X^{-1})\!).$$

Proof. For $\alpha \in \mathbb{K}(X)$, this is was mentioned in the proof of Proposition 5.13. Otherwise, $\alpha \notin \mathbb{K}(X)$, and from Proposition 5.15 we conclude $\alpha = \lim_{n\to\infty} \frac{p_n}{q_n} = \mathbf{CF}(\alpha)$ as $\lim_{n\to\infty} \deg q_n = \infty$. □

With this information about the approximation quality of the canonical convergents, we can now give a complete classification of the best-approximations.

Proposition 5.18 (Classification of best-approximations)**.** *Let $\alpha \in \mathbb{K}(\!(X^{-1})\!)\backslash\mathbb{K}(X)$, and $(p, q) \in \mathcal{B}_\alpha(\mathbb{K})$ a best-approximation. Then there exist a unique $n \in \mathbb{N}_0$ and $r \in \mathbb{K}[X]\backslash\{0\}$ with* $\deg r < \deg a_{n+1}$ *such that*

$$(p, q) = r \cdot (p_n, q_n).$$

In particular, if p and q are coprime, then $r \in \mathbb{K}^\times$.

Moreover, if $(p', q') = r' \cdot (p_{n'}, q_{n'}) \in \mathcal{B}_\alpha(\mathbb{K})$ is another best-approximation with $\deg q < \deg q'$*, then $n \leq n'$.*

Proof. With the sufficient condition for a best-approximation from Proposition 3.15 applied to (p_n, q_n) and $\xi = \deg a_{n+1}$, we see that for every possible $\deg q$ we can produce a best-approximation of the shape $r \cdot (p_n, q_n)$, with any $r \in \mathbb{K}[X]$ satisfying $0 \leq \deg r < \deg a_{n+1}$. Then by Proposition 3.12, all best-approximation have this shape. Because p_n and q_n are always coprime, and $\deg q_n$ is strictly increasing in n, no canonical convergent can be written as a multiple of another, so n must be unique.

Finally, the monotony result is obvious from $\deg q_n \leq \deg (r\, q_n) < \deg q_{n+1}$. □

Remark 5.19. If $\alpha \in \mathbb{K}(X)$, this argument works just as well, except for the last canonical convergent. However, if we put "$\deg a_{n+1} = \infty$", the statement trivially holds even for the last canonical convergent.

For completeness, we also give the analogue for convergents (applying Proposition 3.4 instead of Proposition 3.15):

Corollary 5.20. *Let $\alpha \in \mathbb{K}((X^{-1})) \setminus \mathbb{K}(X)$, and $(p, q) \in \mathcal{C}_\alpha(\mathbb{K})$ a convergent. Then there exist $n \in \mathbb{N}_0$ and $r \in \mathbb{K}[X] \setminus \{0\}$ with $\deg r < \frac{1}{2} \deg a_{n+1}$ such that*

$$(p, q) = r \cdot (p_n, q_n),$$

and if p and q are coprime, then $r \in \mathbb{K}^\times$.

5.6. Multiplication of a continued fraction by a constant

One nice feature of polynomial continued fractions is that it is possible to multiply them with a constant factor. In [Sch00], there is even a generalisation of this identity which holds also for non-constant factors. We limit ourselves to constants, however.

Proposition 5.21. *Let $\mu \in \mathbb{K}^\times$. Then*

$$\mu\, [a_0, a_1, a_2, a_3, \ldots] = [\mu\, a_0, \mu^{-1}\, a_1, \mu\, a_2, \mu^{-1}\, a_3, \ldots].$$

Proof. Again, it is convenient to think of the continued fraction as a product of matrices:

$$\begin{pmatrix} \mu & \\ & 1 \end{pmatrix} \begin{pmatrix} a & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} \mu\, a & \mu \\ 1 & \end{pmatrix} = \begin{pmatrix} \mu\, a & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} 1 & \\ & \mu \end{pmatrix},$$
$$\begin{pmatrix} 1 & \\ & \mu \end{pmatrix} \begin{pmatrix} a & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} a & 1 \\ \mu & \end{pmatrix} = \begin{pmatrix} \mu^{-1}\, a & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} \mu & \\ & 1 \end{pmatrix}.$$

As multiplication by μ corresponds to $\begin{pmatrix} \mu & \\ & 1 \end{pmatrix}$ and division by μ corresponds to $\begin{pmatrix} 1 & \\ & \mu \end{pmatrix}$, we obtain for n even

$$\begin{pmatrix} \mu & \\ & 1 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} \mu\, a_0 & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} \mu^{-1}\, a_1 & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} \mu\, a_n & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} 1 & \\ & \mu \end{pmatrix}$$

and for n odd

$$\begin{pmatrix} \mu & \\ & 1 \end{pmatrix} \begin{pmatrix} a_0 & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} a_n & 1 \\ 1 & \end{pmatrix} = \begin{pmatrix} \mu\, a_0 & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} \mu^{-1}\, a_1 & 1 \\ 1 & \end{pmatrix} \cdots \begin{pmatrix} \mu^{-1}\, a_n & 1 \\ 1 & \end{pmatrix} \begin{pmatrix} \mu & \\ & 1 \end{pmatrix}$$

so the corresponding map would be for n even

$$x \mapsto [\mu\, a_0, \mu^{-1}\, a_1, \ldots, \mu\, a_n, \mu^{-1}\, x]$$

and for n odd

$$x \mapsto [\mu\, a_0, \mu^{-1}\, a_1, \ldots, \mu^{-1}\, a_n, \mu\, x]$$

as desired – because for the empty continued fraction, we have $\mu \cdot [\,] = \frac{\mu}{0} = \frac{1}{0} = [\,]$. □

5.7. Periodic continued fractions

For classical continued fractions, it is a well-known result that continued fractions of quadratics are always periodic. As in the real case, a periodic polynomial continued fraction must be quadratic. However, a continued fraction of a quadratic need not be periodic in the polynomial case. For $\sqrt{D}$ this in fact happens if and only if D is Pellian which we will prove in Section 6.3.

Indeed periodicity gives a solution of (2.2) with $\omega = \pm 1$ (this follows from (6.11)). But if the base field $\mathbb{K}$ is very small, allowing arbitrary ω may give a solution with smaller $\deg q$. So one should not merely study periodicity, but periodicity up to a constant factor. We call this *quasi-periodicity* (sometimes it is also called pseudo-periodicity in the literature). For the continued fraction of $\sqrt{D}$, the period and the quasi-period are tightly linked, and one induces the other.

Later in Chapter 8, we will also see that quasi-periodicity is the more relevant notion for studying reductions of the continued fraction modulo a prime.

5.7.1. Periods

Definition 5.22. The (infinite) continued fraction $\alpha_m = \alpha_{m,\infty} = [a_m, a_{m+1}, \ldots]$ is said to be *periodic*, if for some $m' \geq m$ there exists $l \in \mathbb{N}$ (the minimal such l is called the *period length*) s.t.

$$\forall n \geq m' : \; a_n = a_{n+l}$$

If $m' = m$, the continued fraction is called *pure periodic*, i.e. there is no preperiod. For compact notation, we usually write "$\mathbf{CF}(\alpha_m)$ is (pure) periodic".

From a computational view, this definition is somewhat problematic because there is an infinite number of conditions to check. Fortunately, this can be reduced to a single condition on the complete quotients.

Proposition 5.23. *The following are equivalent:*

1. *The continued fraction $\mathbf{CF}(\alpha_m)$ is periodic.*
2. *There exist $m' \geq m$ and $l \in \mathbb{N}$ s.t. $\alpha_{m'} = \alpha_{m'+l}$.*
3. *There exist $m' \geq m$ and $l \in \mathbb{N}$ s.t. for all $n \geq m' : \; \alpha_n = \alpha_{n+l}$.*

Proof. Because $a_n = \lfloor \alpha_n \rfloor$, 3. directly implies 1.

On the other hand, $\alpha_{m'}$ is uniquely determined by $a_{m'}, a_{m'+1}, \ldots$, and by periodicity of the a_n one obtains

$$\alpha_{m'+l} = [a_{m'+l}, a_{m'+l+1}, \ldots] = [a_{m'}, a_{m'+1}, \ldots] = \alpha_{m'}.$$

so 1. implies 2.

But through the continued fraction process, α_{n+1} is uniquely determined by α_n for every n, so

$$\alpha_n = \alpha_{n+l} \implies \alpha_{n+1} = \alpha_{n+l+1}.$$

and by the induction principle, 2. implies 3. □

5.7.2. Quasi-periods

We now generalise periodicity to quasi-periodicity which is essentially periodicity up to a unit factor. For cleaner notation, we first define

$$\iota(n) = (-1)^n = \begin{cases} 1 & \text{if } n \text{ is even,} \\ -1 & \text{if } n \text{ is odd} \end{cases}.$$

Definition 5.24. The (infinite) continued fraction α_m is called *quasi-periodic*, if there exists $m' \geq m$, $\mu \in \mathbb{K}^\times$ and $l > 0$ (if minimal, called the *quasi-period length*) such that

$$\forall n \geq m' : \; a_n = \mu^{\iota(n)} \, a_{n+l}.$$

If $m' = m$, then it is called *pure quasi-periodic.*

Remark 5.25. Any periodic continued fraction is also quasi-periodic, with $\mu = 1$. See below for a partial converse.

Remark 5.26. It should be obvious that the $l \in \mathbb{Z}$ such that $\alpha_n = \mu^{\pm 1} \, \alpha_{n+l}$ form an ideal, and the (quasi-)period length is the positive generator of it.

In particular, the period length must be a multiple of the quasi-period length.

We also have a complete analogue to Proposition 5.23:

Proposition 5.27. *The following are equivalent:*

1. *The continued fraction* $\mathbf{CF}(\alpha_m)$ *is quasi-periodic.*
2. *There exist* $m' \geq m$*,* $\mu \in \mathbb{K}^\times$ *and* $l > 0$ *such that* $\alpha_{m'} = \mu^{\iota(m')} \, \alpha_{m'+l}$.
3. *There exist* $m' \geq m$*,* $\mu \in \mathbb{K}^\times$ *and* $l > 0$ *such that for all* $n \geq m' : \; \alpha_n = \mu^{\iota(n)} \, \alpha_{n+l}$.

Proof. Using Proposition 5.21, and $\lfloor \mu \, \alpha \rfloor = \mu \, \lfloor \alpha \rfloor$ for $\mu \in \mathbb{K}^\times$, and

$$\alpha_n = \mu \, \alpha_{n+l} \implies \alpha_{n+1} = \mu^{-1} \, \alpha_{n+l+1},$$

the proof is completely analogous to the one of Proposition 5.23. □

Proposition 5.28. *If* $\mathbf{CF}(\alpha_m)$ *is quasi-periodic with* odd *quasi-period length* l *and* $\mu \neq 1$*, then* $\mathbf{CF}(\alpha_m)$ *is also periodic with period length* $2\,l$.

Proof. For all $n \geq m'$, we have $a_n = \mu^{\iota(n)} \, a_{n+l}$ and $a_{n+l} = \mu^{\iota(n+l)} \, a_{n+2l}$. As l is odd, we have $\iota(n+l) = -\iota(n)$ so $a_n = \mu^{\iota(n)+\iota(n+l)} \, a_{n+2l} = a_{n+2l}$. □

Remark 5.29. The (quasi-)period length was above defined as the minimal l, and does not depend on where the (quasi-)period starts, so two complete quotients α_{m_1} and α_{m_2} have the same (quasi-)period length.

Proposition 5.30. *If* $\mathbf{CF}(\alpha)$ *is quasi-periodic, then* $\alpha \in \mathbb{K}((X^{-1}))$ *is quadratic over* $\mathbb{K}(X)$ *(it cannot be in* $\mathbb{K}(X)$ *because it has an infinite continued fraction).*

Proof. From Section 5.2 we know

$$\alpha_m = \begin{pmatrix} p_{m,n} & p_{m,n-1} \\ q_{m,n} & q_{m,n-1} \end{pmatrix} \cdot \alpha_{m+n+1} = \frac{p_{m,n}\,\alpha_{m+n+1} + p_{m,n-1}}{q_{m,n}\,\alpha_{m+n+1} + q_{m,n-1}}$$

so it suffices to treat the case where $\mathbf{CF}(\alpha)$ is pure quasi-periodic, i.e. $\alpha_l = \mu\,\alpha$. Then putting $m = 0$ and $n = l - 1$ the above becomes

$$\alpha = \begin{pmatrix} p_{l-1} & p_{l-2} \\ q_{l-1} & q_{l-2} \end{pmatrix} \cdot \mu\,\alpha = \frac{p_{l-1}\,\mu\,\alpha + p_{l-2}}{q_{l-1}\,\mu\,\alpha + q_{l-2}}.$$

Multiplying with the denominator, we then obtain

$$q_{l-1}\,\mu\,\alpha^2 + (q_{l-2} - p_{l-1}\,\mu)\,\alpha - p_{l-2} = 0$$

and of course $q_{l-1} \neq 0$ so α is quadratic over $\mathbb{K}(X)$. □

5.8. Reduced complete quotients

Our next goal is to understand the continued fraction expansion of $\sqrt{D}$ better. We will explain how we can usually go backwards in this continued fraction. This means we can only have very short preperiods (here just a_0 belongs to the preperiod), and allows to show that for $\mathbf{CF}(\sqrt{D})$, quasi-periodicity is equivalent to periodicity, also in the case of even quasi-period length.

In this case, the complete quotients are contained in the quadratic extension $\mathbb{K}(X, \sqrt{D})$ of $\mathbb{K}(X)$ contained in $\mathbb{K}(\!(X^{-1})\!)$. It has precisely one non-trivial $\mathbb{K}(X)$-automorphism σ which sends $\sqrt{D}$ to $-\sqrt{D}$. As we have chosen $\sqrt{D} \in \mathbb{K}(\!(X^{-1})\!)$, we have an embedding of $\mathbb{K}(X, \sqrt{D})$ into $\mathbb{K}(\!(X^{-1})\!)$.

Definition 5.31. $\alpha \in \mathbb{K}(X, \sqrt{D})$ is said to be *σ-reduced* (with σ as above), if

$$\operatorname{ord}(\sigma(\alpha)) > 0 > \operatorname{ord}(\alpha)\,.$$

Remark 5.32. All elements of $\mathbb{K}(X)$ are invariant under σ, so none of them is σ-reduced.

Proposition 5.33. *Let $\alpha \in \mathbb{K}(X, \sqrt{D}) \setminus \mathbb{K}(X)$. Then there exists at most one $a \in \mathbb{K}[X]$ s.t. $a + \alpha$ is σ-reduced.*

Proof. Clearly $\sigma(a + \alpha) = a + \sigma(\alpha)$. Assume $\operatorname{ord}(\sigma(a + \alpha)) = \operatorname{ord}(a + \sigma(\alpha)) > 0$, then by Remark 2.6 $a = -\lfloor \sigma(\alpha) \rfloor$ so there is at most one choice for a. □

Note that this choice of a does not yet guarantee $\operatorname{ord}(a + \alpha) < 0$.

Proposition 5.34. *If the complete quotient α_m is σ-reduced, then so is $\alpha_{m'}$ for all $m' \geq m$.*

Proof. Using the induction principle, it suffices to treat the case $m' = m+1$. By Remark 5.12, we automatically have $\operatorname{ord}(\alpha_{m+1}) < 0$. Moreover,

$$\sigma(\alpha_{m+1}) = \frac{1}{\sigma(\alpha_m) - a_m}$$

and $\operatorname{ord}(a_m) = \operatorname{ord}(\alpha_m) < \operatorname{ord}(\sigma(\alpha_m))$ implies $\operatorname{ord}(\sigma(\alpha_{m+1})) = -\operatorname{ord}(\alpha_m) > 0$ as desired. □

Lemma 5.35. *α is σ-reduced if and only if $\frac{-1}{\sigma(\alpha)}$ is σ-reduced.*

Proof. This is an immediate consequence of

$$\operatorname{ord}(\alpha) = -\operatorname{ord}\left(\sigma\left(\frac{-1}{\sigma(\alpha)}\right)\right), \qquad \operatorname{ord}(\sigma(\alpha)) = -\operatorname{ord}\left(\frac{-1}{\sigma(\alpha)}\right).$$

□

The most useful property of σ-reduced complete quotients is however that we may go backwards in the continued fraction expansion in a unique way:

Proposition 5.36. *Suppose $\alpha_1 \in \mathbb{K}(X, \sqrt{D})$ is σ-reduced. Then there exists a unique $\alpha_0 \in \mathbb{K}(X, \sqrt{D})$ which is σ-reduced and satisfies*

$$\alpha_1 = \frac{1}{\alpha_0 - \lfloor \alpha_0 \rfloor}.$$

Proof. By Proposition 5.33, there exists at most one $a_0 \in \mathbb{K}[X]$ such that $\alpha_0 = a_0 + \frac{1}{\alpha_1}$ is σ-reduced, namely $a_0 = \left\lfloor \frac{-1}{\sigma(\alpha_1)} \right\rfloor$. Rewriting this to

$$\frac{-1}{\sigma(\alpha_0)} = \frac{1}{\frac{-1}{\sigma(\alpha_1)} - a_0},$$

we see that α_0 is σ-reduced by applying twice Lemma 5.35 and once Proposition 5.34.

Finally, as $\operatorname{ord}(\alpha_1) < 0$ it is also clear that $a_0 = \lfloor \alpha_0 \rfloor$. □

Remark 5.37. Generally, for any n and α_n σ-reduced, we have

$$\frac{-1}{\sigma(\alpha_n)} = \frac{1}{\frac{-1}{\sigma(\alpha_{n+1})} - \left\lfloor \frac{-1}{\sigma(\alpha_{n+1})} \right\rfloor}$$

so also the $\frac{-1}{\sigma(\alpha_n)}$ are the complete quotients of some continued fraction expansion, albeit with n decreasing.

Lemma 5.38. *Suppose α_m is σ-reduced and $\mathbf{CF}(\alpha_m)$ is (quasi-)periodic, then $\mathbf{CF}(\alpha_m)$ is pure (quasi-)periodic.*

Proof. We use Propositions 5.23 and 5.27 here.

Suppose $n > m, l \in \mathbb{N}$ and $\mu \in \mathbb{K}^\times$ (where $\mu = 1$ in the case of periodicity) with $\alpha_n = \mu^{\iota(n)}\,\alpha_{n+l}$. By Proposition 5.34, $\alpha_{n-1}, \alpha_n, \alpha_{n+l-1}, \alpha_{n+l}$ are all σ-reduced, and we have

$$\begin{aligned}\alpha_n = \frac{1}{\alpha_{n-1} - a_{n-1}} &= \mu^{\iota(n)}\,\alpha_{n+l}\\ &= \mu^{\iota(n)}\,\frac{1}{\alpha_{n+l-1} - a_{n+l-1}} = \frac{1}{\mu^{\iota(n-1)}\,\alpha_{n+l-1} - \mu^{\iota(n-1)}\,a_{n+l-1}}.\end{aligned}$$

With $\lfloor \mu^{\iota(n-1)}\,\alpha_{n+l-1}\rfloor = \mu^{\iota(n-1)}\,a_{n+l-1}$, Proposition 5.36 implies $\alpha_{n-1} = \mu^{\iota(n-1)}\,\alpha_{n+l-1}$ as desired, and we may repeat this argument until we arrive at $\alpha_m = \mu^{\iota(m)}\,\alpha_{m+l}$. □

Theorem 5.1. *Suppose $\alpha \in \mathbb{K}(X, \sqrt{D})$ is σ-reduced and has polynomial trace $\alpha + \sigma(\alpha) \in \mathbb{K}[X]$. If $\mathbf{CF}(\alpha)$ is quasi-periodic, it is even pure (quasi-)periodic.*

Proof. Lemma 5.38 already implies that $\mathbf{CF}(\alpha)$ is pure quasi-periodic, and once we prove it is periodic, it is automatically pure periodic. For odd quasi-period length, the general Proposition 5.28 already yields periodicity. For even quasi-period length, a bit more work is required.

From $\lfloor f\rfloor = f$ for $f \in \mathbb{K}[X]$ and $\operatorname{ord}(\sigma(\alpha)) > 0$ we obtain

$$\alpha + \sigma(\alpha) = \lfloor \alpha + \sigma(\alpha)\rfloor = \lfloor \alpha\rfloor = a_0$$

so $\alpha - a_0 = -\sigma(\alpha)$ which implies

$$\alpha_1 = \frac{-1}{\sigma(\alpha_0)} \text{ and thus } \alpha_0 = \frac{-1}{\sigma(\alpha_1)}.$$

In the light of Remark 5.37, the $\frac{-1}{\sigma(\alpha_n)}$, going backwards, are complete quotients of some continued fraction expansion and actually extend $\mathbf{CF}(\alpha)$ for negative n:

$$\begin{array}{cccccccc} \cdots & \frac{-1}{\sigma(\alpha_3)} & \frac{-1}{\sigma(\alpha_2)} & \frac{-1}{\sigma(\alpha_1)} & \frac{-1}{\sigma(\alpha_0)} & & & \\ & & & \alpha_0 & \alpha_1 & \alpha_2 & \alpha_3 & \cdots \end{array}$$

So we can define $\alpha_n = \frac{-1}{\sigma(\alpha_{1-n})}$ for $n \leq 1$, with all α_n σ-reduced, and by Lemma 5.38 the quasi-periodicity extends towards $-\infty$ as well.

Denote by ℓ the quasi-period length of $\mathbf{CF}(\alpha)$, so we may write

$$\alpha_0 = \mu\,\alpha_\ell, \qquad \alpha_\ell = \mu^{\iota(\ell)}\,\alpha_{2\ell}, \qquad \alpha_{1-\ell} = \mu^{\iota(1-\ell)}\,\alpha_1.$$

It follows

$$\alpha_\ell = \frac{-1}{\sigma(\alpha_{1-\ell})} = \frac{1}{\mu^{\iota(1-\ell)}}\,\frac{-1}{\sigma(\alpha_1)} = \mu^{\iota(\ell)}\,\alpha_0$$

and further $\alpha_0 = \mu\,\mu^{\iota(\ell)}\,\alpha_0$. Hence $\mu\,\mu^{\iota(\ell)} = 1$ (if ℓ is even, this means $\mu = \pm 1$), and then $\alpha_0 = \mu\,\alpha_\ell = \mu\,\mu^{\iota(\ell)}\,\alpha_{2\ell} = \alpha_{2\ell}$, so $\mathbf{CF}(\alpha)$ is periodic (with period length ℓ or 2ℓ). □

Remark 5.39. This shows that the involution $x \mapsto \frac{-1}{\sigma(x)}$ acts as a reflection with centre $1/2$ on the $\mathbb{Z}$ -series of α_n ($n \mapsto 1-n$ on the indices).

Remark 5.40. Obviously $\sqrt{D}$ is not σ-reduced. However $\alpha = A + \sqrt{D}$ (recall that $A = \left\lfloor \sqrt{D} \right\rfloor$) is σ-reduced, and $\sqrt{D} - \left\lfloor \sqrt{D} \right\rfloor = \alpha - \lfloor \alpha \rfloor$, so

$$\mathbf{CF}(\sqrt{D}) = [A, a_1, a_2, \dots]$$

differs from $\mathbf{CF}(\alpha)$ only in the first complete (and partial) quotient. This means that if $\mathbf{CF}(\sqrt{D})$ is quasi-periodic, it is almost pure periodic, and the preperiod has length 1 and consists just of a_0.

This reversibility of the continued fraction process also implies that the period must be a palindrome:

Proposition 5.41. *Let $\alpha \in \mathbb{K}(X, \sqrt{D})$ σ-reduced with $\alpha + \sigma(\alpha) \in \mathbb{K}[X]$. Let ℓ the quasi-period length.*

- *If ℓ is even, then $\mathbf{CF}(\alpha)$ has actually period length ℓ, and the period is palindromic, i.e.*
$$\mathbf{CF}(\alpha) = \left[\overline{a_0, a_1, \dots, a_{\ell/2}, \dots, a_1}\right]$$
- *If ℓ is odd, then $\mathbf{CF}(\alpha)$ has a "quasi-palindromic" quasi-period, i.e.*
$$\mathbf{CF}(\alpha) = \left[\overline{a_0, a_1, \dots, a_{(\ell-1)/2}, \mu^{\pm 1}\, a_{(\ell-1)/2}, \mu^{\mp 1}\, a_{(\ell-3)/2}, \dots, \mu\, a_1}\right]$$

Remark 5.42. In the second case, either $\mu = 1$, or the period length 2ℓ is even. Then we can apply the first case for the period instead of the quasi-period to get a palindromic period.

Proof. Recall how we defined the negative complete quotients, whence for any $n \in \mathbb{Z}$

$$\alpha_n = \frac{-1}{\sigma(\alpha_{1-n})} = \sigma\left(-\frac{1}{\alpha_{1-n}}\right) = \sigma\left(\alpha_{-n} - a_{-n}\right) = a_{-n} + \frac{1}{\frac{-1}{\sigma(\alpha_{-n})}} = a_{-n} + \frac{1}{\alpha_{n+1}},$$

the crux of which is $a_n = \lfloor \alpha_n \rfloor = a_{-n}$.

Using quasi-periodicity, we then obtain

$$a_n = a_{-n} = \mu^{\iota(-n)}\, a_{\ell-n} = \mu^{\iota(n)}\, a_{\ell-n}$$

and developing this for $n \leq \ell/2$ we obtain

$$a_0 = \mu\, a_\ell, \quad a_1 = \mu^{-1}\, a_{\ell-1}, \quad a_2 = \mu\, a_{\ell-2}, \quad \dots$$

until for ℓ odd we arrive at $a_{(\ell-1)/2} = \mu^{\iota((\ell-1)/2)}\, a_{(\ell+1)/2}$ and for ℓ even we arrive at $a_{\ell/2} = \mu^{\iota(\ell/2)}\, a_{\ell/2}$ which also implies $\mu = 1$. □

6. Computation of hyperelliptic continued fractions

We now give formulas for computing the continued fraction expansion for quadratic Laurent series. Optimising these formulas is not only useful for computing and studying examples, but it also serve to illustrate the connection between the Pell equation and periodicity of the continued fraction. Of particular interest is also that everything can be expressed as operations on polynomials.

We assume as usual that D is non-square of degree $2d$ and that $\ell c(D)$ is a square in $\mathbb{K}$, a field of characteristic not 2. Recall that we defined $A = \left\lfloor \sqrt{D} \right\rfloor$.

It is well-known that the complete quotients of $\sqrt{D}$ can be written as $\alpha_n = (r_n + \sqrt{D})/s_n$ with $r_n, s_n \in \mathbb{K}[X]$ of bounded degree. We can slightly improve upon this representation by writing $r_n = A +$ terms of lower degree. This seems to be a new result:

Theorem 6.1. *Let $\alpha = \sqrt{D}$. The complete quotients of α can be written as*

$$\alpha_n = \frac{A + t_n + \sqrt{D}}{s_n} \qquad \text{for } n \geq 1 \tag{6.1}$$

where $t_n, s_n \in \mathbb{K}[X]$ with

$$\deg t_n < \deg s_n < \deg A \tag{6.2}$$

for $n \geq 1$. Moreover, there are the following recursion formulas for t_n and s_n:

$$t_n + t_{n+1} = a_n \, s_n - 2\, A, \quad s_n \, s_{n+1} = D - (A + t_{n+1})^2, \tag{6.3}$$

initialised with $t_0 = -A$ and $s_0 = 1$. Finally $\deg s_n = 0$ for $n \geq 1$ if and only if $\mathbf{CF}(\alpha)$ is periodic and the quasi-period length ℓ divides n.

Note that α_n being σ-reduced is equivalent to (6.2) by Proposition 6.6.

Corollary 6.1. *The complete quotients satisfy $\operatorname{ord}(\alpha_n) \geq \operatorname{ord}\left(\sqrt{D}\right)$, so for the partial quotients we have*

$$1 \leq \deg a_n \leq \deg A$$

with equality $\deg a_n = \deg A$ for $n \geq 1$ if and only if the continued fraction $\mathbf{CF}(\sqrt{D})$ is periodic, and the quasi-period length ℓ divides n.

In fact, we show more generally:

Theorem 6.2. *Let $\alpha \in \mathbb{K}(\!(X^{-1})\!)$ any Laurent series quadratic over $\mathbb{K}(X)$. Then for a suitable D depending only on α, the complete quotients α_n may also be written as in* (6.1)*, where t_n and s_n follow the recursion formulas* (6.3)*.*

Moreover, there exists $N \geq 0$, such that t_n and s_n satisfy (6.2) *for $n \geq N$.*

The theorem also gives a more elementary proof of periodicity over finite fields:

Corollary 6.2. *If the base field $\mathbb{K}$ is finite, any Laurent series quadratic over $\mathbb{K}(X)$ has a periodic continued fraction expansion.*

Using this representation of the complete quotients of $\sqrt{D}$, and our accumulated knowledge about the convergents, we also recover Abel's result from [Abe26]:

Theorem 6.3 (Abel 1826)**.** *D is Pellian if and only if $\mathbf{CF}(\sqrt{D})$ is periodic.*

6.1. Representing complete quotients with polynomials

We begin by reiterating the formulas which can (with varying level of detail) be already found in [Abe26], [Ber90] and [vdPT00].

Let $\alpha \in \mathbb{K}(\!(X^{-1})\!)$ quadratic over $\mathbb{K}(X)$, satisfying $s\,\alpha^2 - 2\,r\,\alpha + w = 0$ where $r, s, w \in \mathbb{K}[X]$. The discriminant $4\,D = 4\,(r^2 - s\,w)$ yields D, for which we choose a square root $\sqrt{D}$. Then we write

$$\alpha = \frac{r + \sqrt{D}}{s} \tag{6.4}$$

after possibly multiplying r, s, w with -1 to accommodate our choice of $\sqrt{D}$. Note that $s \mid D - r^2$ holds which is crucial for the following computations. This allows a common factor in r and s which then must divide D as well.

Clearly α is determined by the polynomials r, s, D and our choice of $\sqrt{D}$. For example for $\alpha = \sqrt{D}$ we just put $r = 0$, $s = 1$ and $w = -D$.

All complete quotients of a given α can be written in this way; all of them with the same discriminant D:

Proposition 6.3. *The complete quotients of α as in* (6.4) *have for all $n \geq 0$ the form*

$$\alpha_n = \frac{r_n + \sqrt{D}}{s_n}, \qquad \text{where } s_n \mid (D - r_n^2) \text{ and } r_n, s_n \in \mathbb{K}[X]. \tag{6.5}$$

Proof. We prove this using complete induction. For $n = 0$ we may take $r_0 = r$ and $s_0 = s$ which satisfy the desired conditions by hypothesis.

Suppose (6.5) holds for n. Then write

$$\frac{1}{\alpha_{n+1}} = \alpha_n - a_n = \left(\frac{r_n + \sqrt{D}}{s_n} - a_n\right)\left(\frac{-r_n + \sqrt{D} + a_n\,s_n}{-r_n + \sqrt{D} + a_n\,s_n}\right)$$

$$= \frac{D - (r_n - a_n\,s_n)^2}{s_n\left(a_n\,s_n - r_n + \sqrt{D}\right)}$$

and note that

$$D - r_{n+1}^2 = D - a_n^2\, s_n^2 + 2\, a_n\, s_n\, r_n - r_n^2$$

so by induction hypothesis $s_n \mid D - r_n^2$, this is divisible by s_n and we can set

$$r_{n+1} = a_n\, s_n - r_n, \qquad s_{n+1} = \frac{D - r_{n+1}^2}{s_n}. \tag{6.6}$$

with $r_{n+1}, s_{n+1} \in \mathbb{K}[X]$ and moreover $s_{n+1} \mid D - r_{n+1}^2$. This concludes the induction step. □

Remark 6.4. It should be quite obvious that the discriminant does not change for the complete quotients. After all, the discriminant is invariant under the natural action of $\mathrm{GL}_2(\mathbb{K}(X))$ by linear change of variables on bilinear forms in two variables over $\mathbb{K}(X)$. Such a bilinear form gives of course a minimal polynomial for a quadratic α. But advancing in the continued fraction expansion can exactly be expressed in terms of this action, as seen in Section 5.2.

Berry (and Abel for $\deg D = 4$) give further simplifications of these formulas, see [Ber90] and [Abe26].

We prefer to perform simplifications of a different kind. And we still need to explain how to compute the a_n from our representation.

We may rewrite (6.5) as

$$\alpha_n = \frac{A + t_n + \sqrt{D}}{s_n} \tag{6.7}$$

by setting $t_n = r_n - A$. The recursion formulas (6.6) then obviously change to

$$\begin{aligned} t_0 &= r - A, & t_{n+1} &= a_n\, s_n - 2\, A - t_n, \\ s_0 &= s, & s_{n+1} &= \frac{D - A^2 - 2\, A\, t_{n+1} - t_{n+1}^2}{s_n}. \end{aligned} \tag{6.8}$$

This proves the first half of Theorem 6.2.

Proposition 6.5. *We can compute t_{n+1} and a_n with a single polynomial division, i.e.*

$$(2\, A + t_n) = a_n\, s_n - t_{n+1} \textit{ with } \deg t_{n+1} < \deg s_n.$$

Proof. Recall from (2.4) that $\sqrt{D} = A + \varepsilon$ with $\operatorname{ord}(\varepsilon) > 0$. The equality follows directly from the formula for t_{n+1}, it remains to verify $\deg t_{n+1} < \deg s_n$. Using $\lfloor \varepsilon \rfloor = 0$ and Remark 2.7 ($\lfloor \cdot \rfloor$ is a homomorphism with respect to $+$) we find

$$a_n = \lfloor \alpha_n \rfloor = \left\lfloor \frac{A + t_n + \sqrt{D}}{s_n} \right\rfloor = \left\lfloor \frac{2\, A + t_n + \varepsilon}{s_n} \right\rfloor = \left\lfloor \frac{2\, A + t_n}{s_n} \right\rfloor.$$

So by Remark 2.8 (taking $\lfloor \cdot \rfloor$ of rational functions corresponds to polynomial division) $-t_{n+1}$ must the remainder of the polynomial division of $2\, A + t_n$ by s_n. □

6.2. Complete quotients are eventually σ-reduced

The representation (6.7) also gives a very simple way to check if some complete quotient is σ-reduced:

Proposition 6.6. $\alpha = \frac{A+t+\sqrt{D}}{s}$ *is σ-reduced if and only if*

$$\deg t < \deg s < \deg A, \tag{6.9}$$

and in this case $\operatorname{ord}(\alpha) = \deg s - \deg A$.

Proof. With $A - \sqrt{D} = -\varepsilon$ we note that

$$\operatorname{ord}(\sigma(\alpha)) = \operatorname{ord}\Big(\Big(A + t - \sqrt{D}\Big)/s\Big) = \operatorname{ord}(t - \varepsilon) - \operatorname{ord}(s) = \operatorname{ord}(t - \varepsilon) + \deg s.$$

Hence $0 < \operatorname{ord}(\sigma(\alpha))$ is equivalent to $\deg t < \deg s$: In the case $t = 0$, using $\operatorname{ord}(\varepsilon) > 0$ we have $\operatorname{ord}(\sigma(\alpha)) = \operatorname{ord}(\varepsilon) + \deg s > 0$ if and only if we have $s \neq 0$, i.e. $\deg s > -\infty = \deg 0$. If on the other hand $t \neq 0$, then $\operatorname{ord}(t - \varepsilon) = \operatorname{ord}(t) = -\deg t$, hence $\operatorname{ord}(\sigma(\alpha)) = \deg s - \deg t$.

So for the rest of the proof, we can assume $\operatorname{ord}(\sigma(\alpha)) > 0$.

We may write

$$\begin{aligned}\operatorname{ord}(\alpha) = \operatorname{ord}\Big(\Big(A + t + \sqrt{D}\Big)/s\Big) &= \operatorname{ord}\Big(2\sqrt{D} + t - \varepsilon\Big) - \operatorname{ord}(s) \\ &\geq \min\Big(\operatorname{ord}\Big(2\sqrt{D}\Big), \operatorname{ord}(t - \varepsilon)\Big) + \deg s.\end{aligned}$$

If α is σ-reduced, then $0 > \operatorname{ord}(\alpha) = \operatorname{ord}\Big(2\sqrt{D}\Big) + \deg s$ because $\operatorname{ord}(t - \varepsilon) + \deg s > 0$. Hence $\deg s < \deg A = -\operatorname{ord}\Big(\sqrt{D}\Big)$.

Conversely, if $\deg s < \deg A$, then $\operatorname{ord}(\alpha) = \operatorname{ord}\Big(2\sqrt{D}\Big) + \deg s < 0$ as desired.

As $\operatorname{ord}\Big(2\sqrt{D}\Big) = \operatorname{ord}(A) = -\deg A$, we also showed $\operatorname{ord}(\alpha) = \deg s - \deg A$. □

An immediate and important consequence is that the degrees of the partial quotients of a σ-reduced α are always bounded uniformly – once we show that every continued fraction of a quadratic α eventually becomes σ-reduced, this means all partial quotients have bounded degree.

Corollary 6.7. *Suppose α as above is σ-reduced, and $a = \lfloor \alpha \rfloor$. Then* $0 < \deg a \leq \deg A$. *Moreover, if* $\deg a = \deg A$, *then there exists $\mu \in \mathbb{K}^\times$ such that $\alpha = \mu\,(A + \sqrt{D})$.*

Proof. From α being σ-reduced, the preceding Proposition yields

$$0 > \operatorname{ord}(\alpha) = \deg s - \deg A \geq -\deg A.$$

But $\operatorname{ord}(\alpha) = \operatorname{ord}(a) = -\deg a$, hence $0 < \deg a \leq \deg A$.

Additionally, if $\deg a = \deg A$ this means $\deg s = 0$ and thus $t = 0$. So we get $\mu = s^{-1} \in \mathbb{K}^\times$. □

The second half of Theorem 6.2 follows from

Proposition 6.8. *Let $\alpha \in \mathbb{K}((X^{-1}))$ quadratic over $\mathbb{K}(X)$. Then there exist $N \in \mathbb{N}$ such that for all $n \geq N$, the complete quotient α_n is σ-reduced.*

Proof. Using Proposition 6.6, this boils down to an analysis of the degrees of t_n and s_n.

From Proposition 6.5 follows $\deg t_{n+1} < \deg s_n$. Recall from Remark 5.12 that $\deg a_n \geq 1$ for $n \geq 1$, hence

$$\begin{aligned}\deg t_{n+1} < \deg s_n < \deg (a_n\, s_n) &= \deg (2\,A + t_n + t_{n+1}) \\ &\leq \max(\deg A, \deg t_n, \deg t_{n+1}) = \max(\deg A, \deg t_n).\end{aligned}$$

So if $\deg t_n \geq \deg A$ then $\deg t_{n+1} + 2 \leq \deg t_n$. So after a finite number of steps we must have $\deg t_{n+j} < \deg A$ (actually $\deg t_{n+j} + 2 \leq \deg A$). But if $\deg t_n < \deg A$, then clearly also $\deg t_{n+1} < \deg A$ (actually $\deg t_{n+1} + 2 \leq \deg A$).

So we may now assume $\deg t_n < \deg A$ for all n large enough.

Next, if $t_{n+1} = 0$, then $s_n\, s_{n+1} = D - A^2$ and $\deg \left(D - A^2\right) < \deg A$ (see Proposition 2.11). This implies $\deg s_n + \deg s_{n+1} < \deg A$, so clearly $\deg s_{n+1} < \deg A$, and trivially $-\infty = \deg t_{n+1} < \deg s_{n+1}$, hence α_{n+1} is σ-reduced.

If on the other hand $t_{n+1} \neq 0$, then $s_n\, s_{n+1} = D - A^2 - 2\,A\,t_{n+1} - t_{n+1}^2$ and thus

$$\begin{aligned}\deg s_n + \deg s_{n+1} &= \max(\deg \left(D - A^2\right), \deg A + \deg t_{n+1}, 2\,\deg t_{n+1}) \\ &= \deg A + \deg t_{n+1} < \deg A + \deg s_n\end{aligned}$$

implies $\deg s_{n+1} < \deg A$. If moreover $\deg s_n < \deg A$ (if not, consider s_{n+2} and s_{n+1} instead), we also get $\deg t_{n+1} < \deg s_{n+1}$ and so α_{n+1} is σ-reduced.

All subsequent complete quotients then remain σ-reduced by Proposition 5.34. □

Remark 6.9. From the proof, we easily deduce an effective bound for N. The degree of t_n decreases by at least 2 in every step from t_1, so at most $(\deg t_1 - \deg A)\,/2$ steps are required to arrive at $\deg t_n < \deg A$. From there, we need only one or two additional steps to arrive at a σ-reduced complete quotient. So $N \leq 3 + (\deg t_1 - \deg A)\,/2$. This shows the effectivity in Theorem 6.2.

The σ-reduced case allows also even simpler computation of the partial quotient:

Remark 6.10. If α_n is σ-reduced, then we may use polynomial division of $2\,A$ by s_n to compute t_{n+1} (improving minimally upon 6.5):

$$2\,A = a_n\, s_n - (t_n + t_{n+1}),$$

as both $\deg t_n, \deg t_{n+1} < \deg s_n$.

6.3. Periodicity and Pell equation

Let us now check the theorems given at the beginning of this chapter.

Proof of Theorem 6.1. We expand upon Remark 5.40, and work with $A+\sqrt{D}$ instead of $\sqrt{D}$. This changes only a_0 and α_0. Of course $A+\sqrt{D}$ has $t_0 = 0$ and $s_0 = 1$ which shows again (now using Proposition 6.6) that it is σ-reduced, hence also all complete quotients α_n with $n \geq 1$ are σ-reduced.

Then Theorem 6.1 simply combines (6.7), (6.8) (which follow from Proposition 6.3) and Proposition 6.6.

Additionally, Theorem 5.1 implies that $\mathbf{CF}(\sqrt{D})$ is periodic if and only if $\mathbf{CF}(A+\sqrt{D})$ is pure quasi-periodic, and both continued fraction have the same quasi-period length ℓ. With Proposition 5.27 and Corollary 6.7 it follows that $\alpha_n = \frac{A+\sqrt{D}}{s_n}$ with $s_n \in \mathbb{K}^\times$ (i.e. $\deg s_n = 0$) holds if and only if $\ell \mid n$ from minimality of the quasi-period length ℓ. □

We give a few more details for

Proof of Corollary 6.1. The degree inequalities were stated already in Corollary 6.7 and follow from $\deg a_n = \deg A - \deg s_n$. The corollary also says that $\deg a_n = \deg A$ implies pure quasi-periodicity of $\mathbf{CF}(A+\sqrt{D})$. □

Proof of Theorem 6.3. Set $\alpha = \sqrt{D}$, and recall from Section 5.2 that (for $n \geq 1$)

$$\sqrt{D} = \begin{pmatrix} p_{n-1} & p_{n-2} \\ q_{n-1} & q_{n-2} \end{pmatrix} \alpha_n \iff \alpha_n = (-1)^n \begin{pmatrix} q_{n-2} & -p_{n-2} \\ -q_{n-1} & p_{n-1} \end{pmatrix} \sqrt{D}$$

which we rewrite as

$$\begin{aligned} \alpha_n = \frac{q_{n-2}\sqrt{D} - p_{n-2}}{p_{n-1} - q_{n-1}\sqrt{D}} &= \frac{q_{n-2}\sqrt{D} - p_{n-2}}{p_{n-1} - q_{n-1}\sqrt{D}} \cdot \frac{p_{n-1} + q_{n-1}\sqrt{D}}{p_{n-1} + q_{n-1}\sqrt{D}} \\ &= \frac{D\,q_{n-1}\,q_{n-2} - p_{n-1}\,p_{n-2} + \sqrt{D}\,(p_{n-1}\,q_{n-2} - p_{n-2}\,q_{n-1})}{p_{n-1}^2 - D\,q_{n-1}^2} \\ &= \frac{(-1)^n\,(\dots) + \sqrt{D}}{(-1)^n\,\big(p_{n-1}^2 - D\,q_{n-1}^2\big)} \end{aligned} \tag{6.10}$$

so

$$s_n = (-1)^n\,\big(p_{n-1}^2 - D\,q_{n-1}^2\big). \tag{6.11}$$

Recall Theorem 5.1 which states that periodicity and quasi-periodicity are equivalent in the current situation. So by Corollary 6.7 (proved just above), it follows that $\mathbf{CF}(\sqrt{D})$ is periodic if and only if for some $n \geq 1$ we have $\deg s_n = 0$ which means (p_{n-1}, q_{n-1}) solves (2.2).

On the other hand, we know that Pell solutions are convergents (Proposition 3.9) and from the classification of convergents (Proposition 5.20) follows that every non-trivial solution of (2.2) has the shape $(p,q) = \mu \cdot (p_m, q_m)$ for some $m \geq 0$ with $\mu \in \mathbb{K}^\times$ (because for a Pell solution p, q are coprime). This implies that (p_m, q_m) likewise solves (2.2), and then $\deg s_{m+1} = 0$. □

6.4. Torsion order and period length

Recall the notation from Chapter 4, and assume again that D is square-free. With $2(g+1) = \deg D$, we get the following inequalities between the torsion order and the quasi-period length:

Proposition 6.11. *Suppose $[\mathbf{O}] \in \mathcal{J}$ is torsion of order precisely m, and let ℓ the quasi-period length of $\mathbf{CF}(\sqrt{D})$. Then for $g \geq 1$ we have the inequality*

$$g + \ell \leq m \leq 1 + g\,\ell$$

which for $g = 1$ becomes the equality $m = \ell + 1$.

Proof. Combining the knowledge from the proofs of Theorems 4.1 and 6.3, we know that the minimal n such that (4.2) is satisfied with $r = 0$ by (p_{n-1}, q_{n-1}) is exactly $n = \ell$, with $m = \deg p_{n-1}$. So this m must be the torsion order of $[\mathbf{O}]$.

We then calculate, using $1 \leq \deg a_i \leq g$ for $i = 1, \dots, l-1$,

$$\begin{aligned} m = \deg p_{l-1} = \deg a_0 + \deg q_{l-1} = g + 1 + \deg q_{l-1} &\leq g + 1 + (l-1)g = 1 + l\,g \\ &\geq g + 1 + l - 1 = l + g \end{aligned}$$

which yields the desired inequality. Clearly it collapses to an equality for $g = 1$. □

So bounding the period length is as hard as bounding torsion.

6.5. Period lengths over finite fields

We now give an (elementary) proof of Corollary 6.2, by showing that there are only finitely many possibilities for the σ-reduced complete quotients. These form the tail of the continued fraction of a quadratic Laurent series, and any repetition immediately implies periodicity. Of course we have to avoid characteristic 2 again.

Proposition 6.12. *Let $\mathbb{K} = \mathbb{F}_q$ a finite field of odd characteristic, and recall that $\deg D = 2d$. Then for a fixed D, there are precisely*

$$\frac{q^{2d} - 1}{q + 1} \tag{6.12}$$

σ-reduced expressions of type $\left(A + t + \sqrt{D}\right)/s$.

Proof. For fixed $e = \deg s$, there are $(q-1)\,q^e$ possibilities for s, and as $\deg t < \deg s$, there are q^e possibilities for t. Summing over e, we compute

$$\sum_{e=0}^{d-1} (q-1) q^e\, q^e = (q-1)\,\frac{q^{2d} - 1}{q^2 - 1} = \frac{q^{2d} - 1}{q + 1}$$

using the formula for geometric sums. □

Remark 6.13. Note that the above does not yet take into account that we usually also have the condition $s \mid D - r^2$. This further limits the number of possible complete quotients.

Remark 6.14. The above (6.12) gives an elementary bound for the period length. Using our knowledge about quasi-periods, we could improve it further dividing by $2/(q-1)$.

But anyway we already have a far better bound for for the torsion order in the Jacobian (under the assumption that D is square-free), see Remark 4.13.

Then we can do much better:

Corollary 6.15. *If D is square-free, the quasi-period length is bounded by*

$$\ell \leq m - g \leq (\sqrt{q}+1)^{2g} - g.$$

6.6. Divisors of complete quotients

We now wish to expand upon the results of Section 4.3, and make the connection between the convergent divisors and the continued fraction more explicit. This will be useful later to give an additional viewpoint on the reduction of continued fractions. See also [Ber90], where it is shown that quasi-periodicity of arbitrary elements of $\mathbb{K}(X, \sqrt{D}) \setminus \mathbb{K}(X)$ is equivalent to D being Pellian.

Recall the notation from Chapter 4, and the additional assumption that D is square-free.

Let $\alpha = \frac{r+wY}{s} \in \mathbb{K}(X, Y)$ an arbitrary element of the function field of the (hyper)elliptic curve $\mathcal{C}$ with $r, s, w \in \mathbb{K}[X]$. Put $\alpha_0 = \frac{r+w\sqrt{D}}{s} \in \mathbb{K}((X^{-1}))$. We may assume $\operatorname{ord}(\alpha_0) \leq 0$, otherwise we simply pass to the inverse of α. We also require $w, s \neq 0$ and may of course assume $\gcd(r, s, w) = 1$.

Then the finite poles of α are zeroes of s. So the divisor has the shape

$$\operatorname{div} \alpha = -(Q_1) - \cdots - (Q_h) + \ldots, \quad Q_i \in \mathcal{C}_{\text{aff}}$$

with $h \leq 2 \deg s$ and other poles only at infinity (the points $O_\pm$).

We now generalise Lemma 4.7 about the divisors of convergents of $\sqrt{D}$ to rational functions on $\mathcal{C}$.

Proposition 6.16. *Let $(p, q) \in \mathcal{C}_{\alpha_0}(\mathbb{K})$ a convergent, then*

$$\operatorname{div}(p - \alpha\, q) = -m\,(O_-) - (Q_1) - \cdots - (Q_h) + (m+h-e)\ (O_+) + (P_1) + \cdots + (P_e) \quad (6.13)$$

where $P_i \in \mathcal{C}_{\text{aff}}$, $m \geq 0$ and $0 \leq e < h - \operatorname{ord}(\alpha_0) \leq h + m$.

Proof. Set $\phi = p - \alpha\, q$. Any finite poles (in $\mathcal{C}_{\text{aff}}$) must be among the Q_i because $\operatorname{ord}_P(p) \geq 0, \operatorname{ord}_P(q) \geq 0$ implies

$$\operatorname{ord}_P(\phi) \geq \min\left(\operatorname{ord}_P(p), \operatorname{ord}_P(\alpha) + \operatorname{ord}_P(q)\right) \geq \min(0, \operatorname{ord}_P(\alpha)).$$

From (p, q) being a convergent, we know that $\operatorname{ord}_{O_+}(\phi) = \operatorname{ord}(\phi) > \deg q \geq 0$. As in (3.5), this implies with $\operatorname{ord}(\alpha_0\, q) \leq 0$ that

$$\operatorname{ord}_{O_-}(\phi) = \operatorname{ord}(p + \alpha_0\, q) = \operatorname{ord}(p) = -\deg p = \operatorname{ord}(\alpha_0) + \operatorname{ord}(q)\,.$$

Hence $m = -\operatorname{ord}_{O_-} \geq 0$.

With all possible poles determined, we can write $\operatorname{div}(\phi)$ as in (6.13), where possibly some of the $P_i \in \mathcal{C}_{\text{aff}}$ coincide with some Q_j. The divisor must have degree 0, hence $\operatorname{ord}_{O_+}(\phi) = m + h - e$, and

$$\deg q < m + h - e = \deg q - \operatorname{ord}(\alpha_0) + h - e$$

implies $e < h - \operatorname{ord}(\alpha_0)$. □

We can make this even more precise for the canonical convergents (p_n, q_n):

Corollary 6.17. *Let $\phi_n = p_n - \alpha\, q_n$, then*

$$\operatorname{div}\phi_n = -(\deg p_n)\,(O_-) - (Q_1) - \cdots - (Q_h) + (\deg q_{n+1})\,(O_+) + (P_1^n) + \cdots + \left(P_{e_n}^n\right)$$

where $P_i^n \in \mathcal{C}_{\text{aff}}$ (perhaps some coincide with a Q_j) and $e_n = \deg a_0 - \deg a_{n+1} + h \leq \deg a_0 + h - 1$.

Proof. We obtain the formula for e_n from

$$\deg q_{n+1} = \deg q_n + \deg a_{n+1} = \deg p_n + h - e_n = \deg q_n + \deg a_0 + h - e_n.$$

□

Via (6.10), we can now calculate the divisors of the complete quotients (thinking $Y = \sqrt{D}$):

Corollary 6.18. *Write $\mathbf{P}^n = (P_1^n) + \cdots + \left(P_{e_n}^n\right)$, then*

$$\begin{aligned}\operatorname{div}\alpha_n &= \operatorname{div}\left(-\phi_{n-2}/\phi_{n-1}\right)\\ &= (\deg p_{n-1} - \deg p_{n-2})\ (O_-) + (\deg q_{n-1} - \deg q_n)\ (O_+) + \mathbf{P}^{n-2} - \mathbf{P}^{n-1}\\ &= (\deg a_{n-1})\ (O_-) + (-\deg a_n)\ (O_+) + \mathbf{P}^{n-2} - \mathbf{P}^{n-1}\end{aligned}$$

Note how

$$\begin{aligned}\operatorname{ord}_{O_+}(\alpha_n) &= \operatorname{ord}(\alpha_n) = \operatorname{ord}(a_n) = -\deg a_n,\\ \operatorname{ord}_{O_-}(\alpha_n) &= -\operatorname{ord}\left(\frac{-1}{\sigma(\alpha_n)}\right) = -\operatorname{ord}(a_{n-1}) = \deg a_{n-1}.\end{aligned}$$

This aligns with the observations in Section 5.8, in particular Remark 5.37 about the "conjugate" continued fraction expansion.

So the Q_i can no longer be seen directly in this divisor, but of course they could appear hidden among the P_i^{n-1}, P_i^{n-2}.

Let us now restrict to the case $w = 1$ and $s \mid D - r^2$. This implies $h \leq \deg s$ because now it is impossible for both a point and its conjugate to appear as a pole, and a self-conjugate point can appear at most as a pole of order 1 (assuming that D is square-free).

If $s \in \mathbb{K}^\times$, then there are no finite poles, and we are essentially in the situation of Lemma 4.7.

Remark 6.19. Using $\operatorname{ord}(\alpha_0) \leq 0$, we may also assume that $\deg r \leq d = \frac{1}{2} \deg D$ (otherwise we could subtract some multiple of s from r which does not change the subsequent complete quotients). This implies $\operatorname{ord}(\alpha_0) \geq \operatorname{ord}\left(\sqrt{D}\right) - \operatorname{ord}(s)$, so $e < -\operatorname{ord}(\alpha_0) \leq d + h - \deg s \leq d$ and hence $e \leq g$, so we get

$$j(Q_1) + \cdots + j(Q_h) + m\, j(O_-) = j(P_1) + \cdots + j(P_e). \tag{6.14}$$

We are representing a translate of the multiples of $\mathbf{O}$ as a sum of at most g points in the Jacobian.

Remark 6.20. The divisor $(P_1) + \cdots + (P_e)$ is usually going to be a $\mathbb{K}$-rational divisor. Be aware that this does not mean that the P_i are defined over $\mathbb{K}$. However they are defined over a field extension of degree at most e over $\mathbb{K}$. So if $e = 1$, the single point P_1 is going to be defined over $\mathbb{K}$. We will make use of this later.

7. Specialization of continued fractions

The first goal of this chapter is to explain and recover a theorem of van der Poorten (see Theorem 1 in [vdP98], Theorem 2.1 in [vdP99] and Theorem 6 in [vdP01]) stating that the convergents of some α modulo a prime number $\mathfrak{p}$ all arise by normalising and reducing the original convergents of α (which is a Laurent series with rational coefficients).

Here we actually prove this theorem (as Theorem 7.2) in the general setting of Laurent series defined over a discrete valuation ring (or its fraction field), once some natural conditions are satisfied.

Before we look at the convergents, we however need to understand what we mean by reducing convergents, and or even continued fractions. For the latter, this immediately leads to a notion of good or bad reduction of polynomial continued fractions. In the case of good reduction of a continued fraction, van der Poorten's theorem becomes trivial, using the classification of convergents described in Chapter 5. This already suggests that the bad reduction case is more interesting.

Understanding the reduction of the convergents also helps to understand reduction of the continued fraction better, and we will look at some simple cases at the end of the chapter. This goes already toward the calculation of the Gauss norms of the partial quotients and convergents. These will be further analysed for square roots of polynomials in the next chapter.

7.1. Specialization of Laurent series

7.1.1. Discrete valuation rings

We fix a discrete valuation ring O with its unique (principal) maximal ideal $\mathfrak{m}$. It produces two fields: the *fraction field* $K = \mathsf{Fr}(O)$ and the *residue field* $k = O/\mathfrak{m}$. In order to apply the theory from the preceding chapters, we require that $\operatorname{char} k \neq 2$, which implies $\operatorname{char} K \neq 2$ as well.

We denote the (non-archimedean) valuation of O by $\nu : K \twoheadrightarrow \mathbb{Z} \cup \{\infty\}$, and the corresponding absolute value by $|\cdot|_\nu : K \to \mathbb{R}_{\geq 0}$. Recall that it satisfies

- $\nu(x) = \infty \iff x = 0$,
- $\nu(x\,y) = \nu(x) + \nu(y)$ for all $x, y \in K^\times$,
- $\nu(x + y) \geq \min(\nu(x), \nu(y))$ for all $x, y \in K$.

In the last point, we can replace "$\geq$" with "$=$" if $\nu(x) \neq \nu(y)$.

Moreover we choose an uniformising parameter π (a generator of the maximal ideal $\mathfrak{m}$ in O), with satisfies $\nu(\pi) = 1$. Recall

$$\begin{aligned} O &= \{x \in K \mid \nu\,(x) \geq 0\}, \\ \mathfrak{m} = (\pi) &= \{x \in K \mid \nu\,(x) > 0\}, \\ O^{\times} &= \{x \in K \mid \nu\,(x) = 0\}. \end{aligned} \tag{7.1}$$

We get the reduction map $\rho : O \to O/\mathfrak{m} = k$; we usually write $\overline{x} = \rho\,(x)$ for more compact notation.

Note that by just choosing a discrete non-archimedean valuation ν on a given field K, we get a discrete valuation ring O through (7.1).

For example, starting with $K = \mathbb{Q}$ and some odd integer prime $\mathfrak{p}$ with its corresponding $\mathfrak{p}$-adic valuation $\nu_{\mathfrak{p}}$, one gets the localisation $O = \mathbb{Z}_{(\mathfrak{p})}$ of $\mathbb{Z}$ at $\mathfrak{p}$. In this case, the residue field $k = \mathbb{F}_{\mathfrak{p}}$ is finite.

Another example would be $K = \mathbb{C}(t)$ with a zero-order $\mathrm{ord}_{t=t_0}$ (for some $t_0 \in \mathbb{C}$). Then $O = \mathbb{C}[t]_{(t-t_0)}$ is a localisation of $\mathbb{C}[t]$ at the prime ideal $(t - t_0)$, and $t - t_0$ is a uniformising parameter. The residue field is now $k = \mathbb{C}$, hence infinite. In this example we could then replace $\mathbb{C}$ by any field (of characteristic not 2), even a finite field. The latter would make the residue field finite again.

7.1.2. Gauss norms

It is natural to extend such a valuation to polynomials; for absolute values this is called a *Gauss norm.* In fact, we can extend the valuation even to a subset of Laurent series. By abuse of notation, we use the same symbol ν for the extended valuation.

Definition 7.1. Define $\nu : K(\!(X^{-1})\!) \to \mathbb{Z} \cup \{+\infty, -\infty\}$ by setting for $u \in K(\!(X^{-1})\!)$, with $u_n \in K$:

$$\nu\,(u) = \nu\left(\sum_{n=-\infty}^{N} u_n\, X^n\right) = \inf\{\nu\,(u_n) \mid n \in \mathbb{Z}, n \leq N\}.$$

To avoid $\nu\,(u) = -\infty$, we restrict to the subring

$$K(\!(X^{-1})\!)_\nu = \{u \in K(\!(X^{-1})\!) \mid \text{the } \nu\,(u_n) \text{ are bounded from below}\}.$$

Remark 7.2. If $x \in K$, note that because ν is non-archimedean, $u(x)$ converges if and only if $\nu(u_n\, x^n) = \nu(u_n) + n\,\nu(x)$ *to* $+\infty$ as $n \to \infty$. The boundedness condition ensures that $u(x)$ converges for every $x \in \mathfrak{m}$ (with $\nu(x) > 0$).

Proposition 7.3. *$K(\!(X^{-1})\!)_\nu$ is a ring, and the extended ν is a discrete non-archimedean valuation on it.*

Proof. It suffices to check that ν satisfies the usual properties of an ultrametric valuation on $K(\!(X^{-1})\!)_\nu$. Then $K(\!(X^{-1})\!)_\nu$ is automatically a ring (using the same arguments which show that O defined as in (7.1) is a ring).

It is also obvious that ν is discrete because we take an infimum of a subset of $\mathbb{Z}$ bounded from below.

Clearly, we have $\nu(u) = \infty$ if and only if $u = 0$.

Take $u, v \in K(\!(X^{-1})\!)_\nu$ with

$$u = \sum_{n=-\infty}^{N} u_n X^n, \quad v = \sum_{m=-\infty}^{M} v_m X^m,$$

For the ultrametric inequality, let

$$u + v = w = \sum_{l=-\infty}^{\max(N,M)} w_l X^l.$$

Without loss of generality, one may assume $N = M$, and then $w_n = u_n + v_n$ for all $n \leq N$:

$$\begin{aligned}\nu(w) = \inf\{\nu(u_n + v_n) \mid n \leq N\} &\geq \inf\{\min(\nu(u_n), \nu(v_n)) \mid n \leq N\} \\ &= \min(\inf\{\nu(u_n) \mid n \leq N\}, \inf\{\nu(v_n) \mid n \leq N\}) = \min(\nu(u), \nu(v)).\end{aligned}$$

For multiplicativity, let

$$u\,v = w = \sum_{l=-\infty}^{(N+M)} w_l X^l.$$

As ν is invariant under multiplication with powers of X, we may assume $N = M = 0$. From the definition of the Cauchy product

$$w_l = \sum_{n+m=l} u_n v_m \tag{7.2}$$

it is obvious that $\nu(w) \geq \nu(u) + \nu(v)$ must hold:

$$\begin{aligned}\nu(w) = \inf\{\nu(w_l) \mid l \leq 0\} &\geq \inf\{\min(\nu(u_n) + \nu(v_m) \mid n + m = l) \mid l \leq 0\} \\ &\geq \inf\{\min(\nu(u) + \nu(v) \mid n + m = l) \mid l \leq 0\} \geq \nu(u) + \nu(v).\end{aligned}$$

Because ν is discrete on K, there exist n_0, m_0 such that

$$\nu(u) = \nu(u_{n_0}) \text{ and } \nu(v) = \nu(v_{m_0})$$

and of course, we may choose n_0 and m_0 maximal. Then

$$w_{n_0+m_0} = \sum_{n+m=n_0+m_0} u_n v_m = \sum_{\substack{n+m=n_0+m_0,\\ n>n_0}} u_n v_m + u_{n_0} v_{m_0} + \sum_{\substack{n+m=n_0+m_0,\\ m>m_0}} u_n v_m.$$

We have $\nu(u_n) > \nu(u)$ for all terms with $n > n_0$, hence the absolute value of the left sum is $> \nu(u) + \nu(v)$. And we have $\nu(v_m) > \nu(v)$ for all terms with $m > m_0$, hence the absolute value of the right sum is $> \nu(u) + \nu(v)$.

However, the middle term has absolute value $\nu(u_{n_0}) + \nu(v_{m_0}) = \nu(u) + \nu(v)$, implying $\nu(w_{n_0+m_0}) = \nu(u) + \nu(v)$. It follows $\nu(u\,v) \leq \nu(u) + \nu(v)$, and hence $\nu(u\,v) = \nu(u) + \nu(v)$. □

Remark 7.4. Clearly, $K[X] \subset K(X) \subset K(\!(X^{-1})\!)_\nu$.

Remark 7.5. Of course also $O(\!(X^{-1})\!) \subset K(\!(X^{-1})\!)_\nu$. Applying the reduction map on each coefficient, it extends naturally to

$$\rho : O[X] \to k[X], \qquad \rho : O(\!(X^{-1})\!) \to k(\!(X^{-1})\!).$$

For convenience, we use the same notation, including $\overline{x} = \rho(x)$ for $x \in O(\!(X^{-1})\!)$, and say that we *reduce mod* ν or *specialize at* ν.

In analogue to (7.1), we obviously get

$$\begin{aligned} O[X] &= \{u \in K[X] \mid |u|_\nu \le 1\}, & \mathfrak{m}[X] &= \{u \in K[X] \mid |u|_\nu < 1\}, \\ O(\!(X^{-1})\!) &= \{u \in K(\!(X^{-1})\!) \mid |u|_\nu \le 1\}, & \mathfrak{m}(\!(X^{-1})\!) &= \{u \in K(\!(X^{-1})\!) \mid |u|_\nu < 1\}. \end{aligned}$$

Note that while $\mathfrak{m}(\!(X^{-1})\!)$ is a principal maximal ideal in $O(\!(X^{-1})\!)$ (with uniformising parameter still π), the principal ideal $\mathfrak{m}[X]$ of $O[X]$ is not maximal (because the quotient $k[X]$ is not a field).

Remark 7.6. This shows that $O(\!(X^{-1})\!)$ is a discrete valuation ring, with fraction field $K(\!(X^{-1})\!)$. So we *could* extend ν to a discrete valuation ν' on $K(\!(X^{-1})\!)$, setting $\nu'(u/v) = \nu(u) - \nu(v)$.

On $K(\!(X^{-1})\!)_\nu$, the valuations ν and ν' coincide of course. But outside of $K(\!(X^{-1})\!)_\nu$, only ν' is a valuation. Yet for our purposes, ν is more useful because it gives us information on the valuation of the coefficients. While ν' is a valuation on all of $K(\!(X^{-1})\!)$, it achieves this only by *forgetting* some information.

Definition 7.7. We say for $u \in K(\!(X^{-1})\!)$ that

- u is *unbounded* if $\nu(u) = -\infty$ i.e. $u \notin K(\!(X^{-1})\!)_\nu$,
- u is *bounded* if $\nu(u) \ne -\infty$ i.e. $u \in K(\!(X^{-1})\!)_\nu$,
- u *has a negative valuation* if $\nu(u) < 0$, in particular if it is unbounded,
- u *has a positive valuation* if $\nu(u) > 0$.

For example, if $u \in K[X]$ is a polynomial, it has negative valuation if and only if at least one of its coefficients has negative valuation; and it has positive valuation if and only if all its coefficients are either 0 or have positive valuation. Note the different logical operations: For negative valuation, we have **or**, for positive valuation we have **and**.

Let us now investigate how far away $K(\!(X^{-1})\!)_\nu$ is from being a field. For example, the Laurent polynomial $1 + u_{-1} X^{-1}$ with $\nu(u_{-1}) < 0$ does not have a bounded inverse:

Proposition 7.8. *Let $u \in K(\!(X^{-1})\!)_\nu$ with $u_0 = \ell c(u) \ne 0$ (so $u \ne 0$). Then $u^{-1} \in K(\!(X^{-1})\!)_\nu$ if and only if $\nu(u) = \nu(u_0)$.*

Proof. If u has a bounded inverse, we have $\ell c(u^{-1}) = 1/u_0$ with $\nu(1/u_0) \ge \nu\left(u^{-1}\right) = -\nu(u)$, hence $\nu(u) \ge \nu(u_0)$. But by definition $\nu(u_0) \ge \nu(u)$, so it follows $\nu(u_0) = \nu(u)$.

Conversely, assume $\nu(u_0) = \nu(u)$; again, by dividing by u_0 and some power of X (both are bounded), we may write $u = 1 - v$ for $v \in O[\![X^{-1}]\!]$ as a power series with $\nu(v) \geq 0$ and $\operatorname{ord}(v) > 0$. Then

$$u^{-1} = \frac{1}{1-v} = \sum_{j=0}^{\infty} v^j = \sum_{m=-\infty}^{0} w_m X^m$$

converges in $K(\!(X^{-1})\!)$. Only finitely many v^j (always with $\nu\left(v^j\right) \geq 0$) contribute to each w_m, so clearly $\nu(w_m) \geq 0$ for all m, and $1/u$ is bounded. □

7.1.3. Criterion for bounded square roots

In the next chapter, we will be particularly interested in the specialization of Laurent series which are square roots of polynomials. Proposition 2.9 already describes how to construct square roots that lie in $K(\!(X^{-1})\!)$, we now give additional conditions which are sufficient to have the square root lie in $K(\!(X^{-1})\!)_\nu$.

For a counterexample, take $u = 1 + u_{-1} X^{-1}$ where $u_{-1} \in K$, $|u_{-1}|_\nu > 1$: then it is easy to see that $\nu(u) = -\infty$.

Proposition 7.9. *Let $u \in K(\!(X^{-1})\!)_\nu$ such that $\sqrt{u} \in K(\!(X^{-1})\!)$ and $\nu(u) = \nu(\ell c(u_0))$. Then $\sqrt{u} \in K(\!(X^{-1})\!)_\nu$, i.e. $\sqrt{u}$ is bounded.*

Proof. Recall that $u_0 = \ell c(u)$ must be a square, and $\operatorname{ord}(u)$ must be even. We may thus divide u by u_0 and an appropriate even power of X (because both are squares and bounded), and assume $u = 1 + v$ where $v \in O[\![X^{-1}]\!]$, i.e. $\nu(v) \geq 0$, and $\operatorname{ord}(v) > 0$.

Hence

$$\sqrt{u} = \sqrt{1+v} = \sum_{j=0}^{\infty} \binom{1/2}{j} v^j = \sum_{m=-\infty}^{0} w_m X^m$$

converges in $K(\!(X^{-1})\!)$. By the hypothesis $\operatorname{char} k \neq 2$, we have $\nu(2) = 0$, so $\nu\left(\binom{1/2}{j}\right) \geq 0$ (see also Lemma 9.9). As $\lim_{j\to\infty} \operatorname{ord}\left(v^j\right) = \lim_{j\to\infty} j \operatorname{ord}(v) = -\infty$, only a finite number of $\binom{1/2}{j} v^j$, each having $\nu(\cdot) \geq 0$, influence each w_m. Hence $\nu(w_m) \geq 0$ for all m, and $\sqrt{u}$ is bounded. □

7.2. Specialization of polynomial continued fractions

Given $\alpha \in O(\!(X^{-1})\!)$, we can on the one hand see it as element of $K(\!(X^{-1})\!)$, or reduce it to $\overline{\alpha} \in k(\!(X^{-1})\!)$. For each, one gets a continued fraction over $K[X]$ respectively $k[X]$. If one is *lucky*, then $\mathbf{CF}(\alpha)$ has all "data" defined over O, so one can apply ρ, and ask: do $\mathbf{CF}$ and ρ commute?

The answer is yes, so the obstacle lies in $\mathbf{CF}(\alpha)$ not having all data defined over O.

Let us fix notations for the rest of the chapter:

Let $\alpha \in O(\!(X^{-1})\!)$ with $\ell c(\alpha) \in O^\times$ and $\operatorname{ord}(\alpha) \leq 0$, so that α has a non-trivial polynomial part. It has a continued fraction expansion $\mathbf{CF}(\alpha)$ over $K[X]$, with complete quotients $\alpha_n \in K(\!(X^{-1})\!)$, partial quotients $a_n \in K[X]$ and canonical convergents $(p_n, q_n) \in K[X]^2$, satisfying

$$\alpha = [a_0, a_1, \dots], \qquad \alpha_n = [a_n, a_{n+1}, \dots], \qquad p_n/q_n = [a_0, \dots, a_n].$$

For the *specialization*, we set $\gamma = \overline{\alpha} \in k(\!(X^{-1})\!)$. The condition $\ell c(\alpha) \in O^\times$ ensures $\operatorname{ord}(\gamma) = \operatorname{ord}(\alpha) \leq 0$. Of course γ has a continued fraction expansion $\mathbf{CF}(\gamma)$ with complete quotients denoted $\gamma_n \in k(\!(X^{-1})\!)$ and partial quotients denoted $c_n \in k[X]$. The canonical convergents of γ are written as $(u_n, v_n) \in k[X]^2$ to distinguish them easily, and they satisfy

$$\gamma = [c_0, c_1, \dots], \qquad \gamma_m = [c_m, c_{m+1}, \dots], \qquad u_m/v_m = [c_0, \dots, c_m].$$

7.2.1. Good reduction

To answer the question about "commuting", we want to apply the reduction map on the complete quotients, motivating

Definition 7.10. We say that $\mathbf{CF}(\alpha)$ has *good reduction* at ν if for all $n \geq 0$

$$\alpha_n \in O(\!(X^{-1})\!) \text{ and } \overline{\alpha_n} = \gamma_n.$$

It turns out the second condition is a consequence of the first, and it is also possible to describe good reduction in terms of the partial quotients:

Theorem 7.1. *The following are equivalent:*

1. $\mathbf{CF}(\alpha)$ *has good reduction.*
2. $\alpha_n \in O(\!(X^{-1})\!)$ *for all* $n \geq 0$.
3. $a_n \in O[X]$ *and* $\ell c(a_n) = \ell c(\alpha_n) \in O^\times$ *for all* $n \geq 0$.
4. $\deg a_n = \deg c_n$ *for all* $n \geq 0$.

Remark 7.11. For $n = 0$ we had $\ell c(\alpha_0) \in O^\times$ as a hypothesis.

We begin to prove the various implications of the theorem with the following observation.

Remark 7.12. If $\alpha_n \in O(\!(X^{-1})\!)$, then clearly $a_n = \lfloor \alpha_n \rfloor \in O[X]$.

Next, let us show that $a_n \in O[X]$ cannot be a sufficient condition for good reduction:

Proposition 7.13. *Let* $n \geq 0$. *If* $\alpha_n \in O(\!(X^{-1})\!)$, *then* $\alpha_{n+1} \in O(\!(X^{-1})\!)$ *if and only if* $\ell c(a_{n+1}) = \ell c(\alpha_{n+1}) \in O^\times$.

Proof. By Definition 5.8, we have $\alpha_{n+1}^{-1} = \alpha_n - a_n \in O(\!(X^{-1})\!)$, and of course $\ell c(\alpha_{n+1}^{-1}) = \ell c(\alpha_{n+1})^{-1} \in O$. So if $\alpha_{n+1} \in O(\!(X^{-1})\!)$, then clearly $\ell c(\alpha_{n+1}) \in O^\times$.

Conversely, if $\ell c(\alpha_{n+1}) \in O^\times$, then $\nu\left(\ell c(\alpha_n - a_n)\right) = -\nu\left(\ell c(\alpha_{n+1})\right) = 0$, hence $\nu\left(\alpha_n - a_n\right) = 0$. So Proposition 7.8 implies that $\alpha_n - a_n = \alpha_{n+1}^{-1}$ has a bounded inverse, with $\nu\left(\alpha_{n+1}\right) = 0$. □

This allows to show that the second condition in Definition 7.10 is an automatic consequence of the first condition:

Proposition 7.14. *Let $n \geq 0$. If $\alpha_n, \alpha_{n+1} \in O(\!(X^{-1})\!)$ and $\overline{\alpha_n} = \gamma_n$, then $\overline{\alpha_{n+1}} = \gamma_{n+1}$.*

Proof. Clearly $\overline{\alpha_n} = \gamma_n$ implies $\overline{a_n} = c_n$, and by Propositions 7.8 and 7.13 we have $\alpha_{n+1} \in O(\!(X^{-1})\!)^\times$. Hence

$$\gamma_{n+1}^{-1} = \gamma_n - c_n = \overline{\alpha_n} - \overline{a_n} = \overline{\alpha_{n+1}^{-1}} = \overline{\alpha_{n+1}}^{-1}$$

which implies $\gamma_{n+1} = \overline{\alpha_{n+1}}$ as desired. □

Remark 7.15. If $\ell c(\alpha_n) \in O^\times$ and $\overline{\alpha_n} = \gamma_n$, we have $\operatorname{ord}(\alpha_n) = \operatorname{ord}(\gamma_n) \leq 0$ (< 0 for $n \geq 1$), and hence $\deg a_n = \deg c_n$.

Let us now describe good reduction in terms of the partial quotients; for this we first have a look at the convergents:

Proposition 7.16. *Let $n \geq 0$ and suppose $a_j \in O[X]$ for $j = 0, \dots, n$ and $\ell c(a_j) \in O^\times$ for $j = 1, \dots, n$. Then $p_n, q_n \in O[X]$ and moreover $p_n/q_n \in O(\!(X^{-1})\!)$.*

Proof. The statement $p_n, q_n \in O[X]$ follows directly from the recursion formulas for the canonical convergents (5.3). And the product formula for the leading coefficients (5.5) implies

$$\nu\left(\ell c(q_n)\right) = \sum_{j=1}^{n} \nu\left(\ell c(a_j)\right).$$

But then $\nu\left(\ell c(a_j)\right) = 0$ for $j = 1, \dots, n$ implies $\nu\left(\ell c(q_n)\right) = \nu\left(q_n\right) = 0$. So by Proposition 7.8 the denominator $q_n \in O(\!(X^{-1})\!)^\times$ has bounded inverse, hence $p_n/q_n \in O(\!(X^{-1})\!)$. □

We conclude this section by proving the equivalence of the alternative characterisations of good reduction.

Proof of Theorem 7.1. Equivalence of 1. and 2. is a consequence of Proposition 7.14 above.

Next, 2. implies 3. by Proposition 7.13.

Conversely, 3. implies 2.: Let $m \geq 0$ and recall that $\operatorname{ord}(p_{m,n}/q_{m,n} - \alpha_m) > 2 \deg q_{m,n}$ from Proposition 5.15. Moreover, we have $p_{m,n}/q_{m,n} \in O(\!(X^{-1})\!)$ by Proposition 7.16, so the *first* coefficients of α_m are also in O. As $\lim_{n\to\infty} \deg q_{m,n} = \infty$, we cover all coefficients, and thus $\alpha_m \in O(\!(X^{-1})\!)$.

Next, 1. and 3. imply $\overline{\alpha_n} = \gamma_n$, hence $\overline{a_n} = c_n$ and $\ell c(a_n) \in O^\times$. The latter is equivalent to $\deg a_n = \deg \overline{a_n}$, so we get $\deg a_n = \deg c_n$.

Finally 4. implies 2.: by Proposition 7.20 (below, but independent of this Theorem) there exists n with $\deg a_n < \deg c_n$ 2. if is violated. □

Remark 7.17. So continued fraction expansion and specialization commute as soon as the partial quotients are defined over O and do not "drop degree" on reduction, or even simpler, the degrees of the partial quotients match.

Theorem 1.3 of van der Poorten becomes almost trivial in this case:

Corollary 7.18. *If* $\mathbf{CF}(\alpha)$ *has good reduction, then for all* $n \geq 0$ *we have* $u_n = \overline{p_n}$ *and* $v_n = \overline{q_n}$ *which by the classification of convergents (Proposition 5.20) implies that all convergents of* γ *are obtained by reducing convergents of* α*.*

Proof. We can think of p_n and q_n as polynomials in $\mathbb{Z}[a_0, \dots, a_n]$ (see Proposition 5.1). Of course u_n and v_n are obtained by replacing a_j with c_j in those polynomials. But $c_j = \overline{a_j}$ for all $j \geq 0$, so $(u_n, v_n) = (\overline{p_n}, \overline{q_n})$.

An arbitrary convergent of γ has perhaps an additional polynomial factor in $k[X]$ which we can however lift to a polynomial of same degree in $K[X]$. Because we have $\deg a_{n+1} = \deg c_{n+1}$, multiplying (p_n, q_n) with this polynomial produces a convergent of α. □

7.2.2. Bad reduction

Definition 7.19. The opposite of good reduction of $\mathbf{CF}(\alpha)$ is obviously *bad reduction* of $\mathbf{CF}(\alpha)$, by which we mean that there exists $n \geq 1$ such that $\alpha_n \notin O(\!(X^{-1})\!)$ (i.e. $\nu(\alpha_n) < 0$, so there is a coefficient with negative valuation).

The results for good reduction are still useful in this case, for example Propositions 7.13 and 7.14 can be applied until we arrive at the complete quotient with bad reduction. They should also give an initial idea of what could go wrong in the case of bad reduction.

Proposition 7.20. *Suppose* $\mathbf{CF}(\alpha)$ *has bad reduction and let* n *minimal with* $\alpha_n \notin O(\!(X^{-1})\!)$*. Then in fact* $\nu(\ell c(\alpha_n)) < 0$*, i.e.* α_n *has negative valuation in the leading coefficient.*

If γ_n *is defined, then* $\deg c_n > \deg a_n$ *and* α_n *is unbounded.*

Proof. The first statement is an immediate consequence of Proposition 7.13: by minimality $\alpha_{n-1} \in O(\!(X^{-1})\!)$, so $\nu(\ell c(\alpha_n)) \neq 0$. But $\nu(\ell c(\alpha_n)) > 0$ is impossible because $\ell c(\alpha_n)^{-1} = \ell c(\alpha_n^{-1}) = \ell c(\alpha_{n-1} - a_{n-1}) \in O$.

Now assume γ_n is defined: As we have good reduction up to α_{n-1}, we certainly have $\overline{\alpha_{n-1}} = \gamma_{n-1}$ (use Proposition 7.14 inductively). But $\ell c(\alpha_{n-1} - a_n) \in \mathfrak{m}$, hence

$$\deg a_n = -\operatorname{ord}(\alpha_n) = \operatorname{ord}(\alpha_{n-1} - a_n) < \operatorname{ord}(\gamma_{n-1} - c_{n-1}) = -\operatorname{ord}(\gamma_n) = \deg c_n.$$

In particular $\gamma_{n-1} - c_{n-1} \neq 0$ which implies $\nu(\alpha_{n-1} - a_{n-1}) = 0$. But as the leading coefficient is in $\mathfrak{m}$, Proposition 7.8 implies that the inverse α_n is unbounded. □

Remark 7.21. If we are using the computation scheme with t_n and s_n from Chapter 6 and we are already in the σ-reduced case, the negative valuation in the leading coefficient of α_n corresponds to positive valuation in the leading coefficient of s_n.

Unless γ is rational,[1] γ_n is of course always defined.

7.2.3. Reduction and normalisation of continued fractions

We can extend the reasoning of this section also to an arbitrary Laurent series $\alpha \in K(\!(X^{-1})\!)_\nu$, as long as α is bounded and satisfies $\nu(\alpha) = \nu(\ell c(\alpha))$ and $\mathrm{ord}(\alpha) \leq 0$. If these requirements are met, we can just divide α by $\ell c(\alpha)$, or some $g \in K^\times$ with $\nu(g) = \nu(\alpha)$. For the new series, we can apply the above results.

Of course reduction here must always be preceded by normalisation. But for example the existence of unbounded complete quotients is invariant under normalisation (see Proposition 5.21 about multiplying a continued fraction with a constant), and is characteristic for bad reduction.

We will revisit these issues later, first we need to study the reduction of the convergents closer.

7.3. Normalisation and reduction of convergents

In the case of good reduction of the continued fraction, we were able to simply reduce the canonical convergents. In the case of bad reduction of the continued fraction, we cannot expect the canonical convergents to be polynomials defined over O, so we need to normalise them first.

In other words, we wish to extend the reduction map in a useful way to all of $K[X]$ (or even $K(\!(X^{-1})\!)$) by normalising to valuation 0 before reducing. Of course, extending the reduction map $O \to k$ in this way from O to K is not so useful. But for polynomials and Laurent series, there are usually several coefficients, so thinking projectively makes sense. For obvious reasons, this works only for bounded Laurent series.

Definition 7.22. Let $u \in K(\!(X^{-1})\!)_\nu \setminus \{0\}$, and recall that π is a uniformising parameter of O satisfying $\nu(\pi) = 1$. Define the *normalisation* $\widetilde{u}$ for u as

$$\widetilde{u} = \pi^{-\nu(u)}\, u \in O(\!(X^{-1})\!).$$

Clearly, $\nu(\widetilde{u}) = 0$. For completeness, we also set $\widetilde{0} = 0$.

If $u \in K$, then $\widetilde{u} \in O$, and if $u \in K[X]$, then $\widetilde{u} \in O[X]$.

We denote the composition of reduction and normalisation by

$$\widehat{u} = \rho\left(\widetilde{u}\right).$$

Before we start normalising convergents, we need to check that the normalisation factor is the same for the numerator and the denominator – otherwise we are unable to normalise the convergent as a whole:

[1] In the case $\alpha = \sqrt{D}$ the reduction $\gamma = \sqrt{\overline{D}}$ clearly is rational if and only if $\overline{D}$ is a square.

Proposition 7.23. *Suppose* $\operatorname{ord}(\rho\,(\alpha)) \leq 0$*, and let* $(p,q) \in \mathcal{Q}(K)$ *a rational approximation with* $\operatorname{ord}(p - \alpha\, q) > 0$*. Set* $g = \pi^{\nu(q)} \in K$*.*

Then $(p,q) = g \cdot (\widetilde{p}, \widetilde{q})$ *and in particular* $\nu\,(p) = \nu\,(q) = \nu\,(g)$*.*

Proof. By definition, we have $q = g\,\widetilde{q}$, and $\nu\,(q) = \nu\,(g)$. The condition $\operatorname{ord}(p - \alpha\,q) > 0$ implies $p = -\lfloor \alpha\, q \rfloor = -g\, \lfloor \alpha\, \widetilde{q} \rfloor$. Of course $p' = -\lfloor \alpha\, \widetilde{q} \rfloor \in O[X]$ and $p = g\, p'$.

It remains to show $p' = \widetilde{p}$:

Indeed $\operatorname{ord}(p' - \alpha\, \widetilde{q}) > 0$ implies $\operatorname{ord}(\rho\,(p') - \rho\,(\alpha)\ \rho\,(q)) > 0$. But $\operatorname{ord}(\rho\,(\alpha)\ \rho\,(q)) \leq 0$ by hypothesis, so also $\operatorname{ord}(\rho\,(p')) \leq 0$. This means $\rho\,(p') \neq 0$, or $\nu\,(p') = 0$, hence $p' = \widetilde{p}$ as desired. □

Corollary 7.24. *Every convergent and best-approximation* $(p,q) \in \mathcal{B}_\alpha(\mathbb{K})$ *(in particular the canonical convergents* (p_n, q_n)*) satisfies* $\nu\,(p) = \nu\,(q)$*.*

Setting $g_n = \pi^{\nu(q_n)}$ *we get* $(p_n, q_n) = g_n \cdot (\widetilde{p_n}, \widetilde{q_n})$*.*

Remark 7.25. For $n = -1$ we have $q_{-1} = 0$ and $p_{-1} = 1$. We just set $g_{-1} = 1$, as no normalisation is required.

We finally state and prove the generalised version of Theorem 1.3 on the reduction of convergents by van der Poorten. First we check that convergents remain convergents after reduction.

Proposition 7.26. *Let* $(p,q) \in \mathcal{C}_\alpha(K)$ *a convergent. Then* $\operatorname{ord}(\widehat{p} - \gamma\,\widehat{q}) > \deg q \geq \deg \widehat{q}$*, so* $(\widehat{p}, \widehat{q}) \in \mathcal{C}_\gamma(k)$ *is also a convergent.*

Proof. The important observation is that for $\beta \in O(\!(X^{-1})\!)$ one has $\operatorname{ord}\big(\overline{\beta}\big) \geq \operatorname{ord}(\beta)$, and for $b \in O[X]$ one has $\deg \overline{b} \leq \deg b$, hence

$$\operatorname{ord}(\widehat{p} - \gamma\,\widehat{q}) \geq \operatorname{ord}(p - \alpha\, q) > \deg q \geq \deg \widehat{q}.$$

□

We restrict now to the conveniently enumerated canonical convergents.

Corollary 7.27. *We find that the reduction of a (normalised) convergent remains a convergent. In particular, there exists a (unique) map* $\lambda : \mathbb{N}_0 \to \mathbb{N}_0$ *defined by*

$$\widehat{p_n}/\widehat{q_n} = u_{\lambda(n)}/v_{\lambda(n)}.$$

More precisely, for each n *there exists* $h_n \in k[X] \setminus \{0\}$ *such that*

$$\widehat{p_n} = h_n\, u_{\lambda(n)}, \qquad \widehat{q_n} = h_n\, v_{\lambda(n)}.$$

Proof. The map λ is well defined: every convergent of γ is a multiple of a unique canonical convergent of γ by Corollary 5.20. □

Here one has to be careful, though: the factor h_n need *not be constant*! We will investigate this closer for some special cases later. See also Example 6 in Section 10.3.2, where non-constant h_n in fact occur.

This possibility of non-constant factors make the following less obvious because $\deg \widehat{q_n}$ may not be non-decreasing:

Proposition 7.28. *The map λ is non-decreasing (it need not be increasing).*

Proof. Let $n < n'$ and set $m = \lambda(n), m' = \lambda(n')$, hence $\deg q_n < \deg q_{n'}$.

If $\deg \widehat{q_n} \leq \deg \widehat{q_{n'}}$, Proposition 5.18 (Classification of best-approximations) for γ implies directly $m \leq m'$.

If however $\deg \widehat{q_n} \geq \deg \widehat{q_{n'}}$, then

$$\begin{aligned}\operatorname{ord}(\widehat{p_n} - \gamma\, \widehat{q_n}) &> \deg \widehat{q_n} \geq \deg \widehat{q_{n'}},\\ \operatorname{ord}(\widehat{p_{n'}} - \gamma\, \widehat{q_{n'}}) &> \deg q_{n'} > \deg q_n \geq \deg \widehat{q_n}.\end{aligned}$$

Eliminating γ, one obtains

$$\begin{aligned}\operatorname{ord}(\widehat{p_n}\, \widehat{q_{n'}} - \widehat{p_{n'}}\, \widehat{q_n}) &= \operatorname{ord}((\widehat{p_n} - \gamma\, \widehat{q_n})\, \widehat{q_{n'}} - (\widehat{p_{n'}} - \gamma\, \widehat{q_{n'}})\, \widehat{q_n})\\ &\geq \min\left(\operatorname{ord}(\widehat{p_n} - \gamma\, \widehat{q_n}) + \operatorname{ord}(\widehat{q_{n'}})\, , \operatorname{ord}(\widehat{p_{n'}} - \gamma\, \widehat{q_{n'}}) + \operatorname{ord}(\widehat{q_n})\right) > 0\end{aligned}$$

which implies $\widehat{p_n}/\widehat{q_n} = \widehat{p_{n'}}/\widehat{q_{n'}}$, hence $m = m'$. □

Remark 7.29. If $m < m'$, then Proposition 5.18 immediately implies $\deg \widehat{q_n} < \deg \widehat{q_{n'}}$.

Now we are ready to prove that the map λ is in fact surjective, a result which appeared first [vdP99], and with a slightly different proof in [vdP99] and [vdP01]. [2] Unfortunately, both proofs are somewhat confusing, perhaps because van der Poorten does not include an argument why the map λ should be non-decreasing. He already seems to assume that property in his implicit definition of λ, where he uses an elaborate enumeration scheme.[3]

Theorem 7.2. *All the (coprime) convergents of γ arise as reductions of convergents of α. In other words, the map $\lambda : \mathbb{N}_0 \to \mathbb{N}_0$ is surjective. Moreover, if $n = \min \lambda^{-1}(m)$, then* $\deg v_m = \deg q_n$.

Proof. First, we show that λ has finite fibres. Indeed, for $n \geq 0$ and $m = \lambda(n)$ we have by definition of λ

$$\widehat{p_n} = h_n\, u_m, \quad \widehat{q_n} = h_n\, v_m \text{ where } h_n \in k[X] \setminus \{0\},$$

hence $\deg q_{n+1} \leq \deg v_{m+1}$:

$$\begin{aligned}\deg v_{m+1} \geq \operatorname{ord}(h_n) + \deg v_{m+1} &= \operatorname{ord}(h_n) + \operatorname{ord}(u_m - \gamma\, v_m)\\ &= \operatorname{ord}(\widehat{p_n} - \gamma\, \widehat{q_n}) \geq \operatorname{ord}(p_n - \alpha\, q_n) = \deg q_{n+1} \quad (7.3)\end{aligned}$$

However, we know that $\lim_{n\to\infty} \deg q_{n+1} = \infty$ so for fixed m there can only by finitely many n which satisfy the inequality.

Because we know that λ is monotonous, we can prove its surjectivity by checking that there are no gaps in the image.

[2] Note that van der Poorten speaks of good reduction only for the hyperelliptic curve, not for the continued fraction.

[3] Van der Poorten does not explicitly define the map λ as we do it here.

There is no gap at the start because $v_0 = 1$ and $q_0 = 1$ imply $\lambda(0) = 0$.

For $n \geq 0$, we either have $\lambda(n) = \lambda(n+1)$ in which case there is no gap.

Otherwise $m = \lambda(n) < \lambda(n+1) = m'$, and we need to show $m' = m+1$. Again, by definition of λ

$$\widehat{p_{n+1}} = h_{n+1}\, u_{m'}, \quad \widehat{q_{n+1}} = h_{n+1}\, v_{m'} \text{ where } h_{n+1} \in k[X] \setminus \{0\}.$$

and in particular

$$\deg v_{m'} \leq \deg h_{n+1} + \deg v_{m'} = \deg \widehat{q_{n+1}} \leq \deg q_{n+1}.$$

But from $m + 1 \leq m'$ and (7.3) follows also

$$\deg q_{n+1} \leq \deg v_{m+1} \leq \deg v_{m'},$$

so these are actually equalities, and as desired $m' = \lambda(n+1) = m+1 = \lambda(n)+1$, so there is no gap. Note that $n+1$ is the minimal element of the fibre $\lambda^{-1}(m')$, and we have shown $\deg q_{n+1} = \deg v_{\lambda(n+1)}$. □

Remark 7.30. Observe that $\deg q_{n+1} = \deg v_{m+1}$ implies $\deg h_{n+1} = 0$, and from (7.3) also $\deg h_n = 0$. Hence both for the minimal and maximal fibre element, the reduced convergent remains coprime.

Corollary 7.31. *Suppose that $\lambda^{-1}(m) = \{n, \ldots, n+l\}$. Then*

$$\deg c_{m+1} = \sum_{i=1}^{l+1} \deg a_{n+i} = \deg a_{n+1} + \cdots + a_{n+l+1}. \tag{7.4}$$

Proof. Both n and $n+l+1$ are the minimal elements of their respective fibres, hence $\deg q_n = \deg v_m$ and $\deg q_{n+l+1} = \deg v_{m+1}$. The degree formula for the convergents (5.4) then gives the desired relation between the degrees of the partial quotients. □

If the reduction is not rational, we also get an additional criterion for good reduction:

Proposition 7.32. *If $\gamma \notin k(X)$, the map λ is bijective if and only if* $\mathbf{CF}(\alpha)$ *has good reduction.*

Proof. First observe that by Proposition 7.28, the map λ is bijective if and only if it is the identity.

If $\mathbf{CF}(\alpha)$ has good reduction, Corollary 7.18 implies that λ is the identity.

Conversely, if λ is the identity, then from Theorem 7.2 we obtain $\deg q_n = \deg v_n$ for all n, which in turn implies $\deg a_n = \deg c_n$ for all n. Then by Theorem 7.1, $\mathbf{CF}(\alpha)$ has good reduction. □

We conclude this section by pointing out that while the canonical convergents are usually not normalised, the convergents we get as solutions of the linear system in Section 3.4 are in fact optimally normalised (even independently of the valuation):

Proposition 7.33. *Let $\alpha \in O(\!(X^{-1})\!)$ and suppose that $\gamma = \overline{\alpha} \neq 0$. Let n such that $\mathcal{M}_n$ has full rank, and let (p, q) correspond to an element of the kernel computed from the minors of $\mathcal{M}_n$ as in Remark 3.24.*

Then $p, q \in O[X]$. Moreover, if $\deg q = \deg \widehat{q}$, we have $\nu(q) = 0$.

Proof. By hypothesis, the coefficients of the Laurent series α are in O. The minors of $\mathcal{M}_n$ are polynomials in these coefficients, so clearly the coefficients of p and q are in O too (recall that we need full rank so they do not all vanish).

The coefficients of γ are obtained by reducing those of α, hence the kernel elements of $\overline{\mathcal{M}_n}$ correspond to convergents of γ. For example there is $(\widehat{p}, \widehat{q})$, and then $\deg q = \deg \widehat{q}$ implies that $\overline{\mathcal{M}_n}$ has full rank as well, so we may compute a convergent using the minors. But of course the reduction map ρ is a ring homomorphism, so this convergent is exactly $(\overline{p}, \overline{q})$, with $\overline{q} \neq 0$. Then clearly $\nu(q) = 0$. □

7.4. Calculating valuations

Once we understand the structure of λ and the reduction of convergents thanks to Theorem 7.2, we can go further and attempt to compute the valuations (Gauss norms) for the partial quotients a_n, the canonical convergents q_n and often even for the complete quotients α_n. In the next chapter, we will see how there arise rather simple patterns in the case $\alpha = \sqrt{D}$ with $\deg D = 4$. For now, we remain in the general case which makes things a bit more complicated. However we will thus understand better the obstacles for generalising the degree 4 case.

7.4.1. Relating complete quotients with convergents

In the following, we always assume $\gamma = \overline{\alpha} \notin k(X)$.

Proposition 7.34. *Define for $n \geq -1$*

$$\vartheta_n = \widetilde{p_n} - \alpha\, \widetilde{q_n}. \tag{7.5}$$

Then $\vartheta_n \in O(\!(X^{-1})\!)$ with $\nu(\vartheta_n) = 0$, and $\operatorname{ord}(\vartheta_n) = \deg q_{n+1}$.

With $g_n = \pi^{-\nu(q_n)}$, we may then write

$$\alpha_n = -\frac{g_{n-2}\, \vartheta_{n-2}}{g_{n-1}\, \vartheta_{n-1}} \tag{7.6}$$

as a quotient of elements of $O(\!(X^{-1})\!)$ up to a normalisation factor.

Remark 7.35. Note that $\vartheta_{-1} = 1$ and $\vartheta_0 = a_0 - \alpha$.

Proof. By definition of normalisation, we have $\widetilde{p_n}, \widetilde{q_n} \in O[X]$, and $\widehat{q_n} \neq 0$. Of course $p_n - \alpha\, q_n = g_n\, \vartheta_n$, so $\operatorname{ord}(\vartheta_n) = \deg q_{n+1}$ is immediate from Proposition 5.15 and $\pi \in K$.

As we assume $\alpha \in O(\!(X^{-1})\!)$, this implies $\vartheta_n \in O(\!(X^{-1})\!)$. Moreover, $\gamma = \overline{\alpha} \notin k(X)$ implies $\overline{\vartheta_n} \neq 0$, hence $\nu(\vartheta_n) = 0$.

Finally, from Proposition 5.3 we obtain (see also (5.9))

$$\alpha_n = \frac{q_{n-2}\,\alpha - p_{n-2}}{-q_{n-1}\,\alpha + p_{n-1}} = -\frac{g_{n-2}\,(\widetilde{p_{n-2}} - \alpha\,\widetilde{q_{n-2}})}{g_{n-1}\,(\widetilde{p_{n-1}} - \alpha\,\widetilde{q_{n-1}})} = -\frac{g_{n-2}\,\vartheta_{n-2}}{g_{n-1}\,\vartheta_{n-1}}.$$

□

So in order to understand whether α_n is bounded, we need a criterion for when the ϑ_n have a bounded inverse:

Proposition 7.36. *The following are equivalent:*

- $\vartheta_n^{-1} \in K(\!(X^{-1})\!)_\nu$,
- $\vartheta_n \in O(\!(X^{-1})\!)^\times$,
- $\operatorname{ord}(\vartheta_n) = \operatorname{ord}\big(\overline{\vartheta_n}\big)$,
- $\nu\,(\ell c(\vartheta_n)) = 0$,

Proof. By the previous Proposition, we have $\vartheta_n \in O(\!(X^{-1})\!)$ and $\nu\,(\vartheta_n) = 0$. So by Proposition 7.8 the inverse is bounded if and only if

$$\nu\,(\ell c(\vartheta_n)) = 0 \iff \overline{\ell c(\vartheta_n)} \neq 0 \iff \operatorname{ord}(\vartheta_n) = \operatorname{ord}\big(\overline{\vartheta_n}\big)\,.$$

Finally, it is clear that if the inverse is bounded, then $\nu(\vartheta_n^{-1}) = 0$, so it is in $O(\!(X^{-1})\!)$. □

Remark 7.37. Of course $\vartheta_n^{-1} \in K(\!(X^{-1})\!)_\nu$ implies via (7.6) that also $\alpha_{n+1} \in K(\!(X^{-1})\!)_\nu$.

We use this to show that there are always infinitely many bounded complete quotients:

Proposition 7.38. *Let $m \in \mathbb{N}$, and set $n = \min \lambda^{-1}(m)$. Then $\vartheta_{n-1}^{-1} \in K(\!(X^{-1})\!)_\nu$, hence $\alpha_n \in K(\!(X^{-1})\!)_\nu$.*

Proof. With Theorem 7.2 follows from n being minimal in the fibre that $\deg q_n = \deg v_m$, and $\lambda(n-1) = m-1$. By Remark 7.30, we moreover know $\widehat{q_{n-1}} = h_{n-1}\,v_{m-1}$ with $h_{n-1} \in k$, hence

$$\operatorname{ord}\big(\overline{\vartheta_{n-1}}\big) = \operatorname{ord}\big(\widehat{p_{n-1}} - \gamma\,\widehat{q_{n-1}}\big) = \operatorname{ord}(u_{m-1} - \gamma\,v_{m-1}) = \deg v_m = \deg q_n = \operatorname{ord}\vartheta_{n-1},$$

so Proposition 7.36 implies ϑ_{n-1} has bounded inverse. Then (7.6) implies that α_n is bounded. □

Note that the condition for α_n bounded we give here is only sufficient, but not necessary.

7.4.2. Fibre analysis of λ

Using the Lemmata for estimating valuations in quotients of Laurent/power series from Section A.2 in the appendix, we now attack the problem of computing valuations by doing case analysis for the different sizes of the fibres of λ, and the degrees of the partial quotients. This is successful mostly when we can read off the valuations (Gauss norms) from the leading coefficients.

The simplest case is the following, we get information on everything (recall that $\nu(g_n) = \nu(q_n)$ for all $n \geq 0$):

Proposition 7.39 (Single element fibre)**.** *Let $m \in \mathbb{N}$ such that $\lambda^{-1}(m) = \{n\}$ has a single element. Then α_{n+1} is bounded and*

$$\nu(\alpha_{n+1}) = \nu(\ell c(\alpha_{n+1})) = \nu(a_{n+1}) = \nu(g_{n-1}) - \nu(g_n). \tag{7.7}$$

The normalised complete quotient reduces to

$$\widehat{\alpha_{n+1}} = \frac{h_{n-1}}{h_n} \gamma_{m+1} \qquad \textit{with } h_{n-1}, h_n \in k^\times, \tag{7.8}$$

hence $\deg a_{n+1} = \deg c_{m+1}$.

For the corresponding convergent we have

$$\nu(g_{n+1}) = \nu(q_{n+1}) = \nu(\ell c(q_{n+1})) = \nu(g_{n-1}).$$

Proof. Both n and $n+1$ are the minimal elements of their fibres, so Proposition 7.38 implies that both $\vartheta_{n-1}, \vartheta_n \in O(\!(X^{-1})\!)^\times$. Hence α_{n+1} is bounded, and (7.7) follows from (7.6) and $\nu(\vartheta_{n-1}) = \nu(\vartheta_n) = 0$.

Normalising and reducing α_{n+1}, we get

$$\widehat{\alpha_{n+1}} = \rho\left(\frac{g_n}{g_{n-1}}\,\alpha_{n+1}\right) = -\frac{\overline{\vartheta_{n-1}}}{\overline{\vartheta_n}} = -\frac{h_{n-1}\,(u_{m-1} - \gamma\, v_{m-1})}{h_n\,(u_m - \gamma\, v_m)} = \frac{h_{n-1}}{h_n}\,\gamma_{m+1}.$$

Here $h_{n-1}, h_n \in k^\times$ by Remark 7.30.

Again using that n and $n+1$ are minimal in their fibres, Theorem 7.2 implies $\deg \widehat{q_n} = \deg q_n$ and $\deg \widehat{q_{n+1}} = \deg q_{n+1}$. This means $\nu(g_n) = \nu(q_n) = \nu(\ell c(q_n))$ and

$$\nu(g_{n+1}) = \nu(q_{n+1}) = \nu(\ell c(q_{n+1})) = \nu(\ell c(a_{n+1})) + \nu(\ell c(q_n)) = \nu(g_{n-1}).$$

For $\deg a_{n+1} = \deg c_{m+1}$ see also Corollary 7.31. □

If there is more than one element in the fibre, we can say a few things in general. However boundedness of the complete quotients cannot be determined a priori, except for the first and last complete quotient. But even if the complete quotients are bounded, the reduction of the normalisation is *never* a complete quotient of γ as in the single element case of Proposition 7.7 above.

Proposition 7.40 (Multiple element fibre)**.** *Let $m \in \mathbb{N}$ such that $\lambda^{-1}(m) = \{n, n+1, \dots, n+l\}$ has $l \geq 2$ elements. Then α_{n+1} is unbounded and α_{n+l+1} is bounded. The α_{n+i+1} for $1 \leq i < l$ can be bounded* or *unbounded.*

If some α_{n+l+1} (for $1 \leq i \leq l$) is bounded, the reduction of the normalised complete quotient is a rational function (and a polynomial for $i = l$, as $h_{n+l} \in k^{\times}$):

$$\widehat{\alpha_{n+i+1}} = -\frac{h_{n+i-1}}{h_{n+i}}.$$

In particular $\operatorname{ord}\bigl(\widehat{\alpha_{n+l+1}}\bigr) = -\deg h_{n+l-1}$*. In this case, we also get*

$$\nu(\ell c(\alpha_{n+i+1})) \geq \nu(a_{n+i+1}) \geq \nu(\alpha_{n+i+1}) = \nu(g_{n+i-1}) - \nu(g_{n+i}), \tag{7.9}$$

and thus

$$\nu(\ell c(q_{n+i+1})) \geq \nu(q_{n+i+1}) \geq \nu(g_{n+i-1}). \tag{7.10}$$

Proof. Here n and $n+l+1$ are minimal in their fibre, so $\vartheta_{n-1}, \vartheta_{n+l} \in O(\!(X^{-1})\!)^{\times}$ by Proposition 7.38; and h_{n-1}, h_n, h_{n+l} are constant by Remark 7.30. Moreover, Theorem 7.2 tells that $\deg q_n = \deg v_m$ and $\deg q_{n+l+1} = \deg v_{m+1}$, from which we deduce

$$\deg a_{n+1} + \dots + \deg a_{n+l+1} = \deg c_{m+1}$$

as in Corollary 7.31.

Observe that ϑ_n has an unbounded inverse because

$$\operatorname{ord}(\vartheta_n) = \deg q_{n+1} = \deg q_n + \deg a_{n+1} < \operatorname{ord}\bigl(\overline{\vartheta_n}\bigr) = \deg v_{m+1} = \deg q_n + \deg c_{m+1}.$$

Hence α_{n+1} is unbounded. But α_{n+l+1} is of course bounded by Proposition 7.38, even if it need not have a bounded inverse. For the complete quotients in between, we cannot a priori say anything.

But assume that α_{n+i+1} (where $1 \leq i \leq l$) *is bounded.* Then it follows

$$\widehat{\alpha_{n+i+1}} = \rho\left(\frac{g_{n+i}}{g_{n+i-1}}\,\alpha_{n+i+1}\right) = -\frac{\overline{\vartheta_{n+i-1}}}{\overline{\vartheta_{n+i}}} = -\frac{h_{n+i-1}\,(u_m - \gamma\, v_m)}{h_{n+i}\,(u_m - \gamma\, v_m)} = -\frac{h_{n+i-1}}{h_{n+i}}.$$

As always $\nu(\vartheta_i) = 0$, we may deduce (7.9) directly from (7.6), with the inequalities obvious from the definition of ν on polynomials and Laurent series as infimum over the coefficients. With the recurrence relation (5.3), we then get (again only in the bounded case)

$$\nu(\ell c(q_{n+i+1})) \geq \nu(q_{n+i+1}) \geq \min\left(\nu(a_{n+i+1}) + \nu(q_{n+i}), \nu(q_{n+i-1})\right) \geq \nu(g_{n+i-1}).$$

□

Observe that $\operatorname{ord}\bigl(\widehat{\alpha_{n+l+1}}\bigr) = -\deg h_{n+l-1}$, while $\operatorname{ord}(\alpha_{n+l+1}) = -\deg a_{n+l+1} \neq 0$. Its inverse, and hence α_{n+l}, can be bounded only if h_{n+l-1} is non-constant.

For a fibre with just two elements, we can under the most simple conditions precisely calculate the valuations.

Proposition 7.41 (Two element fibre)**.** *Let $m \in \mathbb{N}$ such that $\lambda^{-1}(m) = \{n, n+1\}$ has two elements. Then α_{n+1} is unbounded, but α_{n+2} is bounded, with*

$$\widehat{\alpha_{n+2}} = -\frac{h_{n+1}}{h_n}, \qquad \text{where } h_n, h_{n+1} \in k^\times .$$

If moreover $\deg a_{n+2} = 1$, then for the partial quotients we have

$$\nu(a_{n+1}) = \nu(g_{n-1}) - \nu(g_n) - (1 + \deg a_{n+1})\, \nu(\ell c(\vartheta_n)), \tag{7.11}$$

$$\nu(\ell c(a_{n+1})) = \nu(g_{n-1}) - \nu(g_n) - \nu(\ell c(\vartheta_n)), \tag{7.12}$$

$$\nu(\alpha_{n+2}) = \nu(a_{n+2}) = \nu(g_n) - \nu(g_{n+1}), \tag{7.13}$$

$$\nu(\ell c(a_{n+2})) = \nu(g_n) - \nu(g_{n+1}) + \nu(\ell c(\vartheta_n)), \tag{7.14}$$

and for the convergents we have

$$\nu(g_{n+1}) = \nu(q_{n+1}) = \nu(g_{n-1}) - (1 + \deg a_{n+1})\, \nu(\ell c(\vartheta_n)), \tag{7.15}$$

$$\nu(\ell c(q_{n+1})) = \nu(g_{n-1}) - \nu(\ell c(\vartheta_n)), \tag{7.16}$$

$$\nu(g_{n+2}) = \nu(q_{n+2}) = \nu(\ell c(q_{n+2})) = \nu(g_n) + (1 + \deg a_{n+1})\, \nu(\ell c(\vartheta_n)). \tag{7.17}$$

Proof. The first part follows from Proposition 7.40. Here n and $n+2$ are minimal in their fibre, so $\vartheta_{n-1}, \vartheta_{n+1} \in O(\!(X^{-1})\!)^\times$, and h_{n-1}, h_n, h_{n+1} are all constant.

Now $\operatorname{ord}(\vartheta_n) = \deg q_{n+1}$, but $\operatorname{ord}(\overline{\vartheta_n}) = \deg v_{m+1} = \operatorname{ord}(\vartheta_n) + \deg a_{n+2}$ because $\deg c_{m+1} = \deg a_{n+1} + \deg a_{n+2}$. So the first $\deg a_{n+2}$ coefficients of ϑ_n vanish after reduction, and when assuming $\deg a_{n+2} = 1$ we can apply the results of section A.2 to compute the valuations. In particular note that $\nu(\ell c(\vartheta_n)) > 0$, while the next coefficient of ϑ_n is in $O^\times$.

With Proposition A.5 on the valuations of a quotient of Laurent series, we easily compute (7.11) and (7.12) from the quotient presentation (7.6) of α_{n+1}. Of course a_{n+1} contains precisely the first $1 + \deg a_{n+1}$ coefficients of α_{n+1}.

Then (7.11) allows to compute

$$\nu(a_{n+1}\, q_n) = \nu(g_{n-1}) - (1 + \deg a_{n+1})\, \nu(\ell c(\vartheta_n)) < \nu(q_{n-1}).$$

This implies (7.15) via $q_{n+1} = a_{n+1}\, q_n + q_{n-1}$ and the ultrametric “equality”. As n is minimal in the fibre, we have $\deg q_n = \deg \widehat{q_n}$ and hence $\nu(g_n) = \nu(q_n) = \nu(\ell c(q_n))$, so

$$\nu(\ell c(q_{n+1})) = \nu(\ell c(q_n)) + \nu(\ell c(a_{n+1})) = \nu(g_{n-1}) - \nu(\ell c(\vartheta_n)).$$

On the other hand, the first part of Lemma A.3 applied to (7.6) gives (7.13) and (7.14) – there are just two coefficients in a_{n+2}. By Theorem 7.2, we also know that

$\deg \widehat{q_{n+2}} = \deg q_{n+2}$, so we can compute the valuation of the convergent via the leading coefficient:

$$\begin{aligned}\nu(g_{n+2}) = \nu(q_{n+2}) = \nu(\ell c(q_{n+2})) &= \nu(\ell c(q_{n+1})) + \nu(\ell c(a_{n+2})) \\ &= \nu(g_{n-1}) - \nu(\ell c(\vartheta_n)) + \nu(g_n) - \nu(g_{n+1}) + \nu(\ell c(\vartheta_n)) \\ &= \nu(g_n) + (1 + \deg a_{n+1})\, \nu(\ell c(\vartheta_n)).\end{aligned}$$

□

Remark 7.42. The h_n and also the quotients h_{n-1}/h_n do not seem to follow any larger (obvious) patterns. If they are all constants, we locally – in "areas" with only single element fibres – observe patterns as in Proposition 5.21. But that is an unsurprising consequence of (7.8).

Remark 7.43. For $\deg a_{n+2} > 1$, there is more than one coefficient of ϑ_n that vanishes, and our reasoning which essentially boils down to geometric series arguments, breaks down. If we wanted to treat for example fibres $\lambda^{-1}(m) = \{n, n+1, n+2\}$ with three elements, we get additional complications, as h_{n+1} can now be non-constant.

We have seen that the reduction of the normalisation of a bounded complete quotient of α yields a complete quotient of γ if and only if we are at a single element fibre of λ. Otherwise, it becomes a rational (or even polynomial) function.

We have also seen that the g_n do not change at the single element fibres. We will later investigate this closer for $\mathbf{CF}(\sqrt{D})$ with $\deg D = 4$ (see Theorem 8.2).

8. Specialization of hyperelliptic continued fractions

We now apply and extend the reduction theory for continued fractions from the previous chapter to square roots. After briefly treating reduction of periodic continued fractions, we finally prove Theorem 1.1 from the introduction, after rephrasing it to include number fields.

We go on to study the valuations more closely for $\deg D = 4$ which leads to Theorem 1.2 about unbounded valuations, also from the introduction.

We also explain how reduction of abelian varieties is related with the reduction of continued fractions via the reduction of the divisors of the convergents. This leads to a well-known effective method for testing if D is Pellian by reducing modulo two primes.

We conclude with a discussion of specialization of continued fractions, i.e. when the base field is $\mathbb{C}(t)$.

We continue using the notation from the previous chapter. From now on, let $D \in O[X]$ non-square with even degree and $\ell c(D) \in O^\times$ a square, so that $\alpha = \sqrt{D} \in O(\!(X^{-1})\!)$, and $\gamma = \overline{\alpha} = \sqrt{\overline{D}}$ (see Proposition 7.9). For example for $D \in \mathbb{Z}[X]$ we can ask that D is monic to ensure that $\sqrt{D} \in \mathbb{Q}(\!(X^{-1})\!)_{\nu_\mathfrak{p}}$ for every prime number $\mathfrak{p}$.

8.1. Reduction of periodic quadratic continued fractions

In this section, we discuss reduction of periodic $\mathbf{CF}(\sqrt{D})$. Then all necessary information is contained in finitely many partial quotients, and we can study reduction by looking at this finite data.

First, we check that nothing strange can happen – we should not be able to reduce to a non-periodic continued fraction. Recall that periodicity of $\mathbf{CF}(\sqrt{D})$ is equivalent to D being Pellian (see Theorem 4.1).

Proposition 8.1. *If D is Pellian, then either $\overline{D}$ is a square, or it is also Pellian.*

Proof. Let $(p,q) \in \mathcal{P}^\times(D)$. Then normalising it, we have also $(\widetilde{p},\widetilde{q}) \in \mathcal{P}^\times(D)$, with reduction $\widehat{q} \neq 0$ in $k[X]$. Of course

$$\widetilde{p}^2 - D\,\widetilde{q}^2 = \omega$$

where $\omega \in O$. If $\omega \in \mathfrak{m}$, then

$$\widehat{p}^2 - \overline{D}\,\widehat{q}^2 = 0$$

which implies $\overline{D}$ is a square.

Otherwise we have $\omega \in O^{\times}$, hence $\overline{\omega} \in k^{\times}$. Then clearly $(\widehat{p}, \widehat{q}) \in \mathcal{P}^{\times}(\overline{D})$ and $\overline{D}$ is Pellian. □

This proof shows that the degree $\deg q$ of the minimal solution can only decrease under reduction. This has been exploited by Platonov [Pla14] to produce Jacobians of hyperelliptic curves over $\mathbb{Q}$ with torsion points of various order. In a previous article together with Petrunin [PP12], he gives $\mathbb{Q}$-rational torsion points of orders 36 and 48. It seems they employ a refined brute force approach for searching Pellian polynomials by checking that D is Pellian only modulo several primes which speeds up the necessary calculations sufficiently (see also Example 4 in Section 10.2).

The following does not even require that D is Pellian:

Proposition 8.2. *If $\overline{D}$ is a square, then $\mathbf{CF}(\sqrt{D})$ has bad reduction, with $\alpha_1 \notin O((X^{-1}))$.*

Proof. This is rather obvious because now $\gamma = \sqrt{\overline{D}} \in k[X]$, so $c_0 = \gamma_0$ and already γ_1 does not exist. So we have bad reduction of $\mathbf{CF}(\sqrt{D})$ by Proposition 7.20. □

Remark 8.3. If $\overline{D}$ is square, the map λ has image $\{0\}$, so there is a single infinite fibre. We neglected this case in Section 7.4.2. As already $\overline{a_0 - \alpha_0} = 0$, we do not get much information about the valuations. So we do not know whether α_1 should be bounded or not.

Remark 8.4. Suppose that $\mathbf{CF}(\sqrt{D})$ is quasi-periodic, with $\mu \in K^{\times}$ such that $\alpha_\ell = \mu\,(A + \sqrt{D})$. Then $\deg a_\ell = d$ being maximal implies by Lemma 7.20 that bad reduction of the continued fraction cannot start at ℓ. As $\nu(A) = \nu(\sqrt{D}) = 0$, we have $\nu(\mu) = 0$ unless bad reduction of $\mathbf{CF}(\sqrt{D})$ occurred already before α_ℓ, i.e. somewhere inside the quasi-period.

In particular, this means that $\mu \in K$ cannot have too many different factors.

Remark 8.5. If k has positive characteristic, it is possible that the (quasi-)period length shortens. This is best understood using the geometric viewpoint from Chapter 4 and will be analysed in Section 8.4.3 later.

Anyway, we can easily determine whether we have good or bad reduction of periodic $\mathbf{CF}(\alpha)$ by checking whether any of $\nu(\ell c(a_1)), \ldots, \nu(\ell c(a_\ell))$ is negative (with ℓ the quasi-period length).

The quasi-period being palindromic (see Proposition 5.41) also implies that the bad reduction of the continued fraction must start at the latest at $\frac{\ell}{2}$ for ℓ even, or $\frac{\ell-1}{2} + 2$ for ℓ odd (in the latter case, we have to account for $\nu(\mu) \neq 0$).

8.1.1. Reduction for $\deg D = 2$

Let us briefly describe what happens in the case $\deg D = 2$.

Suppose for simplicity that D is monic, then we can write $D = (X + b)^2 + \omega$ with $b, \omega \in O$ and $\omega \neq 0$ so that D is not a square. Of course $A = X + b$, and one easily computes

$$\alpha_0 = \sqrt{D}, \quad \alpha_{2i+1} = \frac{A + \sqrt{D}}{\omega}, \quad \alpha_{2i} = A + \sqrt{D}.$$

So bad reduction occurs if and only if $\overline{\omega} = 0$, in which case $\overline{D}$ is a square. Obviously we can reduce the α_{2i} directly, but the α_{2i+1} only after normalising. If $\overline{\omega} = 0$, then the map λ has a single infinite fibre and clearly the $\widehat{\alpha_n}$ are all polynomials.

The partial quotients are

$$a_0 = A, \quad a_{2i+1} = \frac{2\,A}{\omega}, \quad a_{2i} = 2\,A$$

with Gauss norms

$$\nu(a_0) = 0, \quad \nu(a_{2i+1}) = -\nu(\omega), \quad \nu(a_{2i}) = 0$$

which of course remain bounded.

As to the convergents, it is easy to see that ($\lceil\cdot\rceil_{\mathbb{Z}}$ is the ceiling function)

$$\nu(p_n) = \nu(q_n) \geq -\left\lceil\frac{n}{2}\right\rceil_{\mathbb{Z}} \nu(\omega) \text{ and } \nu(\ell c(q_n)) = -\left\lceil\frac{n}{2}\right\rceil_{\mathbb{Z}} \nu(\omega)$$

hence $\nu(q_n) = -\left\lceil\frac{n}{2}\right\rceil_{\mathbb{Z}} \nu(\omega)$ and $\deg q_n = \deg \widehat{q_n}$. Indeed we expect this in the case of good reduction of $\mathbf{CF}(\sqrt{D})$.

Otherwise we have bad reduction of $\mathbf{CF}(\sqrt{D})$, hence $\overline{A}^2 = \overline{D}$ and $\nu(\omega) > 0$. Then we also know that $(\widehat{p_n}, \widehat{q_n}) = h_n\,(\overline{A}, 1)$ for all $n \geq 0$. We can even calculate this: set $\eta = \pi^{-\nu(\omega)}$ (so that $\eta/\omega \in O^\times$). Then for even n we get $g_n^{-1} = \eta^{n/2}$ and for odd n we get $g_n^{-1} = \eta^{(n+1)/2}$. We calculate for even n:

$$\begin{aligned}&\widetilde{p_n} = \eta^{n/2} p_n = 2A\,\eta^{n/2}\,p_{n-1} + \eta^{n/2}\,p_{n-2} = 2A\,\widetilde{p_{n-1}} + \eta\,\widetilde{p_{n-2}}, \text{ and similarly}\\ &\widetilde{q_n} = 2A\,\widetilde{q_{n-1}} + \eta\,\widetilde{q_{n-2}}\end{aligned}$$

which yields

$$\widehat{p_n} = 2\overline{A}\,\widehat{p_{n-1}}, \quad \widehat{q_n} = 2\overline{A}\,\widehat{q_{n-1}}.$$

On the other hand, we get for n odd

$$\begin{aligned}&\widetilde{p_n} = \eta^{(n+1)/2} p_n = \frac{2A}{\omega}\,\eta^{(n+1)/2} p_{n-1} + \eta^{(n+1)/2} p_{n-2} = 2A\,\frac{\eta}{\omega}\,\widetilde{p_{n-1}} + \eta\,\widetilde{p_{n-2}}, \text{ and}\\ &\widetilde{q_n} = 2A\,\frac{\eta}{\omega}\,\widetilde{q_{n-1}} + \eta\,\widetilde{q_{n-2}}\end{aligned}$$

so

$$\widehat{p_n} = 2\overline{A}\,\overline{\eta/\omega}\,\widehat{p_{n-1}}, \quad \widehat{q_n} = 2\overline{A}\,\overline{\eta/\omega}\,\widehat{q_{n-1}}.$$

It follows that h_n is $\overline{A}^n$ times some constant factor depending on n.

8.2. Reduction of non-periodic quadratic continued fractions

As before, let $D \in O[X]$ non-square with even degree, and $\ell c(D) \in O^\times$ a square, so that $\alpha = \sqrt{D} \in O((X^{-1}))$, and $\gamma = \overline{\alpha} = \sqrt{\overline{D}}$. But now, we assume that $\mathbf{CF}(\sqrt{D})$ is non-periodic. Recall that this requires $\deg D \geq 4$ (Corollary 2.16 and Theorem 6.3).

8.2.1. Reduction to square

If $\overline{D}$ is a square, this implies bad reduction of $\mathbf{CF}(\sqrt{D})$ by Proposition 8.2. Then $\lambda : \mathbb{N}_0 \to \mathbb{N}_0$ has image $\{0\}$, so we do not get a lot of information from it.

Anyway, for a fixed D, this can happen only for finitely many valuations ν. From Proposition 2.11 about completion of the square and Proposition 7.9 about boundedness of the square root, it follows that $\overline{D}$ is a square if and only if $\nu(D - A^2) > 0$ holds. So this can be checked easily, and concerns only finitely many valuations.

See Example 2 in Section 10.1 for a non-periodic $\mathbf{CF}(\sqrt{D})$ where $\overline{D}$ is a square.

8.2.2. Reduction to periodic and denominators

We now study the case where $\mathbf{CF}(\gamma)$ becomes periodic. This happens automatically if k is finite, for example with $D \in \mathbb{Z}[X]$ and reduction modulo some odd prime (see e.g. Corollary 6.2). Instead of talking about denominators which is rather vague, we are looking for negative valuation.

Lemma 8.6. *If $\mathbf{CF}(\gamma)$ is periodic, then infinitely many fibres of λ have at least 2 elements. Hence $\mathbf{CF}(\alpha)$ has bad reduction at ν, and there exists $n > 0$ where α_n has negative valuation in the leading coefficient.*

Proof. Corollary 6.1 implies that for all $n \geq 1$ holds $\deg a_n < \frac{1}{2} \deg D$, and that there are infinitely many (because of pure periodicity) $m \geq 1$ such that $\deg c_m = \frac{1}{2} \deg \overline{D}$.

However, $\deg D = \deg \overline{D} = 2d$, and good reduction of $\mathbf{CF}(\alpha)$ would by Remark 7.15 imply that $\deg a_n = \deg c_n$ for all n. In fact, by Corollary 7.31, for every $m \geq 1$ with $\deg c_m = d$, the fibre $\lambda^{-1}(m-1)$ has more than a single element, so λ is certainly not bijective.

So $\mathbf{CF}(\alpha)$ must have bad reduction. The statement about negative valuation then follows directly from Proposition 7.20. □

Remark 8.7. Proposition 7.40 implies that in the case of bad reduction of $\mathbf{CF}(\alpha)$ each fibre with more than one element yields an unbounded complete quotient. It follows that there are infinitely many unbounded complete quotients. Compare also Proposition 7.20.

If $\deg D = 4$, the statement of the Lemma becomes an equivalence:

Proposition 8.8. *Suppose $\deg D = 4$ and $\overline{D}$ non-square. Then $\mathbf{CF}(\alpha)$ has bad reduction if and only if $\mathbf{CF}(\gamma)$ is periodic.*

Proof. This extends Lemma 8.6, it only remains to prove that bad reduction of $\mathbf{CF}(\alpha)$ implies periodicity. As $\deg D = 4$, we have $\deg a_n = 1$ for all $n \geq 1$. By Proposition 7.20, there is a minimal complete quotient α_n with $\nu(\alpha_n) < 0$. Because $\overline{D}$ is non-square, $\mathbf{CF}(\gamma)$ is infinite and hence the proposition also implies $\deg c_n > 1$.

Then from Corollary 6.1 follows $\deg c_n \leq \frac{1}{2} \deg \overline{D} = 2$, so $\deg c_n = 2$ and thus $\mathbf{CF}(\gamma)$ must be periodic. □

Remark 8.9. If the residue field k is finite (and K is obviously infinite), then unless $\overline{D}$ is square, one always has periodic $\mathbf{CF}(\gamma)$ (see Corollary 6.2). So it is impossible to avoid bad reduction of $\mathbf{CF}(\sqrt{D})$ in that case, for example if the base field K is a number field.

Remark 8.10. Lemma 8.6 works only if α is a square root of a polynomial (or shares a complete quotient with some $\sqrt{D}$). As we are interested in periodicity, we may assume that α is σ-reduced (the complete quotients eventually have this property, see Proposition 6.8). But then $\deg a_n = \deg A$ is equivalent to $\alpha_n = \mu\,(A+\sqrt{D})$, so we would necessarily end up in the continued fraction expansion of $\sqrt{\mu^2\,D}$.

8.2.3. Primes occurring in infinitely many denominators

We are now ready to attack the proof of Theorem 1.1 from the introduction about a prime occurring in infinitely many denominators of the a_n. We first give a more technical version for a fixed prime, which holds in full generality.

Proposition 8.11. *Suppose that there are infinitely many fibres of λ with one element, and infinitely many fibres with at least two elements. Then there exist infinitely many n with $\nu(\ell c(\alpha_n)) < 0$, i.e. infinitely many complete (and partial) quotients have the "prime" as a factor in the denominator of the leading coefficient.*

Proof. Let $N \geq 0$. By assumption, there exists $n \geq N$ such that $\{n\} = \lambda^{-1}(m)$ for some m. Then Proposition 7.39 implies

$$\nu(\ell c(\alpha_{n+1})) = \nu(\alpha_{n+1}) = \nu(g_{n-1}) - \nu(g_n) = \nu(g), \quad \nu(g_{n+1}) = \nu(g_{n-1}), \tag{8.1}$$

where we set $g = g_{n-1}/g_n$. Hence $g^{-1}\,\alpha_{n+1} \in O(\!(X^{-1})\!)$ with leading coefficient in $O^\times$, so from Theorem 7.1 and Proposition 7.20 we deduce that the first complete quotient $g\,\alpha_{n+2}$ cannot have positive valuation in the leading coefficient:

$$\nu(\ell c(g\,\alpha_{n+2})) \leq 0 \implies \nu(\ell c(\alpha_{n+2})) \leq -\nu(g) = -\nu(\ell c(\alpha_{n+1})).$$

So if $\nu(\ell c(\alpha_{n+1})) \neq 0$, either α_{n+1} or α_{n+2} has the desired negative valuation in the leading coefficient.

Otherwise $\nu(\ell c(\alpha_{n+1})) = 0$, so $\alpha_{n+1} \in O(\!(X^{-1})\!)$ and we can reproduce the argument from Proposition 7.20: Let $n' > n$ minimal such that $\lambda^{-1}(\lambda(n'))$ has multiple elements. In this case we know

$$\nu(\ell c(\alpha_{n'+1})) < \nu(\alpha_{n'+1}) = \nu(g_{n'-1}) - \nu(g_{n'}) = \pm(\nu(g_{n-1}) - \nu(g_n)) = 0$$

by the minimality of n'. So the desired pole is in the leading coefficient of $\alpha_{n'+1}$. □

Remark 8.12. The proof also illustrates that there are infinitely many partial quotients with negative valuation, even if we multiply α with π^e for some $e \in \mathbb{Z}$ (or some other constant). This merely changes g in in (8.1); and recall Proposition 5.21 about multiplying a continued fraction with a constant factor. We do not even need to assume $\alpha \in O(\!(X^{-1})\!)$ here, if we define λ appropriately – it is determined by the sequences $\deg a_n$ and $\deg c_m$ which do not change under this multiplication.

Our results from Section 7.4.2 tell us that multiple element fibres correspond to unbounded complete quotients, and hence bad reduction of the continued fraction. If we want this to occur repeatedly, it is very natural to ask for infinitely many such fibres.

On the other hand, asking also for infinitely many fibres with just a single element is a more technical condition. Right now we cannot avoid this because we do not understand how the valuations behave for multiple element fibres (except for $\deg D = 4$, to be treated in Theorem 8.2). Recall that the complete quotients belonging to multiple element fibres, if at all, reduce to rational functions, about which we have hardly any information (see Proposition 7.40).

At least we have a simple criterion that guarantees the existence of infinitely many fibres of λ with just a single element:

Proposition 8.13. *With $\alpha = \sqrt{D}$, suppose that $\mathbf{CF}(\alpha)$ is non-periodic, but $\mathbf{CF}(\gamma)$ is periodic. Let $\delta = \min\{\deg a_n \mid n \geq 0\}$. If there exists m such that $\deg c_m = \delta$, then $\lambda : \mathbb{N}_0 \to \mathbb{N}_0$ has infinitely many fibres with a single element.*

Proof. Recall that $\deg D = 2d$. From Corollary 6.1 we know that $\deg a_n < d = \deg a_0 = \deg c_0$ for $n \geq 1$, so certainly $m \geq 1$. But the quasi-period of $\mathbf{CF}(\gamma)$ begins at c_1; this means there are actually infinitely many m with $\deg c_m = \delta$.

Then Corollary 7.31 implies for $\lambda^{-1}(m-1) = \{n-1, \dots, n-1+l\}$ that $\deg a_n + \cdots + \deg a_{n+l} = \deg c_m = \delta$. Minimality of δ forces $l = 0$, hence the fibre has a single element. It follows are infinitely many fibres of λ with a single element. □

Remark 8.14. Note that along the way, we have proved that there are infinitely many partial quotients with $\deg a_n = \delta$, so the minimal degree must be assumed infinitely often (however, we used periodicity of $\mathbf{CF}(\gamma)$ which is in general a rather strong hypothesis).

Now we restrict ourselves to K being a number field. The ring of integers $\mathcal{O}_K$ of a number field, while it need not be a unique factorisation domain, has unique factorisations of ideals into prime ideals. Every $x \in K$ can be written as $x = \frac{a}{b}$ with $a \in \mathcal{O}_K$ and $b \in \mathbb{N}$. In the theorem below, we refer to b as the denominator.

Each prime ideal $\mathfrak{P}$ of $\mathcal{O}_K$ corresponds to a non-archimedean valuation $\nu_{\mathfrak{P}}$ on K. By localising $\mathcal{O}_K$ at $\mathfrak{P}$, one obtains a discrete valuation ring with a finite residue field. The latter is a finite extension of $\mathbb{F}_{\mathfrak{p}}$, where $\mathfrak{p}$ is the unique prime number (of $\mathbb{Z}$) contained in $\mathfrak{P}$.[1]

The following generalises Theorem 1.1 from the introduction:

Theorem 8.1. *Let K a number field, suppose that $D \in K[X]$ is monic, non-square and has even degree, but is not Pellian.*

Then for all but finitely prime numbers (in $\mathbb{Z}$), the prime $\mathfrak{p}$ appears in infinitely many a_n (actually $\ell c(a_n)$) in a (the) denominator.

The primes excluded are 2 (because the residue field would have characteristic 2), those which already appear in a denominator in D, and those which make $D \pmod{\mathfrak{P}}$

[1] See for example [Neu99], Chapter I §8. Or any other decent textbook on algebraic number theory.

a square polynomial. Additionally, we may need to exclude a finite number of primes, depending on where the first partial quotient of minimal degree δ occurs. This can be made effective, as discussed below in Remark 8.15.

The Theorem relies on α being a square root. For other elements of $K(X, \sqrt{D})$, we may in fact have good reduction of the continued fraction at infinitely many primes, see Section 8.5 for an example.

Proof. Removing the finitely many primes with $\nu_{\mathfrak{P}}(D) < 0$ (i.e. $\mathfrak{p}$ is in the denominator of D), and ignoring the primes $\mathfrak{P}$ above 2, the conditions on D ensure that $\sqrt{D} \in O(\!(X^{-1})\!)$ for $\nu = \nu_{\mathfrak{P}}$. Let us also ignore the $\mathfrak{P}$ for which $D - A^2 \in \mathfrak{P}[X]$, i.e. the reduction $\overline{D}$ is a square (there are only finitely many, as $D - A^2 \in K[X]$ is a polynomial).

Of course D not Pellian means that $\mathbf{CF}(\sqrt{D})$ is non-periodic. However, the residue field k is finite for every prime, so $\mathbf{CF}(\sqrt{\overline{D}})$ must necessarily be periodic. Then Lemma 8.6 implies that we are always in the case of bad reduction (of the continued fraction), and there are infinitely many fibres of λ which have at least two elements.

In order to apply Proposition 8.11, we use Proposition 8.13, so we need to check that there exists m with $\deg c_m = \delta = \min\{\deg a_n \mid n \geq 0\}$.

Let n_0 the minimal n with $\deg a_n = \delta$ (obviously $n_0 \geq 1$). We restrict to primes $\mathfrak{P}$ for which we have good reduction of $\mathbf{CF}(\sqrt{D})$ up to n_0, i.e. $\alpha_0, \alpha_1, \ldots, \alpha_{n_0} \in O(\!(X^{-1})\!)$. This excludes only finitely many $\mathfrak{P}$: we can factor each $\ell c(a_n) \in K$ for $n = 0, \ldots, n_0$ into a product (with possibly negative exponents) of prime ideals of $\mathcal{O}_K$. Of course $\nu_{\mathfrak{P}}(\ell c(a_n)) < 0$ happens if and only if $\mathfrak{P}$ appears with a negative exponent in the factorisation. Of these there are obviously just finitely many, and by Proposition 7.20 the bad reduction of $\mathbf{CF}(\sqrt{D})$ starts only later for all other primes.

For the remaining primes $\mathfrak{P}$, the complete quotients up to α_{n_0} are thus contained in $O(\!(X^{-1})\!)$. By Proposition 7.13 and Remark 7.15 this implies $\deg c_{n_0} = \deg a_{n_0} = \delta$.

Hence $\nu_{\mathfrak{P}}(\ell c(a_n)) < 0$ for infinitely many n. If we write $\ell c(a_n) = a/b$ with $a \in \mathcal{O}_K$ and $b \in \mathbb{N}$, then naturally $\nu_{\mathfrak{P}}(a) \geq 0$, hence $\nu_{\mathfrak{P}}(b) > 0$. Applying the Norm, $b \in \mathbb{N}$ must have $\mathfrak{p} \mid b$ as desired. □

Let us briefly discuss effectivity of $\delta = \min\{\deg a_n \mid n \geq 0\}$. In [Zan16], it is shown how a Skolem-Mahler-Lech theorem for algebraic groups implies that the sequence of the $\deg a_n$ is eventually periodic (even if $\mathbf{CF}(\sqrt{D})$ is not periodic!). While in certain cases it seems possible to obtain from this a bound for the period length (of the degrees), there is unfortunately no information on the pre-period. Summing upper bounds for the pre-period and the period would of course produce a upper bounds for δ.

However, the issues with effectivity are actually related to finding $\max\{\deg a_n \mid n \geq 1\}$. For finding the minimal degree, we do not need to know the entire period (of degrees):

Remark 8.15. Let $\mathcal{V}$ the Zariski closure of $\{n\,\mathbf{O} \mid n \in \mathbb{Z}\}$ in the Jacobian of $\mathcal{C}$; we need to find the maximal r such that $\mathcal{V} \subset W_r$ but $\mathcal{V} \not\subset W_{r-1}$. Using the divisor relations coming from the convergents, explained in Sections 4.3 and 6.6, this implies $\delta = g - r + 1$ for said maximal r (recall from Section 4.2.3 that W_r is the r fold symmetric sum of $\mathcal{C}$ embedded in its Jacobian variety).

We can effectively compute $\mathcal{V}$ from a factorisation of the Jacobian as in Theorem 1.2 of [GR14] (which extends a deep result of Masser and Wüstholz, [MW14]). As the W_r can also be effectively represented, we can determine in which of the W_r, $(r = 1, \dots, g-1)$ our subvariety $\mathcal{V}$ is not contained. Certainly $\mathcal{V}$ is contained in $W_g = \mathcal{J}(\mathcal{C})$.

In practice finding δ is not a big issue because we usually immediately find a partial quotient with $\deg a_n = 1$ (which in fact *must* occur for $\deg D = 4$ or 6, see Theorem 1.2 of [Zan16], stated below as Theorem 8.7), and then we know that $\delta = 1$.

Theorem 8.2 below gives another (similar) proof for the occurrence of a prime in the denominators of infinitely many a_n in the case $\deg D = 4$. This relies on being able to control cancellation issues sufficiently, so we do not need the single element fibres to estimate the Gauss norms.

8.3. Genus 1 valuation patterns

We analyse the case $\deg D = 4$ much closer now, and will describe how the valuations of the complete quotients, partial quotients and convergents behave in the case of bad reduction at ν. When studying examples (see Tables 10.1 and 10.2 in Section 10.3.1), one notes that the valuations (Gauss norms) of the partial quotients a_n are often divisible by 4, with alternating signs, while the valuations of the convergents q_n are always divisible by 2, again with alternating signs. Both also exhibit an almost pseudo-periodic behaviour. The theorem below aims to explain these patterns:

Theorem 8.2. *Suppose* $\deg D = 4$, *and that* $\mathbf{CF}(\sqrt{D})$ *is non-periodic, while* $\mathbf{CF}(\sqrt{\overline{D}})$ *is periodic with quasi-period* ℓ, *so that we have bad reduction as shown in Proposition 8.8. Then we observe the following:*

- *The unbounded complete quotients* α_n *are exactly those with*
$$n \in \mathcal{U} = \{j\,(\ell+1) - 1 \mid j \geq 1\}. \tag{8.2}$$
- *Defining*
$$f_n = \nu(\ell c(\vartheta_{n-1})) \geq 0$$
we have $f_n > 0$ *if and only if* $n \in \mathcal{U}$ *(so* $f_n = 0$ *otherwise).*
- *Recursively defining* $F_0 = 0$ *and* $F_n = -(F_{n-1} + f_n)$, *we get formulas for the valuations*
$$\nu(a_n) = 2\,(F_{n-2} + F_n), \qquad \nu(\ell c(a_n)) = \nu(a_n) + f_{n-1} + f_n,$$
$$\nu(q_n) = 2\,F_n, \qquad \nu(\ell c(q_n)) = \nu(q_n) + f_n.$$

Remark 8.16. Note that if $n-1, n \notin \mathcal{U}$, we have $F_{n-2} = F_n$ and thus $\nu(a_n) = 4\,F_n$, explaining the divisibility by 4.

Remark 8.17. For higher genus, one probably has to consider other coefficients besides f_n. But it is not at all clear how this generalises.

The following is the general version of Theorem 1.2 (recall that $\ell+1$ is the torsion order of $[\mathbf{O}_{\mathrm{red}}]$, see Proposition 6.11):

Corollary 8.18. *Under the same hypotheses as Theorem 8.2, and additionally assuming the quasi-period ℓ of $\mathbf{CF}(\sqrt{D})$ is odd, the Gauss norms grow at least linearly (in particular they are unbounded):*

$$(-1)^n\nu(a_n) \geq 2\left(\left\lfloor\frac{n-1}{\ell+1}\right\rfloor_{\mathbb{Z}} + \left\lfloor\frac{n+1}{\ell+1}\right\rfloor_{\mathbb{Z}}\right), \qquad (-1)^n\nu(q_n) \geq 2\left\lfloor\frac{n+1}{\ell+1}\right\rfloor_{\mathbb{Z}}.$$

Proof. We can easily write F_n as an alternating sum of the f_n:

$$F_n = \sum_{j=0}^{n}(-1)^{n-j+1}\,f_j = \sum_{\substack{j\in\mathcal{U},\\ j\leq n}}(-1)^{n-j+1}\,f_j = (-1)^n\sum_{\substack{j\in\mathcal{U},\\ j\leq n}}(-1)^{j+1}\,f_j.$$

In case ℓ is odd, for every $j \in \mathcal{U}$ we have $j+1 = i\,(\ell+1)$ even. For $j+1 \leq n+1$, we have $1 \leq i \leq \frac{n+1}{\ell+1}$. As every $f_j \geq 1$, this implies $(-1)^n\,F_n \geq \left\lfloor\frac{n+1}{\ell+1}\right\rfloor_{\mathbb{Z}}$. With the formulas from the theorem, we get the desired estimates for the Gauss norm. □

Remark 8.19. For even ℓ, it is completely unclear if the F_n could be bounded. In example calculations, we sometimes observe cancellation, but not always (see Table 10.2) . As we currently have almost no control over the f_n for $n \in \mathcal{U}$, any result in this direction would be quite surprising.

In some examples, we see almost periodic patterns in the values of the f_n. Usually though, there comes a disturbance in these patterns at some point. We will revisit this issue briefly in Section 8.4.4.

However, we can check that the valuations are negative for infinitely many n (giving another proof of Theorem 8.1 for $\deg D = 4$):

Corollary 8.20. *Under the hypotheses of the Theorem 8.2, there are infinitely many n with $\nu(a_n) < 0$, and infinitely many n with $\nu(q_n) < 0$.*

Proof. The previous Corollary 8.18 gives a stronger statement when the quasi-period length ℓ of $\mathbf{CF}(\gamma)$ is odd, so we only need to check the case where ℓ is even. In particular, this means $\ell \geq 2$.

So if we take $n \in \mathcal{U}$, this implies (from the structure of $\mathcal{U}$) that $f_n > 0$ but $f_{n+1} = f_{n+2} = 0$, hence

$$F_n = -(F_{n-1}+f_n), \quad F_{n+1} = -F_n = F_{n-1} - f_n, \quad F_{n+2} = -F_{n+1} = F_n.$$

If both $F_{n-1} \geq 0$ and $F_n \geq 0$, then also $F_{n-1} + F_n = -f_n \geq 0$; but that contradicts our choice of n. So one of $\nu(q_{n-1}) < 0$ or $\nu(q_n) < 0$ must be satisfied.

Similarly, if both $F_{n-1} + F_{n+1} = 2\,F_{n-1} - f_n \geq 0$ and $F_n + F_{n+2} = 2\,F_n = -2\,F_{n-1} - 2\,f_n \geq 0$, then also $F_{n-1} + F_{n+1} + F_n + F_{n+2} = -3\,f_n \geq 0$; this is again a contradiction. Hence at least one of $\nu(a_{n+1}) < 0$ or $\nu(a_{n+2}) < 0$ is satisfied.

As $\mathcal{U}$ is infinite, we find infinitely many of these partial quotients and convergents. □

Remark 8.21. For all n, we have $\widehat{p_n}$ and $\widehat{q_n}$ coprime because for single and two element fibres we observed that all the h_n must be constant (see Propositions 7.39 and 7.41).

We begin the proof of Theorem 8.2 by analysing the fibres of λ. For the rest of this section, assume the hypotheses on D from the Theorem are satisfied.

Proposition 8.22. *The fibres of λ have at most 2 elements. The fibres with 2 elements are given by*

$$\lambda^{-1}(j\,\ell-1)=\{j(\ell+1)-2, j(\ell+1)-1\}, \quad j\geq 1,$$

all other fibres have just one element.

Proof. The degrees of the a_n are given by the sequence $2,1,1,1,\ldots$, while the degrees of the c_n are given by the sequence $2,1,\ldots,1,2,1,\ldots,1,2,1,\ldots$ with precisely $\ell-1$ "1" between the "2" (the quasi-period is determined by the degrees of the partial quotients, see Corollary 6.7). Recall that for a fibre $\{n,n+1,\ldots,n+l\}=\lambda^{-1}(m)$ we always have $\deg a_{n+1}+\cdots+\deg a_{n+l+1}=\deg c_{m+1}$ (see Corollary 7.31). So clearly, we can have at most two elements in a fibre.

The first fibre with two elements, due to $\deg c_\ell = 2$ by the properties of the quasi-period, is

$$\lambda^{-1}(\ell-1)=\{\ell-1,\ell\}.$$

In fact, we generally have $\deg c_{j\,\ell}=2$. In between, there are always $\ell-1$ fibres with a single element, so the minimal element increases by $\ell+1$ each time:

$$\lambda^{-1}(j\,\ell-1)=\{\ell-1+(j-1)\,(\ell+1),\ell+(j-1)\,(\ell+1)\}=\{j\,(\ell+1)-2,j\,(\ell+1)-1\}.$$

□

Proof of Theorem 8.2. With our analysis of fibres with one or two elements (Proposition 7.39 and 7.41), this implies directly (8.2): the unbounded complete quotients come only from the minimal element of the two element fibres (index of course shifted by 1). These results also show that $\operatorname{ord}\big(\overline{\vartheta_n}\big)>\operatorname{ord}(\vartheta_n)$, i.e. $f_{n+1}=\nu(\ell c(\vartheta_n))>0$, happens just for the minimal element of the two element fibres. As the definition of f_n corrects for the index shift, it is clear that $f_n>0$ happens precisely if α_n is unbounded.

It remains to check the valuation formulas, for which we use a complete induction. Recall that $\nu(g_n)=\nu(q_n)$.

For $n=0$, by our assumption on D we have $\nu(\ell c(a_0))=\nu(a_0)=0$. As $q_0=1$, the valuation formulas are clearly satisfied.

Actually, we should also check $n=1$. But the careful reader will find that we use the induction hypothesis for "$n-2$" only for $\nu(q_{n-2})$. So we can check $n=-1$ instead of $n=1$.

By convention, we have $p_{-1}=1, q_{-1}=0$, so $\vartheta_{-1}=1$. So $\nu(q_{-1})=\infty$ looks like a problem, but in fact we only need $\nu(g_{-1})=0$. Recall that $g_{-1}=1$ is the normalisation factor of the "canonical convergent" $(p_{-1},q_{-1})=(1,0)$.

For the induction step, we first check the single element fibre case:

Suppose $\{n\} = \lambda^{-1}(m)$, so we refer to Proposition 7.39. In this case, $f_n = f_{n+1} = 0$. Hence

$$\nu(a_{n+1}) = \nu(\ell c(a_{n+1})) = \nu(\alpha_{n+1}) = \nu(g_{n-1}) - \nu(g_n) = 2(F_{n-1} - F_n) = 2(F_{n-1} + F_{n+1})$$

which covers a_n and its leading coefficient.

We also get

$$\nu(\ell c(q_{n+1})) = \nu(q_{n+1}) = \nu(g_{n-1}) = 2\,F_{n-1} = 2\,F_{n+1}.$$

Then we verify the valuation formulas for the two element fibre:

If on the other hand $\{n, n+1\} = \lambda^{-1}(m)$, we use Proposition 7.41. As observed above, $f_n = f_{n+2} = 0$, but $f_{n+1} > 0$. We already computed in (7.11)

$$\nu(a_{n+1}) = \nu(g_{n-1}) - \nu(g_n) - 2\,f_{n+1} = 2(F_{n-1} - F_n - f_{n+1}) = 2\,(F_{n-1} + F_{n+1}),$$

and also $\nu(\ell c(a_{n+1})) = \nu(a_{n+1}) + f_{n+1}$ as desired. For the convergent, we had

$$\nu(q_{n+1}) = \nu(g_{n-1}) - 2f_{n+1} = 2(F_{n-1} - f_{n+1}) = 2\,F_{n+1}.$$

Moreover, using $F_n = -F_{n-1}$, we find

$$\begin{aligned}\nu(\ell c(q_{n+1})) &= \nu(\ell c(a_{n+1})) + \nu(\ell c(q_n))\\ &= 2\,(F_{n-1} + F_{n+1}) + f_{n+1} + 2\,F_n + f_n = 2\,F_{n+1} + f_{n+1}.\end{aligned}$$

For the second element of the fibre, we have (analogous to the calculation for the single element fibre, but using only $f_{n+2} = 0$)

$$\nu(a_{n+2}) = \nu(g_n) - \nu(g_{n+1}) = 2(F_n + F_{n+2})$$

and $\nu(\ell c(a_{n+2})) = \nu(a_{n+2}) + f_{n+1}$, as desired. For the convergent, we have

$$\nu(\ell c(q_{n+2})) = \nu(q_{n+2}) = \nu(q_n) + 2\,f_{n+1} = 2\,(F_n + f_{n+1}) = -2\,F_{n+1} = 2\,F_{n+2}.$$

This concludes the proof of Theorem 8.2. □

To visualise this, have a look at the three following tables, with horizontal lines directly before the unbounded complete quotients:

n	$\deg a_n$	$\deg q_n$	$\operatorname{ord}(\vartheta_n)$	m	$\deg c_m$	$\deg v_m$	$\operatorname{ord}\big(\overline{\vartheta_n}\big)$
0	2	0	1	0	2	0	1
1	1	1	2	1	1	1	2
⋮							
$l-1$	1	$l-1$	l	$l-1$	1	$l-1$	$l+1$
l	1	l	$l+1$	$l-1$		$l-1$	$l+1$
$l+1$	1	$l+1$	$l+2$	l	2	$l+1$	$l+2$
$l+2$	1	$l+2$	$l+3$	$l+1$	1	$l+2$	$l+3$
⋮							
$2l$	1	$2l$	$2l+1$	$2l-1$	1	$2l$	$2l+2$
$2l+1$	1	$2l+1$	$2l+2$	$2l-1$		$2l$	$2l+2$
$2l+2$	1	$2l+2$	$2l+3$	$2l$	2	$2l+2$	$2l+3$
$2l+3$	1	$2l+3$	$2l+4$	$2l+1$	1	$2l+3$	$2l+4$
⋮							

For simplicity, we assume that all $f_n \leq 1$. For l odd, we get

n	f_n	$\nu(a_n)$	$\nu(\alpha_n)$	$\nu(\ell c(a_n))$	$\nu(q_n)$	$\nu(\ell c(q_n))$
0	0	0	0	0	0	0
1	0	0	0	0	0	0
$\vdots$					0	0
$l-1$	0	0	0	0	0	0
l	1	-2	$-\infty$	-1	-2	-1
$l+1$	0	2	3	2	2	2
$l+2$	0	-4	-4	-4	-2	-2
$l+3$	0	4	4	4	2	2
$\vdots$						
$2l-1$	0	-4	-4	-4	-2	-2
$2l$	0	4	4	4	2	2
$2l+1$	1	-6	$-\infty$	-5	-4	-3
$2l+2$	0	6	6	7	4	5
$2l+3$	0	-8	-8	-8	-4	-4
$2l+4$	0	8	8	8	4	4
$\vdots$						

But for l even, we get

n	f_n	$\nu(a_n)$	$\nu(\alpha_n)$	$\nu(\ell c(a_n))$	$\nu(q_n)$	$\nu(\ell c(q_n))$
0	0	0	0	0	0	0
1	0	0	0	0	0	0
$\vdots$					0	0
$l-1$	0	0	0	0	0	0
l	1	-2	$-\infty$	-1	-2	-1
$l+1$	0	2	3	2	2	2
$l+2$	0	-4	-4	-4	-2	-2
$l+3$	0	4	4	4	2	2
$\vdots$						
$2l-1$	0	4	4	4	2	2
$2l$	0	-4	-4	-4	-2	-2
$2l+1$	1	2	$-\infty$	3	0	1
$2l+2$	0	-2	-2	-1	0	0
$2l+3$	0	-8	-8	-8	0	0
$2l+4$	0	8	8	8	0	0
$\vdots$						

See also the tables of the Example 5 in Section 10.3.1

8.4. Reduction of abelian varieties

To understand how the quasi-period length may change under reduction, and hence to understand bad reduction to a periodic continued fraction, it serves to study reduction of torsion points on the Jacobian of the (hyper)elliptic curve.

8.4.1. Reduction of curve and its Jacobian

Our first step is to define a model of $\mathcal{C}$ over O. Here we can mostly retrace the steps from Section 4.1: instead of over the field $\mathbb{K}$, we are working over the discrete valuation ring O. Note that $\operatorname{Spec} O$ is a local affine Dedekind scheme of dimension 1 with just two points: the generic point and a closed point corresponding to $\mathfrak{m}$.[2]

For this section, we assume that both D and $\overline{D}$ are square-free to ensure that the curve $\mathcal{C}$ (and its Jacobian) has good reduction at ν.

Gluing together

$$\operatorname{Spec} O[X,Y]/\left(Y^2 - D(X)\right) \text{ and } \operatorname{Spec} O[U,V]/\left(V^2 - D^\flat(U)\right)$$

via the morphisms given by $X\,U = 1$ and $U^{g+1}\,Y = V$ (respectively $X^{g+1}\,V = Y$) we get a scheme $\mathcal{X}$ of dimension 2 which is our model of $\mathcal{C}$ over O. Note that the coefficients of $D^\flat$ are those of D in reverse order, so $D^\flat \in O[U]$.

Think of the surface $\mathcal{X}$ as containing two curves: the fibre $\mathcal{X}_0$ over the generic point of $\operatorname{Spec} O$ which is essentially our curve $\mathcal{C}$, and the fibre $\mathcal{X}_\mathfrak{m}$ over the closed point of $\operatorname{Spec} O$ which is the curve $\mathcal{C}_{\text{red}}$ defined over k, with D replaced by $\overline{D}$.

Proposition 8.23. *The fibered surface $\mathcal{X} \to \operatorname{Spec} O$ is normal, regular, projective and flat, in other words it is a normal arithmetic surface.*

Proof. Normal and regular are local conditions and may be checked at the stalks. Hence this follows from $\mathcal{C}$ and $\mathcal{C}_{\text{red}}$ being normal and smooth (and thus regular), see Propositions 4.1 and 4.3.

Flatness follows from surjectivity via Proposition 4.3.9 of [Liu02]: clearly the generic point of $\mathcal{X}$ maps to the generic point of $\operatorname{Spec} O$.

As the fibre $\mathcal{X}_0$ is proper, and both fibres are geometrically connected, Remark 3.3.28 of [Liu02] with surjectivity implies that $\mathcal{X} \to \operatorname{Spec} O$ is proper.

Then we can apply the second part of Remark 9.3.5 in [Liu02] to obtain that $\mathcal{X} \to \operatorname{Spec} O$ is projective. □

In order to properly define the reduction map, we need our field K to be Henselian, i.e. complete with respect to the valuation ν. See also Section 10.1.3 in [Liu02] for further details.

[2] Instead of $\operatorname{Spec} O$, we could also work with any Dedekind scheme, for example $\operatorname{Spec} \mathbb{Z}$ if $D \in \mathbb{Z}[X]$. We stick to the local case for simplicity and consistency of notation.

Definition 8.24. Let $\hat{K}$ the completion of K, and $\hat{O} = \{x \in \hat{K} \mid \nu(x) \geq 0\}$. This remains a discrete valuation ring with residue field still k. We now consider $\mathcal{X}$ as a scheme over $\hat{O}$.

For a closed point $P \in \mathcal{X}_0 = \mathcal{C}$, the Zariski closure $\overline{\{P\}}$ in $\mathcal{X}$ is irreducible and has a unique closed point, the point of $\overline{\{P\}} \cap \mathcal{X}_{\mathfrak{m}}$. This defines a reduction map $\rho : \mathcal{C}(\hat{K}) \to \mathcal{C}(k)$ which extends linearly to Weil divisors.

Remark 8.25. For a point $P = (x, y) \in \mathcal{C}_{\text{aff}}$ we easily see that $\nu(x) \geq 0$ implies $\nu(y) \geq 0$, so in this case we set $\overline{P} = (\overline{x}, \overline{y})$. Otherwise $\nu(x) < 0$, but then write $P = (u, v) \in \mathcal{C}_\infty$ where now $\nu(u) \geq 0$, so we may set $\overline{P} = (\overline{u}, \overline{v})$. This also covers $O_\pm$, note that $\overline{O_+} = O_+$ and $\overline{O_-} = O_-$.

Remark 8.26. Actually, for rational points $P \in \mathcal{C}(K)$ we do not have to worry about K being Henselian because the minimal polynomial of x remains irreducible after reduction.

Remark 8.27. The reduction map extends to the algebraic closure $\mathbb{K}$ of $\hat{K}$. Namely, for x algebraic over $\hat{K}$, we define the valuation

$$\nu(x) = [\hat{K}(x) : \hat{K}]^{-1} \, \nu(\mathrm{Nm}_{\hat{K}(x)/\hat{K}}(x)).$$

In fact the integral closure $\mathbb{O}$ of $\hat{O}$ in $\mathbb{K}$ is still a valuation ring (but no longer discrete).

We also get a reduction map for the Jacobian: it is defined as a quotient of divisors of degree 0 modulo principal divisors, and the latter are preserved by the reduction map:

Proposition 8.28. *Let* $\mathbf{D}$ *a principal divisor over* $\mathcal{C}(\mathbb{K})$*. Then also the divisor* $\overline{\mathbf{D}}$ *over* $\mathcal{C}_{\text{red}}(\overline{k})$ *is principal.*

Proof. Let

$$\mathbf{D} = \operatorname{div} f = \sum_{P \in \mathcal{C}(\overline{\mathbb{K}})} n_P \, (P) \,,$$

a principal divisor over $\mathcal{C}(\mathbb{K})$ with $f \in \mathbb{K}(X, Y)$. As only finitely many $n_P \neq 0$, we may assume all $P \in \mathcal{C}(\hat{K})$, and $f \in \hat{K}(X, Y)$ by passing to a finite extension of $\hat{K}$.

Of course we may write $f = g/h$ with $g, h \in \hat{O}[X, Y]$, and multiply with a suitable power of the uniformiser π such that $\nu(g) = \nu(h) = 0$ (recall that $\hat{O}[X, Y] \subset \hat{O}((X^{-1})))$. Thus f is also a rational function on $\mathcal{X}$ which does not vanish nor has a pole on all of $\mathcal{X}_{\mathfrak{m}}$ (so there is no vertical component), because both $\overline{g} \neq 0$ and $\overline{h} \neq 0$.

Moreover, note that for every $P \in \mathcal{C}(\hat{K})$ we have that $\overline{\{P\}}$ is a zero or pole (with multiplicity n_P) of f on $\mathcal{X}$. This follows from zeroes and poles being Zariski-closed, or the stalks being isomorphic $\mathcal{O}_{\mathcal{X}_0, P} \simeq \mathcal{O}_{\mathcal{X}, \overline{\{P\}}}$ (see the proof of Lemma 8.3.3 and Definition 7.1.27 of multiplicities in [Liu02]). So

$$\mathbf{D}_{\mathcal{X}} = \operatorname{div} f_{\mathcal{X}} = \sum_{P \in \mathcal{C}(\overline{\mathbb{K}})} n_P \left(\overline{\{P\}} \right),$$

which we intersect with $\mathcal{X}_{\mathfrak{m}}$ to get the divisor

$$\mathbf{D}_{\mathcal{X}_{\mathfrak{m}}} = \sum_{P \in \mathcal{C}(\overline{\mathbb{K}})} n_P \, \left(\overline{P}\right).$$

We wish to show that this is the divisor of the function $f_{\mathfrak{m}} = \overline{g}/\overline{h} \in k(X,Y)$. Let $P \in \mathcal{X}$ a closed point (i.e. a closed point in $\mathcal{X}_{\mathfrak{m}}$), and consider the intersection number $i_P(\cdot,\cdot)$. Without loss of generality, we may assume that $n_P \geq 0$ (otherwise pass to $-\mathbf{D}$ and $1/f$). Now by Corollary 9.1.32 in [Liu02], we have $i_P(\left(\overline{\{P\}}\right), \mathcal{X}_{\mathfrak{m}}) = 1$. This implies (using Definitions 7.1.27 and 9.1.1 in [Liu02]) that

$$n_P = i_P(\operatorname{div} f_{\mathcal{X}}, \mathcal{X}_{\mathfrak{m}}) = \operatorname{length} \mathcal{O}_{\mathcal{X},P}/((f_{\mathcal{X}}) + \mathfrak{m}\,\mathcal{O}_{\mathcal{X},P}) = \operatorname{length} \mathcal{O}_{\mathcal{X}_{\mathfrak{m}},P}/(f_{\mathfrak{m}}) = \operatorname{ord}_P\,(f_{\mathfrak{m}})$$

and hence $\mathbf{D}_{\mathcal{X}_{\mathfrak{m}}} = \operatorname{div} f_{\mathfrak{m}}$ is principal as desired. □

8.4.2. Reduction of torsion points and periodicity test

While it is not so clear how the period length changes when reducing a periodic continued fraction, it is quite well understood how the torsion order of $[\mathbf{O}]$ can change:

Theorem 8.3 (Serre-Tate)**.** *Suppose D and $\overline{D}$ are square-free. Let $P \in \mathcal{J}$ torsion of order n, and suppose that $\rho\,(P) \in \mathcal{J}_{\text{red}}$ has order m. If $\operatorname{char} k = 0$, then $n = m$, otherwise there exists $e \in \mathbb{N}$ such that $n = \mathfrak{p}^e\, m$ with $\mathfrak{p} = \operatorname{char} k$.*

Proof. As $0 = \rho\,(n\,P) = n\,\rho\,(P)$, we see that $\rho : \mathcal{J} \to \mathcal{J}_{\text{red}}$ restricts to a homomorphism of groups $\rho : \mathcal{J}[n] \to \mathcal{J}_{\text{red}}[n]$. But the conditions we pose on D and ν ensure that $\mathcal{J}$ has good reduction at ν. So by Theorem 1 and Lemma 2 of [ST68], for $\operatorname{char} k \nmid m$, this map $\mathcal{J}[n] \simeq \mathcal{J}_{\text{red}}[n]$ is actually an isomorphism of groups; this is always the case in zero characteristic.

For positive characteristic $\operatorname{char} k = \mathfrak{p}$, we may write $n = \mathfrak{p}^e\, n'$ with $\mathfrak{p} \nmid n'$, and assume that $\mathfrak{p} \nmid m$: Because $m \,|\, n$, we can remove any common power of $\mathfrak{p}$ and go to a multiple of P.

Now $\mathfrak{p}^e\, P$ has order precisely n' not divisible by $\mathfrak{p}$, so $\rho\,(\mathfrak{p}^e\, P)$ has the same order n'. However, $\mathfrak{p}^e$ is coprime with the order of $\rho\,(P)$, implying that $\mathfrak{p}^e\,\rho\,(P) = \rho\,(\mathfrak{p}^e\, P)$ has likewise order m. So $n' = m$, and we are done. □

The above theorem enables an old trick to effectively test if a point is torsion, mentioned already in [Dav81], and described in [Yu99] for hyperelliptic continued fractions.

Remark 8.29 (Reduction modulo two primes). Given a square-free $D \in K[X]$ with K some number field, $\deg D$ even and $\ell c(D)$ a square as usual, we can always find two prime ideals $\mathfrak{P}_1$ and $\mathfrak{P}_2$ such that $D \in O_{\mathfrak{P}_i}[X]$, $\nu_{\mathfrak{P}_i}(\ell c(D)) = 0$ and D is square-free modulo $\mathfrak{P}_i$, for $i = 1, 2$. Of course the residue fields are finite, and we may assume they are of different characteristics $\mathfrak{p}_1$ and $\mathfrak{p}_2$. Then $[\mathbf{O}_{\mathfrak{P}_i}]$ is torsion, of order m_i. Assuming that also $[\mathbf{O}]$ is torsion, of order m, we can write

$$m = {\mathfrak{p}_1}^{e_1}\, m_1 = {\mathfrak{p}_2}^{e_2}\, m_2, \quad e_i \geq 0.$$

This implies $e_1 \leq e_1' = \nu_{\mathfrak{p}_1}(m_2)$ and $e_2 \leq e_2' = \nu_{\mathfrak{p}_2}(m_1)$, and moreover

$$m \,|\, \gcd({\mathfrak{p}_1}^{e_1'}\, m_1, {\mathfrak{p}_2}^{e_2'}\, m_2) \,|\, \operatorname{lcm}(m_1, m_2).$$

This already gives a bound for the torsion order m which translates into a bound for the period length via Proposition 6.11. So we can test for periodicity effectively.

Indeed as $m_1, m_2 \mid m$, often it is even possible to immediately find a contradiction if m_1 and m_2 have too many different prime factors.

Also if K is finitely generated over a number field, we can specialize D to be defined over a number field. If $[\mathbf{O}]$ is already torsion, this should not alter the torsion order, so we can lift the torsion bound as obtained above, and still determine effectively if $\mathbf{CF}(\sqrt{D})$ is periodic.

However, we need to be careful to avoid bad reduction of the continued fraction: it might happen that specializing a non-Pellian D, we end up with a Pellian D. Usually, it should however not be a problem to find a specialization where this does not happen. See for example Proposition 8.35 below.

8.4.3. Shortening of quasi-period

Theorem 8.3 also gives a little bit of information on how the quasi-period may change in the case of bad reduction of the continued fraction.

Suppose that $\mathbf{CF}(\sqrt{D})$ has quasi-period length ℓ. Set $d_i = \deg a_i < d = \deg D$, $i = 1, \ldots, \ell - 1$. The torsion order of $[\mathbf{O}]$ is

$$m = \deg p_{\ell-1} = d + d_1 + \cdots + d_{\ell-1}. \tag{8.3}$$

Because the quasi-period is palindromic (see Proposition 5.41) we have $d_i = d_{\ell-i}$.

Assuming that $\overline{D}$ is not a square, with $\mathbf{CF}(\sqrt{\overline{D}})$ having quasi-period length ℓ', we set $d_i' = \deg c_i < d$, $i = 1, \ldots, \ell' - 1$, with $d_i' = d_{\ell-i}'$. The torsion order of $[\mathbf{O}_{\mathrm{red}}]$ is then

$$m' = \deg u_{\ell'-1} = d + d_1' + \cdots + d_{\ell'-1}', \tag{8.4}$$

where $m = \mathfrak{p}^e m'$ for some non-negative integer e if $\operatorname{char} k = \mathfrak{p}$, and $m = m'$ if $\operatorname{char} k = 0$.

If $m' \neq m$, then (8.4) has to be repeated $\mathfrak{p}^e$ times to make up (8.3). Recall from Corollary 7.31 that each $d_i' = d_{i_1} + \cdots + d_{i_j}$.

For example

$$\begin{aligned} d_1' &= d_1 + \cdots + d_{j_1}, & d_{\ell-1}' &= d_{\ell-1} + \cdots + d_{\ell-j_1} \\ d_2' &= d_{j_1+1} + \cdots + d_{j_2}, & & \\ d_3' &= d_{j_2+1} + \cdots + d_{j_3}, & & \end{aligned}$$

and so on. Of course $m' = m$ does not prevent bad reduction of $\mathbf{CF}(\sqrt{D})$, as the d_i' might just be larger than the d_i – a better criterion is to check if $\ell' = \ell$.

Unfortunately, we do not get a lot more information about these degrees in general. But in some special cases, we can at a glance exclude the possibility of bad reduction of the continued fraction:

- If the sequence of $\deg a_n$ starts with $2, 1, 1, 1, 2, 1, 1, 1, 2, 1, \ldots$, then bad reduction of $\mathbf{CF}(\sqrt{D})$ is impossible because $\deg c_n$ cannot follow the sequences $2, 2, 1, 2, 2, 1, \ldots$ or $2, 1, 2, 2, 1, 2, \ldots$. Then some complete quotients γ_n would have quasi-period length 1, but others would have quasi-period length 2, which is impossible.

- Similarly, if the $\deg a_n$ start with $3, 1, 2, 1, 3, 1, 2, 1, 3, 1, \ldots$, then bad reduction of $\mathbf{CF}(\sqrt{D})$ would make $\deg c_n$ start with $3, 3, 1, 3, 3, 1, \ldots$ or $3, 1, 3, 3, 1, 3, \ldots$. As above, this is not possible.

8.4.4. Reduction of convergent divisors

We now attempt to give a geometric description for the reduction of a hyperelliptic continued fraction, in terms of the divisors associated to convergents.

Recall from Section 6.6 that we can write the divisors of the canonical convergents of $\alpha \in K(X, Y)$ as

$$\begin{aligned}&\operatorname{div}(p_n - \alpha\, q_n) = \operatorname{div}(\vartheta_n) = \\ &\qquad -(\deg p_n)\,(O_-) - (Q_1) - \cdots - (Q_h) + (\deg q_{n+1})\,(O_+) + (P_1^n) + \cdots + \left(P_{e_n}^n\right)\end{aligned}$$

where $e_n = \deg a_0 - \deg a_{n+1} + h$. If $\alpha = Y\ (= \sqrt{D})$, then this divisor satisfies $P_i^n \neq \sigma(P_j^n)$ if $i \neq j$ because p_n and q_n are coprime.

What happens when we reduce this divisor, and pass to

$$\begin{aligned}&\operatorname{div}(\widehat{p_n} - \gamma\, \widehat{q_n}) = \operatorname{div}(\overline{\vartheta_n}) = \\ &\qquad -(\deg p_n)\left(\overline{O_-}\right) - \left(\overline{Q_1}\right) - \cdots - \left(\overline{Q_h}\right) + (\deg q_{n+1})\left(\overline{O_+}\right) + \left(\overline{P_1^n}\right) + \cdots + \left(\overline{P_{e_n}^n}\right)\end{aligned}$$

as in the proof of Proposition 8.28?

1. Of course $\overline{O_\pm} = O_\pm$.
2. The $\overline{Q_i}$ are always the same, and we can control them from the start.
3. It is possible that $\overline{P_i^n} = O_+$ which means $\operatorname{ord}\left(\overline{\vartheta_n}\right) > \operatorname{ord}(\vartheta_n)$.
4. Or $\overline{P_i^n} = O_-$ which means $\deg\left(\widehat{q_n}\right) < \deg q_n$.
5. Of if $\alpha = Y$, then possibly $\overline{P_i^n} = \sigma(\overline{P_j^n})$ for some $i \neq j$. This corresponds to $\widehat{p_n}$ and $\widehat{q_n}$ sharing a common factor.
6. Otherwise, $\overline{P_i^n}$ is just a finite point.

In the case $\alpha = Y$ with $g = 1$, we have $e_n \leq 1$, so there is at most P_1^n and we do not have to worry about case 5. We also know that P_1^n must be K-rational. But for higher genus, we may need to work over an algebraic extension of K (not necessarily the algebraic closure because the degree of the equations defining the P_i^n is uniformly bounded in terms of $\deg D$).

Let us have a closer look at the genus 1 case, and study how it is related to the valuation analysis from Theorem 8.2.

Proposition 8.30. *Under the same hypotheses as for Theorem 8.2 and additionally $D, \overline{D}$ square-free, we have for $n \geq 0$*

$$-(n+2)\ [\mathbf{O}] = j(P_n) \qquad \textit{where } P_n = (x_n, y_n) \in \mathcal{C}_{\mathrm{aff}}.$$

Let $[\mathbf{O}_{\mathrm{red}}]$ have torsion order $m = \ell + 1$, then

1. *If $m \,|\, n+2$, then $\overline{P_n} = O_+$ and $\nu(x_n) = -f_{n+1} < 0$.*
2. *If $m \,|\, n+1$, then $\overline{P_n} = O_-$ and $\nu(x_n) = -f_n < 0$.*
3. *Otherwise $\overline{P_n}$ is a finite point and $\nu(x_n) \geq 0$.*

Remark 8.31. Notice that P_n reduces to infinity precisely when n is in a two element fibre (compare Proposition 8.22).

Proof. We have $\alpha = Y$, $g = 1$, and $\mathbf{CF}(\sqrt{D})$ non-periodic. This implies $\deg p_n = n+2$ and

$$\operatorname{div}(\vartheta_n) = -(n+2)\ (O_-) + (n+1)\ (O_+) + (P_n)$$

where $P_n = (x_n, y_n) \in \mathcal{C}_{\mathrm{aff}}(K)$ for all $n \geq 0$ because $\deg a_n = 1$ for $n \geq 1$.

The reduction of this divisor is

$$\operatorname{div}(u_{\lambda(n)} - Y\, v_{\lambda(n)}) = -(n+2)\ (O_-) + (n+1)\ (O_+) + \left(\overline{P_n}\right).$$

With $\mathbf{CF}(\gamma)$ periodic, the point $\mathbf{O}_{\mathrm{red}}$ over k has torsion order $m = \ell + 1$.

1. If $m \,|\, n+2$, this forces $\overline{P_n} = O_+$.
2. If $m \,|\, n+1$, this forces $\overline{P_n} = O_-$.
3. Otherwise $\overline{P_n} \in \mathcal{C}_{\mathrm{aff}}(k)$.

Recall from Proposition 7.34 and (6.11) that

$$\vartheta_n\, \sigma(\vartheta_n) = {g_n}^{-2}\, (-1)^{n+1}\, s_{n+1} = b_n(X - x_n)$$

for some $b_n \in K^\times$. From the normalisation of ϑ_n with $\nu\left(\vartheta_n\right) = 0$ (and analogously $\nu\left(\sigma(\vartheta_n)\right) = 0$), we get $\nu\left(b_n(X - x_n)\right) = 0$.

1. In the case $m \,|\, n+2$, we have

   $$\operatorname{ord}\left(\overline{\vartheta_n}\right) = 1 + \operatorname{ord}(\vartheta_n)\,, \quad \operatorname{ord}\left(\overline{\sigma(\vartheta_n)}\right) = \operatorname{ord}(\sigma(\vartheta_n))$$

   which means $f_{n+1} > 0$. In fact,

   $$f_{n+1} = \nu\left(\ell c(\vartheta_n)\right) + \nu\left(\ell c(\sigma(\vartheta_n))\right) = \nu\left(b_n\right).$$

   This forces $\nu\left(x_n\right) = -\nu\left(b_n\right) = -f_{n+1} < 0$ because $\nu\left(b_n(X - x_n)\right) = 0$, so x_n has negative valuation as expected.

2. In the case $m \,|\, n+1$, we have

$$\operatorname{ord}\big(\overline{\vartheta_n}\big) = \operatorname{ord}(\vartheta_n)\,, \quad \operatorname{ord}\Big(\overline{\sigma(\vartheta_n)}\Big) = 1 + \operatorname{ord}(\sigma(\vartheta_n))\,.$$

   so $f_{n+1} = 0$. But $\nu\left(\ell c(\sigma(\vartheta_n))\right) = \nu(\ell c(\widetilde{q_n})) = f_n > 0$, and similarly as above $\nu\left(x_n\right) = -\nu\left(\ell c(\sigma(\vartheta_n))\right) = -f_n < 0$ we get that x_n has negative valuation as expected.

3. Otherwise, we have

$$\operatorname{ord}\big(\overline{\vartheta_n}\big) = \operatorname{ord}(\vartheta_n)\,, \quad \operatorname{ord}\Big(\overline{\sigma(\vartheta_n)}\Big) = \operatorname{ord}(\sigma(\vartheta_n))$$

   and hence $f_{n+1} = 0$. This implies $\nu\left(b_n\right) = 0$ and thus $\nu\left(x_n\right) \geq 0$, i.e. $x_n \in O$.

□

Observe how this matches Theorem 8.2 and that $-f_{n+1}$ is the valuation of both x_n and x_{n+1} at the two element fibre.

So we have found a second description of the f_n from Theorem 8.2. Unfortunately, this still does not suggest what type of patterns they might follow, or whether they might be bounded. Generally, we should not expect the f_n to be bounded, see for example the proposition on page 55 of [ST15]. It suggests that for an elliptic curve defined over $\mathbb{Q}$, there are rational points $P = (x, y)$ with arbitrarily low $\nu_{\mathfrak{p}}(x)$ for any prime $\mathfrak{p}$.

8.5. Good reduction at infinitely many primes

We mentioned before that Theorem 8.1 holds only for $\mathbf{CF}(\sqrt{D})$, but not for other elements of $K(X, \sqrt{D})$.

Theorem 8.4. *Let $D = X^4 + 16\,X^2 + 24\,X + 9$ which is not Pellian (the torsion orders of $\mathbf{O}_3$ and $\mathbf{O}_{17}$ differ just by 1). Set $\alpha = \frac{\sqrt{D}-3}{X}$.*

There are infinitely many primes $\mathfrak{p}$ for which $\mathbf{CF}(\alpha)$ has good reduction (hence $\mathfrak{p}$ never divides a denominator of a partial quotient a_n).

This is related to questions treated in [CRS97], and earlier in [Sch60]. Here we present an explicit proof for our particular example, to illustrate these arguments more concretely. In fact, the given problem boils down to an analogue for elliptic curves (more generally abelian varieties) of

Proposition 8.32. *There exist infinitely many prime numbers $\mathfrak{p}$ such that for all $n \in \mathbb{Z}$ we have $2^n \not\equiv 5$.*[3]

[3] The reader might find it enjoyable to try and prove this exercise for himself.

Proof of Theorem 8.4. Note that D non Pellian implies that $\mathbf{CF}(\alpha)$ is not quasi-periodic by Theorem A in [Ber90].

Recall from Proposition 6.16 that the divisors induced by the convergents have the shape

$$\operatorname{div}(p_n - \alpha\, q_n) = (\deg p_n)\, \mathbf{O} - (Q) + (P_n)$$

because α has a single pole at $Q = (0, ?)$. Note that $X \mid D - 3^2$, so we are in the situation of Theorem 6.2. In principle, we could have $P_n = O_+$ for a single n, but the reduction arguments below imply that P_n must always be a finite point.

Recall from Section 8.4.4 that here bad reduction of $\mathbf{CF}(\alpha)$ is equivalent to P_n reducing to a point at infinity $O_\pm$ which means

$$j(\overline{Q}) + (\deg p_n)\, j(O_-) = j(O_\pm) = 0 \text{ or } j(O_-) \qquad (\operatorname{mod} \mathfrak{p}).$$

So if we ensure that for all $m \in \mathbb{Z}$

$$j(\overline{Q}) + m\, j(O_-) \neq 0 \qquad (\operatorname{mod} \mathfrak{p}) \tag{8.5}$$

then we know that we must have good reduction of $\mathbf{CF}(\alpha)$ at this prime $\mathfrak{p}$ (and additionally we confirm that $P_n \neq O_\pm$ for all n).

We deduce from the Čebotarev density theorem (see Theorem 13.4 and Lemma 13.5 in [Neu99]) that this holds for infinitely many primes $\mathfrak{p}$. For reasons of space, we will assume the reader is already familiar with this famous theorem, and also ramification of prime ideals, Galois theory and the Frobenius automorphism.

Our curve $\mathcal{C}$ is an elliptic curve, isomorphic to its Jacobian. We write it in Weierstrass form

$$\mathcal{E} : V^2 = U^3 + 16\, U^2 - 36\, U = U(U-2)(U+18)$$

using the transformation

$$U = 2X^2 + 2Y, \qquad V = 2X(U+16) + 24$$

which sends $j(O_-)$ to $R_1 = (-16, -24)$ and $j(Q)$ to $R_2 = (6, 24)$ (over $\mathbb{Q}$). These are non-torsion rational points. Note that the 2 torsion points of $\mathcal{E}$ are by design rational. This implies that for any of the four choices for a point $R_i' \in \mathcal{C}$ with $2\, R_i' = R_i$, the point R_i' is defined over the same number field K_i (for $i = 1, 2$). Here the fields are

$$K_1 = \mathbb{Q}(\zeta), \text{ where } \zeta^4 + 1 = 0, \qquad K_2 = \mathbb{Q}(\sqrt{6}).$$

The composite field K_1K_2 has degree 8 and is Galois, with abelian Galois group $G = \operatorname{Gal}(K_1K_2/\mathbb{Q})$. We denote by H_i the subgroup of G whose fixed field is K_i.

Ignoring the finitely many primes where the curve $\mathcal{E}$ has bad reduction (just $2, 3, 5$, the discriminant of $\mathcal{E}$ is $2^{12} \cdot 3^4 \cdot 5^2$), we now wish to find infinitely many primes $\mathfrak{p}$ such that $\overline{R_1'}$ is a $\mathbb{F}_\mathfrak{p}$-rational point, but $\overline{R_2'}$ is not. In that case, the point $m\, \overline{R_1'} + \overline{R_2'}$ cannot be $\mathbb{F}_\mathfrak{p}$-rational for every $m \in \mathbb{Z}$. In particular it is not 0 or torsion of order 2. This implies that for all $m \in \mathbb{Z}$

$$m\, R_1 + R_2 \neq 0 \quad (\operatorname{mod} \mathfrak{p})$$

which is equivalent to (8.5).

If we restrict to the (infinitely many) primes $\mathfrak{p}$ which are unramified over K_1K_2 (and hence over K_1 and K_2), the condition on rationality of $\overline{R_1'}$ and $\overline{R_2'}$ amounts to saying that $\mathfrak{p}$ has a prime divisor $\mathfrak{P}_1$ of degree 1 over K_1 (i.e. the residue field $k(\mathfrak{P}_1)$ has degree 1 over $\mathbb{F}_\mathfrak{p}$), but over K_2 all the prime divisors of $\mathfrak{p}$ have degree > 1 (i.e. $[k(\mathfrak{P}_2) : \mathbb{F}_\mathfrak{p}] > 1$ for any prime divisor $\mathfrak{P}_2$).

As described in Lemma 13.5 of [Neu99], this happens if and only if the conjugacy class of the Frobenius automorphisms of prime ideals of K_1K_2 lying over $\mathfrak{p}$ intersects H_1, but not H_2. Here $H_1 = \{\mathrm{id}, \sigma_1\}$ where σ_1 is defined by $\sigma_1(\zeta) = \zeta$ and $\sigma_1(\sqrt{6}) = -\sqrt{6}$. The conjugacy classes are trivial because G is abelian, so we are looking precisely for the primes $\mathfrak{p}$ for which σ_1 is a Frobenius automorphism of some prime $\mathfrak{P}$ of K_1K_2 over $\mathfrak{p}$.

But the set of these primes has positive density $\geq \frac{1}{8}$ by the Čebotarev density theorem (Theorem 13.4 in [Neu99]), so in particular there are infinitely many of them. □

Remark 8.33. In our computations, we observed that among the first 100 odd prime numbers, there are 62 prime numbers $\mathfrak{p}$ for which $\mathbf{CF}(\alpha)$ has good reduction at $\mathfrak{p}$. This is a much higher density than predicted by our Čebotarev density estimate, but of course the latter gives only a sufficient condition.

Moreover, we argued with R_i' satisfying $2\,R_i' = R_i$. Instead we could argue with $m\,R_i' = R_i$ for any m; then we probably get additional primes with the desired property.

8.6. Complex functions case

Let us now discuss the situation of specialization, i.e. where $K = \mathbb{C}(t)$ and the reduction map corresponds to assigning a special value $t_0 \in \mathbb{C}$ to t. This corresponds to the valuation $\nu = \mathrm{ord}_{t_0}$ measuring the zero-order at t_0, with uniformising parameter $t - t_0$, and the discrete valuation ring $O = \mathbb{C}[t]_{(t-t_0)}$, the localisation of the maximal ideal $(t - t_0)$ in $\mathbb{C}[t]$. We also write $\overline{\alpha} = \alpha_{t=t_0}$ to distinguish different specializations.

8.6.1. Results of Masser and Zannier

This is precisely the situation found in the article of Masser and Zannier on the connection between the Pell equation and Unlikely intersections [MZ15]. Let me restate their results in our language of specialization of continued fractions.

For genus 1, they mention the following result:

Proposition 8.34 (Masser-Zannier)**.** *Let $D = X^4+X+t$, then $\mathbf{CF}(\sqrt{D})$ is non-periodic. The set of t_0 such that $\mathbf{CF}(\sqrt{D_{t=t_0}})$ is periodic (i.e. $\mathbf{CF}(\sqrt{D})$ has bad reduction at $t-t_0$ by Proposition 8.8), is infinite and denumerable.*

For genus 2 however, their Theorem P1 says:

Theorem 8.5 (Masser-Zannier)**.** *Let $D = X^6 + X + t$, then $\mathbf{CF}(\sqrt{D})$ is non-periodic. The set of t_0 such that $\mathbf{CF}(\sqrt{D_{t=t_0}})$ is periodic is* finite.

$D = X^6 + X + t$	basefield $\mathbb{C}(t)$
Discriminant of D: $(-46656) \cdot (t^5 - \frac{3125}{46656})$	D never reduces to a square.
D is not Pellian	
Partial quotients	
$a_0 = X^3$	
$a_1 = 2X^2 - 2tX + 2t^2$	
$a_2 = \frac{-\frac{1}{2}X - \frac{1}{2}t}{t^3}$	

For example $\mathbf{CF}(\sqrt{D_{t=0}})$ is periodic. But because $\deg D = 6$, it is now possible that $\mathbf{CF}(\alpha)$ has bad reduction at $t - t_0$, even if $\mathbf{CF}(\sqrt{D_{t=t_0}})$ is non-periodic.

To describe the t_0 in the Theorem, we need to search for an increase in the degree of the partial quotients when specialising, as seen in Lemma 7.20. Clearly $\deg c_0 = 3, \deg c_1 = 2$ for the partial quotients of any specialization. However Theorem P2 says:

Theorem 8.6 (Masser-Zannier)**.** *Let $D = X^6 + X + t$, with $\mathbf{CF}(\sqrt{D})$ non-periodic. The set of t_0 such that for $\gamma = \sqrt{D_{t=t_0}}$ there exists $n \geq 2$ with $\deg c_n = 2$, is an infinite and denumerable subset of $\overline{\mathbb{Q}}$.*

This set also includes the t_0 with $\mathbf{CF}(\sqrt{D_{t=t_0}})$ periodic: because the period always begins at c_1, there are then infinitely many n with $\deg c_n = 2$ for each of these t_0. By Theorem 7.1 the increase of degrees is necessary for bad reduction of $\mathbf{CF}(\sqrt{D})$, so this infinite set is actually the set of all t_0 with bad reduction of $\mathbf{CF}(\alpha)$ at $t - t_0$.

The hard part in the proof of this theorem is showing that this set of t_0 is *infinite* which is done in Section 11 of [MZ15].

With the theory of Chapter 7, it is however not so hard to show:

Proposition 8.35. *Let $\alpha \in \mathbb{C}(t)((X^{-1}))$ with $\ell c(\alpha) = 1$. Then there exist at most countably many $t_0 \in \mathbb{C}$ such that $\mathbf{CF}(\alpha)$ has bad reduction at $t - t_0$ (the valuation being $\nu = \operatorname{ord}_{t=t_0}$).*

Proof. For every $n \geq 0$, let d_n the denominator of $\ell c(\alpha_n) \in \mathbb{C}(t)$ which is a monic polynomial in $\mathbb{C}[t]$. Clearly $\operatorname{ord}_{t=t_0}(\ell c(\alpha_n)) < 0$ holds if and only if $d_n(t_0) = 0$.

Then if $\mathbf{CF}(\alpha)$ has bad reduction at $t - t_0$, there exists by Proposition 7.20 at least one n such that $d_n(t_0)$.

Of course there are only countably many polynomials d_n, each with finitely many zeroes (even though $\deg d_n$ might increase with n). Hence there are only countably many possibilities for bad reduction of $\mathbf{CF}(\alpha)$. □

Masser and Zannier also give another example in degree 6, with different behaviour (see also Section 3.4.5 in [Zan12]):

Proposition 8.36 (Masser-Zannier)**.** *Let $D = X^6 + X^2 + t$, then $\mathbf{CF}(\sqrt{D})$ is non-periodic. The set of t_0 such that $\mathbf{CF}(\sqrt{D_{t=t_0}})$ is periodic, is infinite and denumerable.*

8.6.2. Repeated occurrences of $t - t_0$

Also for the specialisation case we can say something about the occurrence of "primes" in infinitely many denominators of partial quotients a_n. In degree 4, we can simply use Corollary 8.20 to deduce this from Proposition 8.34 above:

Corollary 8.37. *Let $D = X^4 + X + t$ as in Proposition 8.34. For each of the infinitely many t_0 where $\mathbf{CF}(\sqrt{D})$ has bad reduction, there exist infinitely many n such that $t - t_0$ appears in a denominator of a_n.*

In degree 6, the situation becomes more subtle, and we need to use Proposition 8.11. If $D = X^6 + X + t$, then for $t_0 = 0$, there is the problem that the only fibre of $\lambda : \mathbb{N}_0 \to \mathbb{N}_0$ (the map describing the reduction of convergents, introduced in Section 7.2.3) with a single element is $\lambda^{-1}(0) = \{0\}$, so we cannot apply the proposition. This happens because for this specialization we have $\deg c_n \neq 1$ for all n.

$D = X^6 + X$	basefield $\mathbb{Q}$
Discriminant of D: 5^5	D never reduces to a square.
period length 2 for $\mathbf{CF}(\sqrt{D})$	
Minimal Pell solution	
$p_1 = 2X^5 + 1$	$q_1 = 2X^2$
Partial quotients of $\sqrt{D}$	
$a_0 = X^3$	
$a_1 = 2X^2$	
$a_2 = 2X^3$	

For all other t_0, we can use Proposition 8.13 because clearly $\deg c_2 = 1$ for any specialization to $t_0 \neq 0$, and the minimal degree δ is of course 1. This implies infinitely many fibres with a single element, and from Lemma 8.6 and Proposition 8.11 then follows:

Corollary 8.38. *Let $D = X^6 + X + t$ as in Theorem 8.5. For each of the finitely many $t_0 \neq 0$ where $\mathbf{CF}(\sqrt{D_{t=t_0}})$ is periodic, there exist infinitely many n such that $t - t_0$ appears in a denominator of a_n.*

It is likely this property also holds for $t_0 = 0$, but this would require a different argument, for example an analogue to Theorem 8.2 for degree 6 which is unfortunately not in sight.

For $D = X^6 + X^2 + t$, we have however $a_1 = 2\,X$, so Proposition 8.13 implies infinitely many fibres of λ with a single element for every t_0, so again by Lemma 8.6 and Proposition 8.11 follows:

Corollary 8.39. *Let $D = X^6 + X^2 + t$ as in Proposition 8.36. For each of the finitely many $t_0 \neq 0$ where $\mathbf{CF}(\sqrt{D_{t=t_0}})$ is periodic, there exist infinitely many n such that $t - t_0$ appears in a denominator of a_n.*

Now let us return to $D = X^6 + X + t$ and consider the remaining t_0 with non-periodic bad reduction of the continued fraction. To understand this better, we need Theorem 1.3 from [Zan16] (which we state for arbitrary base field $\mathbb{K}$):

Theorem 8.7 (Zannier). *Let $D \in \mathbb{K}[X]$ of degree $2d$, non-square, but $\ell c(D)$ square. Suppose further that if $D = E^2 D'$ with D' Pellian, then $\deg D \leq \frac{3}{2}d$ (for example assume D is square-free). Then there are only finitely many n with $\deg a_n > \frac{d}{2}$.*

This is a consequence of a Skolem-Mahler-Lech theorem for algebraic groups (for example the Jacobian of $\mathcal{C}$) explained in the same article.

For $\deg D = 6$, in particular for $D = X^6 + X + t$ and its specializations, this means that if $\mathbf{CF}(\sqrt{D})$ is not periodic, only finitely many partial quotients have degree $\deg a_n > 1$; and the same property holds for $\mathbf{CF}(\sqrt{D_{t=t_0}})$. This means that for the t_0 with non-periodic $\mathbf{CF}(\sqrt{D_{t=t_0}})$, the corresponding map λ has only finitely many fibres with more than one element. Again, we cannot apply Proposition 8.11. This does not mean there might not be infinitely many a_n with $t - t_0$ in the denominator, but for every n large enough, we can normalise the complete quotient α_n to

$$\begin{aligned}\widetilde{\alpha_n} &= (t - t_0)^{-\operatorname{ord}_{t_0}(\alpha_n)}\, \alpha_n = \mu\, \alpha_n \\ &= [\mu\, a_n, \mu^{-1}\, a_{n+1}, \mu\, a_{n+2}, \mu^{-1}\, a_{n+3}, \dots] = [b_0, b_1, b_2, b_3, b_4, \dots]\end{aligned}$$

such that none of the b_i has $t - t_0$ in a denominator. So $\mathbf{CF}(\widetilde{\alpha_n})$ has good reduction at $t - t_0$.

In fact, we cannot even exclude the possibility that $\mu = 1$, in which case only finitely many a_n have $t - t_0$ in the denominator, despite there being bad reduction of $\mathbf{CF}(\alpha)$ at $t - t_0$ (compare Remark 8.12).

9. Heights

While the valuations used in the previous chapters give a local estimate for the complexity of the partial quotients, affine and projective heights provide a global measure of complexity. For the convergents, and more generally Padé approximations, the projective logarithmic height of the convergents has known lower bounds (in the non-Pellian case), see [BC97]: they should increase at least quadratically. A lower bound for the height of the partial quotients is more delicate, and has been found only recently: [Zan16] gives a lower bound for affine logarithmic height, with at least quadratic growth for a frame of fixed length of partial quotients.

Upper bounds for the projective heights of the partial quotients follow from those of the convergents.

9.1. Heights

For the convenience of the reader, I will list some definitions and properties of heights to be used in this chapter, following mainly [BG06] in notation and normalisation. For the Weil height machine, I follow [HS00].

9.1.1. Places and Product formula

For a number field K, one defines the set of *places* M_K as the equivalence classes of non-trivial absolute values on K, where two absolute values are equivalent if they induce the same topology. A place contains either only archimedean absolute values, and then is called *infinite*, or only non-archimedean absolute values, in which case we call it *finite*. The infinite places correspond to embeddings of K into $\mathbb{C}$ up to complex conjugation, hence there is only a finite number. The finite places correspond to prime ideals of the ring of integers of K, so there are infinitely many.

In order to define heights, one carefully chooses and fixes an absolute value to represent a place. For $\mathbb{Q}$, there is one infinite place, the restriction of the standard complex absolute value, and countably many finite places corresponding to the prime numbers. We represent these places by

$$\begin{aligned} M_{\mathbb{Q}} &= \{\mathfrak{p} \in \mathbb{N} \text{ prime } \} \cup \{\infty\}, \\ |x|_\infty &= \max(x, -x), \\ |x|_{\mathfrak{p}} &= \mathfrak{p}^{-n} \text{ for } x \neq 0 \text{ where } x = \mathfrak{p}^n \frac{a}{b} \text{ with } n \in \mathbb{Z}, a, b \in \mathbb{Z} \setminus \mathfrak{p}\mathbb{Z} \end{aligned}$$

This ensures $K = \mathbb{Q}$ satisfies the *product formula*

$$\prod_{\nu \in M_{\mathbb{Q}}} |x|_\nu = 1 \text{ for } x \in \mathbb{Q}^\times .$$

The product on the left is well defined because for a fixed x, only for finitely many $\nu \in M_{\mathbb{Q}}$ have $|x|_\nu \neq 1$.

On any number field K, a place ω on K restricts to a unique place ν on $\mathbb{Q}$, written $\omega \mid \nu$, and we can choose a cleverly normalised representative ω such that

$$\prod_{\omega \mid \nu} |x|_\omega = |x|_\nu \text{ for all } x \in \mathbb{Q}^\times .$$

In consequence, on any number field K, there is a *product formula*

$$\prod_{\omega \in M_K} |x|_\omega = 1 \text{ for } x \in K^\times ,$$

where on the left side only finitely ω have $|x|_\omega \neq 1$.

We also remark that ω normalised in this way satisfies an *improved triangle equality*: Let $x_1, \dots, x_r \in K$, then

$$|x_1 + \cdots + x_r|_\omega \leq \max(1, |r|_\omega) \max\left(|x_1|_\omega, \dots, |x_r|_\omega\right) . \tag{9.1}$$

9.1.2. Height on projective and affine space

The product formula allows defining an *exponential absolute projective height* on $\mathbb{P}^n(K)$ for a number field K by setting

$$H_{\text{proj}}(x_0 : \cdots : x_n) = \prod_{\nu \in M_K} \max\left(|x_0|_\nu, \dots, |x_n|_\nu\right)$$

By considering $\mathbb{A}^n(K) \subset \mathbb{P}^n(K)$, this also defines an *affine height* on $\mathbb{A}^n(K)$,

$$H_{\text{aff}}(x_1, \dots, x_n) = H_{\text{proj}}(1 : x_1 : \cdots : x_n)$$

and in particular we get a height H on $K = \mathbb{A}^1(K)$. Note that the affine height is always larger than the projective height, i.e.

$$H_{\text{proj}}(x_1, \dots, x_n) \leq H_{\text{aff}}(x_1, \dots, x_n).$$

It can be shown this definition does not depend on the number field K, and thus extends uniquely to $\overline{\mathbb{Q}}$.

Often it is convenient to work with the *logarithmic heights*, $h_{\text{proj}} = \log \circ H_{\text{proj}}$, $h_{\text{aff}} = \log \circ H_{\text{aff}}$ and $h = \log \circ H$.

9.1.3. Height of polynomials

A non-zero polynomial in $K[X]$ of degree $\leq n$ can be considered both as a point in $\mathbb{P}^n(K)$, or in $\mathbb{A}^{n+1}(K)$. So we define

$$\begin{aligned} H_{\text{proj}}(a_n\,X^n + \cdots + a_0) &= H_{\text{proj}}(a_n : \cdots : a_0) \\ H_{\text{aff}}(a_n\,X^n + \cdots + a_0) &= H_{\text{aff}}(a_n, \ldots, a_0) \end{aligned}$$

and likewise the logarithmic heights. Note that this means the projective height of a polynomial depends only on its zeroes, while the affine height coincides with the height on K for constant polynomials.

Similarly as in Section 7.1.2, we define the *Gauss norm* for ν on $K[X]$ by

$$|a_n\,X^n + \cdots + a_0|_\nu = \max\left(|a_n|_\nu\,, \ldots, |a_0|_\nu\right).$$

If ν is non-archimedean, this is even an non-archimedean absolute value (see Proposition 7.3). Nevertheless, the notation is useful also for archimedean absolute values.

In the following, K is always a number field which has at most $[K : \mathbb{Q}]$ archimedean places.

Proposition 9.1. *Let $f_1, \ldots, f_r \in K[X]$ and $f = f_1 \cdots f_r$. Then*

$$-\deg f\;\log 2 + \sum_{i=1}^{r} h_{\text{proj}}(f_i) \leq h_{\text{proj}}(f) \leq \deg f\;\log 2 + \sum_{i=1}^{r} h_{\text{proj}}(f_i)$$

For a proof see [BG06], Theorem 1.6.13.

Proposition 9.2. *Let $f_1, \ldots, f_r \in K[X]$ and $f = f_1 + \cdots + f_r$. Then*

$$h_{\text{aff}}(f_1 + \cdots + f_r) \leq h_{\text{aff}}(f_1) + \cdots + h_{\text{aff}}(f_r) + \log r.$$

This follows from Proposition 1.5.15 in [BG06].

Proposition 9.3. *Let $a, b, q, r \in K[X] \setminus \{0\}$ with $a = q\,b + r$ and $\deg r < \deg b$. Set $N = \deg q = \deg a - \deg b$. Then*

$$h_{\text{proj}}(q) \leq h_{\text{proj}}(a) + N\left(\log 2 + h_{\text{proj}}(b)\right), \tag{9.2}$$

$$h_{\text{proj}}(r) \leq h_{\text{proj}}(a) + (N+1)\left(\log 2 + h_{\text{proj}}(b)\right). \tag{9.3}$$

Remark 9.4. The bound for $h_{\text{proj}}(q)$ holds also if $r = 0$.

Proof. We can assume $\ell c(b) = 1$, as the projective height for polynomials is invariant under multiplication with a constant factor. This conveniently implies $|b|_\nu = \max(1, |b|_\nu)$ for every $\nu \in M_K$.

Using the standard algorithm for division, we define a sequence of polynomials, beginning with $a_0 = a$ and continuing for $i \geq 0$ via

$$a_i = \ell c(a_i)\,X^{N_i}\,b + a_{i+1} \text{ where } \deg a_{i+1} < \deg a_i \text{ and } N_i = \deg a_i - \deg b.$$

Using (9.1) on the coefficients, we estimate

$$\begin{aligned} |a_{i+1}|_\nu &= |\ell c(a_i)b - a_i|_\nu \\ &\leq \max(1,|2|_\nu)\,\max\left(|\ell c(a_i)|_\nu\,|b|_\nu\,, |a_i|_\nu\right) \leq \max(1,|2|_\nu)\;|a_i|_\nu\;|b|_\nu \end{aligned}$$

and obtain

$$|a_i|_\nu \leq |a|_\nu\,\left(\max(1,|2|_\nu)\,|b|_\nu\right)^i .$$

There are at most $N+1$ steps necessary in the algorithm to reach $\deg a_i < \deg b$ (because in every step, the degree decreases by at least 1, hence $\deg a_i < \deg a - i$, so $i \leq N+1$), at which point $a_i = r$. Consequently

$$|r|_\nu \leq |a|_\nu\,\left(\max(1,|2|_\nu)\;|b|_\nu\right)^i \leq |a|_\nu\,\left(\max(1,|2|_\nu)\;|b|_\nu\right)^N$$

which implies (9.2). The coefficients of q are precisely $\ell c(a_0), \dots, \ell c(a_{i-1})$, so

$$|q|_\nu \leq \max\left(|a_0|_\nu\,, \dots, |a_{i-1}|_\nu\right) \leq |a|_\nu\,\left(\max(1,|2|_\nu)\;|b|_\nu\right)^{i-1} \leq |a|_\nu\,\left(\max(1,|2|_\nu)\;|b|_\nu\right)^N$$

whence (9.3). □

Proposition 9.5. *Let $f \in K[X]$ a polynomial of degree r with roots $\alpha_1, \dots, \alpha_r \in \overline{\mathbb{Q}}$ (accounted for multiplicities). Then*

$$-r\,\log 2 + h_{\mathrm{proj}}(f) \leq h_{\mathrm{aff}}(\alpha_1) + \cdots + h_{\mathrm{aff}}(\alpha_r) \leq r\,\log 2 + h_{\mathrm{proj}}(f).$$

Proof. This follows directly from Proposition 9.1, noting that f factors as

$$f = \mu\,(X-\alpha_1)\cdots(X-\alpha_r)$$

with $\mu \in K$ having $h_{\mathrm{proj}}(\mu) = 0$, while $h_{\mathrm{proj}}(X-\alpha_i) = h_{\mathrm{proj}}(1:\alpha_i) = h_{\mathrm{aff}}(\alpha_i)$. □

9.1.4. Weil's Height Machine and Néron-Tate height

On varieties defined over a number field K, there is a plethora of different height functions. However, many of them differ only by a bounded function, so they produce essentially the same height. We capture this notion of *quasi-equivalence of heights* in the following notation:

Definition 9.6. Let V a variety defined over a number field K, and let $f_1, f_2 : V(\overline{\mathbb{Q}}) \to \mathbb{R}$ two function. We write $f_1 \approx f_2$ if there exists $C \in \mathbb{R}$ such that

$$|f_1(P) - f_2(P)| \leq C \text{ for all } P \in V(\overline{\mathbb{Q}}).$$

We reproduce the following from Theorem B.3.2 in [HS00]

Theorem 9.1 (Weil's Height Machine)**.** *Let K a number field. For every smooth projective variety V/K there exists a map*

$$h_V : \mathrm{Div}(V) \longrightarrow \{\textit{functions } V(\overline{\mathbb{Q}}) \to \mathbb{R}\}$$

satisfying the following:

1. *(Normalisation) For* $\mathbf{H} \subset \mathbb{P}^n$ *a hyperplane holds* $h_{\mathbb{P}^n,\mathbf{H}} \approx h_{\text{proj}}$.
2. *(Functoriality) For* $\phi : V \to W$ *a morphism,* $D \in \operatorname{Div}(W)$ *holds* $h_{V,\phi^*(\mathbf{D})} \approx h_{W,\mathbf{D}} \circ \phi$.
3. *(Additivity) For* $\mathbf{D}, \mathbf{E} \in \operatorname{Div}(V)$ *holds* $h_{V,\mathbf{D}+\mathbf{E}} \approx h_{V,\mathbf{D}} + h_{V,\mathbf{E}}$
4. *(Linear Equivalence) For* $\mathbf{D}, \mathbf{E} \in \operatorname{Div}(V)$ *with* $\mathbf{D} \sim \mathbf{E}$ *holds* $h_{V,\mathbf{D}} \approx h_{V,\mathbf{E}}$.

There are further properties which we omit because we will not use them directly.

Theorem 9.2 (Néron-Tate height)**.** *Let* K *a number field, and* $\mathcal{A}/K$ *an abelian variety. Let* $\mathbf{D} \in \operatorname{Div}(\mathcal{A})$ *have symmetric divisor class (i.e.* $[-1]^*\mathbf{D} \sim \mathbf{D}$*). Then there exists the (unique)* canonical height on $\mathcal{A}$ relative to $\mathbf{D}$, *a height function* $\hat{h}_{\mathcal{A},\mathbf{D}} : \mathcal{A}(\overline{\mathbb{Q}}) \longrightarrow \mathbb{R}$ *satisfying the following:*

1. *It is equivalent to the height from the height machine:* $\hat{h}_{\mathcal{A},\mathbf{D}} \approx h_{\mathcal{A},\mathbf{D}}$.
2. *For all integers* m *and* $P \in \mathcal{A}(\overline{\mathbb{Q}})$ *we have* $\hat{h}_{\mathcal{A},\mathbf{D}}(m\,P) = m^2\,\hat{h}_{\mathcal{A},\mathbf{D}}(P)$.
3. *It is a quadratic form.*

Proposition 9.7. *With* $\mathcal{A}$ *and* $\mathbf{D}$ *as in Theorem 9.2, and* $\mathbf{D}$ *moreover ample, then for all* $P \in \mathcal{A}(\overline{\mathbb{Q}})$ *we have* $\hat{h}_{\mathcal{A},\mathbf{D}}(P) \geq 0$*, and* $\hat{h}_{\mathcal{A},\mathbf{D}}(P) = 0$ *if and only if* P *is torsion on* $\mathcal{A}$.

The preceding Theorem and Proposition are adapted from Theorem B.5.1 and Proposition B.5.3 in [HS00], respectively.

9.1.5. Heights on the Jacobian

With these tools, we are finally able to setup our heights on our (hyper)elliptic curve $\mathcal{C}$ and its Jacobian $\mathcal{J}$. On the Jacobian, we use the height corresponding to the Theta divisor. The Theta divisor induced by the map $j : \mathcal{C} \to \mathcal{J}$ defined in Section 4.2.3 is unfortunately not symmetric for $g > 1$. So we use a different embedding, which differs only by a translation on the Jacobian.

Indeed let $P_0 \in \mathcal{C}$ one of the Weierstrass points, i.e. $P_0 = (\xi, 0)$ where ξ is one of the roots of D. This implies $2\,(P_0) \sim (O_+) + (O_-)$ (via the function $X - \xi$), hence the canonical divisor is a multiple of P_0: $\mathbf{K}_\mathcal{C} \sim 2(g-1)\ (P_0)$.

Now embed the curve into the Jacobian via

$$j_0 : \mathcal{C} \longrightarrow \mathcal{J}, \quad P \mapsto [P] - [P_0]$$

and note that $j_0(\mathbf{K}_\mathcal{C}) = 0$ (recall that j_0 extends naturally to divisors).

Now Theorem A.8.2.1. in [HS00] implies that

$$\Theta_0 = j_0(\mathcal{C}) + \cdots + j_0(\mathcal{C}) \quad (g-1 \text{ copies})$$

is symmetric, i.e. $[-1]^*\Theta_0 = \Theta_0$. Then Theorem 9.2 implies that the Néron-Tate height $\hat{h} = \hat{h}_{\mathcal{J},\Theta_0}$ associated to the height $h_{\mathcal{J},\Theta_0}$ is a quadratic form.

Theorem A.8.2.1. also says that $j_0^* \Theta_0 \sim g\ (P_0)$, so $\hat{h} \circ j_0 \approx g\, h_{\mathcal{C},P_0}$.

Recall that the hyperelliptic curve comes with a degree 2 map $\pi : \mathcal{C} \to \mathbb{P}^1$, with the hyperplane $H = \{(\xi : 1)\}$ in $\mathbb{P}^1$ having $\pi^*(H) = 2\ (P_0)$, hence $h_{\mathrm{proj}}(\pi(P)) \approx 2\, h_{\mathcal{C},P_0}$. So we get $2\,\hat{h} \circ j_0 \approx g\, h_{\mathrm{proj}}(\pi(P))$, or more precisely for $P = (x, y) \in \mathcal{C}_{\mathrm{aff}}$:

$$\hat{h}(j_0(P)) \approx \frac{g}{2}\, h_{\mathrm{aff}}(x)$$

9.2. Height of convergents

We are now ready to study the height of the convergents. We begin by analysing the coefficients of the Laurent series $\sqrt{D}$, and comparing the heights of the numerator and denominator of a convergent.

9.2.1. Height bounds for series coefficients of $\sqrt{D}$

We need some bounds for the absolute values (and height) of the coefficients of the power series $\sqrt{D}$.

Proposition 9.8. *Recall that* $\deg D = 2\,d$ *and write*

$$\sqrt{D} = X^d \sum_{n=0}^{\infty} w_n\, X^{-n}. \tag{9.4}$$

For any place ν on a number field, we have

$$|w_n|_\nu \le \left|\sqrt{\ell c(D)}\right|_\nu \cdot \left(\max(1, |1/4|_\nu) \cdot \max(1, \left|(2d)^2\right|_\nu) \cdot |D|_\nu \,/\, |\ell c(D)|_\nu\right)^n,$$

which implies

$$h(w_n) \le \frac{1}{2} h(\ell c(D)) + n\ (\log 4 + 2\ \log(2d) + h_{\mathrm{proj}}(D))\,.$$

Before we can prove this, we need an estimate for the growth of the binomial coefficient:

Lemma 9.9. *For $n \ge 1$, there exists an integer $b_n \in \mathbb{Z}$ with $|b_n|_{\mathbb{R}} \le 2^{2n-3}$ such that the binomial coefficient $\binom{1/2}{n} = b_n / 2^{2n-1}$.*

For ν a place on a number field, not over 2, this implies $\left|\binom{1/2}{n}\right|_\nu \le 1$ for all $n \ge 0$.

But if ν represents a place over 2 (with the normalisations introduced at the beginning of the chapter), we find $\left|\binom{1/2}{n}\right|_\nu \le 2^{2n}$ for all $n \ge 0$.

This is an easy exercise. Note that the b_n are closely related to the Catalan numbers (see for example [Aig07], pages 101, 102). See also Theorem 5 in [Sie14] for a generalisation of this lemma.

Proof of Proposition 9.8. We write

$$D = d_{2d}\,X^{2d} + d_{2d-1}\,X^{2d-1} + \cdots + d_0 = d_{2d}\,X^{2d}(1 + f(X)) \text{ with } f(X) \in K[X^{-1}].$$

We may then compute

$$\sqrt{D} = \sqrt{d_{2d}}\,X^d \sum_{n=0}^{\infty} \binom{1/2}{n} f(X)^n$$

which converges in $K(\!(X^{-1})\!)$ because $\operatorname{ord}(f(X)) > 0$. Now let ν any place on K, and write $f(X) = f_1 X^{-1} + \cdots + f_{2d} X^{-2d}$, to define

$$C_\nu = \max(1, |f_1|_\nu, \ldots, |f_{2d}|_\nu) = |1 + f|_\nu = |D|_\nu \,/\, |\ell c(D)|_\nu\,.$$

Studying for $i \geq 0$ the power

$$\binom{1/2}{i} f(X)^i = \sum_{j_1+\cdots+j_{2d}=i} \binom{1/2}{i} \binom{i}{j_1, \ldots, j_{2d}} \left(\prod_{l=1}^{2d} f_l{}^{j_l} \right) X^{-(j_1+2\,j_2+\cdots+(2d)\,j_{2d})},$$

we note that $\binom{i}{j_1,\ldots,j_{2d}} \leq 2d^i$ is an integer, so the coefficient of every summand is bounded in $|\cdot|_\nu$ by $(\max(1, |1/4|_\nu \max(1, |m|_\nu)\, C_\nu)^i$.

Now observe that every w_i/w_0 (clearly $w_0 = \sqrt{\ell c(D)}$) is a sum of at most $(2d)^i$ of these, so with the improved triangle inequality (9.1) we obtain the desired result

$$|w_n|_\nu \leq |w_0|_\nu \max(1, |(2d)|_\nu)^n \,(\max(1, |1/4|_\nu \max(1, |m|_\nu)\, C_\nu)^n$$

With $h_{\text{proj}}(1 + f) = h_{\text{proj}}(D)$ and $C_\nu \geq 1$ the inequality for the height follows after we replace $|w_i|_\nu$ with $\max(1, |w_i|_\nu)$ (also for $i = 0$). □

We can now bound the projective height of the numerator of a convergent in terms of the height of the denominator.[1]

Proposition 9.10. *For every convergent $(p, q) \in \mathcal{C}_{\sqrt{D}}(K)$ holds*

$$h_{\text{proj}}(p) \leq h_{\text{proj}}(q) + (\deg p)\,(\log 2 + \log 4 + \log(2d) + h_{\text{proj}}(D))\,.$$

Proof. Recall from Proposition 3.8 that $p = \left\lfloor \sqrt{D}\,q \right\rfloor$. However we did not define a height for $\sqrt{D}$. To workaround this problem, let $m = \deg q$ and write $D = A_m + \varepsilon_m$ with A_m a Laurent polynomial and $\operatorname{ord}(\varepsilon) > m$. This ensures $p = \lfloor A_m\, q \rfloor$, so $h_{\text{proj}}(p) \leq h_{\text{proj}}(A_m\, q)$. With Proposition 9.8, we bound the projective height

$$h_{\text{proj}}(A_m) \leq (d + m)\ (\log 4 + 2\ \log(2d) + h_{\text{proj}}(D))\,.$$

Note that A_m has precisely $d + m = d + \deg q = \deg p$ coefficients. The overall bound then comes from the bound for the product (Proposition 9.1). □

[1] It is not quite clear if there is a similar bound in the other direction.

9.2.2. Lower bound

In [BC97], a general result about the height of Padé approximations predicts that the projective height of the convergents of a square root should grow quadratically in the degree of the convergent.

If we just want to prove this for square root, a shorter and simpler proof by Zannier suffices. It is explained briefly (for example) in [Zan16], here we give a bit more detailed version.

Theorem 9.3. *If D is not Pellian, there exists a constant $C = C(D) > 0$ such that for every convergent $(p, q) \in \mathcal{C}_{\sqrt{D}}(\overline{Q})$ we have for $\deg q$ large enough.*

$$C \cdot (\deg q)^2 \leq h_{\text{proj}}(q).$$

Proof. Recall that by Lemma 4.7 and subsequent remarks, the convergents produce an equality on the Jacobian (recall $[\mathbf{O}] = -j(O_-) = [O_+] - [O_-]$)

$$-m\ [\mathbf{O}] = j(P_1) + \cdots + j(P_r)$$

with $r \leq g$ and $P_i = (x_i, y_i) \in \mathcal{C}_{\text{aff}}$. The x_i are precisely the zeroes of $\Omega = p^2 - D\,q^2$ (accounted for multiplicities). And we have $m = \deg p = \deg q + g + 1 \geq \deg q$.

Applying the Néron-Tate height, and using Lemma A.6 from the Appendix, we obtain

$$m^2\hat{h}([\mathbf{O}]) = \hat{h}(-m\ [\mathbf{O}]) = \hat{h}(j(P_1) + \cdots + j(P_r)) \leq g\left(\hat{h}(j(P_1)) + \cdots + \hat{h}(j(P_r))\right).$$

Next, there is a constant $C_1 \geq 0$ depending only on D such that $\hat{h}(j(P_i)) \leq C_1 + h_{\text{aff}}(x_i)$ (see Section 9.1.5). Moreover, we know that

$$h(x_1) + \cdots + h(x_r) \leq g\ \log 2 + h_{\text{proj}}(\Omega).$$

We combine these estimates to

$$m^2\hat{h}([\mathbf{O}]) \leq g\,C_1 + g^2\ \log 2 + g\,h_{\text{proj}}(\Omega).$$

Because D is not Pellian, the point $[\mathbf{O}]$ is not torsion in the Jacobian, so $\hat{h}([\mathbf{O}]) > 0$. We get a constant $C_2 > 0$ such that

$$C_2\,(\deg q)^2 \leq h_{\text{proj}}(\Omega).$$

As we are only interested in the projective height of q, we may without restriction normalise the convergent (p, q) such that p is monic, so that both p^2 and $D\,q^2$ have to be monic. Of course, for a monic polynomial, affine and projective height coincide, and using Proposition 9.2 we can estimate

$$h_{\text{proj}}(\Omega) \leq h_{\text{aff}}(\Omega) \leq \log 2 + h_{\text{aff}}(p^2) + h_{\text{aff}}(D\,q^2) = \log 2 + h_{\text{proj}}(p^2) + h_{\text{proj}}(D\,q^2).$$

By Proposition 9.10, we have $h_{\text{proj}}(p) \leq h_{\text{proj}}(q) + \deg p \, C_3$ for some $C_3 \geq 0$ depending only on D. With Proposition 9.1 we get

$$h_{\text{proj}}(D\,q^2) \leq 2 \,\deg p\, \log 2 + h_{\text{proj}}(D) + 2\, h_{\text{proj}}(q),$$
$$h_{\text{proj}}(p^2) \leq 2 \,\deg p\, \log 2 + 2\, h_{\text{proj}}(p) \leq 2 \,\deg p\, \log 2 + 2 \deg p\, C_3 + 2\, h_{\text{proj}}(q).$$

Combining these, and noting $\deg p = \deg q + d$, we find

$$C_2\, (\deg q)^2 \leq C_4 + C_5 \,\deg q + 4\, h_{\text{proj}}(q)$$

with $C_4, C_5 \geq 0$, so for example $C = C_2/8$ yields the desired constant. □

This does not yet give a lower bound for the height of partial quotients. This seems more challenging, especially for the projective height – see also Example 8 in Section 10.5. But there are new results for the affine height if we take some type of average, see Theorem 1.4 in [Zan16]:

Theorem 9.4 (Zannier)**.** *There exist $M \in \mathbb{N}$ and $C > 0$ s.t. for n large enough*

$$C\, n^2 \leq \max(h_{\text{aff}}(a_{n-i}) \mid i = 0, \ldots, M)$$

9.2.3. Upper bound

An upper bound for the height of the convergents can be deduced with more elementary tools, using the Toeplitz determinants from Section 3.4. It is then straightforward to deduce an upper bound also for the height of the partial quotients.

Theorem 9.5. *For the canonical convergents (p_m, q_m) of $\sqrt{D}$ we obtain the height bounds*

$$h_{\text{proj}}(p_m) \leq ((n+1)\, d + \tfrac{3}{2}(n^2 + n))\,(\log 4 + 2\, \log(2d) + h_{\text{proj}}(D)), \tag{9.5}$$
$$h_{\text{proj}}(q_m) \leq (n\, d + \tfrac{1}{2}(3\, n^2 + n))\,(\log 4 + 2\, \log(2d) + h_{\text{proj}}(D)), \tag{9.6}$$

where $\deg D = 2d$ and $n = \deg q_m$.

Proof. We wish to apply the results of Section 3.4 for $\alpha = \sqrt{D}$. Connecting the notations of (9.4) and (3.8), we have $N = d$ and $A_j = w_{d-j}$. As we chose $n = \deg q_m$ and the canonical convergents are coprime, Proposition 3.23 tells us that the matrix $\mathcal{M}_n$ has full rank. So the kernel has dimension 1, and we can compute a solution (p, q) using (3.9) which differs from (p_m, q_m) only by a constant factor, and thus has the same projective height. The coefficients of p and q are (up to signs) the minors of

$$\mathcal{M}_n = \left(\begin{array}{ccccc|ccc} -1 & & & & & w_0 & & \\ & \ddots & & & & w_1 & \ddots & \\ & & \ddots & & & \vdots & \ddots & w_0 \\ & & & \ddots & & \vdots & \ddots & \vdots \\ & & & & -1 & w_{d+n} & \cdots & w_d \\ \hline & & & & & w_{d+n+1} & \cdots & w_{d+1} \\ & & 0 & & & \vdots & \ddots & \vdots \\ & & & & & w_{d+2n} & \cdots & w_{d+n} \end{array}\right).$$

If we strike any column (to get the minor), we obtain a $(d+2n+1)\times(d+2n+1)$ matrix. Now set

$$C_\nu = \max(1, |1/4|_\nu) \cdot \max(1, \left|(2d)^2\right|_\nu) \cdot |D|_\nu \,/\, |\ell c(D)|_\nu$$

so that $|w_j|_\nu \leq |w_0|_\nu \, {C_\nu}^j$.

If we strike a column in the right block (to compute the coefficients of q), by using Laplace development we get (up to sign) a minor $\mathcal{M}'$ of the lower right block of dimensions $n \times n$, with determinant

$$\det \mathcal{M}' = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, \mathcal{M}'_{1\,\sigma(1)} \cdots \mathcal{M}'_{n\,\sigma(n)}$$

with $\left|\mathcal{M}'_{ij}\right|_\nu \leq |w_0|_\nu \, {C_\nu}^{d+n+i}$, hence

$$|q|_\nu \leq \prod_{i=1}^{n} |w_0|_\nu \, {C_\nu}^{d+n+i} \leq {|w_0|_\nu}^n {C_\nu}^{n\,d+n^2+n(n+1)/2}.$$

When taking the product over all ν, the first term with w_0 vanishes by the product formula, and likewise the term with $\ell c(D)$. We arrive at (9.6) by a straightforward calculation.

If we strike a column in the left block (to compute the coefficients of p), we can use similar arguments, with Laplace development we only get a $(n+1)\times(n+1)$ matrix $\mathcal{M}''$, with $\left|\mathcal{M}''_{ij}\right|_\nu \leq |w_0|_\nu \, {C_\nu}^{d+n+i-1}$, so

$$|p|_\nu \leq \prod_{i=1}^{n+1} |w_0|_\nu \, {C_\nu}^{d+n+i-1} \leq {|w_0|_\nu}^{n+1} {C_\nu}^{(n+1)\,d+(n+1)n+n(n+1)/2}.$$

Again, (9.5) follows by a straightforward calculation. □

Corollary 9.11. *The projective height of the convergents grows at most quadratically:*

$$h_{\text{proj}}(p_m) = O(m^2), \qquad h_{\text{proj}}(p_m) = O(m^2).$$

Proof. The partial quotients have bounded degree $1 \leq \deg a_i \leq d$, hence $m \leq \deg q_m = n \leq d\,m$, and the above theorem gives $h_{\text{proj}}(p_m) = O(n^2)$ and $h_{\text{proj}}(q_m) = O(n^2)$. □

Corollary 9.12. *The projective height of the partial quotients also grows at most quadratically:*

$$h_{\text{proj}}(a_m) = O(m^2)$$

Proof. We can compute the partial quotients from subsequent convergents as in

$$a_m = \left\lfloor \frac{p_m}{p_{m-1}} \right\rfloor \qquad a_m = \left\lfloor \frac{q_m}{q_{m-1}} \right\rfloor$$

and then Proposition 9.3 yields

$$h_{\text{proj}}(a_m) \leq h_{\text{proj}}(q_m) + (\deg a_m)(\log 2 + h_{\text{proj}}(q_{m-1})) = O(m^2).$$

□

Remark 9.13. An explicit bound is

$$h_{\mathrm{proj}}(a_m) \leq \big(((a+1)n+a)\,d + a\,\log 2 + \tfrac{1}{2}\big(3(a+1)n^2 + (7a+1)n + 3a^2 + a\big)\big) \\ \cdot (\log 4 + 2\,\log(2d) + h_{\mathrm{proj}}(D))$$

where $a = \deg a_m$, $n = \deg q_{m-1}$.

9.3. Connecting heights and valuations

From the definitions of the height of polynomials in Section 9.1.3 and the Gauss norms in Chapter 7 it is clear that there is a direct connection between the height and the valuations computations, as for example in Theorem 8.2. Note that for a polynomial $f \in K[X]$, we have

$$h_{\mathrm{proj}}(f) = \sum_{\nu \in M_K} \log |f|_\nu\,, \tag{9.7}$$

$$h_{\mathrm{aff}}(f) = \sum_{\nu \in M_K} \max\left(0, \log |f|_\nu\right). \tag{9.8}$$

For ν non-archimedean, $\log |f|_\nu$ is essentially $c \cdot \nu(f)$ for some $c < 0$.

However, in Chapters 7 and 8 we do not treat the places over 2, and certainly not the archimedean places. So this connection must remain incomplete. Still, we try to point out some phenomena relating the global picture of heights and the local picture of Gauss norms.

We restrict to the genus 1 case with $\deg D = 4$ and D non-Pellian, so that we may use Theorem 8.2.

Corollary 8.18 says that for places with $[\mathbf{O}_{\mathrm{red}}]$ having even torsion order, the Gauss norms of q_n grow at least linearly in n. But $h_{\mathrm{proj}}(q_n)$ should grow quadratically, which suggests that either the Gauss norms grow faster than linearly, or the number of places with bad reduction of the continued fraction before n also grows linearly.

Computational evidence suggests that the latter is the case. Also, if we work over $\mathbb{Q}$, the Hasse-Weil interval (see Remark 4.13) predicts that $\mathcal{J}_{\mathrm{red}}(\mathbb{F}_\mathfrak{p})$ grows about linearly in $\mathfrak{p}$, so the quasi-period length of $\mathbf{CF}(D_\mathfrak{p})$ and hence the first occurrence of $\mathfrak{p}$ in a denominator of a_n grows linearly in $\mathfrak{p}$. However, the number of primes $\mathfrak{p} \leq n$ grows only as $n/\log n$.

The valuations $\nu(q_n)$ mostly alternate between positive and negative signs, so for the projective height they might cancel each other out. But for the affine height, there is no such cancellation, and in fact computations for examples suggest that $h_{\mathrm{aff}}(q_n)$ grows more or less cubically. This is in line with Remark 4.8 (ii) in [Zan16], which says that $h_{\mathrm{aff}}(q_n)$ should at most grow cubically in n.

Similar observations can be made for the partial quotients.

10. Examples

We now apply the theory developed in this thesis to some examples. Hopefully, this illustrates our theorems and their limitations. To this end, we include examples also for the corner cases that have been somewhat neglected in the theoretical part.

For D defined over the rationals (or perhaps a number field), and $\mathfrak{p}$ some prime (in the ring of integers), we use the notations $D_\mathfrak{p}$ for $\overline{D}$ in $\mathbb{F}_\mathfrak{p}[X]$, $\nu_\mathfrak{p}$ for the $\mathfrak{p}$-adic valuation, we denote by $\mathbf{O}_\mathfrak{p}$ the torsion divisor $(O_+) - (O_-)$ on $\mathcal{C}_{\text{red}}$ over $\mathbb{F}_\mathfrak{p}$.

For D defined over $\mathbb{C}(t)$, we use analogous notation, with $t - t_0$ or $t = t_0$ instead of $\mathfrak{p}$.

10.1. Reduction to a square

We begin with some examples where D reduces (or specializes) to a square.

Example 1

$D = (X-1) \cdot X \cdot (X-t) \cdot (X+t-1)$	basefield $\mathbb{C}(t)$
Discriminant of D: $(4) \cdot (t-\frac{1}{2})^2 \cdot (t-1)^4 \cdot t^4$	Primes with $\overline{D}$ square: $t-1, t$
period length 2 for $\mathbf{CF}(\sqrt{D})$	quasi-period length 1 for $\mathbf{CF}(\sqrt{D})$
Minimal Pell solution	
$p_0 = X^2 - X - \frac{1}{2}t^2 + \frac{1}{2}t$	$q_0 = 1$
Partial quotients of $\sqrt{D}$	
$a_0 = X^2 - X - \frac{1}{2}t^2 + \frac{1}{2}t$	
$a_1 = \frac{-8X^2+8X+4t^2-4t}{(t-1)^2 \cdot t^2}$	
$a_2 = 2X^2 - 2X - t^2 + t$	

Example 1 is very simple, but it illustrates already that bad reduction of the continued fraction is not the same as bad reduction of the elliptic curve. Only $D_{t=0}$ and $D_{t=1}$ are square, and $t, t-1$ are the only irreducible/prime factors appearing in the coefficient denominators (of a_1).

By the way, this means we can specialise t to say an integer $t_0 \in \mathbb{Z} \setminus \{0, 1\}$ to get a periodic continued fraction over $\mathbb{Q}$. This continued fraction has bad reduction precisely at the prime numbers dividing $t_0\,(t_0 - 1)$. In this way we obtain an example also for the reduction modulo $\mathfrak{p}$ case.

Example 2 clearly reduces to a square at $t = 0$. We check that $\mathbf{CF}(\sqrt{D})$ is non-periodic by specializing to $t = 3$. Then reduction of the continued fraction $\mathbf{CF}(D_{t=3})$ modulo 5 and 7 yields torsion orders 5 respectively 10 which implies non-periodicity for both $\mathbf{CF}(\sqrt{D_{t=3}})$ and $\mathbf{CF}(\sqrt{D})$. As we are in the degree 4 case, this implies good reduction

Example 2

$D = X^4 + 2X^2 + tX + 1$	basefield $\mathbb{C}(t)$
Discriminant of D: $(-27) \cdot t^2 \cdot (t^2 - \frac{256}{27})$	Primes with $\overline{D}$ square: t
D is not Pellian	
Partial quotients of $\sqrt{D}$	
$a_0 = X^2 + 1$	
$a_1 = \frac{2X}{t}$	
$a_2 = \frac{1}{2}tX - \frac{1}{8}t^2$	
$\deg a_n = 2, 1, 1, 1, \ldots$	

of $\mathbf{CF}(\sqrt{D})$ at $t - 3$ (by Proposition 8.8). Of course $D_{t=3}$ reduces then to a square modulo 3, so the example works for the reduction modulo $\mathfrak{p}$ case too.

So both in the periodic and non-periodic case, it is possible that D reduces to a square.

10.2. Reduction periodic to periodic

For $\deg D = 4$, we have seen that torsion order m and quasi-period length ℓ satisfy $m = \ell + 1$ (see Proposition 6.11). Together with the discussion of Section 8.4.3 on how the quasi-period may shorten, and rational torsion on elliptic curves being bounded by 12, there cannot be many examples of bad reduction of a continued fraction for $D \in \mathbb{Q}[X]$.

Example 3

$D = X^4 - 8X^3 - 42X^2 + 424X - 119$	basefield $\mathbb{Q}$
Discriminant of D: $-1 \cdot 2^{29} \cdot 3^5$	Primes with $\overline{D}$ square: 3
period length 8 for $\mathbf{CF}(\sqrt{D})$	
Minimal Pell solution	
$\deg p_7 = 9$	$\deg q_7 = 7$
Partial quotients of $\sqrt{D}$	
$a_0 = X^2 - 4X - 29$	$a_4 = \frac{4}{3}X - \frac{44}{3}$
$a_1 = \frac{1}{96}X + \frac{1}{96}$	$a_5 = \frac{1}{32}X + \frac{5}{32}$
$a_2 = -4X + 12$	$a_6 = -4X + 12$
$a_3 = \frac{1}{32}X + \frac{5}{32}$	$a_7 = \frac{1}{96}X + \frac{1}{96}$

Indeed this is the case in Example 3, where we see only 2 and 3 in the denominators. As $\mathbf{O}$ has order 9, the only candidate for bad reduction of $\mathbf{CF}(D)$ is 3, but D_3 is already a square. Everywhere else we have good reduction of $\mathbf{CF}(\sqrt{D})$.

But if we are working over number fields, and can increase the torsion orders, then in principle one should be able to construct example with bad reduction of the continued fraction (employing Theorem 8.3).

We also analysed an example with $\deg D = 6$ given in [Pla14] (f_{33} in Section 6). In Example 4, we have torsion order 33, so both 3 and 11 are good (but a priori not the only) candidates for bad reduction of $\mathbf{CF}(\sqrt{D})$. But again D_3 is already square. And we do not see 11 in the denominators (it suffices to check the first half of the palindromic quasi-period as mentioned in Section 8.4.3). We have good reduction of $\sqrt{D}$ at all other primes which is not surprising given that the example seems to have been constructed by testing for this. Also observe that the factor for the quasi-period is $\mu = 3$, as predicted by Remark 8.4.

Example 4

$D = 4X^6 + 28X^5 + 37X^4 - 30X^3 + 87X^2 - 54X + 9$	basefield $\mathbb{Q}$
Discriminant of D: $-1 \cdot 2^{22} \cdot 3^{14} \cdot 127$	Primes with $\overline{D}$ square: 3
period length 54 for $\mathbf{CF}(\sqrt{D})$	quasi-period length 27 for $\mathbf{CF}(\sqrt{D})$
Minimal Pell solution	
$\deg p_{26} = 33$	$\deg q_{26} = 30$
Partial quotients of $\sqrt{D}$	
$a_0 = 2X^3 + 7X^2 - 3X + 3$	$a_9 = 6X$
$a_1 = \frac{1}{9}X + \frac{1}{2}$	$a_{10} = \frac{1}{9}X + \frac{7}{18}$
$a_2 = 3X - \frac{9}{2}$	$a_{11} = -3X - \frac{3}{2}$
$a_3 = \frac{2}{27}X^2 + \frac{1}{3}X + \frac{2}{9}$	$a_{12} = -\frac{1}{3}X - \frac{5}{6}$
$a_4 = 6X^2 + 21X - 9$	$a_{13} = \frac{2}{3}X + 1$
$a_5 = \frac{1}{9}X - \frac{1}{6}$	$a_{14} = 2X + 3$
$a_6 = X + \frac{11}{2}$	$a_{15} = -\frac{1}{9}X - \frac{5}{18}$
$a_7 = -2X + 2$	$a_{16} = -9X - \frac{9}{2}$
$a_8 = -\frac{1}{9}X - \frac{1}{2}$	$a_{17} = \frac{1}{27}X + \frac{7}{54}$

10.3. Reduction non-periodic to periodic

10.3.1. Genus 1

For a polynomial of degree 4, our Theorem 8.2 describes the behaviour of the valuations. We now give an example to illustrate this, both for odd and even quasi-period length of $\mathbf{CF}(\sqrt{D_{\mathfrak{p}}})$.

Example 5

$D = X^4 + 5X^2 - 3X + 19$	basefield $\mathbb{Q}$
Discriminant of D: $3^2 \cdot 7^2 \cdot 11^2 \cdot 17$	Primes with $\overline{D}$ square: 3
D is not Pellian because of incompatible torsion orders	torsion order 8 modulo 5
	torsion order 3 modulo 7
Partial quotients of $\sqrt{D}$	
$a_0 = X^2 + \frac{5}{2}$	
$a_1 = -\frac{2}{3}X - \frac{17}{6}$	
$a_2 = -\frac{24}{329}X + \frac{33270}{108241}$	
$\deg a_n = 2, 1, 1, 1, 1, \ldots$	

Example 5 is chosen randomly. Note again that while the discriminant has a finite number of prime divisors, only D_3 is a square polynomial, and of course $\mathbf{CF}(\sqrt{D})$ has bad reduction for every odd prime number $\mathfrak{p}$ by Lemma 8.6 and Corollary 6.2.

modulo 5

Table 10.1 lists the 5-adic valuations (Gauss norms). Note how the changes in the patterns, and the unbounded α_n are aligned with the 2-element fibres of λ. We can also read off the quasi-period length of $\mathbf{CF}(\sqrt{D_5})$ from the first occurrence of non-zero valuations, and determine it to be 7 (this works only in degree 4).

As the quasi-period length is odd, we can observe (as predicted by Corollary 8.18) that the valuations increase in absolute value. We also see that the sign of the exponent is alternating, and that almost all the $\nu(a_n)$ are divisible by 4 (as predicted by Theorem 8.2). Pay attention to the valuations of the leading coefficients being larger which for q_n indicates that $\widehat{q_n}$ has a lower degree.

modulo 19

Table 10.2 is for the 19-adic valuations. The patterns are very similar to the table for 5. We can also read off the quasi-period of $\mathbf{CF}(\sqrt{D_{19}})$: it is 6, hence even. The alternating signs of the valuations then lead to cancellation of exponents at the 2-element fibres. However it remains an open question whether these valuations are eventually periodic. If so, our computations suggest that their period length must be significantly larger than the quasi-period length of $\mathbf{CF}(\sqrt{D_{\mathfrak{p}}})$.

Table 10.1.: 5-adic valuations for Example 5

n	$\lambda(n)$	$\nu(\alpha_n)$	$\nu(a_n)$	$\nu(\ell c(a_n))$	$\nu(q_n)$	$\nu(\ell c(q_n))$
0	0	0	0	0	0	0
1	1	0	0	0	0	0
2	2	0	0	0	0	0
3	3	0	0	0	0	0
4	4	0	0	0	0	0
5	5	0	0	0	0	0
6	6	0	0	0	0	0
7	6	$-\infty$	-2	-1	-2	-1
8	7	2	2	3	2	2
9	8	-4	-4	-4	-2	-2
10	9	4	4	4	2	2
11	10	-4	-4	-4	-2	-2
12	11	4	4	4	2	2
13	12	-4	-4	-4	-2	-2
14	13	4	4	4	2	2
15	13	$-\infty$	-6	-5	-4	-3
16	14	6	6	7	4	4
17	15	-8	-8	-8	-4	-4
18	16	8	8	8	4	4
19	17	-8	-8	-8	-4	-4
20	18	8	8	8	4	4
21	19	-8	-8	-8	-4	-4
22	20	8	8	8	4	4
23	20	$-\infty$	-10	-9	-6	-5
24	21	10	10	11	6	6
25	22	-12	-12	-12	-6	-6
26	23	12	12	12	6	6
27	24	-12	-12	-12	-6	-6
28	25	12	12	12	6	6
29	26	-12	-12	-12	-6	-6
30	27	12	12	12	6	6
31	27	$-\infty$	-14	-13	-8	-7
32	28	14	14	15	8	8
33	29	-16	-16	-16	-8	-8
34	30	16	16	16	8	8

Table 10.2.: 19-adic valuations for Example 5

n	$\lambda(n)$	$\nu(\alpha_n)$	$\nu(a_n)$	$\nu(\ell c(a_n))$	$\nu(q_n)$	$\nu(\ell c(q_n))$
0	0	0	0	0	0	0
1	1	0	0	0	0	0
2	2	0	0	0	0	0
3	3	0	0	0	0	0
4	4	0	0	0	0	0
5	5	0	0	0	0	0
6	5	$-\infty$	-2	-1	-2	-1
7	6	2	2	3	2	2
8	7	-4	-4	-4	-2	-2
9	8	4	4	4	2	2
10	9	-4	-4	-4	-2	-2
11	10	4	4	4	2	2
12	11	-4	-4	-4	-2	-2
13	11	$-\infty$	2	3	0	1
14	12	-2	-2	-1	0	0
15	13	0	0	0	0	0
16	14	0	0	0	0	0
17	15	0	0	0	0	0
18	16	0	0	0	0	0
19	17	0	0	0	0	0
20	17	$-\infty$	-2	-1	-2	-1
21	18	2	2	3	2	2
22	19	-4	-4	-4	-2	-2
23	20	4	4	4	2	2
24	21	-4	-4	-4	-2	-2
25	22	4	4	4	2	2
26	23	-4	-4	-4	-2	-2
27	23	$-\infty$	2	3	0	1
28	24	-2	-2	-1	0	0
29	25	0	0	0	0	0
30	26	0	0	0	0	0
31	27	0	0	0	0	0
32	28	0	0	0	0	0
33	29	0	0	0	0	0
34	29	$-\infty$	-2	-1	-2	-1
35	30	2	2	3	2	2
36	31	-4	-4	-4	-2	-2
37	32	4	4	4	2	2

Note that the torsion order of $\mathbf{O}_{19}$ is 7, while the torsion order of $\mathbf{O}_5$ is 8 (from Proposition 6.11). This implies that $\mathbf{CF}(\sqrt{D})$ cannot be periodic, using the arguments from Remark 8.29 in Section 8.4.2 with reduction modulo two primes.

10.3.2. Genus 2

We also give an example of degree 6, to illustrate the difficulties arising in higher genus, and the more complicated patterns of the valuations in this case. Moreover, we will find convergents where $\widehat{p_n}$ and $\widehat{q_n}$ share a common linear factor.

Example 6

$D = X^6 + 7X^4 + 8X^3 + 9X^2 + 5$	basefield $\mathbb{Q}$
Discriminant of D: $-1 \cdot 2^{10} \cdot 5 \cdot 7^2 \cdot 353^2$	D never reduces to a square.
D is not Pellian	
Partial quotients of $\sqrt{D}$	
$a_0 = X^3 + \frac{7}{2}X + 4$	
$a_1 = -\frac{8}{13}X + \frac{896}{169}$	
$a_2 = -\frac{2197}{100508}X - \frac{112744970}{631366129}$	
$\deg a_n = 3, 1, 1, 1, 1, \ldots$	

Example 6 is again a random non-periodic example. See how the coefficient size explodes worse than in the genus 1 example. And observe that the prime 13 appears already in a_1. However, it turns out that $\mathbf{CF}(\sqrt{D_{13}})$ has a quite long quasi-period length: it is 126.

So unlike in genus 1, the first occurrence of a prime $\mathfrak{p}$ in a denominator of the a_n does not give so much information on the quasi-period length of $\mathbf{CF}(\sqrt{D_{\mathfrak{p}}})$.

modulo 3

In Table 10.3, we compare the degrees of partial quotients between $\mathbf{CF}(\sqrt{D})$ and its reduction $\mathbf{CF}(\sqrt{D_3})$. We put $m = \lambda(n)$, and be aware that the columns depending on m contain *repeated entries.*

Note particularly that the sequence of the $\deg \widehat{q_n}$ is also decreasing, and sometimes is larger than the corresponding $\deg v_m$. This means that $\widehat{p_n}, \widehat{q_n}$ have a common linear factor. This of course happens here only in the 3-element fibres of λ (in the table $m = \lambda(n)$). For example

$$\widehat{p_7} = (X+1) \cdot (2X^9 + 2X^8 + X^7 + 2X^6 + 2X^5 + 2X^4 + 2),$$
$$\widehat{q_7} = (X+1) \cdot (2X^6 + 2X^5 + 2X^3 + X^2).$$

The patterns for the valuations in Table 10.4 are now more interesting, as there are fibres of λ with 2 or 3 elements. But at least these are still isolated. Observe the differences between the valuation of the entire polynomial (the Gauss norm) and of the

Table 10.3.: Degrees for reduction mod 3 in Example 6

n	m	$\deg a_n$	$\deg c_m$	$\deg q_n$	$\deg \widehat{q_n}$	$\deg v_m$
0	0	3	3	0	0	0
1	1	1	1	1	1	1
2	2	1	1	2	2	2
3	2	1	1	3	2	2
4	3	1	2	4	4	4
5	4	1	1	5	5	5
6	5	1	1	6	6	6
7	5	1	1	7	7	6
8	5	1	1	8	6	6
9	6	1	3	9	9	9
10	7	1	1	10	10	10
11	8	1	1	11	11	11
12	8	1	1	12	11	11
13	9	1	2	13	13	13
14	10	1	1	14	14	14
15	11	1	1	15	15	15
16	11	1	1	16	16	15
17	11	1	1	17	15	15
18	12	1	3	18	18	18
19	13	1	1	19	19	19
20	14	1	1	20	20	20
21	14	1	1	21	20	20
22	15	1	2	22	22	22
23	16	1	1	23	23	23
24	17	1	1	24	24	24
25	17	1	1	25	25	24
26	17	1	1	26	24	24
27	18	1	3	27	27	27
28	19	1	1	28	28	28

Table 10.4.: 3-adic valuations for Example 6

n	$\lambda(n)$	$\nu(\alpha_n)$	$\nu(a_n)$	$\nu(\ell c(a_n))$	$\nu(q_n)$	$\nu(\ell c(q_n))$
0	0	0	0	0	0	0
1	1	0	0	0	0	0
2	2	0	0	0	0	0
3	2	$-\infty$	-2	-1	-2	-1
4	3	2	2	3	2	2
5	4	-4	-4	-4	-2	-2
6	5	4	4	4	2	2
7	5	$-\infty$	-5	-5	-3	-3
8	5	5	6	6	2	3
9	6	-5	-5	-5	-2	-2
10	7	4	4	4	2	2
11	8	-4	-4	-4	-2	-2
12	8	$-\infty$	0	2	-2	0
13	9	0	0	2	2	2
14	10	-4	-4	-4	-2	-2
15	11	4	4	4	2	2
16	11	$-\infty$	-5	-5	-3	-3
17	11	5	6	6	2	3
18	12	-5	-5	-5	-2	-2
19	13	4	4	4	2	2
20	14	-4	-4	-4	-2	-2
21	14	$-\infty$	2	3	0	1
22	15	-2	-2	-1	0	0
23	16	0	0	0	0	0
24	17	0	0	0	0	0
25	17	$-\infty$	-2	-2	-2	-2
26	17	2	4	4	0	2
27	18	-2	-2	-2	0	0
28	19	0	0	0	0	0

leading coefficient between 2-element fibres and 3-element fibres. Note that now odd valuations are occurring.

Also, we cannot read off the quasi-period length of $\mathbf{CF}(\sqrt{D_3})$ just by counting the rows with only zero valuations. From Table 10.3, we know that it is actually 6, not 3 (by looking for c_m of degree 3 which first occurs for $m = 6$). This corresponds to torsion order $9 = \deg p_5$ of $\mathbf{O}_3$ (via Theorem 4.1 and Remark 4.9).

modulo 19

Another interesting prime would be 19. There $\mathbf{CF}(\sqrt{D_{19}})$ has (quasi-)period length 6, with degrees of the a_n having the periodic pattern $\deg a_n = \overline{3, 1, 1, 1, 1, 1}$.

This degree pattern implies that λ has only fibres with 1 or 3 elements.

In table 10.5, note how the valuations are reset to 0 after the 3-element fibres. This illustrates nicely why we require infinitely many fibres of λ with multiple elements in Proposition 8.11.

modulo 5

So far, the regularity of these valuation patterns has been deceiving, so let us look at the 5-adic valuations too. The quasi-period length of $\mathbf{CF}(\sqrt{D_5})$ is just 6. The partial quotients period is $\deg a_n = \overline{3, 1, 1, 2, 1, 1}$.

So compared to $\mathfrak{p} = 3$, there are also 2-element fibres. This makes the patterns much more complicated, as seen in Table 10.6 (and in other examples, this might be even worse).

Table 10.5.: 19-adic valuations for Example 6

n	$\lambda(n)$	$\nu(\alpha_n)$	$\nu(a_n)$	$\nu(\ell c(a_n))$	$\nu(q_n)$	$\nu(\ell c(q_n))$
0	0	0	0	0	0	0
1	1	0	0	0	0	0
2	2	0	0	0	0	0
3	3	0	0	0	0	0
4	4	0	0	0	0	0
5	5	0	0	0	0	0
6	5	$-\infty$	-1	-1	-1	-1
7	5	1	2	2	0	1
8	6	-1	-1	-1	0	0
9	7	0	0	0	0	0
10	8	0	0	0	0	0
11	9	0	0	0	0	0
12	10	0	0	0	0	0
13	11	0	0	0	0	0
14	11	$-\infty$	-1	-1	-1	-1
15	11	1	2	2	0	1
16	12	-1	-1	-1	0	0
17	13	0	0	0	0	0
18	14	0	0	0	0	0
19	15	0	0	0	0	0
20	16	0	0	0	0	0
21	17	0	0	0	0	0
22	17	$-\infty$	-1	-1	-1	-1
23	17	1	2	2	0	1
24	18	-1	-1	-1	0	0
25	19	0	0	0	0	0
26	20	0	0	0	0	0
27	21	0	0	0	0	0
28	22	0	0	0	0	0
29	23	0	0	0	0	0
30	23	$-\infty$	-1	-1	-1	-1
31	23	1	2	2	0	1
32	24	-1	-1	-1	0	0
33	25	0	0	0	0	0
34	26	0	0	0	0	0
35	27	0	0	0	0	0

Table 10.6.: 5-adic valuations for Example 6

n	$\lambda(n)$	$\nu(\alpha_n)$	$\nu(a_n)$	$\nu(\ell c(a_n))$	$\nu(q_n)$	$\nu(\ell c(q_n))$
0	0	0	0	0	0	0
1	1	0	0	0	0	0
2	2	0	0	0	0	0
3	2	$-\infty$	-6	-3	-6	-3
4	3	6	6	9	6	6
5	4	-12	-12	-12	-6	-6
6	5	12	12	12	6	6
7	5	$-\infty$	-13	-13	-7	-7
8	5	13	14	14	6	7
9	6	-13	-13	-13	-6	-6
10	7	12	12	12	6	6
11	8	-12	-12	-12	-6	-6
12	8	$-\infty$	6	9	0	3
13	9	-6	-6	-3	0	0
14	10	0	0	0	0	0
15	11	0	0	0	0	0
16	11	$-\infty$	-1	-1	-1	-1
17	11	1	2	2	0	1
18	12	-1	-1	-1	0	0
19	13	0	0	0	0	0
20	14	0	0	0	0	0
21	14	$-\infty$	-8	-4	-8	-4
22	15	8	8	12	8	8
23	16	-16	-16	-16	-8	-8
24	17	16	16	16	8	8
25	17	$-\infty$	-17	-17	-9	-9
26	17	17	18	18	8	9
27	18	-17	-17	-17	-8	-8
28	19	16	16	16	8	8
29	20	-16	-16	-16	-8	-8
30	20	$-\infty$	10	13	2	5
31	21	-10	-10	-7	-2	-2
32	22	4	4	4	2	2
33	23	-4	-4	-4	-2	-2
34	23	$-\infty$	3	3	1	1
35	23	-3	-2	-2	-2	-1
36	24	3	3	3	2	2
37	25	-4	-4	-4	-2	-2

10.4. Non-constant degrees

The following example was constructed in collaboration with Prof. Zannier and Francesca Malagoli, to answer a question raised during preparation of [Zan16]: In the article, it is a consequence of the Skolem-Mahler-Lech Theorem for algebraic groups mentioned before that the sequence of the $\deg a_n$ (for $\alpha = \sqrt{D}$) becomes eventually periodic. However, in any non-periodic examples known previously, these degrees stabilised on a single value. Of course, in that case periodicity of the degrees is not very interesting.

So we searched for an non-periodic example where the degrees assume multiple values infinitely often.

We found Example 7 which has infinitely many partial quotients a_n both of degree 1 and of degree 2 (we remark that this would be impossible for $\deg D = 4$ or 6, so we cannot do better than $\deg D = 8$).

Example 7

$D = X^8 - X^7 - \frac{3}{4}X^6 + \frac{7}{2}X^5 - \frac{21}{4}X^4 + \frac{7}{2}X^3 - \frac{3}{4}X^2 - X + 1$	basefield $\mathbb{Q}$
Discriminant of D: $-1 \cdot 2^2 \cdot 3 \cdot 13 \cdot 173^2$	Primes in denominators of D: 2
D never reduces to a square.	
D is not Pellian because of incompatible torsion orders	torsion order 10 modulo 3
	torsion order 40 modulo 11
Partial quotients of $\sqrt{D}$	
$a_0 = X^4 - \frac{1}{2}X^3 - \frac{1}{2}X^2 + \frac{3}{2}X - 2$	
$a_1 = \frac{2}{3}X + \frac{7}{9}$	
$a_2 = -\frac{27}{4}X - \frac{45}{8}$	
$a_3 = \frac{16}{81}X^2 - \frac{64}{81}X + \frac{200}{81}$	
$a_4 = -\frac{27}{200}X - \frac{171}{400}$	
$\deg a_n = 4, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, 1, 1, 1, 1, 1, 1, 1, 2, 1, \ldots$	

This is related to the fact that the Jacobian in Example 7 is not simple. It contains an elliptic curve, and infinitely many multiples of the point $\mathbf{O}$ lie on a certain translate of it. This causes the degrees of the a_n to follow the pattern $4, 1, 1, \overline{2, 1, 1, 1, 1, 1, 1, 1, 1}$. For details, we refer to an article in preparation together with Malagoli and Zannier.

Here we remark only that if we reduce modulo 3, we actually get a square-free polynomial in $(X+1)^2$ (and divisible by $(X+1)^2$ too, hence the 3 in the discriminant). So all the partial quotients have at least degree 2 (see also Table 10.7).

Note that λ has still infinitely many fibres with a single element. Observe the sequence $\deg \widehat{q_n}$ is sometimes decreasing, so there are again convergents with a common factor between $\widehat{p_n}$ and $\widehat{q_n}$.

Table 10.7.: Degrees modulo 3 for Example 7

n	m	$\deg a_n$	$\deg c_m$	$\deg q_n$	$\deg \widehat{q_n}$	$\deg v_m$
0	0	4	4	0	0	0
1	0	1	4	1	0	0
2	1	1	2	2	2	2
3	2	2	2	4	4	4
4	2	1	2	5	4	4
5	3	1	2	6	6	6
6	3	1	2	7	7	6
7	3	1	2	8	7	6
8	3	1	2	9	6	6
9	4	1	4	10	10	10
10	4	1	4	11	10	10
11	5	1	2	12	12	12
12	6	2	2	14	14	14
13	6	1	2	15	14	14
14	7	1	2	16	16	16
15	7	1	2	17	17	16
16	7	1	2	18	17	16
17	7	1	2	19	16	16
18	8	1	4	20	20	20
19	8	1	4	21	20	20
20	9	1	2	22	22	22
21	10	2	2	24	24	24
22	10	1	2	25	24	24
23	11	1	2	26	26	26
24	11	1	2	27	27	26
25	11	1	2	28	27	26
26	11	1	2	29	26	26

10.5. Recurring partial quotients

Recall that [Zan16] gave a lower bound for an average of the affine heights of partial quotients (see the end of Section 9.2.2). A strengthening of this would be a bound like

$$C\,n^2 \leq h_{\mathrm{proj}}(a_n)$$

for the non-Pellian case.

However, together with Prof. Zannier and Francesca Malagoli, and some assistance from Solomon Vishkautsan for the computations, we have found Example 8 below. There for $n = 7 + 17\,j \pm 1,\ j \in \mathbb{N}_0$ the partial quotients have the shape

$$a_n = C_n\,(X-2), \quad C_n \in \overline{\mathbb{Q}}$$

so in particular $h_{\mathrm{proj}}(a_n)$ remains constant on this subsequence and the above lower bound is *impossible* in general.

Example 8

<table>
<tr><td>$D = X^{12}+(-8\tau^4+6\tau^3-28\tau^2+22\tau+22)X^{10}+(-8\tau^4+6\tau^3-28\tau^2+22\tau+22)X^9+(83\tau^4-62\tau^3+291\tau^2-225\tau-309)X^8+(166\tau^4-124\tau^3+582\tau^2-450\tau-618)X^7+(-127\tau^4+92\tau^3-447\tau^2+327\tau+529)X^6+(-630\tau^4+462\tau^3-2214\tau^2+1656\tau+2514)X^5+(-538\tau^4+398\tau^3-1893\tau^2+1434\tau+2115)X^4+(158\tau^4-102\tau^3+546\tau^2-336\tau-758)X^3+(552\tau^4-384\tau^3+1926\tau^2-1332\tau-2394)X^2+(368\tau^4-256\tau^3+1284\tau^2-888\tau-1596)X+92\tau^4-64\tau^3+321\tau^2-222\tau-399$</td><td>basefield $K = \mathbb{Q}(\tau)$, where τ has minimal polynomial t^5+3t^3-6t-3</td></tr>
<tr><td></td><td>D never reduces to a square.</td></tr>
<tr><td>D is not Pellian because of incompatible torsion orders</td><td>torsion order 42 modulo τ</td></tr>
<tr><td></td><td>torsion order 861 modulo $3\tau^4-2\tau^3+11\tau^2-8\tau-11$</td></tr>
<tr><td colspan="2">Partial quotients of $\sqrt{D}$</td></tr>
<tr><td colspan="2">$a_0 = X^6+(-4\tau^4+3\tau^3-14\tau^2+11\tau+11)X^4+(-4\tau^4+3\tau^3-14\tau^2+11\tau+11)X^3+(16\tau^4-12\tau^3+56\tau^2-43\tau-65)X^2+(32\tau^4-24\tau^3+112\tau^2-86\tau-130)X+28\tau^4-22\tau^3+98\tau^2-79\tau-119$</td></tr>
<tr><td colspan="2">$a_1 = (\frac{7}{18}\tau^4-\frac{5}{18}\tau^3+\frac{25}{18}\tau^2-\tau-\frac{5}{3})X-\frac{118}{81}\tau^4+\frac{55}{54}\tau^3-\frac{140}{27}\tau^2+\frac{98}{27}\tau+\frac{115}{18}$</td></tr>
<tr><td colspan="2">$a_2 = (\frac{54453438756}{1411680971}\tau^4+\frac{15324049962}{1411680971}\tau^3+\frac{225993869532}{1411680971}\tau^2+\frac{96839726742}{1411680971}\tau-\frac{49873817664}{1411680971})X-\frac{233866366289476431120}{1992843163883502841}\tau^4-\frac{33996626533462319838}{1992843163883502841}\tau^3-\frac{951199485540229734912}{1992843163883502841}\tau^2-\frac{278019658364688097056}{1992843163883502841}\tau+\frac{30559163428385941771 8}{1992843163883502841}$</td></tr>
<tr><td colspan="2">$\deg a_n = 6,1,1,1,1,1,1,3,1,1,1,1,1,1,1,2,2,1,1,1,1,1,1,1,3,1,1,1,1,1,\ldots$</td></tr>
</table>

This also gives and example where $\deg a_n$ assumes three different values infinitely often, and again this is related to the Jacobian containing an elliptic curve. We hope to describe this example in much more detail in the article in preparation together with Malagoli and Zannier mentioned above.

A. Appendix

A.1. Polynomial Pell equation in characteristic 2

Let us quickly have a look at the polynomial Pell equation in characteristic 2 and give a criterion which allows to easily test for and construct solutions in this case.

Theorem A.1. *Let $\mathbb{K}$ a field of characteristic 2 and $D \in \mathbb{K}[X]$. There exists a non-trivial solution (with $q \neq 0$) of*

$$p^2 - D\,q^2 = \eta, \qquad p, q \in \mathbb{K}[X], \eta \in \mathbb{K}^\times$$

if and only if there exist $E \in \mathbb{K}[X], r \in \mathbb{K}$ such that $D = E^2 + r$.

Moreover, $r = 0$ happens if and only if there exists a non-trivial solution with η a square.

Proof. Let us first treat the second part with $r = 0$. Suppose $D = E^2$, and choose $\mu \in \mathbb{K}^\times$, $p = E - \mu$, $q = 1$. This yields

$$p^2 - D\,q^2 = (E-\mu)^2 - E^2 = \mu^2 = \eta,$$

hence η can be chosen a square.

On the other hand, suppose $(p, q) \in \mathbb{K}[X]^2$ with $q \neq 0$ is a solution with $\eta = \mu^2$ a square, then

$$D\,q^2 = p^2 - \mu^2 = (p-\mu)^2$$

implies D is a square because $\mathbb{K}[X]$ is a unique factorisation domain.

For the general case, note that if $D = E^2 + r$ with $r \neq 0$, then

$$p = E, q = 1, \eta = r \implies p^2 - D\,q^2 = -r = \eta$$

gives the desired non-trivial solution.

Conversely, if there exists with a solution $(p, q) \in \mathbb{K}[X]^2$ with $q \neq 0$, set $K = \mathbb{K}(\sqrt{\eta})$ and reduce to the case with $r = 0$ – we now write $D = E^2$ with $E \in K[X]$, or rather $E = E_0 + \mu\,E_1$ with $E_0, E_1 \in \mathbb{K}[X]$ (here again $\mu = \sqrt{\eta}$). We obtain

$$D = E^2 = E_0^2 + \mu^2\,E_1^2 = E_0^2 + \eta\,E_1^2$$

and plugging it into the Pell equation we have

$$0 = p^2 - q^2\,\left(E_0^2 - \eta\,E_1^2\right) + \eta = (p - q\,E_0)^2 - \eta\,(q\,E_1 + 1)^2$$

If $(q\,E_1 + 1) \neq 0$, then $\mu \in K \cap \mathbb{K}(X) = \mathbb{K}$ and we are actually in the first case. Otherwise, $q\,E_1 = 1$, so $E_1 \in \mathbb{K}^\times$ (because $q \in \mathbb{K}[X]$), hence $D = E_0^2 + \eta\,E_1^2 = E_0^2 + r$ with $r = \eta\,E_1^2 \in \mathbb{K}$. □

Remark A.1. So if we require $\eta = 1$, we see that in characteristic 2 non-trivial solutions to the Pell equations only exists if D is actually a square.

Remark A.2. The proof also yields a classification of the Pell solutions:

For the first case with $D = E^2$, the solutions always have the shape $p = q\,E - \mu$. And obviously, we are free to choose q here, so there are a lot of non-trivial solutions in this case.

In the second case with $D = E^2 + r$, we need to expand this observation. But note that we actually showed $q \in \mathbb{K}^\times$ in the above proof, so essentially $q = 1$ after multiplying η with a square factor. Hence there is *only one* non-trivial solution up to a constant factor.

A.2. Valuations in Laurent series quotients

The problem that arises with bad reduction is that we can no longer read off $\nu(\alpha_n)$ from the leading coefficient. This also means that $\nu(a_n)$ could be different, so we need to compute the valuations of the coefficients of α_n. As the latter can be written as a quotient of ϑ_i, we naturally need to study quotients of Laurent series.

Indeed we may work with quotients of power series, as multiplying with powers of X only shifts coefficient indices. For convenience, we work in $K[\![Z]\!]$ (think $Z = X^{-1}$) to avoid negative indices.

As in Chapters 7 and 8, K is the fraction field of a discrete valuation ring O with maximal ideal $\mathfrak{m}$ and valuation ν.

Let $a_n, c_n \in O, b_n \in K$, and consider the Cauchy product

$$\left(\sum_{n=0}^{\infty} a_n\, Z^n\right)\left(\sum_{n=0}^{\infty} b_n\, Z^n\right) = \sum_{n=0}^{\infty} c_n\, Z^n.$$

For the coefficients, we get the relations

$$c_n = \sum_{i+j=n} a_i\, b_j$$

which we can recursively solve to b_n as

$$b_n = \frac{1}{a_0}\left(c_n - \sum_{\substack{i+j=n,\\ i\neq 0}} a_i\, b_j\right). \tag{A.1}$$

For the first couple of indices, we compute

$$\begin{aligned} b_0 &= \frac{c_0}{a_0} \\ b_1 &= \frac{1}{a_0^2}\,(a_0\, c_1 - a_1\, c_0) \\ b_2 &= \frac{1}{a_0^3}\,\left(a_1^2 c_0 - a_0 a_2 c_0 - a_0 a_1 c_1 + a_0^2 c_2\right) \end{aligned}$$

So we can try to calculate or estimate the valuations of the coefficient with these formulas.

The following Lemma addresses the simplest case (sufficient to treat $\deg D = 4$).

Lemma A.3.

- *Suppose $\nu(c_0) > 0$, but $\nu(c_1) = \nu(a_0) = 0$. Then $\nu(b_0) = \nu(c_0) > 0$ and $\nu(b_1) = 0$.*
- *Suppose $\nu(a_0) > 0$ and $\nu(c_0) = \nu(a_1) = 0$. Then $\nu(b_0) = -\nu(a_0) < 0$ and $\nu(b_1) = -2\,\nu(a_0) < 0$.*

Proof. The valuation of b_0 is obvious. In the first situation, we deduce from $\nu(c_0) > 0$ and $\nu(a_1) \geq 0$

$$\nu(b_1) = \nu(c_1\,a_0 - c_0\,a_1) = \min(0, \nu(c_0) + \nu(a_1)) = 0.$$

In the second situation, $\nu(c_1) \geq 0, \nu(a_0) > 0$ implies

$$\nu(b_1) = -2\,\nu(a_0) + \nu(c_1\,a_0 - c_0\,a_1) = -2\,\nu(a_0) + \min(\nu(c_1) + \nu(a_0), 0) = -2\,\nu(a_0).$$

□

We can actually generalise this somewhat, but first we need a better description of the formulas for the b_n:

Proposition A.4. *Define $B_n = -(-a_0)^{n+1}\,b_n$. Then we find*

$$B_n = \sum_{\substack{i_0+\cdots+i_l=n \\ 0\leq i_0\leq n, 1\leq i_1,\ldots,i_l\leq n}} c_{i_0}\,a_{i_1}\cdots a_{i_l}\,(-a_0)^{n-l}.$$

Essentially, we are summing over integer partitions of n with (at most) $n+1$ parts. However, except for from the parts which are 0, the ordering of the parts matters.

Proof. We prove this by induction. Clearly $B_0 = c_0$, precisely what the formula produces as no a_i appears in the sum.

For the induction step, we use the recursion formula (A.1)

$$\begin{aligned} B_n &= (-a_0)^n\,c_n + \sum_{\substack{i+j=n, \\ i\neq 0}} a_i\,(-a_0)^{i-1}\,B_j \\ &= (-a_0)^n\,c_n + \sum_{\substack{i+j=n, \\ i\neq 0}} a_i\,(-a_0)^{i-1} \sum_{\substack{i_0+\cdots+i_l=l \\ 0\leq i_0\leq j, 1\leq i_1,\ldots,i_l\leq j}} c_{i_0}\,a_{i_1}\cdots a_{i_l}\,(-a_0)^{j-l} \\ &= \sum_{\substack{i_0+\cdots+i_l+i=n \\ 0\leq i_0\leq n, 1\leq i_1,\ldots,i_l,i\leq n}} c_{i_0}\,a_{i_1}\cdots a_{i_l}\,a_i\,(-a_0)^{n-l-1}. \end{aligned}$$

Essentially, we are recursing by fixing the last (or first) a_i. □

We can now generalise the second part of Lemma A.3:

Proposition A.5. *If $c_0, a_1 \in O^\times$ and $a_0 \in \mathfrak{m}$, then for all $n \geq 0$ we have $B_n \in O^\times$. This implies $\nu(b_n) = -(n+1)\,\nu(a_0)$.*

Proof. It is clear that $B_n \in O$, as all the summands are in O (recall that $a_i, c_i \in O$). We show that precisely one summand lies in $O^\times$, while the other are in $\mathfrak{m}$.

Of course, with $i_0 = 0$ and $i_j = 1$ for the rest, we get $c_0\, a_1^n \in O^\times$.

For all other summands, we show that $l < n$ which implies that a_0 appears in $c_{i_0}\, a_{i_1} \cdots a_{i_l}\, (-a_0)^{n-l}$, so the product is in $\mathfrak{m}$.

If still $i_0 = 0$, but one of the $i_j \neq 1$, i.e. $i_j \geq 2$, then clearly $l < i_1 + \cdots + i_l = n$.

If on the other hand $i_0 > 0$, then immediately $l \leq i_1 + \cdots + i_l < n$. □

A.3. A lemma for a quadratic form

Let G a $\mathbb{Z}$-module (an abelian group) and $q : G \to \mathbb{R}$ a quadratic form. By abuse of notation, we also denote the corresponding $\mathbb{Z}$-bilinear form by $q : G \times G \to \mathbb{R}$.

Lemma A.6. *Suppose that q is positive (i.e. $q(g) \geq 0$ for all $g \in G$). Let $g_1, \ldots, g_r \in G$. Then*

$$q(g_1 + \cdots + g_r) \leq r \cdot (q(g_1) + \cdots + q(g_r)) \leq r^2 \max\{q(g_i) \mid i = 1, \ldots, r\}. \tag{A.2}$$

Proof. Because q is a quadratic form, we have

$$q(g_1 + \cdots + g_r) = \sum_{i=1}^{r} q(g_i) + 2 \sum_{1 \leq i < j \leq r} q(g_i, g_j).$$

Moreover q positive implies that

$$0 \leq q(g_i - g_j) = q(g_i) + q(g_j) - 2\,q(g_i, g_j)$$

so we deduce

$$q(g_1 + \cdots + g_r) \leq \sum_{i=1}^{r} q(g_i) + 2 \sum_{1 \leq i < j \leq r} q(g_i) + q(g_j) = \sum_{1 \leq i,j \leq r} q(g_i) = r \sum_{i=1}^{r} q(g_i).$$

The second inequality in (A.2) is then obvious. □

Bibliography

[Abe26] Niels Henrik Abel, *Über die Integration der Differential-Formel $\rho dx/\sqrt{r}$, wenn r und ρ ganze Functionen sind*, J. reine angew. Math. **1** (1826), 185–221.

[Aig07] Martin Aigner, *A course in enumeration*, Graduate Texts in Mathematics, vol. 238, Springer, Berlin, 2007. MR 2339282

[AR80] William W. Adams and Michael J. Razar, *Multiples of points on elliptic curves and continued fractions*, Proc. London Math. Soc. (3) **41** (1980), no. 3, 481–498. MR 591651 (82c:14031)

[BC97] Enrico Bombieri and Paula B. Cohen, *Siegel's lemma, Padé approximations and Jacobians*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **25** (1997), no. 1-2, 155–178 (1998), With an appendix by Umberto Zannier, Dedicated to Ennio De Giorgi. MR 1655513 (99k:11092)

[Ber90] T. G. Berry, *On periodicity of continued fractions in hyperelliptic function fields*, Arch. Math. (Basel) **55** (1990), no. 3, 259–266. MR 1075050 (91h:11049)

[BG06] Enrico Bombieri and Walter Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs, vol. 4, Cambridge University Press, Cambridge, 2006. MR 2216774 (2007a:11092)

[BKP07] V. V. Benyash-Krivets and V. P. Platonov, *Groups of S-units in hyperelliptic fields*, Dokl. Akad. Nauk **417** (2007), no. 4, 446–450. MR 2458904

[Cas57] J. W. S. Cassels, *An introduction to Diophantine approximation*, Cambridge Tracts in Mathematics and Mathematical Physics, No. 45, Cambridge University Press, New York, 1957. MR 0087708

[Che57] Pafnuty Lvovich Chebyshev, *Sur l'intégration des différentielles qui contiennent une racine carré d'un polynôme du troisième ou du quatrième degré.*, J. Math. Pures Appl. (2) **2** (1857), 1–42.

[CRS97] Capi Corrales-Rodrigáñez and René Schoof, *The support problem and its elliptic analogue*, J. Number Theory **64** (1997), no. 2, 276–290. MR 1453213

[Dav81] James Harold Davenport, *On the integration of algebraic functions*, Lecture Notes in Computer Science, vol. 102, Springer-Verlag, Berlin-New York, 1981. MR 617377 (84k:14024)

[GM90] Rajiv Gupta and M. Ram Murty, *Cyclicity and generation of points mod p on elliptic curves*, Invent. Math. **101** (1990), no. 1, 225–235. MR 1055716

[GR14] Éric Gaudron and Gaël Rémond, *Polarisations et isogénies*, Duke Math. J. **163** (2014), no. 11, 2057–2108. MR 3263028

[GW10] Ulrich Görtz and Torsten Wedhorn, *Algebraic geometry I*, Advanced Lectures in Mathematics, Vieweg + Teubner, Wiesbaden, 2010, Schemes with examples and exercises. MR 2675155 (2011f:14001)

[Har77] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York-Heidelberg, 1977, Graduate Texts in Mathematics, No. 52. MR 0463157 (57 #3116)

[Has36a] Helmut Hasse, *Zur Theorie der abstrakten elliptischen Funktionenkörper I. Die Struktur der Gruppe der Divisorenklassen endlicher Ordnung*, J. Reine Angew. Math. **175** (1936), 55–62. MR 1581496

[Has36b] ______, *Zur Theorie der abstrakten elliptischen Funktionenkörper II. Automorphismen und Meromorphismen. Das Additionstheorem*, J. Reine Angew. Math. **175** (1936), 69–88. MR 1581499

[Has36c] ______, *Zur Theorie der abstrakten elliptischen Funktionenkörper III. Die Struktur des Meromorphismenrings. Die Riemannsche Vermutung*, J. Reine Angew. Math. **175** (1936), 193–208. MR 1581508

[Haz97] Fumio Hazama, *Pell equations for polynomials*, Indag. Math. (N.S.) **8** (1997), no. 3, 387–397. MR 1622236

[HMPLR87] Y. Hellegouarch, D. L. McQuillan, and R. Paysant-Le Roux, *Unités de certains sous-anneaux des corps de fonctions algébriques*, Acta Arith. **48** (1987), no. 1, 9–47. MR 893459 (89g:11112)

[HS00] Marc Hindry and Joseph H. Silverman, *Diophantine geometry*, Graduate Texts in Mathematics, vol. 201, Springer-Verlag, New York, 2000, An introduction. MR 1745599 (2001e:11058)

[Khi56] A. Khintchine, *Kettenbrüche*, B. G. Teubner Verlagsgesellschaft, Leipzig, 1956. MR 0080630 (18,274f)

[Liu02] Qing Liu, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics, vol. 6, Oxford University Press, Oxford, 2002, Translated from the French by Reinie Erné, Oxford Science Publications. MR 1917232 (2003g:14001)

[Mir95] Rick Miranda, *Algebraic curves and Riemann surfaces*, Graduate Studies in Mathematics, vol. 5, American Mathematical Society, Providence, RI, 1995. MR 1326604 (96f:14029)

[MW14] David Masser and Gisbert Wüstholz, *Polarization estimates for abelian varieties*, Algebra Number Theory **8** (2014), no. 5, 1045–1070. MR 3263135

[MZ15] David Masser and Umberto Zannier, *Torsion points on families of simple abelian surfaces and Pell's equation over polynomial rings*, J. Eur. Math. Soc. (JEMS) **17** (2015), no. 9, 2379–2416, With an appendix by E. V. Flynn. MR 3420511

[Neu99] Jürgen Neukirch, *Algebraic number theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 322, Springer-Verlag, Berlin, 1999, Translated from the 1992 German original and with a note by Norbert Schappacher, With a foreword by G. Harder. MR 1697859

[Per54] Oskar Perron, *Die Lehre von den Kettenbrüchen. Bd I. Elementare Kettenbrüche*, B. G. Teubner Verlagsgesellschaft, Stuttgart, 1954, 3te Aufl. MR 0064172 (16,239e)

[Per57] ———, *Die Lehre von den Kettenbrüchen. Dritte, verbesserte und erweiterte Aufl. Bd. II. Analytisch-funktionentheoretische Kettenbrüche*, B. G. Teubner Verlagsgesellschaft, Stuttgart, 1957. MR 0085349 (19,25c)

[Pla14] V. P. Platonov, *Number-theoretic properties of hyperelliptic fields and the torsion problem in Jacobians of hyperelliptic curves over the rational number field*, Uspekhi Mat. Nauk **69** (2014), no. 1(415), 3–38. MR 3222877

[PP12] V. P. Platonov and M. M. Petrunin, *On the torsion problem in Jacobians of curves of genus 2 over the rational number field*, Dokl. Akad. Nauk **446** (2012), no. 3, 263–264. MR 3052248

[Sch60] A. Schinzel, *On the congruence* $a^x \equiv b \pmod{p}$, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. **8** (1960), 307–309. MR 0125070

[Sch00] Wolfgang M. Schmidt, *On continued fractions and Diophantine approximation in power series fields*, Acta Arith. **95** (2000), no. 2, 139–166. MR 1785412 (2001j:11063)

[Ser88] Jean-Pierre Serre, *Algebraic groups and class fields*, Graduate Texts in Mathematics, vol. 117, Springer-Verlag, New York, 1988, Translated from the French. MR 918564

[Sie14] Carl Ludwig Siegel, *On some applications of Diophantine approximations*, Quaderni/Monographs, vol. 2, Edizioni della Normale, Pisa, 2014, A translation of Carl Ludwig Siegel's "Über einige Anwendungen diophantischer Approximationen" by Clemens Fuchs, With a commentary and the article "Integral points on curves: Siegel's theorem after Siegel's proof" by Fuchs and Umberto Zannier, Edited by Zannier. MR 3309332

[ST68] Jean-Pierre Serre and John Tate, *Good reduction of abelian varieties*, Ann. of Math. (2) **88** (1968), 492–517. MR 0236190 (38 #4488)

[ST15] Joseph H. Silverman and John T. Tate, *Rational points on elliptic curves*, second ed., Undergraduate Texts in Mathematics, Springer, Cham, 2015. MR 3363545

[vdP98] Alfred J. van der Poorten, *Formal power series and their continued fraction expansion*, Algorithmic number theory (Portland, OR, 1998), Lecture Notes in Comput. Sci., vol. 1423, Springer, Berlin, 1998, pp. 358–371. MR 1726084 (2000m:11009)

[vdP99] ———, *Reduction of continued fractions of formal power series*, Continued fractions: from analytic number theory to constructive approximation (Columbia, MO, 1998), Contemp. Math., vol. 236, Amer. Math. Soc., Providence, RI, 1999, pp. 343–355. MR 1665378 (2000i:11111)

[vdP01] ———, *Non-periodic continued fractions in hyperelliptic function fields*, Bull. Austral. Math. Soc. **64** (2001), no. 2, 331–343. MR 1860070 (2002f:11087)

[vdPT00] Alfred J. van der Poorten and Xuan Chuong Tran, *Quasi-elliptic integrals and periodic continued fractions*, Monatsh. Math. **131** (2000), no. 2, 155–169. MR 1798560 (2002b:11093)

[Wei49] André Weil, *Numbers of solutions of equations in finite fields*, Bull. Amer. Math. Soc. **55** (1949), 497–508. MR 0029393

[Yu99] Jing Yu, *On arithmetic of hyperelliptic curves*, Aspect of Mathematics (1999), 1–21.

[Zan12] Umberto Zannier, *Some problems of unlikely intersections in arithmetic and geometry*, Annals of Mathematics Studies, vol. 181, Princeton University Press, Princeton, NJ, 2012, With appendixes by David Masser. MR 2918151

[Zan14] ———, *Trends in contemporary mathematics*, Springer INdAM Series, vol. 8, ch. Unlikely Intersections and Pell's Equations in Polynomials, pp. 151–169, Springer International Publishing, 2014.

[Zan16] ———, *Hyperelliptic Continued Fractions and Generalized Jacobians*, ArXiv e-prints (2016), 1–29.