
Hilbert's tenth problem and some generalizations

Esther Bod
January 23, 2009

10. ENTSCHEIDUNG DER LÖSBARKEIT EINER DIOPHANTISCHEN GLEICHUNG.

Eine diophantische Gleichung mit irgendwelchen Unbekannten und mit ganzen rationalen Zahlkoeffizienten sei vorgelegt: man soll ein Verfahren angeben, nach welchen sich mittels einer endlichen Anzahl von Operationen entscheiden läßt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist.

Abstract

Hilbert's tenth problem asks for an algorithm to determine the solvability in integers of diophantine equations over \mathbb{Z} . We prove that such an algorithm does not exist, and prove analogous statements for equations over polynomial rings, for equations over rings of integers of quadratic extensions of \mathbb{Q} and for equations over rational function fields defined over either formally real fields or finite fields. This requires a proof that the positive existential theories of several languages with divisibility relations are undecidable. We also try to prove that the diophantine theory of rational function fields over infinite fields of positive characteristic is undecidable, i.e. that no analogous algorithm exists. However, we are only able to show that the first order theory is undecidable.

Acknowledgements

First of all, I would like to thank my supervisor Gunther Cornelissen. You were always willing to help me and spend time thinking about the problems we came across. Your enthusiasm was inspiring and helped me to keep trying when I was stuck.

I am also grateful to Frits Beukers for being the second reader of my thesis. In the next four years, I will write my PhD-thesis under your supervision. I look forward to these years and a pleasant co-operation.

A word of gratitude is also due to Bjorn Poonen (MIT) and Thanases Pheidas (University of Crete) for their help in the last part of my research.

My fellow students have made the past few years much more enjoyable. Your company made the task of writing a thesis less lonely. In particular I would like to mention Lotte van der Zalm, Marte Koning, Amarins van de Voorde, Erik Leppen, Sander Wolters, Ralph van Gelderen, Willem Maat and Roeland Warringa. Now we all go our various ways, and I wish you all good luck. I will fondly remember these years and hope that we will stay in touch.

Joachim, of course I also want to thank you for the co-operation in the last few years, but above all, I thank you for your unconditional support. Your love and confidence in me helped me a lot. Especially in the last few months, it has been very important for me to have you on my side.

I would also like to thank my grandfather Marius and my grandmother Tonie. Your interest in me and my studies has been a great support.

Marianne, despite our different interests, we have lots of fun together. After a holiday or weekend together, I was always refreshed and ready to go on.

Finally, I would like to express my gratitude towards my parents. You always loved me, supported me and were interested in everything I did. That made me who I am now and gave me self-confidence. That has been very important for me, also in writing this thesis.

Contents

Abstract	iii
Acknowledgements	v
1 Introduction	1
1.1 Hilbert's tenth problem	1
1.2 Diophantine equations and sets	2
1.3 Positive existential sets and relations	3
2 Diophantine relations on \mathbb{N}	6
2.1 Reduction to solutions in natural numbers	6
2.2 Exponentiation is diophantine	7
2.3 Coding	11
2.4 Concatenation of positional codes	13
2.5 Functions on tuples	16
3 Undecidability of Hilbert's tenth problem	18
3.1 Universal equations	18
3.2 Turing machines	24
3.3 Recursively enumerable and diophantine sets	26
4 Languages with divisibility relations	31
4.1 Positive existential models	31
4.2 The theories of \mathbb{Z} in $(0, 1; +; ^n)$ and \mathbb{Z} in $(0, 1; +; , _p)$	32
4.3 The theory of \mathbb{N} in $(0, 1; +; _p)$	35
5 Polynomial rings	38
5.1 Rings of characteristic zero	38
5.2 Rings of odd characteristic	41
5.3 Rings of characteristic two	43
6 Quadratic rings	47
6.1 The Pell equation	47
6.2 Real rings	49
6.3 Imaginary rings	50
6.4 Proof of the negative solution to Hilbert's tenth problem	53
7 Function fields over finite fields	55
7.1 Defining the order of a rational function	55
7.2 Finite fields of odd characteristic	57
7.3 Finite fields of characteristic two	60

7.4	The undecidability of the diophantine theory of $(K(t); +, \cdot)$	64
8	Formal groups	66
8.1	Definition and basic properties	66
8.2	Multiplication in formal groups	69
8.3	Formal groups over valuation rings	70
8.4	The formal group of an elliptic curve	73
9	Function fields over formally real fields	77
9.1	The rational points on the twist of the elliptic curve	77
9.2	Proof of the negative solution to Hilbert's tenth problem	80
10	Divisibility sequences and exponentiation	85
10.1	Elliptic divisibility sequences	85
10.2	Exponentiation revisited (incomplete proof)	89
11	Function fields over fields of odd characteristic	92
11.1	Endomorphisms of an ordinary elliptic curve	92
11.2	A positive existential model of $(\mathbb{I}, L_{\mathbb{I}})$ in $(K(t), (0, 1, t; +, \cdot))$	94
11.3	A positive existential model of $(\mathbb{N}, (0, 1; +;))$ in $(\mathbb{I}, L_{\mathbb{I}})$	96
11.4	Defining multiplication in \mathbb{N} in $(0, 1; +;)$	99
11.5	Other approaches	101
11.6	Supersingular curves	104
A	Computing ranks using SAGE	108
	Bibliography	111
	List of tables	113
	List of notation	114
	List of diophantine relations	116
	Index	118

Chapter 1

Introduction

1.1 Hilbert's tenth problem

In this thesis, we will consider various problems based on Hilbert's tenth problem. Hilbert stated this problem on the Second International Congress of Mathematicians, held in Paris in 1900, as the tenth problem on a list of 23 problems. He considered these problems to be the most important mathematical problems that were to be solved in the twentieth century. The question of the tenth problem was the following:

10. ENTSCHEIDUNG DER LÖSBARKEIT EINER DIOPHANTISCHEN GLEICHUNG.

Eine diophantische Gleichung mit irgendwelchen Unbekannten und mit ganzen rationalen Zahlkoeffizienten sei vorgelegt: man soll ein Verfahren angeben, nach welchem sich mittels einer endlichen Anzahl von Operationen entscheiden läßt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist.¹

Notice that Hilbert did not ask whether such a process exists, but asks to give it, so he seems to have expected it to exist. Nowadays, we call such a process an algorithm, but in 1900, there was no mathematical definition of 'algorithm'. This doesn't matter for this problem: any algorithm that solves Hilbert's tenth problem would have been recognized as an algorithm, despite the lack of a formal definition.

During the 1930's, the notion of 'algorithm' was formalized by Gödel, Church and Turing. They showed that certain problems can not be solved by an algorithm. In 1970, Matiyasevich proved that deciding whether a diophantine equation has an integral solution is a problem that cannot be solved by an algorithm, so he gave a negative solution to Hilbert's tenth problem. In his proof, he used earlier work of Davis, Putnam and Robinson.

After Matiyasevich gave his proof of the unsolvability of Hilbert's tenth problem, it has been asked whether such an algorithm does exist if we take other coefficients and other possible solutions. For example, we can take equations with complex coefficients and ask whether they have a solution in the complex numbers. In this case, the algorithm to determine this is easy: since \mathbb{C} is algebraically closed, every equation of degree at least 1 has a solution. If we consider equations over finite fields, and ask for solutions in the field, we can just try them all. Hence in this case the algorithm also exists.

In many other cases, the situation is different. As we will prove in this thesis, there is no algorithm to decide the solvability in polynomial rings of equations with coefficients in $\mathbb{Z}[t]$. The same holds for some rational function fields.

¹**10. DETERMINATION OF THE SOLVABILITY OF A DIOPHANTINE EQUATION.**

Given a diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: to devise a process according to which it can be determined by a finite number of operations whether the equation is solvable in rational integers

There are also many open problems. For example, the answer to Hilbert's tenth problem is not known if the solutions must be rational instead of integral. It is also not known whether it is algorithmically decidable whether an equation with coefficients in $\mathbb{Z}[t]$ has a solution in $\mathbb{C}(t)$. This is quite surprising, since the answer is known if we take $\mathbb{R}(t)$ or $K(t)$ (for any finite field K) instead of $\mathbb{C}(t)$, or take equations with coefficients and solutions in $\mathbb{Z}[t, u]$. In all these cases, such an algorithm doesn't exist.

We will not try to solve these problems in this thesis, but we will try to prove that there exists no algorithm to decide whether an equations with coefficients in $\mathbb{Z}[t]$ has solutions in $K(t)$, if K is an infinite field of positive characteristic.

In the first three chapters, we will give a proof of the undecidability of Hilbert's tenth problem based on the proof of Matiyasevich. After that, we will consider whether there exists an algorithm to determine the solvability of equations with other coefficients than elements of \mathbb{Z} . To this end, we will always use a reduction to the original problem over \mathbb{Z} . Up to chapter 7, we will use elementary methods. In the last chapters, the methods of proof will involve elliptic curves.

1.2 Diophantine equations and sets

Hilbert's tenth problem asks for a decision procedure for diophantine equations with coefficients and solutions in \mathbb{Z} . By a diophantine equation, Hilbert means an equation of the form

$$P(x_1, \dots, x_m) = 0$$

with $P \in \mathbb{Z}[x_1, \dots, x_m]$.

Since we will also consider variants on Hilbert's tenth problem, we will define diophantine equations more generally.

Definition 1.1. A language L consists of a set of constants, a set of function symbols and a set of relation symbols. A term in L is recursively defined to be a constant or a variable, usually denoted $x, x_1, x_2, \dots, y, z, \dots$, or a function symbol applied to terms. A *diophantine equation* in L is an equality between two terms.

Definition 1.2. Let a language L and a set S be given, such that the set of constants of L is a subset of S and the function and relation symbols of L can be applied to elements of S . A *diophantine formula over S in L* is a formula of the form $\exists x_1, \dots, x_n \in S : P(x_1, \dots, x_n, y_1, \dots, y_m)$ where P is a diophantine equation in L . Sometimes we will omit the condition $x_1, \dots, x_n \in S$ from the notation, and simply write $\exists x_1, \dots, x_n : P(x_1, \dots, x_n, y_1, \dots, y_m)$. A *diophantine sentence* is a diophantine formula in which all variables are bound by quantifiers, i.e. $m = 0$. The *diophantine theory* of S in L is the set

$$\{\phi \mid \phi \text{ diophantine sentence in } L; \phi \text{ is true}\}$$

The theory is *decidable* if exists an algorithm that takes as input a diophantine sentence and decides whether it is an element of the theory.

Example 1.3. Let R be a commutative ring, and consider the language $L = (R; +, \cdot)$. A diophantine equation in L is of the form $P(x_1, \dots, x_m) = 0$, where P is a polynomial with coefficients in R . Notice that the definition over \mathbb{Z} , as given above, is a special case of this.

Definition 1.4. Let R be a ring, and let S be a subring. If S is not mentioned explicitly, we will assume that $S = R$. *Hilbert's tenth problem over R with coefficients in S* asks to give an algorithm that, when given a polynomial $f \in S[x_1, \dots, x_n]$ as input, outputs YES or NO, depending on whether f has a root in R^n . Equivalently, it asks for a decision procedure for the diophantine theory of R in $L = (S; +, \cdot)$. To make it possible for such an algorithm to exist, it must be possible to decide of every element of R whether or not it an element of S (in a finite amount of time). Hilbert's tenth problem is *undecidable* if the algorithm it asks for doesn't exist.

Definition 1.5. Let S and L be as in definition 1.2. A subset of S^n is *diophantine over S in L* if it is of the form $\{(a_1, \dots, a_n) \in S^n \mid \phi(a_1, \dots, a_n)\}$ where $\phi(a_1, \dots, a_n)$ is a diophantine formula over S in L with free variables a_1, \dots, a_n . ϕ is a *diophantine representation* of this set.

Notice that a diophantine representation of a diophantine set is not necessarily unique.

Definition 1.6. Let S and L be as in definition 1.2. A relation \mathcal{R} on S^n is *diophantine* if $\{(a_1, \dots, a_n) \in S^n \mid \mathcal{R}(a_1, \dots, a_n)\}$ is a diophantine subset of S^n . The representation of this set is also called the representation of \mathcal{R} . A function is diophantine if its graph is diophantine.

For subsets of the natural numbers² in the language $L = (0, 1; +, \cdot)$, we have another characterization of diophantine sets:

Lemma 1.7. *A set \mathcal{M} of natural numbers is diophantine in $(0, 1; +, \cdot)$ if and only if there exists a polynomial $P \in \mathbb{Z}[X_0, X_1, \dots, X_m]$ such that*

$$\mathcal{M} = \{P(x_0, x_1, \dots, x_m) \mid x_0, x_1, \dots, x_m \in \mathbb{N}, P(x_0, x_1, \dots, x_m) \in \mathbb{N}\}$$

Proof. Assume that \mathcal{M} is diophantine with representation $\exists x_1, \dots, x_m : D(a, x_1, \dots, x_m) = 0$. Define P by

$$P(x_0, x_1, \dots, x_m) = (x_0 + 1)(1 - D(x_0, x_1, \dots, x_m)^2) - 1$$

If $a = P(x_0, x_1, \dots, x_m)$, then $a + 1 = (x_0 + 1)(1 - D(x_0, x_1, \dots, x_m)^2)$, which is only possible if $x_0 + 1 = a + 1$ and $1 - D(x_0, x_1, \dots, x_m)^2 = 1$. Hence $D(a, x_1, \dots, x_m) = 0$, so $a \in \mathcal{M}$. On the other hand, if $a \in \mathcal{M}$, then there exist x_1, \dots, x_m such that $D(a, x_1, \dots, x_m) = 0$. Hence $P(a, x_1, \dots, x_m) = (a + 1)(1 - 0) - 1 = a$. This shows that if \mathcal{M} is diophantine, then such a polynomial exists.

Conversely, suppose that

$$\mathcal{M} = \{P(x_0, x_1, \dots, x_m) \mid x_0, x_1, \dots, x_m \in \mathbb{N}, P(x_0, x_1, \dots, x_m) \in \mathbb{N}\}$$

Then $\mathcal{M} = \{a \mid \exists x_0, \dots, x_m : a - P(x_0, \dots, x_m) = 0\}$, so \mathcal{M} is diophantine with representation $D(a, x_0, x_1, \dots, x_m) = a - P(x_0, \dots, x_m)$. \square

1.3 Positive existential sets and relations

We defined a diophantine set to be the set of parameters for which a diophantine equation has a solution. By taking unions and intersection of diophantine sets, we get a system of equations instead of a single equation. The goal of this section is to give a criterion to decide whether the union and intersection of diophantine sets are still diophantine.

Definition 1.8. A *positive existential formula* is inductively defined to be an equality between terms, a relation symbol applied to terms, a conjunction of two positive existential formulas (by using the connective \wedge), a disjunction between two positive existential formulas (by using the connective \vee) or a formula of the form $\exists x : \phi(x)$, where ϕ is a positive existential formula. The definitions of *positive existential sentences* and the *positive existential theory* are similar to definition 1.2, with ‘diophantine’ replaced by ‘positive existential’.

Example 1.9. Consider the language $L = (0, 1; +; |^n)$ (the meaning of $|^n$ will be defined in definition 4.3, but for now, we can just consider it to be a relation between two terms). A positive existential formula over \mathbb{Z} in L is of the form

$$\exists x_1, \dots, x_m \in \mathbb{Z} : \bigwedge_{i=1}^s \bigvee_{j=1}^{t_i} \phi_{i,j}(x_1, \dots, x_m)$$

where $\phi(i, j)$ is either of the form $F_{i,j}(x_1, \dots, x_m) = G_{i,j}(x_1, \dots, x_m)$, with $F_{i,j}$ and $G_{i,j}$ constant or linear polynomials over \mathbb{Z} , or of the form $F_{i,j}(x_1, \dots, x_m) |^n G_{i,j}(x_1, \dots, x_m)$.

²We use the conventions $0 \in \mathbb{N}$ and $0|0$.

Definition 1.10. As in definition 1.5, we define a subset of S^n to be *positive existential* if it is of the form $\{(a_1, \dots, a_n) \in S^n \mid \phi(a_1, \dots, a_n)\}$ where $\phi(x_1, \dots, x_n)$ is a positive existential formula over S in the language L with free variables x_1, \dots, x_n .

Remark 1.11. The union and intersection of two positive existential subsets of R^n are again positive existential, since we can combine the defining sets of equations using \vee and \wedge , respectively.

Lemma 1.12. *Let R be a ring, and let L be a language $(S; +, \cdot)$ with $S \subseteq R$. Let S' be the ring generated by the elements of S .*

If there are polynomials $f(x, y), g(x, y) \in S'[x, y]$ such that for all $a, b \in R$

$$f(a, b) = 0 \iff a = 0 \text{ and } b = 0$$

$$g(a, b) = 0 \iff a = 0 \text{ or } b = 0$$

then a subset of R^n is diophantine over R in L if and only if it is positive existential over R in L .

Proof. By definition, all diophantine sets are positive existential.

Let a positive existential subset of R^n be given, with corresponding formula $\exists x_1, \dots, x_m : \phi$. Then ϕ is a combination of polynomial equations, and the equations are combined using \vee and \wedge . First we use reduce all equations to the form $p = 0$, by changing $r = s$ into $r - s = 0$. A conjunction of equations $p = 0 \wedge q = 0$ has a solution if and only if $f(p, q) = 0$ has a solution, and a disjunction $p = 0 \vee q = 0$ has a solution if and only if $g(p, q) = 0$ has a solution. By repeatedly replacing conjunctions and disjunction by expressions with f and g , we can transform ϕ into a single polynomial equation with coefficients in S' . It follows that ϕ is equivalent to a diophantine equation in L . Hence $\exists x_1, \dots, x_m : \phi$ is equivalent to a diophantine formula, so the subset given by $\exists x_1, \dots, x_m : \phi$ is diophantine. \square

Theorem 1.13. *Let R be an integral domain, with field of fractions K and let $L = (S; +, \cdot)$ be a language with $S \subseteq R$. Let S' be the ring generated by the elements of S , and let $Q(S')$ be the field of fractions of S' . Suppose that K is not algebraically closed and has an extension $K(\alpha)$ such that the norm map $N_{K(\alpha)/K}$ is given by a polynomial over $Q(S')$. Then a subset of R^n is diophantine over R in the language $(S; +, \cdot)$ if and only if it is positive existential over R in this language.*

Proof. We apply lemma 1.12. Since R is an integral domain, $ab = 0$ implies that $a = 0$ or $b = 0$. Hence we can take $g(x, y) = xy \in S'[x, y]$.

There exists $r \in S'$ such that $rN_{K(\alpha)/K}(x + \alpha y)$ is given by a polynomial with coefficients in S' . Notice that

$$rN_{K(\alpha)/K}(x + \alpha y) = 0 \iff x + \alpha y = 0 \iff x = 0 \text{ and } y = 0$$

since 1 and α are independent over K . Hence we can take $f(x, y) = rN_{K(\alpha)/K}(x + \alpha y) \in S'[x, y]$. \square

Lemma 1.14. *All positive existential sets over \mathbb{Z} or \mathbb{N} in $(0, 1; +, \cdot)$ are diophantine.*

Proof. Apply theorem 1.13 to the set \mathbb{Z} and the language $(0, 1; +, \cdot)$. $\mathbb{Q}(i)$ is a non-trivial extension of the field of fractions \mathbb{Q} , and the norm map is $N(x + iy) = x^2 + y^2$. Hence if \mathcal{M}_1 and \mathcal{M}_2 are diophantine sets and $\exists x_1, x_2, \dots, x_{m_i} : D_i(a_1, \dots, a_n, x_1, \dots, x_{m_i}) = 0$ is a representation of \mathcal{M}_i , then representations of $\mathcal{M}_1 \cup \mathcal{M}_2$ and $\mathcal{M}_1 \cap \mathcal{M}_2$ are given by

$$\exists x_1, x_2, \dots, x_{m_1}, \exists y_1, y_2, \dots, y_{m_2} : D_1(a_1, \dots, a_n, x_1, \dots, x_{m_1}) \cdot D_2(a_1, \dots, a_n, y_1, \dots, y_{m_2}) = 0$$

and

$$\exists x_1, x_2, \dots, x_{m_1}, \exists y_1, y_2, \dots, y_{m_2} : (D_1(a_1, \dots, a_n, x_1, \dots, x_{m_1}))^2 + (D_2(a_1, \dots, a_n, y_1, \dots, y_{m_2}))^2 = 0$$

respectively.

The same representations of $\mathcal{M}_1 \cup \mathcal{M}_2$ and $\mathcal{M}_1 \cap \mathcal{M}_2$ show that positive existential sets over \mathbb{N} are diophantine over \mathbb{N} . \square

Lemma 1.15. *Let R be an integral domain. Then all positive existential sets over $R[t]$ are diophantine over $R[t]$ in the language $(0, 1, t; +, \cdot)$. The same holds for $R(t)$.*

Proof. By theorem 1.13, it suffices to show that $R(t)$ has a non-trivial extension $R(t)(\alpha)$ such that the norm map $N_{R(t)(\alpha)/R(t)}$ is given by a polynomial over $\mathbb{Z}(t)$. Take $\alpha = \sqrt{t}$. Then $\alpha \notin R(t)$, since if $\alpha = \frac{p(t)}{q(t)}$ with $p, q \in R[t]$, then $p(t)^2 = tq(t)^2$, but $p(t)^2$ has even degree, while $tq(t)^2$ has odd degree.

The norm map is

$$N_{R(\sqrt{t})/R(t)}(a + b\sqrt{t}) = \det \begin{pmatrix} a & b \\ bt & a \end{pmatrix} = a^2 - b^2t$$

Hence all positive existential sets over $R[t]$ are diophantine over $R[t]$, and all positive existential sets over $R(t)$ are diophantine over $R(t)$. \square

Chapter 2

Diophantine relations on \mathbb{N}

In the next chapter we will prove that Hilbert's tenth problem, as stated originally, is undecidable. To prove this, we will need many diophantine relations. In this chapter, we will consider some of these relations, and prove that there are diophantine. The most difficult step, and historically the last step in the proof of the undecidability of Hilbert's tenth problem, is to show that exponentiation is a diophantine relation.

To prove that Hilbert's tenth problem is undecidable, we will need to encode tuples of numbers as one number. Therefore, we will also introduce two codings, and show that one of these codings is diophantine, as well as concatenating tuples in this coding.

2.1 Reduction to solutions in natural numbers

It will be easier to work with natural numbers (including 0) than with integers. To show that this suffices, we need the following lemma:

Lemma 2.1. *Hilbert's tenth problem over \mathbb{Z} in the language $L = (0, 1; +, \cdot)$ is equivalent to Hilbert's tenth problem over \mathbb{N} in L .*

Proof. Suppose that we have an algorithm as in Hilbert's tenth problem, and let $P(x_1, \dots, x_m) = Q(x_1, \dots, x_m)$ be a diophantine equation in which P and Q have natural numbers as coefficients. Consider the following system of equations:

$$\begin{aligned} P(x_1, \dots, x_m) &= Q(x_1, \dots, x_m) \\ x_1 &= y_{1,1}^2 + y_{1,2}^2 + y_{1,3}^2 + y_{1,4}^2 \\ &\vdots \\ x_m &= y_{m,1}^2 + y_{m,2}^2 + y_{m,3}^2 + y_{m,4}^2 \end{aligned}$$

By Lagrange's theorem, every natural number can be written as the sum of four squares, so the last m equations have a solution $(y_{1,1}, \dots, y_{m,4})$, if and only if x_1, \dots, x_m are non-negative. Hence this system has a solution in integers if and only if $P(x_1, \dots, x_m) = Q(x_1, \dots, x_m)$ has a solution in natural numbers. By lemma 1.14, we can reduce the system to a single equation. Since we assumed there is a decision procedure to decide whether this equation has a solution in integers, then we can also decide whether $P(x_1, \dots, x_m) = Q(x_1, \dots, x_m)$ has a solution in natural numbers.

Conversely, suppose we have an algorithm to decide whether an equation has a solution in natural numbers. Take any diophantine sentence over \mathbb{Z} in L , and replace the variables x_i by $y_i - z_i$. This gives an equivalent diophantine sentence over \mathbb{N} in L , of which we can decide whether it is true. Hence we have an algorithm to decide whether an equation has a solution in integers. \square

By this lemma, we can restrict ourselves to using only natural numbers as variables. In the remaining of this chapter and the next chapter, variables will always run over the natural numbers,

unless stated otherwise. The language will always be $(0, 1; +, \cdot)$.

In the table below, we give some important diophantine relations with a representation. $\text{rem}(b, c)$ is the function that gives the remainder on dividing b by c , taken between 0 and $c - 1$, and $\text{arem}(b, c)$ gives the absolute value of the remainder on dividing b by c , taken between 0 and $c/2$.

$a \leq b$	$\exists x : a + x = b$
$a < b$	$\exists x : a + x + 1 = b$
$a b$	$\exists x : ax = b$
$a = \text{rem}(b, c)$	$a < c \wedge c b - a$
$a = \text{arem}(b, c)$	$2a < c \wedge (c b - 1 \vee c b + 1)$
$a \nmid b$	$\text{rem}(b, a) > 0$
$a = \lfloor \frac{b}{c} \rfloor$	$ac + \text{rem}(b, c) = b$
$a \equiv b \pmod{c}$	$\text{rem}(a, c) = \text{rem}(b, c)$
$a = \text{gcd}(b, c)$	$(b > 0 \wedge c > 0 \wedge \exists x, y : b x \wedge c y \wedge (a = x - y \vee a = y - x)) \vee (b = 0 \wedge c > 0 \wedge a = c) \vee (b > 0 \wedge c = 0 \wedge a = b)$
$a = \max(b, c)$	$(a = b \wedge b > c) \vee (a = c \wedge c \geq b)$

Table 2.1: Some examples of diophantine relations

2.2 Exponentiation is diophantine

In this section, we will show that exponentiation is diophantine:

Theorem 2.2. *The sets $\{(a, b, c) \mid a, b, c \in \mathbb{N}, a = b^c\}$ and $\{(a, b) \mid a, b \in \mathbb{N}, \exists n \in \mathbb{N} : a = b^n\}$ are diophantine.*

Remark 2.3. We define $0^0 = 1$.

After proving this theorem, we can use exponentiation in definitions of diophantine relations. This will allow us to prove us that certain relations, such as being a prime number, are diophantine.

We prove theorem 2.2 by use of a recurrence relation:

Definition 2.4. Let $b \geq 2$. Define $\alpha_b(n)$ by

$$\alpha_b(n) = \begin{cases} 0 & \text{if } n = 0 \\ 1 & \text{if } n = 1 \\ b\alpha_b(n-1) - \alpha_b(n-2) & \text{if } n \geq 2 \end{cases} \quad (2.1)$$

Remark 2.5. Notice that if $\alpha_b(n+1) > \alpha_b(n)$, then $\alpha_b(n+2) = b\alpha_b(n+1) - \alpha_b(n) > (b-1)\alpha_b(n+1) \geq \alpha_b(n+1)$. Hence by induction, the sequence $(\alpha_b(n))_n$ is strictly increasing, so in particular, $\alpha_b(n) \geq n$.

Definition 2.6. Define the matrices $A_b(n)$ and Ξ_b by

$$A_b(n) = \begin{pmatrix} \alpha_b(n+1) & -\alpha_b(n) \\ \alpha_b(n) & -\alpha_b(n-1) \end{pmatrix} \quad \Xi_b = \begin{pmatrix} b & -1 \\ 1 & 0 \end{pmatrix}$$

Lemma 2.7. *If we define $\alpha_b(-1) = -1$, then $A_b(n) = \Xi_b^n$ for all $n \in \mathbb{N}$.*

Proof. It is clear that $A_b(0) = I$ and $A_b(n+1) = A_b(n)\Xi_b$. By induction, it follows immediately that $A_b(n) = \Xi_b^n$. \square

First we show that $\{(a, b) \mid a, b \in \mathbb{N}, b \geq 2 \wedge \exists n : a = \alpha_b(n)\}$ is diophantine. After that, we will show that $\{(a, b, c) \mid a, b, c \in \mathbb{N}, b \geq 4 \wedge a = \alpha_b(c)\}$ is diophantine. Finally, we show that $\{(a, b, c) \mid a, b, c \in \mathbb{N}, a = b^c\}$ is diophantine.

Lemma 2.8. $\mathcal{S}_1 := \{(a, b) \mid a, b \in \mathbb{N}, b \geq 2 \wedge \exists n \in \mathbb{N} : a = \alpha_b(n)\}$ is diophantine, since it equals

$$\mathcal{S}_2 := \{(a, b) \mid a, b \in \mathbb{N}, b \geq 2 \wedge \exists x \in \mathbb{N} : x^2 - abx + a^2 = 1\}$$

Furthermore, if $a = \alpha_b(n)$, then $x^2 - abx + a^2 = 1$ holds if and only if $x = \alpha_b(n \pm 1)$

Proof. First we prove that $\mathcal{S}_1 \subseteq \mathcal{S}_2$. Since $A_b(n) = \Xi_b^n$ by lemma 2.7, $\det(A_b(n)) = 1$, so

$$\begin{aligned} \alpha_b(n-1)^2 - b\alpha_b(n-1)\alpha_b(n) + \alpha_b(n)^2 &= \\ \alpha_b(n)^2 - \alpha_b(n-1)(b\alpha_b(n) - \alpha_b(n-1)) &= \\ \alpha_b(n)^2 - \alpha_b(n-1)\alpha_b(n+1) &= 1 \end{aligned}$$

So $\alpha_b(n-1)^2 - b\alpha_b(n-1)\alpha_b(n) + \alpha_b(n)^2 = 1$, and hence, if $(a, b) \in \mathcal{S}_1$, then $a = \alpha_b(n)$ for some n , and taking $x = \alpha_b(n-1)$ shows that $(a, b) \in \mathcal{S}_2$.

Now we show the converse inclusion. Let $x^2 - abx + a^2 = 1$, and assume $a < x$. Then we will show by induction on a that there exists n such that $x = \alpha_b(n+1)$ and $a = \alpha_b(n)$. If $a = 0$, then $x = 1$ and $n = 0$ satisfies the claim. Now suppose that $a > 0$. Then

$$ba - a \leq ba - \frac{a^2}{x} < ba + \frac{1}{x} - \frac{a^2}{x} = x = \frac{1 + bxa - a^2}{x} = ba + \frac{1 - a^2}{x} < ba$$

so $ba - a < x < ba$. Define $x' = a$ and $a' = ba - x$. Then $0 < a' \leq a = x'$ and

$$x'^2 - bx'a' + a'^2 = a^2 - ba(ba - x) + (ba - x)^2 = x^2 - bxa - a^2 = 1$$

so we can apply the induction hypothesis: there exists n such that $x' = \alpha_b(n+1)$ and $a' = \alpha_b(n)$. But now $x = bx' - a' = b\alpha_b(n+1) - \alpha_b(n) = \alpha_b(n+2)$ and $a = x' = \alpha_b(n+1)$. Hence $n+1$ satisfies the claim.

If $a > x$, then a similar reasoning shows that $x = \alpha_b(n)$ and $a = \alpha_b(n+1)$. In both cases we get that $(a, b) \in \mathcal{S}_1$. Notice that $a = x$ cannot occur, since this would give $(2-b)a^2 = 1$, but $b \geq 2$, so $2-b \leq 0$.

The second statement follows from the above and the fact that a quadratic equation has at most 2 solutions. \square

The next step is proving that the following set is diophantine:

Theorem 2.9. $\mathcal{S}_3 = \{(a, b, c) \mid a, b, c \in \mathbb{N}, b \geq 4 \wedge a = \alpha_b(c)\}$ is diophantine.

Below, we will give the system of equations that defines this set explicitly. We start by proving three lemmas, the first of which is trivial:

Lemma 2.10. $\alpha_2(n) = n$ and if $b_1 \equiv b_2 \pmod{q}$, then $\alpha_{b_1}(n) \equiv \alpha_{b_2}(n) \pmod{q}$. In particular, $\alpha_b(n) \equiv n \pmod{b-2}$.

Proof. Definition 2.1 gives that $\alpha_2(n+2) = 2\alpha_2(n+1) - \alpha_2(n)$. By induction, it easily follows that $\alpha_2(n) = n$.

If $b_1 \equiv b_2 \pmod{q}$, then it also follows by induction that $\alpha_{b_1}(n) \equiv \alpha_{b_2}(n) \pmod{q}$. In particular, taking $b_1 = b$, $b_2 = 2$ and $q = b-2$, this implies that $\alpha_b(n) \equiv n \pmod{b-2}$. \square

Lemma 2.11. Suppose that $\alpha_b(k) \mid \alpha_b(m)$. Then $k \mid m$.

Proof. Write $m = lk + n$, with $0 \leq n < k$. Then

$$\begin{pmatrix} \alpha_b(m+1) & -\alpha_b(m) \\ \alpha_b(m) & -\alpha_b(m-1) \end{pmatrix} = A_b(m) = \Xi_b^m = \Xi_b^n \cdot (\Xi_b^k)^l = \\ A_b(n)A_n(k)^l = \begin{pmatrix} \alpha_b(n+1) & -\alpha_b(n) \\ \alpha_b(n) & -\alpha_b(n-1) \end{pmatrix} \begin{pmatrix} \alpha_b(k+1) & -\alpha_b(k) \\ \alpha_b(k) & -\alpha_b(k-1) \end{pmatrix}^l$$

Modulo $\alpha_b(k)$, this reduces to

$$\begin{pmatrix} \alpha_b(m+1) & -\alpha_b(m) \\ \alpha_b(m) & -\alpha_b(m-1) \end{pmatrix} \equiv \begin{pmatrix} \alpha_b(n+1) & -\alpha_b(n) \\ \alpha_b(n) & -\alpha_b(n-1) \end{pmatrix} \begin{pmatrix} \alpha_b(k+1)^l & 0 \\ 0 & (-\alpha_b(k-1))^l \end{pmatrix}$$

Hence $\alpha_b(m) \equiv \alpha_b(n)\alpha_b^l(k+1) \pmod{\alpha_b(k)}$. Since $\alpha_b(k)^2 - \alpha_b(k-1)\alpha_b(k+1) = \det(A_b(k)) = 1$, $\alpha_b(k)$ and $\alpha_b(k+1)$ are coprime. Since $\alpha_b(k)$ divides $\alpha_b(m)$, it also divides $\alpha_b(n)\alpha_b^l(k+1)$, so $\alpha_b(k)|\alpha_b(n)$. But since $(\alpha_b(n))_{n \in \mathbb{N}}$ is strictly increasing (remark 2.5) and the fact that $n < k$, we have that $\alpha_b(n) < \alpha_b(k)$. Hence we must have $\alpha_b(n) = 0$, so $n = 0$. This implies that $m = lk$, so $k|m$. \square

Lemma 2.12. *Suppose that $\alpha_b^2(k)|\alpha_b(m)$. Then $\alpha_b(k)|\frac{m}{k}$, so in particular $\alpha_b(k)|m$.*

Proof. From lemma 2.11, we get $k|m$. Write $m = lk$.

We compute $\alpha_b(m)$ modulo $\alpha_b(k)^2$:

$$A_b(m) = A_b^l(k) = \begin{pmatrix} \alpha_b(k+1) & -\alpha_b(k) \\ \alpha_b(k) & -\alpha_b(k-1) \end{pmatrix}^l = \begin{pmatrix} b\alpha_b(k) - \alpha_b(k-1) & -\alpha_b(k) \\ \alpha_b(k) & \alpha_b(k-1) \end{pmatrix}^l = \\ (\alpha_b(k)\Xi_b - \alpha_b(k-1)I)^l = \sum_{i=0}^l (-1)^{l-i} \binom{l}{i} \alpha_b(k)^i \alpha_b(k-1)^{l-i} \Xi_b^i$$

All terms, except for the first and the second, are divisible by $\alpha_b(k)^2$. Hence

$$\begin{pmatrix} \alpha_b(m+1) & -\alpha_b(m) \\ \alpha_b(m) & -\alpha_b(m-1) \end{pmatrix} = A_b(m) \equiv (-1)^l \alpha_b^l(k-1)I + (-1)^{l-1} l \alpha_b(k) \alpha_b^{l-1}(k-1) \Xi_b$$

so

$$\alpha_b(m) \equiv (-1)^{l-1} l \alpha_b(k) \alpha_b^{l-1}(k-1) \pmod{\alpha_b^2(k)} \quad (2.2)$$

$\alpha_b^2(k)$ divides $\alpha_b(m)$, so it divides $l \alpha_b(k) \alpha_b^{l-1}(k-1)$. Hence $\alpha_b(k)$ divides $l \alpha_b^{l-1}(k-1)$. Since $\alpha_b(k)$ and $\alpha_b(k-1)$ are coprime, it follows that $\alpha_b(k)|l$. \square

Theorem 2.13. *$(a, b, c) \in \mathcal{S}_3$ if and only if the following system of equations has a solution (r, s, t, u, v, w, x, y) :*

$$\begin{array}{lll} b \geq 4 & u^2 - but + t^2 = 1 & v|w - b \\ 2a < u & s^2 - bsr + r^2 = 1 & u|w - 2 \\ r < s & x^2 - wxy + y^2 = 1 & u^2|s \\ w > 2 & \text{arem}(x, v) = a & \\ v = bs - 2r & \text{arem}(x, u) = c & \end{array}$$

Proof. Suppose that this system has a solution. Then we have to show that $a = \alpha_b(c)$. By lemma 2.8 and the fact the $(\alpha_b(n))_{n \in \mathbb{N}}$ is increasing (remark 2.5), it follows that there exist k, m and n such that $u = \alpha_b(k)$, $r = \alpha_b(m-1)$, $s = \alpha_b(m)$ and $x = \alpha_w(n)$ (notice the subscript). $u^2|s$, so $\alpha_b(k)^2|\alpha_b(m)$. By lemma 2.12, it follows that $\alpha_b(k)|m$, i.e. $u|m$. Since $v = bs - 2r$, we have $v = b\alpha_b(m) - 2\alpha_b(m-1) = \alpha_b(m+1) - \alpha_b(m-1)$. Since $v|w - b$ and $u|w - 2$, it follows from lemma

2.10 that $x = \alpha_w(n) \equiv \alpha_b(n) \pmod{v}$ and $x = \alpha_w(n) \equiv n \pmod{u}$. Now write $n = 2lm \pm j$ with $0 \leq j \leq m$. Then

$$A_b(n) = ((\Xi_b^m)^2)^l \Xi^{\pm j} = ((A_b(m))^2)^l (A_b(j))^{\pm 1}$$

and modulo v

$$\begin{aligned} A_b(m) &= \begin{pmatrix} \alpha_b(m+1) & -\alpha_b(m) \\ \alpha_b(m) & -\alpha_b(m-1) \end{pmatrix} \equiv \begin{pmatrix} \alpha_b(m-1) & -\alpha_b(m) \\ \alpha_b(m) & -\alpha_b(m+1) \end{pmatrix} = \\ &= - \begin{pmatrix} -\alpha_b(m-1) & \alpha_b(m) \\ -\alpha_b(m) & \alpha_b(m+1) \end{pmatrix} = -(A_b(m))^{-1} \end{aligned}$$

This implies that $(A_b(m))^2 \equiv -I$, so $A_b(n) \equiv \pm(A_b(j))^{\pm 1} \pmod{v}$. Since $x \equiv \alpha_b(n) \pmod{v}$ we have that $x \equiv \pm\alpha_b(j) \pmod{v}$. Since the sequence $(\alpha_b(n))_n$ is strictly increasing, $j \leq m$ and $b \geq 4$, we have that

$$2\alpha_b(j) \leq 2\alpha_b(m) \leq (b-2)\alpha_b(m) < b\alpha_b(m) - 2\alpha_b(m-1) = v$$

so $a = \text{arem}(x, v) = \text{arem}(\alpha_b(n), v) = \text{arem}(\pm\alpha_b(j), v) = \alpha_b(j)$. Hence $2j \leq 2\alpha_b(j) = 2a < u$. Using $u|m$ and $x \equiv n \pmod{u}$, we also have $c = \text{arem}(x, u) = \text{arem}(n, u) = \text{arem}(2lm \pm j, u) = \text{arem}(\pm j, u) = j$. Hence $\alpha_b(c) = \alpha_b(j) = a$, as was to be shown.

For the converse, we assume that $(a, b, c) \in \mathcal{S}_3$, i.e. $b \geq 4$ and $a = \alpha_b(c)$. We are going to construct a solution to the system of equations. Since $\alpha_b(n)$ and $\alpha_b(n+1)$ are coprime for all n , at least one of them is odd. Furthermore, since $(\alpha_b(n))_n$ is strictly increasing, we can choose k such that $\alpha_b(k) > 2a$. Choose k such that $\alpha_b(k)$ is odd, and $\alpha_b(k) > 2a$ holds, and define $u = \alpha_b(k)$. Define $t = \alpha_b(k+1)$; then by lemma 2.8 $u^2 - but + t^2 = 1$. Choose $m = uk$, and let $r = \alpha_b(m-1)$, $s = \alpha_b(m)$. Then $s^2 - bsr + r^2 = 1$ and $r < s$. By equation (2.2)

$$\alpha_b(m) \equiv (-1)^{u-1} u \alpha_b(k) \alpha_b^{u-1}(k-1) \pmod{\alpha_b^2(k)}$$

Since $u = \alpha_b(k)$ and $s = \alpha_b(m)$, we have that $s \equiv (-1)^{u-1} u^2 \alpha_b^{u-1}(u-1) \equiv 0 \pmod{u^2}$, so $u^2|s$. From $bs - 2r = b\alpha_b(m) - 2\alpha_b(m-1) \geq 4\alpha_b(m) - 2\alpha_b(m-1) > 2\alpha_b(m) \geq 0$ it follows that there exists a natural number v such that $v = bs - 2r$.

Now suppose that d divides both u and v . Then, since $u^2|s$, d divides s . Hence d divides $2r$. Since u was chosen to be odd, d is odd, so $d|r$. Now $s^2 - bsr + r^2 = 1$ implies that $d|1$. It follows that u and v are coprime.

By the Chinese remainder theorem, there exists $w > 2$ such that $w \equiv 2 \pmod{u}$ and $w \equiv b \pmod{v}$, i.e. $v|w-b$ and $u|w-2$. Define $x = \alpha_w(c)$ and $y = \alpha_w(c+1)$. Then $x^2 - wxy + y^2 = 1$ by lemma 2.8. Since $w \equiv b \pmod{v}$, we have $x = \alpha_w(c) \equiv \alpha_b(c) = a \pmod{v}$. Furthermore, by choosing k large enough, we can get $m \geq c$ so $v = bs - 2r > 2\alpha_b(m) \geq 2\alpha_b(c) = 2a$, so $v > 2a$, and hence $a = \text{arem}(a, v) = \text{arem}(x, v)$. Finally, $x = \alpha_w(c) \equiv \alpha_2(c) = c \pmod{w-2}$, and since $u|w-2$, this gives $x \equiv c \pmod{u}$. Since $2c \leq 2\alpha_b(c) = 2a < u$ by the choice of u , we have $c = \text{arem}(c, u) = \text{arem}(x, u)$. Hence we have constructed a solution to the system of equations. \square

Proof of theorem 2.9. This follows immediately from theorem 2.13. \square

Lemma 2.14. For all $b > 0$ and all $n \in \mathbb{N}$ we have $(b-1)^n \leq \alpha_b(n+1) \leq b^n$.

Proof. We prove this by induction to n .

For $n = 0$, $\alpha_b(1) = 1$ so $(b-1)^0 = \alpha_b(1) = b^0$.

For $n = 1$, $\alpha_b(2) = b\alpha_b(1) - \alpha_b(0) = b$ so $(b-1)^1 = b-1 < \alpha_b(1) = b = b^1$.

Suppose that $(b-1)^n \leq \alpha_b(n+1) \leq b^n$ for all $n < k$ with $k \geq 2$. Then

$$\alpha_b(k) = b\alpha_b(k-1) - \alpha_b(k-2) \leq b^k - (b-1)^{k-2} - 2 \leq b^k$$

By remark 2.5, $\alpha_b(k-1) > \alpha_b(k-2)$ so

$$\alpha_b(k) = b\alpha_b(k-1) - \alpha_b(k-2) \geq (b-1)\alpha_b(k-1) \geq (b-1)^{k-1}$$

The statement follows by induction. \square

Now we can finally prove that exponentiation is diophantine:

Proof of theorem 2.2. Consider the number $d_{b,c,x} := \frac{\alpha_{bx+4}(c+1)}{\alpha_x(c+1)}$.

Suppose that $b > 0$ and put $x = 16(c+1)\alpha_{b+4}(c+1)$. Then $x > 16c$, so

$$\begin{aligned} d_{b,c,x} &= \frac{\alpha_{bx+4}(c+1)}{\alpha_x(c+1)} \leq \frac{(bx+4)^c}{(x-1)^c} = \frac{(b+\frac{4}{x})^c}{(1-\frac{1}{x})^c} \leq \frac{(1+\frac{4}{x})^c}{(1-\frac{1}{x})^c} b^c = \\ &\frac{(1+\frac{4}{x})^c(1-\frac{4}{x})^c}{(1-\frac{1}{x})^c(1-\frac{4}{x})^c} b^c \leq \frac{b^c}{(1-\frac{1}{x})^c(1-\frac{4}{x})^{2c}} \leq \frac{b^c}{(1-\frac{4}{x})^c} \leq \frac{b^c}{1-\frac{8c}{x}} \leq b^c(1+\frac{16c}{x}) \end{aligned}$$

By lemma 2.14, $x \geq 16(c+1)(b+3)^c \geq 16(c+1)b^c$. This implies

$$\frac{16c}{x} \leq \frac{16c}{16(c+1)b^c} < \frac{1}{b^c}$$

so $d_{b,c,x} \leq b^c(1+\frac{16c}{x}) < b^c + 1$. Since we also have

$$d_{b,c,x} = \frac{\alpha_{bx+4}(c+1)}{\alpha_x(c+1)} \geq \frac{(bx+3)^c}{x^c} = (b+\frac{3}{x})^c \geq b^c$$

it follows that

$$b^c = \lfloor d_{b,c,x} \rfloor = \left\lfloor \frac{\alpha_{bx+4}(c+1)}{\alpha_x(c+1)} \right\rfloor \quad (2.3)$$

Now suppose $b = c = 0$. Then $d_{0,0,16} = \frac{\alpha_4(1)}{\alpha_x(1)} = 1 = 0^0$.

Finally, if $b = 0$ and $c > 0$, then $x = 16(c+1)\alpha_4(c+1) > 16\alpha_4(c+1) \geq 16 \cdot 3^c$ by lemma 2.14. Hence $x > 48$, and

$$d_{0,c,x} = \frac{\alpha_4(c+1)}{\alpha_x(c+1)} < \frac{4^c}{47^c} < 1$$

So in these cases equation (2.3) also holds. It follows that

$$\{(a, b, c) \mid a, b, c \in \mathbb{N}, a = b^c\} = \left\{ (a, b, c) \mid a, b, c \in \mathbb{N}, a = \left\lfloor \frac{\alpha_{bx+4}(c+1)}{\alpha_x(c+1)} \right\rfloor \right\}$$

By theorem 2.9, this is a diophantine definition, so $\{(a, b, c) \mid a, b, c \in \mathbb{N}, a = b^c\}$ is a diophantine set.

It follows immediately that $\{(a, b) \mid a, b \in \mathbb{N}, \exists n \in \mathbb{N} : a = b^n\}$ is also diophantine. \square

2.3 Coding

In this section, we will introduce two ways to encode a tuple of natural numbers into a single number. The first coding is the *Cantor numbering*. It is given by a polynomial over \mathbb{Q} , and is defined inductively by

$$\begin{aligned} \text{Cantor}_1(a_1) &= a_1 \\ \text{Cantor}_2(a_1, a_2) &= \frac{(a_1 + a_2)(a_1 + a_2 + 1) + 2a_1}{2} \\ \text{Cantor}_{n+1}(a_1, \dots, a_{n+1}) &= \text{Cantor}_n(a_1, \dots, a_{n-1}, \text{Cantor}(a_n, a_{n+1})) \end{aligned}$$

Lemma 2.15. *For all $n \geq 1$, Cantor_n is a bijection $\mathbb{N}^n \rightarrow \mathbb{N}$.*

Proof. We prove this by induction to n .

For $n = 1$, the statement is clearly true.

Suppose that $n = 2$, and suppose that $\text{Cantor}_2(a, b) = \text{Cantor}_2(c, d)$. If $a + b \geq c + d + 1$, then

$$\begin{aligned} \text{Cantor}_2(a, b) &= \frac{(a+b)(a+b+1) + 2a}{2} \geq \frac{(c+d+1)(c+d+2) + 2a}{2} = \\ &= \frac{(c+d)(c+d+1) + 2(c+d+1)}{2} > \text{Cantor}_2(c, d) \end{aligned}$$

Contradiction, so $a + b \leq c + d$. By symmetry, $c + d \leq a + b$ so $a + b = c + d$. Now $\frac{(a+b)(a+b+1)+2a}{2} = \frac{(c+d)(c+d+1)+2c}{2}$ implies $a = c$, so $b = d$. Hence Cantor_2 is injective.

To show that Cantor_2 is surjective, let $c \in \mathbb{N}$ and choose $n \in \mathbb{N}$ such that $\frac{n(n+1)}{2} \leq c < \frac{(n+1)(n+2)}{2}$. Define $a = c - \frac{n(n+1)}{2}$. Then a is an integer, and $a \geq 0$. Now take $b = n - a = n - c + \frac{n(n+1)}{2} = \frac{(n+1)(n+2)}{2} - 1 - c$. Then $b \in \mathbb{N}$, and

$$\text{Cantor}_2(a, b) = \frac{(a+b)(a+b+1) + 2a}{2} = \frac{n(n+1)}{2} + a = c$$

This shows that Cantor_2 is a surjection, and hence a bijection.

Now let $n \geq 2$ and assume that Cantor_n is a bijection for some n . If $\text{Cantor}_{n+1}(a_1, \dots, a_{n+1}) = \text{Cantor}_{n+1}(b_1, \dots, b_{n+1})$, then by definition we have $\text{Cantor}_n(a_1, \dots, a_{n-1}, \text{Cantor}_2(a_n, a_{n+1})) = \text{Cantor}_n(b_1, \dots, b_{n-1}, \text{Cantor}_2(b_n, b_{n+1}))$ so $a_1 = b_1, \dots, a_{n-1} = b_{n-1}$ and $\text{Cantor}_2(a_n, a_{n+1}) = \text{Cantor}_2(b_n, b_{n+1})$ which implies $a_n = b_n$ and $a_{n+1} = b_{n+1}$. Hence Cantor_{n+1} is injective.

For $c \in \mathbb{N}$, take $a_1, \dots, a_{n-1}, a \in \mathbb{N}$ such that $\text{Cantor}_n(a_1, \dots, a_{n-1}, a) = c$. Now choose $a_n, a_{n+1} \in \mathbb{N}$ such that $\text{Cantor}_2(a_n, a_{n+1}) = a$. Then $\text{Cantor}_{n+1}(a_1, \dots, a_{n+1}) = c$. Hence Cantor_{n+1} is surjective, so it is a bijection. \square

Given m and n , the function that gives the m^{th} element of the n -tuple with Cantor number b is diophantine, since it is given by

$$\begin{aligned} a = \text{Elem}_{n,m}(b) &\iff \exists x_1, \dots, x_{m-1}, x_{m+1}, \dots, x_n : \\ &2^{2^n} \text{Cantor}_n(x_1, \dots, x_{m-1}, a, x_{m+1}, \dots, x_n) = 2^{2^n} b \end{aligned}$$

The factors 2^{2^n} are needed to ensure that this equation has natural numbers as coefficients.

Unfortunately, we cannot show that Elem is a diophantine function of n, m and b . Therefore, although the Cantor numbering will be useful in the next chapter, we will also need another way to encode tuples. Positional coding is based on the idea that a tuple can be seen as the b -ary representation of a number.

Definition 2.16. A triple (a, b, c) is called a *positional code* of tuple (a_1, \dots, a_n) if $b > a_i$ for all i , $a = a_n b^{n-1} + a_{n-1} b^{n-2} + \dots + a_1$, and $c = n$. a the *cipher*, b the *base* and c the *length*.

This means that a_1, \dots, a_n are the coefficients of a in the b -ary expansion of a , and c is the length of the tuple. Note that not every triple is the positional code of a tuple. The predicate 'is a positional code' is diophantine:

$$\text{Code}(a, b, c) \iff b \geq 2 \wedge a < b^c$$

The function that gives the d^{th} element of a tuple is also diophantine:

$$e = \text{Elem}(a, b, d) \iff \exists x, y, z : d = z + 1 \wedge a = xb^d + eb^z + y \wedge e < b \wedge y < b^z$$

The concatenation of the tuples (a_1, \dots, a_n) and (b_1, \dots, b_m) is the tuple $(a_1, \dots, a_n, b_1, \dots, b_m)$. Denote the corresponding relation on codes by Concat , so $\text{Concat}(a, b, c, a_1, b_1, c_1, a_2, b_2, c_2)$ holds if and only if (a, b, c) is a code of the concatenation of the tuples encoded by (a_1, b_1, c_1) and (a_2, b_2, c_2) .

Remark 2.17. When the tuples (a_1, b_1, c_1) and (a_2, b_2, c_2) have the same base, this relation is diophantine:

$$\text{Concat}(a, b, c, a_1, b, c_1, a_2, b, c_2) \iff \text{Code}(a_1, b, c_1) \wedge \text{Code}(a_2, b, c_2) \wedge a = a_2 b^{c_1} + a_1 \wedge c = c_1 + c_2$$

We will show in theorem 2.24 that it is also diophantine when the bases b_1 and b_2 are distinct.

2.4 Concatenation of positional codes

Using positional coding, we can show that the property of being a prime number is diophantine, as well as some other relations.

Lemma 2.18. (i) *The relation $a = \binom{b}{c}$ is diophantine.*

(ii) *The relation $a = b!$ is diophantine.*

(iii) *The relation ‘ a is a prime number’ is diophantine.*

Proof. (i) Notice that $(x+1)^b = \sum_{c=0}^b \binom{b}{c} x^c$ for all x . Take $x = 2^b + 1$, then

$$a = \binom{b}{c} \iff a = \text{Elem}((2^b + 2)^b, 2^b + 1, c + 1)$$

(ii) Since $\binom{c}{b} = \frac{c!}{b!(c-b)!}$, we have

$$\frac{c^b}{\binom{c}{b}} = \frac{c^b b! (c-b)!}{c!} = \frac{c^b b!}{c \cdot \dots \cdot (c - (b+1))} = b! \cdot \frac{c}{c} \cdot \frac{c}{c-1} \cdot \dots \cdot \frac{c}{c - (b+1)} \quad (2.4)$$

From this expression, it is clear that $\frac{c^b}{\binom{c}{b}} \geq b!$ for all c .

Consider the functions $f(x) = (1+x)^{b-1}$ and $g(x) = 1 + 2(b-1)x$. Notice that $f(0) = g(0) = 0$. The derivatives are $f'(x) = (b-1)(1+x)^{b-2}$ and $g'(x) = 2(b-1)$, so if $(1+y)^{b-2} \leq 2$ for all $y \in [0, x]$, then $f(x) \leq g(x)$. Since $y \mapsto (1+y)^{b-2}$ is an increasing function, to show that $f(x) \leq g(x)$ it suffices to check that $(1+x)^{b-2} \leq 2$. Take $c = (b+1)^{b+2}$ and $x = \frac{b+1}{c - (b+1)}$. Then it is clear that $c \geq (b+1)^{b-2\sqrt{2}} = (b+1)^{(b-2\sqrt{2}-1) + (b+1)}$, so $x \leq b-2\sqrt{2}-1$ and hence $(1+x)^{b-2} \leq 2$. It follows that

$$\left(\frac{c}{c - (b+1)} \right)^{b-1} = \left(1 + \frac{b+1}{c - (b+1)} \right)^{b-1} \leq 1 + 2(b-1) \frac{b+1}{c - (b+1)}$$

From equation 2.4, we have

$$\frac{c^b}{\binom{c}{b}} = b! \cdot \frac{c}{c-1} \cdot \dots \cdot \frac{c}{c - (b+1)} < b! \cdot \left(\frac{c}{c - (b+1)} \right)^{b-1} \leq b! \left(1 + 2(b-1) \frac{b+1}{c - (b+1)} \right)$$

Since $c = (b+1)^{b+2}$, c certainly satisfies $c > 2(b-1)(b+1)b! + (b+1)$ so

$$2(b-1) \frac{b+1}{c - (b+1)} < \frac{2(b-1)(b+1)}{2(b-1)(b+1)b!} = \frac{1}{b!}$$

Hence

$$\frac{c^b}{\binom{c}{b}} < b! \left(1 + \frac{1}{b!} \right) = b! + 1$$

This implies that

$$b! \leq \frac{c^b}{\binom{c}{b}} < b! + 1$$

so $b! = \left\lfloor \frac{c^b}{\binom{c}{b}} \right\rfloor$ with $c = (b+1)^{b+2}$. This gives a diophantine definition for $b!$.

(iii) Denote ‘ a is a prime number’ by $\text{Prime}(a)$. Then

$$\text{Prime}(a) \iff a > 1 \wedge \text{gcd}(a, (a-1)!) = 1$$

which is a diophantine definition by part (ii), using the fact that the greatest common divisor is a diophantine function (see table 2.1). \square

Corollary 2.19. *There exists a polynomial such that the set of natural numbers it assumes, when the variables range over the natural numbers, is exactly the set of prime numbers.*

Proof. This follows immediately from lemma 1.7 and lemma 2.18(iii). \square

Using this, we can show that concatenation is also diophantine when the bases are distinct:

Definition 2.20. Define the following relations:

$$\begin{aligned} \text{Equal}(a_1, b_1, c_1, a_2, b_2, c_2) &\iff (a_1, b_1, c_1) \text{ and } (a_2, b_2, c_2) \text{ are codes of the same tuple.} \\ \text{NotGreater}(a_1, b_1, a_2, b_2) &\iff \forall k : \text{Elem}(a_1, b_1, k) \leq \text{Elem}(a_2, b_2, k) \\ \text{Small}(a, b, c, e) &\iff \text{Code}(a, b, c) \wedge \forall k : \text{Elem}(a, b, k) \leq e \end{aligned}$$

In the definition of NotGreater, we use the convention that $\text{Elem}(a, b, k) = 0$ if k is greater than the length of the tuple induced by a and b .

Theorem 2.21. *The relations Equal, NotGreater and Small are diophantine.*

In the proof, we need some auxiliary relations, of which we will first show that they are diophantine:

Definition 2.22.

$$\begin{aligned} \text{PNotGreater}(a_1, a_2, b) &\iff \text{Prime}(b) \wedge \text{NotGreater}(a_1, b, a_2, b) \\ \text{PSmall}(a, b, c, e) &\iff \text{Prime}(b) \wedge \text{Small}(a, b, c, e) \\ \text{Eq}(a_1, b_1, c_1, a_2, b_2, c_2) &\iff \text{Code}(a_1, b_1, c_1) \wedge c_1 = c_2 \wedge \text{PSmall}(a_2, b_2, c_2, b_1 - 1) \wedge \\ &\quad b_1^{c_1} + b_1 < b_2 \wedge a_1 \equiv a_2 \pmod{b_2 - b_1} \end{aligned}$$

Lemma 2.23. *The relations PNotGreater, PSmall and Eq are diophantine.*

Proof. Suppose that b is a prime number. Since $\binom{a_2}{a_1} = \frac{a_2!}{a_1!(a_2-a_1)!}$, and $\deg_b(n!) = \sum_{k=1}^{\infty} \lfloor \frac{n}{b^k} \rfloor$, we have, assuming $a_2 \geq a_1$,

$$\text{ord}_b \left(\binom{a_2}{a_1} \right) = \sum_{k=1}^{\infty} \left\lfloor \frac{a_2}{b^k} \right\rfloor - \left\lfloor \frac{a_1}{b^k} \right\rfloor - \left\lfloor \frac{a_2 - a_1}{b^k} \right\rfloor \quad (2.5)$$

Let (a_1, b, c_1) be a code for $(a_{1,1}, \dots, a_{1,c_1})$, let (a_2, b, c_2) be a code for $(a_{2,1}, \dots, a_{2,c_2})$ and define $a_3 = a_2 - a_1$ and $c_3 = \max(c_1, c_2)$. Let (a_3, b, c_3) be a code for $(a_{3,1}, \dots, a_{3,c_3})$. Since

$$\left\lfloor \frac{a_i}{b^k} \right\rfloor = \left\lfloor \frac{\sum_{j=1}^c a_{i,j} b^{j-1}}{b^k} \right\rfloor = \left\lfloor \sum_{j=1}^c a_{i,j} b^{j-k-1} \right\rfloor = \sum_{j=k+1}^c a_{i,j} b^{j-k-1}$$

equation (2.5) is equivalent to

$$\text{ord}_b \left(\binom{a_2}{a_1} \right) = \sum_{k=1}^{\infty} \sum_{j=k+1}^c (a_{2,j} - a_{1,j} - a_{3,j}) b^{j-k-1}$$

Now suppose that $\text{PNotGreater}(a_1, a_2, b)$ holds. Then $a_{1,j} \leq a_{2,j}$ for all j , so $a_{3,j} = a_{2,j} - a_{1,j}$. Hence equation (2.4) gives $\text{ord}_b \left(\binom{a_2}{a_1} \right) = 0$.

Conversely, suppose that $\text{ord}_b \left(\binom{a_2}{a_1} \right) = 0$. Since $\lfloor \frac{a_1}{b^k} \rfloor + \lfloor \frac{a_2 - a_1}{b^k} \rfloor \leq \lfloor \frac{a_2}{b^k} \rfloor$, every term in the sum in equation (2.5) is non-negative. Hence all terms must be zero, i.e.

$$\left\lfloor \frac{a_2}{b^k} \right\rfloor = \left\lfloor \frac{a_1}{b^k} \right\rfloor + \left\lfloor \frac{a_2 - a_1}{b^k} \right\rfloor$$

so

$$\sum_{j=k+1}^c a_{2,j} b^{j-k-1} = \sum_{j=k+1}^c (a_{1,j} + a_{3,j}) b^{j-k-1}$$

for all k . By subtracting two consecutive terms, it follows that $a_{1,j} \leq a_{1,j} + a_{3,j} = a_{2,j}$ for all j , so $\text{NotGreater}(a_1, b, a_2, b)$ holds.

It follows that

$$\text{PNotGreater}(a_1, a_2, b) \iff \text{Prime}(b) \wedge b \nmid \binom{a_2}{a_1}$$

Hence PNotGreater is a diophantine relation.

Notice that

$$\text{PSmall}(a, b, c, e) \iff \text{Prime}(b) \wedge (e \geq b \vee \text{PNotGreater}(a, e(b^{c-1} + \dots + b + 1), b))$$

It follows that PSmall is also diophantine.

Finally, Eq is diophantine by definition (since PSmall is diophantine). \square

Proof of theorem 2.21. We claim that the following equivalence holds:

$$\text{Equal}(a_1, b_1, c_1, a_2, b_2, c_2) \iff \exists x, y, z : \text{Eq}(a_1, b_1, c_1, x, y, z) \wedge \text{Eq}(a_2, b_2, c_2, x, y, z) \quad (2.6)$$

Suppose that $\text{Equal}(a_1, b_1, c_1, a_2, b_2, c_2)$ holds. Then (a_1, b_1, c_1) and (a_2, b_2, c_2) are codes for the same tuple (d_1, \dots, d_c) , so $c_1 = c_2 = c$ for some $c \in \mathbb{N}$. Choose a prime number y satisfying $b_1^c + b_1 < y$ and $b_2^c + b_2 < y$. Let $z = c$, and $x = \sum_{j=1}^c d_j y^{j-1}$. Then (x, y, z) is again a code of the tuple (d_1, \dots, d_c) . Since (a_1, b_1, c_1) is a code of this tuple, we have $d_k \leq b_1 - 1$ for all k , so $\text{Elem}(x, y, k) \leq b_1 - 1$ for all k . This implies that $\text{PSmall}(x, y, z, b_1 - 1)$ holds. Similarly, $\text{PSmall}(x, y, z, b_2 - 1)$ holds. Finally,

$$a_1 = \sum_{j=1}^c d_j b_1^{j-1} \equiv \sum_{j=1}^c d_j y^{j-1} = x \pmod{y - b_1}$$

and similarly $a_2 \equiv x \pmod{y - b_2}$. Hence $\text{Eq}(a_1, b_1, c_1, x, y, z) \wedge \text{Eq}(a_2, b_2, c_2, x, y, z)$ holds.

Conversely, suppose that there exist x, y, z such that $\text{Eq}(a_1, b_1, c_1, x, y, z) \wedge \text{Eq}(a_2, b_2, c_2, x, y, z)$ holds. Then (a_1, b_1, c_1) is the code of a tuple (d_1, \dots, d_z) of length $c_1 = z$, and (x, y, z) is a code of a tuple of the same length. Since

$$a_1 = \sum_{j=1}^z d_j b_1^{j-1} \leq \sum_{j=1}^z (b_1 - 1) b_1^{j-1} = \sum_{j=1}^z b_1^j - \sum_{j=1}^z b_1^{j-1} = b_1^z - 1 < b_1^z < y - b_1$$

we have $0 \leq a_1 < y - b_1$. Given x , this determines a_1 uniquely by $a_1 \equiv x \pmod{y - b_1}$. Since $\text{Elem}(x, y, k) \leq b_1 - 1$ for all k , (x, y, z) is a code for a tuple (e_1, \dots, e_z) satisfying $e_j < b_1$ for all j . Now consider $x' = \sum_{j=1}^z e_j b_1^{j-1}$. Similar to a_1 , we have $0 \leq x' < y - b_1$, and

$$x' = \sum_{j=1}^z e_j b_1^{j-1} \equiv \sum_{j=1}^z e_j y^{j-1} = x \pmod{y - b_1}$$

Hence x' satisfies the conditions that determine a_1 uniquely, so $a_1 = x'$. It follows that the b_1 -ary expansions of a_1 and x' are also the same, so $d_j = e_j$ for all j . Hence (a_1, b_1, c_1) and (x, y, z) are codes for the same tuple.

Similarly, (a_2, b_2, c_2) and (x, y, z) are codes for the same tuple. Hence $\text{Equal}(a_1, b_1, c_1, a_2, b_2, c_2)$ holds.

This proves equivalence (2.6). It follows immediately that Equal is a diophantine relation.

For prime numbers, we already showed that PNotGreater and PSmall are diophantine. Therefore, NotGreater and Small are also diophantine:

$$\begin{aligned} \text{NotGreater}(a_1, b_1, a_2, b_2) &\iff \exists x_1, x_2, y, z : \text{Equal}(a_1, b_1, z, x_1, y, z) \wedge \text{Equal}(a_2, b_2, z, x_2, y, z) \wedge \\ &\quad \text{PNotGreater}(x_1, x_2, y) \\ \text{Small}(a, b, c, e) &\iff \exists x, y : \text{Equal}(a, b, c, x, y, c) \wedge \text{PSmall}(x, y, c, e) \end{aligned}$$

□

Theorem 2.24. *The relation Concat is diophantine.*

Proof. Let (a_1, b_1, c_1) and (a_2, b_2, c_2) be codings for two tuples. By using Equal, we change both codes to codes with the same base, and then we can easily concatenate them:

$$\begin{aligned} \text{Concat}(a, b, c, a_1, b_1, c_1, a_2, b_2, c_2) &\iff \exists x_1, x_2 : \text{Equal}(a_1, b_1, c_1, x_1, b, c_1) \wedge \\ &\quad \text{Equal}(a_2, b_2, c_2, x_2, b, c_2) \wedge a = x_2 b^{c_1} + x_1 \wedge c = c_1 + c_2 \end{aligned}$$

□

2.5 Functions on tuples

Finally, we define functions on tuples:

Definition 2.25. Let $(a_1, b, c), \dots, (a_m, b, c)$ be codes of m tuples $(a_{i,1}, \dots, a_{i,c})$ ($1 \leq i \leq m$), and let $F : \{0, \dots, b-1\}^m \rightarrow \{0, \dots, b-1\}$ be a function. Then we define the function $F[b]$ by defining $F[b](a_1, \dots, a_m; c)$ to be the cipher of the tuple $(F(a_{1,1}, \dots, a_{m,1}), \dots, F(a_{1,c}, \dots, a_{m,c}))$, on base b . So if $a_i = \sum_{j=1}^c a_{i,j} b^{j-1}$, then $F[b](a_1, \dots, a_m; c) = \sum_{j=1}^c F(a_{1,j}, \dots, a_{m,j}) b^{j-1}$.

Lemma 2.26. *If F is a diophantine function, then $F[b]$ is also diophantine.*

Proof. For $1 \leq i \leq m$, $0 \leq k_i \leq b-1$ and $1 \leq j \leq c$, define $h_{k_1, \dots, k_m, j} = \delta_{a_{1,j}, k_1} \cdots \delta_{a_{m,j}, k_m}$. Then

$$\begin{aligned} F[b](a_1, \dots, a_m; c) &= \sum_{j=1}^c F(a_{1,j}, \dots, a_{m,j}) b^{j-1} = \\ &= \sum_{j=1}^c \sum_{k_1=0}^{b-1} \cdots \sum_{k_m=0}^{b-1} F(k_1, \dots, k_m) \delta_{a_{1,j}, k_1} \cdots \delta_{a_{m,j}, k_m} b^{j-1} = \\ &= \sum_{j=1}^c \sum_{k_1=0}^{b-1} \cdots \sum_{k_m=0}^{b-1} F(k_1, \dots, k_m) h_{k_1, \dots, k_m, j} \end{aligned}$$

Hence to show that $F[b]$ is diophantine, it suffices to find diophantine expressions for $h_{k_1, \dots, k_m, j}$.

We claim that a diophantine definition of $h_{k_1, \dots, k_m, j}$ is given by

$$\begin{aligned} \forall i \in \{1, \dots, m\}, \forall k_1, \dots, k_m, k'_1, \dots, k'_m \in \{0, \dots, b-1\}, \forall j \in \{1, \dots, c\} \exists a_{k_1, \dots, k_m} : \\ h_{k_1, \dots, k_m, j} &= \text{Elem}(a_{k_1, \dots, k_m}, b, j) \wedge \text{Small}(a_{k_1, \dots, k_m}, b, m, 1) \wedge \\ ((k_1, \dots, k_m) \neq (k'_1, \dots, k'_m)) &\Rightarrow (\text{Elem}(a_{k_1, \dots, k_m}, b, j) = 0 \vee \text{Elem}(a_{k'_1, \dots, k'_m}, b, j) = 0) \wedge \\ &\quad \sum_{k_1, \dots, k_m} k_i a_{k_1, \dots, k_m} = a_i \wedge \sum_{k_1, \dots, k_m} a_{k_1, \dots, k_m} = \sum_j b^{j-1} \quad (2.7) \end{aligned}$$

First suppose that $h_{k_1, \dots, k_m, j}$ satisfies (2.7), and put $a_{k_1, \dots, k_m} = \sum_{j=1}^c h_{k_1, \dots, k_m, j} b^{j-1}$. Then $h_{k_1, \dots, k_m, j} = \text{Elem}(a_{k_1, \dots, k_m}, b, j)$. Notice that all $h_{k_1, \dots, k_m, j}$ are pairwise orthonormal, in the sense that all elements of the tuples they represent are 0 or 1, and there is no position in which both

tuples have a 1. This implies that $\text{Small}(a_{k_1, \dots, k_m}, b, m, 1)$ and $((k_1, \dots, k_m) \neq (k'_1, \dots, k'_m)) \Rightarrow (\text{Elem}(a_{k_1, \dots, k_m}, b, j) = 0 \vee \text{Elem}(a_{k'_1, \dots, k'_m}, b, j) = 0)$. Furthermore,

$$\begin{aligned} \sum_{k_1=0}^{b-1} \dots \sum_{k_m=0}^{b-1} k_1 a_{k_1, \dots, k_m} &= \sum_{j=1}^c \sum_{k_1=0}^{b-1} \dots \sum_{k_m=0}^{b-1} k_1 \delta_{a_1, j, k_1} \cdot \dots \cdot \delta_{a_m, j, k_m} b^{j-1} = \\ &= \sum_{j=1}^c b^{j-1} \sum_{k_1=0}^{b-1} k_1 \delta_{a_1, j, k_1} \cdot \sum_{k_2=0}^{b-1} \delta_{a_2, j, k_2} \cdot \dots \cdot \sum_{k_m=0}^{b-1} \delta_{a_m, j, k_m} = \\ &= \sum_{j=1}^c b^{j-1} \sum_{k_1=0}^{b-1} a_{1, j} \cdot 1 \cdot \dots \cdot 1 = \sum_{j=1}^c a_{1, j} b^{j-1} = a_1 \end{aligned}$$

Similarly,

$$\sum_{k_1=0}^{b-1} \dots \sum_{k_m=0}^{b-1} k_i a_{k_1, \dots, k_m} = a_i$$

for all $1 \leq i \leq m$.

Finally, we compute in the same way that

$$\sum_{k_1=0}^{b-1} \dots \sum_{k_m=0}^{b-1} a_{k_1, \dots, k_m} = \sum_{j=1}^c b^{j-1}$$

It follows that $h_{k_1, \dots, k_m, j}$ indeed satisfies definition (2.7).

Conversely, these properties of $h_{k_1, \dots, k_m, j}$ determine them uniquely: for every j , there is at most one (k_1, \dots, k_m) such that $h_{k_1, \dots, k_m, j} = 1$, and $h_{k'_1, \dots, k'_m, j} = 0$ for all other (k'_1, \dots, k'_m) . Furthermore, since $\sum_{k_1, \dots, k_m} a_{k_1, \dots, k_m} = \sum_j b^{j-1}$, such (k_1, \dots, k_m) exists for all j . Since

$$\sum_{k_1=0}^{b-1} \dots \sum_{k_m=0}^{b-1} k_i h_{k_1, \dots, k_m, j} b^{j-1} = a_i$$

and $a_i = \sum_{j=1}^c a_{i, j} b^{j-1}$, we must have

$$\sum_{k_1=0}^{b-1} \dots \sum_{k_m=0}^{b-1} k_i h_{k_1, \dots, k_m, j} = a_{i, j}$$

This implies that the tuple (k_1, \dots, k_m) with $h_{k_1, \dots, k_m, j} = 1$ satisfies $k_i = a_{i, j}$. This determines uniquely for which (k_1, \dots, k_m) we have $h_{k_1, \dots, k_m, j} = 1$, so all $h_{k_1, \dots, k_m, j}$ are determined uniquely by the conditions in (2.7). Hence this is a diophantine definition of $h_{k_1, \dots, k_m, j}$, so $F[b]$ is a diophantine function. \square

Chapter 3

Undecidability of Hilbert's tenth problem

In the previous chapter, we considered some relations on \mathbb{N} and showed that they are diophantine in $(0, 1; +, \cdot)$. We also introduced positional coding, and showed that concatenation is diophantine. In this chapter, we will show that the set of all solvable equations is diophantine, while its complement is not. After that, we will introduce a formalization of the notion of algorithm: the Turing machine. By considering which kind of sets can be recognized by Turing machines, we will show that Hilbert's tenth problem is undecidable.

3.1 Universal equations

In this section, we will introduce universal equations, and show that they exist. We will also introduce a way to encode equations. We will show that the set of all solvable diophantine equations equals the set of parameters for which a certain universal equation has a solution. This implies that the set of codes of solvable equations is diophantine. We will also show that the complement of this set is not diophantine.

Definition 3.1. A *universal equation* is an equation

$$U(a_1, \dots, a_n, k_1, \dots, k_l, y_1, \dots, y_w) = 0$$

with *element parameters* a_1, \dots, a_n and *code parameters* k_1, \dots, k_l , satisfying the following condition: for any family of diophantine equations $D(a_1, \dots, a_n, x_1, \dots, x_m) = 0$ with n parameters there are numbers k_1^D, \dots, k_l^D such that $D(a_1, \dots, a_n, x_1, \dots, x_m) = 0$ has a solution (x_1, \dots, x_m) if and only if $U(a_1, \dots, a_n, k_1^D, \dots, k_l^D, y_1, \dots, y_w) = 0$ has a solution (y_1, \dots, y_w) .

The equation $D(a_1, \dots, a_n, x_1, \dots, x_m) = 0$ defines a diophantine set, consisting of all tuples (a_1, \dots, a_n) such that the equation has a solution (x_1, \dots, x_m) . The same set is determined by the equation $U(a_1, \dots, a_n, k_1^D, \dots, k_l^D, y_1, \dots, y_w) = 0$. Therefore, we can view (k_1^D, \dots, k_l^D) as a code for this set.

We will prove the following theorem:

Theorem 3.2. *There exists a natural number w such that for all $n \geq 1$, there exists a universal diophantine equation $U_n(a_1, \dots, a_n, k, y_1, \dots, y_w)$ with n element parameters, one code parameter and w unknowns.*

We start by making some reductions:

Lemma 3.3. *If there exists a universal equation with n element parameters, l code parameters and w unknowns, then there also exists a universal equation with n element parameters, one code parameter and $l + w$ unknowns.*

Proof. Let $U(a_1, \dots, a_n, k_1, \dots, k_l, y_1, \dots, y_w) = 0$ be a universal equation with n element parameters, l code parameters and w unknowns. For any diophantine equation $D(a_1, \dots, a_n, x_1, \dots, x_m) =$

0, there exist k_1^D, \dots, k_l^D such that $D(a_1, \dots, a_n, x_1, \dots, x_m) = 0$ has a solution if and only if $U(a_1, \dots, a_n, k_1^D, \dots, k_l^D, y_1, \dots, y_w) = 0$ has a solution. Now let $k^D = 2^{2^l} \text{Cantor}_l(k_1^D, \dots, k_l^D)$. It is clear that $D(a_1, \dots, a_n, x_1, \dots, x_m) = 0$ has a solution (x_1, \dots, x_m) if and only if the equation

$$U(a_1, \dots, a_n, k_1, \dots, k_l, y_1, \dots, y_w)^2 + (k - 2^{2^l} \text{Cantor}_l(k_1, \dots, k_l))^2 = 0$$

has a solution $(k_1, \dots, k_l, y_1, \dots, y_w)$. Hence this is a universal equation with n element parameters, 1 code parameter and $l + w$ unknowns. \square

Lemma 3.4. *Suppose that there exists a universal equation with one element parameter, l code parameters and w unknowns. Then for every $n \geq 1$ there exists a universal equation with n element parameters, l code parameters and w unknowns.*

Proof. Let $U_1(a, k_1, \dots, k_l, y_1, \dots, y_w) = 0$ be a universal equation. Define

$$U_n(a_1, \dots, a_n, k_1, \dots, k_l, y_1, \dots, y_w) = U_1(2^{2^n} \text{Cantor}_n(a_1, \dots, a_n), k_1, \dots, k_l, y_1, \dots, y_w)$$

We claim that $U_n = 0$ is a universal equation.

Consider the equation $D(a_1, \dots, a_n, x_1, \dots, x_l) = 0$, for fixed a_1, \dots, a_n . This equation has a solution (x_1, \dots, x_l) if and only if

$$D(z_1, \dots, z_n, x_1, \dots, x_l)^2 + (a - 2^{2^n} \text{Cantor}_n(z_1, \dots, z_n))^2 = 0 \quad (3.1)$$

has a solution $(z_1, \dots, z_n, x_1, \dots, x_l)$, where $a = 2^{2^n} \text{Cantor}_n(a_1, \dots, a_n)$. This equation has only one element parameter a , so there exists k_1^D, \dots, k_l^D such that $U_1(a, k_1^D, \dots, k_l^D, y_1, \dots, y_w) = 0$ has a solution (y_1, \dots, y_w) if and only if equation (3.1) has a solution $(z_1, \dots, z_n, x_1, \dots, x_l)$. By definition

$$\begin{aligned} U_n(a_1, \dots, a_n, k_1^D, \dots, k_l^D, y_1, \dots, y_w) &= \\ U_1(2^{2^n} \text{Cantor}_n(a_1, \dots, a_n), k_1^D, \dots, k_l^D, y_1, y_w) &= \\ U_1(a, k_1^D, \dots, k_l^D, y_1, \dots, y_w) & \end{aligned}$$

It follows that $U_n(a_1, \dots, a_n, k_1^D, \dots, k_l^D, y_1, \dots, y_w) = 0$ has a solution if and only if the equation $D(a_1, \dots, a_n, x_1, \dots, x_l) = 0$ has a solution. This shows that U_n is a universal equation, with n element parameters, one code parameter and w unknowns. \square

Corollary 3.5. *If there exists a universal equation with 1 element parameter, 6 code parameters and $w - 6$ unknowns, then theorem 3.2 holds.*

Proof. Suppose that such a universal equation exists. Then by lemma 3.3, there also exists a universal equation with 1 element parameter, 1 code parameter and w unknowns. Now lemma 3.4 shows that for every n there exists a universal equation with n element parameters, 1 code parameter and w unknowns. \square

To construct a universal equation, we need to define a coding for equations. We also need a code for the possible solutions, and a way to check whether a tuple is indeed a solution of the coded equation.

We start by constructing a coding for equations. In the next part, we will always consider the equation $D(a, x_1, \dots, x_m) = 0$. The sub- and superscripts D will be omitted from the notation.

Definition 3.6. A *format* of the equation $D = 0$ is a triple (d, e, f) , such that d is greater than the degree of D , $e = m$ is the numbers of unknowns and $f = 2^{d^1} + \dots + 2^{d^e}$.

Of course, a format does not determine the equation uniquely. Therefore, we will extend it by three other numbers b, c_L and c_R .

Since we only defined codes for natural numbers, and the coefficients of D can be arbitrary integers, we write D as a difference of two polynomials with natural numbers as coefficients:

$$\begin{aligned} D(x_0, \dots, x_m) &= C_L(x_0, \dots, x_m) - C_R(x_0, \dots, x_m) \\ C_L(x_0, \dots, x_m) &= \sum_{i_0 + \dots + i_m < d} c_{L, i_0, \dots, i_m} x_0^{i_0} \cdots x_m^{i_m} \\ C_R(x_0, \dots, x_m) &= \sum_{i_0 + \dots + i_m < d} c_{R, i_0, \dots, i_m} x_0^{i_0} \cdots x_m^{i_m} \end{aligned}$$

where the coefficients c_{L, i_0, \dots, i_m} and c_{R, i_0, \dots, i_m} are natural numbers.

Definition 3.7. An *extended code* of D is a sextuple (b, c_L, c_R, d, e, f) , where (d, e, f) is the format of D , b satisfies $b > d! \max_{S \in \{L, R\}, i_0 + \dots + i_m < d} \{c_{S, i_0, \dots, i_m}\}$ and

$$\begin{aligned} c_L &= \sum_{i_0 + \dots + i_m < d} i_0! i_m! (d - 1 - i_0 - \dots - i_m)! c_{L, i_0, \dots, i_m} b^{d^{m+1} - i_0 d^0 - \dots - i_m d^m} \\ c_R &= \sum_{i_0 + \dots + i_m < d} i_0! i_m! (d - 1 - i_0 - \dots - i_m)! c_{R, i_0, \dots, i_m} b^{d^{m+1} - i_0 d^0 - \dots - i_m d^m} \end{aligned}$$

Lemma 3.8. *An extended code determines D uniquely.*

Proof. Since b satisfies $b > d! \max_{S \in \{L, R\}, i_0 + \dots + i_m < d} \{c_{S, i_0, \dots, i_m}\}$, we have

$$i_0! i_m! (d - 1 - i_0 - \dots - i_m)! c_{L, i_0, \dots, i_m} < b$$

Furthermore, $i_j < d$ for all j . This implies that every power of b occurs at most once in the sum that defines c_L , and that i_0, \dots, i_m are determined by the exponent. Hence the coefficients c_{L, i_0, \dots, i_m} are also uniquely determined by b, c_L and d . The same holds for c_R . Since D is uniquely determined by the coefficients c_{L, i_0, \dots, i_m} and c_{R, i_0, \dots, i_m} , the extended code determines D uniquely. \square

Definition 3.9. Define the relation $\text{ECode}(b, c_L, c_R, d, e, f)$ to hold if and only if there exist a polynomial of which (b, c_L, c_R, d, e, f) is an extended code.

Definition 3.10. (d, e, f, g, h) is a code of the tuple (x_1, \dots, x_m) with respect to D if (d, e, f) is a format of D , g satisfies $g > \max\{1, x_1, \dots, x_m\}$ and $h = x_1 g^{d^1} + \dots + x_m g^{d^m}$.

Definition 3.11. Let (d, e, f) be the format of an equation D . Define the relation SCode by: $\text{SCode}(d, e, f, g, h)$ holds if and only if there exist a tuple (x_1, \dots, x_m) of which (d, e, f, g, h) is a code.

Lemma 3.12. *Given that (d, e, f) is a format of an equation, the relation $\text{SCode}(d, e, f, g, h)$ is diophantine.*

Proof. Assuming that (d, e, f) is the format of an equation, (d, e, f, g, h) is the code of a possible solution if and only if $(h, g, d^e + 1)$ is the positional code of a tuple, with zeros except at the d^i th position, and coefficients $x_i < g$ at the d^i th position for all $i \leq e$. This means that element-wise, $(h, g, d^e + 1)$ is not greater than $(0, \dots, 0, g - 1, 0, \dots, 0, g - 1, 0, \dots, 0, \dots, 0, g - 1)$. By definition of (d, e, f) , the triple $(f, 2, d^e + 1)$ is a positional code of $(0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots, 0, \dots, 0, 1)$ with the ones at the same positions as the x_i . Now define t such that $(t, g, d^e + 1)$ is a positional code of the same tuple, i.e. $t = g^{d^1} + \dots + g^{d^e}$. Then $(g - 1)t = (g - 1)g^{d^1} + \dots + (g - 1)g^{d^e}$, so $((g - 1)t, g, d^e + 1)$ is a code of $(0, \dots, 0, g - 1, 0, \dots, 0, g - 1, 0, \dots, 0, \dots, 0, g - 1)$. So (d, e, f, g, h) is the code of an equation if $(h, g, d^e + 1)$ is element-wise not greater than $((g - 1)t, g, d^e + 1)$.

Hence the following equivalence holds:

$$\text{SCode}(d, e, f, g, h) \iff \exists t : \text{Equal}(f, 2, d^e + 1, t, f, d^e + 1) \wedge \text{NotGreater}(h, (g - 1)t, g)$$

This shows that SCode is diophantine. \square

We have defined extended codes of polynomials (b, c_L, c_R, d, e, f) and codes of possible solutions (d, e, f, g, h) . Now we want to determine whether (d, e, f, g, h) is actually a solution of (b, c_L, c_R, d, e, f) , without decoding the polynomial and the possible solution.

Lemma 3.13. *Let (b, c_L, c_R, d, e, f) be a code of $D(a, x_1, \dots, x_m) = 0$, and let (d, e, f, g, h) be a code of a possible solution. Let C be either C_L or C_R , and c either c_L or c_R (with the same subscript as C) and write $C(x_0, x_1, \dots, x_m) = \sum_{i_0+\dots+i_m < d} c_{i_0\dots i_m} x_0^{i_0} x_1^{i_1} \dots x_m^{i_m}$. Suppose that $u > (1+a+h)d^{d-1}(c_L+c_R)$. Then there exists s and t such that $\text{Equal}(c, b, d^{e+1}+1, s, u, d^{e+1}+1)$ and $\text{Equal}(h, g, d^e+1, t, u, d^e+1)$ hold. Furthermore,*

$$C(a, x_1, \dots, x_m) = \frac{\text{Elem}((1+aw+t)^{d-1}s, w, d^{e+1}+1)}{(d-1)!}$$

Proof. First we show that s and t exist.

By definition 3.7, we have

$$c = \sum_{i_0+\dots+i_m < d} i_0! \dots i_m! (d-1-i_0-\dots-i_m)! c_{i_0\dots i_m} b^{d^{m+1}-i_0d^0-\dots-i_md^m}$$

Hence $(c, b, d^{e+1}+1)$ represents the tuple with elements $i_0! \dots i_m! (d-1-i_0-\dots-i_m)! c_{i_0\dots i_m}$. Since u satisfies the inequality $u > (1+a+h)d^{d-1}(c_L+c_R)$, in particular we have

$$u > c \geq \sum_{i_0+\dots+i_m < d} i_0! \dots i_m! (d-1-i_0-\dots-i_m)! c_{i_0\dots i_m}$$

so we can also write this tuple on the basis u . Hence

$$s = \sum_{i_0+\dots+i_m < d} i_0! \dots i_m! (d-1-i_0-\dots-i_m)! c_{i_0\dots i_m} u^{d^{m+1}-i_0d^0-\dots-i_md^m}$$

satisfies $\text{Equal}(c, b, d^{e+1}+1, s, u, d^{e+1}+1)$.

(h, g, d^e+1) is a code of a tuple whose elements are all smaller than g . Since $u > g \geq h$, we can write this tuple on the basis u , and $t = x_1 u^{d^1} + \dots + x_m u^{d^m}$ satisfies $\text{Equal}(h, g, d^e+1, t, u, d^e+1)$.

Next, we compute $(1+au+t)^{d-1}$. Notice that

$$\begin{aligned} (1+au+t)^{d-1} &= (1+au^{d^0} + x_1 u^{d^1} + \dots + x_m u^{d^m})^{d-1} = \\ &= \sum_{i_0+\dots+i_m < d} \binom{d-1}{i_0} \binom{d-1-i_0}{i_1} \dots \binom{d-1-i_0-\dots-i_{m-1}}{i_m} a^{i_0} x_1^{i_1} \dots x_m^{i_m} u^{i_0d^0+\dots+i_md^m} = \\ &= \sum_{i_0+\dots+i_m < d} \frac{(d-1)!}{i_0! i_1! \dots i_m! (d-1-i_0-i_1-\dots-i_m)!} a^{i_0} x_1^{i_1} \dots x_m^{i_m} u^{i_0d^0+\dots+i_md^m} \end{aligned}$$

Notice that $(1+au+t)^{d-1}$ has degree at most $(d-1)d^m$ (as a polynomial in u).

Now we compute $(1+au+t)^{d-1}s$. Since the degree of s is d^{m+1} , the degree of $(1+au+t)^{d-1}s$ is at most $(d-1)d^m + d^{m+1} = 2d^{m+1} - d^m$. By the expressions for $(1+au+t)^{d-1}$ and s given above, we have

$$(1+au+t)^{d-1}s = \sum \frac{(d-1)! j_0! \dots j_m! (d-1-j_0-\dots-j_m)!}{i_0! \dots i_m! (d-1-i_0-\dots-i_m)!} c_{j_0\dots j_m} a^{i_0} x_1^{i_1} \dots x_m^{i_m} u^{d^{m+1}+(i_0-j_0)d^0+\dots+(i_m-j_m)d^m}$$

where the sum runs over all tuples (i_0, \dots, i_m) and (j_0, \dots, j_m) such that $i_0 + \dots + i_m < d$ and $j_0 + \dots + j_m < d$.

Notice that

$$\sum_{\substack{i_0+\dots+i_m < d \\ j_0+\dots+j_m < d}} \frac{(d-1)!j_0!\dots j_m!(d-1-j_0-\dots-j_m)!}{i_0!\dots i_m!(d-1-i_0-\dots-i_m)!} c_{j_0\dots j_m} a^{i_0} x_1^{i_1} \dots x_m^{i_m} = \\ \sum_{j_0+\dots+j_m < d} j_0!\dots j_m!(d-1-j_0-\dots-j_m)! c_{j_0\dots j_m} (1+a+x_1+\dots+x_m)^{d-1}$$

Since

$$u > (1+a+h)^{d-1}(c_L+c_R) \geq (1+a+x_1g^{d^1}+\dots+x_mg^{d^m})^{d-1}c \geq \\ (1+a+x_1+\dots+x_m)^{d-1} \sum_{j_0+\dots+j_m < d} j_0!\dots j_m!(d-1-j_0-\dots-j_m)! c_{j_0\dots j_m}$$

u is greater than the sum of the coefficients of $(1+au+t)^{d-1}s$, written on the base u . Hence, if we define C_k by

$$(1+au+t)^{d-1}s = \sum_{k=0}^{2d^{m+1}-d^m} C_k u^k$$

then

$$C_k = \sum \frac{(d-1)!j_0!\dots j_m!(d-1-j_0-\dots-j_m)!}{i_0!\dots i_m!(d-1-i_0-\dots-i_m)!} c_{j_0\dots j_m} a^{i_0} x_1^{i_1} \dots x_m^{i_m}$$

where the sum runs over all tuples (i_0, \dots, i_m) and (j_0, \dots, j_m) with $d^{m+1} + (i_0 - j_0)d^0 + \dots + (i_m - j_m)d^m = k$.

If we take $k = d^{e+1} = d^{m+1}$, then we must have $i_0 = j_0, \dots, i_m = j_m$ because all i_k and j_k are non-negative and at most $d-1$. Hence

$$C_{d^{e+1}} = \sum_{i_0+\dots+i_m < d} \frac{(d-1)! \cdot i_0! \cdot \dots \cdot i_m!(d-1-i_0-\dots-i_m)!}{i_0!i_1!\dots i_m!(d-1-i_0-\dots-i_m)!} c_{i_0\dots i_m} a^{i_0} x_1^{i_1} \dots x_m^{i_m} = \\ \sum_{i_0+\dots+i_m < d} (d-1)! c_{i_0\dots i_m} a^{i_0} x_1^{i_1} \dots x_m^{i_m} = (d-1)! C(a, x_1, \dots, x_m)$$

This implies that $((1+au+t)^{d-1}s, u, 2d^{m+1}-d^m)$ is the code of a tuple of which the $d^{e+1}+1$ -th element is $C_{d^{e+1}}$. Hence $C(a, x_1, \dots, x_m) = \frac{elem((1+au+t)^{d-1}s, u, d^{e+1}+1)}{(d-1)!}$. \square

Definition 3.14. Define the following relation:

$$\text{Solution}(a, b, c_L, c_R, d, e, f, g, h) \iff \text{SCode}(d, e, f, g, h) \wedge \exists s_L, s_R, t, u : \\ u > (1+a+h)^{d-1}(c_L+c_R) \wedge \\ \text{Equal}(c_L, b, d^{e+1}+1, s_L, u, d^{e+1}+1) \wedge \\ \text{Equal}(c_R, b, d^{e+1}+1, s_R, u, d^{e+1}+1) \wedge \\ \text{Equal}(h, g, d^e+1, t, u, d^e+1) \wedge \\ \text{Elem}((1+au+t)^{d-1}s_L, u, d^{e+1}+1) = \\ \text{Elem}((1+au+t)^{d-1}s_R, u, d^{e+1}+1)$$

Then it is clear that Solution is diophantine, and the following statement holds:

$$\text{ECode}(b, c_L, c_R, d, e, f) \Rightarrow ((\exists g, h : \text{SCode}(d, e, f, g, h) \wedge \text{Solution}(a, b, c_L, c_R, d, e, f, g, h)) \iff \\ \exists x_1, \dots, x_m : D_{b, c_L, c_R, d, e, f}(a, x_1, \dots, x_m) = 0)$$

where $D_{b, c_L, c_R, d, e, f}$ is the polynomial with code (b, c_L, c_R, d, e, f) .

Now we can finally prove theorem 3.2:

Proof of theorem 3.2. $\text{Solution}(a, b, c_L, d_R, e, f, g, h)$ is a diophantine relation, so it is given by some polynomial $U(a, b, c_L, c_R, d, e, f, g, h, y_9, \dots, y_w)$. If we consider g and h as unknowns (as well as y_9, \dots, y_w), then, for every extended code (b, c_L, c_R, d, e, f) , $U(a, b, c_L, c_R, d, e, f, g, h, y_9, \dots, y_w) = 0$ has a solution if and only if $D_{b, c_L, c_R, d, e, f}(a, x_1, \dots, x_m) = 0$ has a solution. Hence this is a universal equation with 1 element parameter, 6 code parameters and $w - 6$ unknowns. Now the theorem follows from corollary 3.5. \square

Using this coding of equations, not every number is a code of an equation. If k is not the code of an equation, then we define it to be the code of the equation $U_1(a, k, x_1, \dots, x_m) = 0$. Then we have for all k that the equation with code k is solvable if and only if $U_1(a, k, x_1, \dots, x_m) = 0$ is solvable.

For $n = 0$, we define U_0 to be

$$U_0(k, y_1, \dots, y_w) = U_1(0, k, y_1, \dots, y_w)$$

Using these universal equations, we show that the set of codes of solvable equations is diophantine, while its complement is not:

Definition 3.15. Define the set \mathfrak{h}_0 by

$$\mathfrak{h}_0 = \{a \mid U_0(a, y_1, \dots, y_m) \text{ has a solution}\}$$

i.e. \mathfrak{h}_0 is the set of all parameter-free equations that have a solution.

Theorem 3.16. \mathfrak{h}_0 is a diophantine set with a non-diophantine complement.

Proof. It is clear from the definition that \mathfrak{h}_0 is diophantine.

Define \mathfrak{h}_1 by

$$\mathfrak{h}_1 = \{a \mid U_1(a, a, y_1, \dots, y_w) \text{ has a solution}\}$$

Suppose that $\overline{\mathfrak{h}_1}$ is diophantine. Then it has a defining polynomial D_c , so

$$\overline{\mathfrak{h}_1} = \{a \mid \exists z_1, \dots, z_l : D_c(a, z_1, \dots, z_l) = 0\}$$

Since U_1 is a universal equation, for every diophantine equation $D(a, x_1, \dots, x_m) = 0$ there exists k^D such that $D(a, x_1, \dots, x_m) = 0$ has a solution if and only if $U_1(a, k^D, y_1, \dots, y_m) = 0$ has a solution. In particular, this holds for $D = D_c$, so

$$\overline{\mathfrak{h}_1} = \{a \mid \exists y_1, \dots, y_w : U_1(a, k^{D_c}, y_1, \dots, y_w) = 0\}$$

Suppose that $k^{D_c} \in \mathfrak{h}_1$. Then $U_1(k^{D_c}, k^{D_c}, y_1, \dots, y_w) = 0$ has a solution, so $k^{D_c} \in \overline{\mathfrak{h}_1}$. Contradiction, so $k^{D_c} \notin \mathfrak{h}_1$. But then $k^{D_c} \in \overline{\mathfrak{h}_1}$, so $U_1(k^{D_c}, k^{D_c}, y_1, \dots, y_w) = 0$ has a solution, and hence $k^{D_c} \in \mathfrak{h}_1$. This again gives a contradiction. It follows that $\overline{\mathfrak{h}_1}$ is not diophantine.

Now consider the equation $W_{p,q}(y_1, \dots, y_m) = U_1(p, q, y_1, \dots, y_m)$ for fixed p, q , and let its extended code be (b, c_L, c_R, d, e, f) . We can choose the format (d, e, f) independent of p and q . The coefficients $c_{L, i_0 \dots i_m}$ and $c_{R, i_0 \dots i_m}$ are polynomials in p and q , so we can choose b and c to be a polynomial expression in p and q . By construction of the universal equations, the code k will also be a polynomial in p and q , say $k = K(p, q)$. Now $a \in \mathfrak{h}_1$ holds if and only if $U_1(a, a, y_1, \dots, y_w) = 0$ has a solution. This is equivalent to the statement that $W(a, a) = 0$ has a solution, which holds if and only if the equation with code $k = K(a, a)$ is solvable, i.e. $K(a, a) \in \mathfrak{h}_0$. Hence $K(a, a) \in \overline{\mathfrak{h}_0}$ if and only if $a \in \overline{\mathfrak{h}_1}$. This implies that if $\overline{\mathfrak{h}_0}$ would be diophantine, then $\overline{\mathfrak{h}_1}$ would be diophantine. Hence $\overline{\mathfrak{h}_0}$ is not diophantine. \square

3.2 Turing machines

In his tenth problem, Hilbert asks for a process to decide whether or not a diophantine equation has a solution. Usually, ‘process’ is interpreted as ‘algorithm’. Although the meaning of ‘algorithm’ is intuitively clear, it doesn’t have a rigorous definition. In the first half of the twentieth century, many definitions for ‘algorithm’ were proposed. They all turned out to be equivalent to ‘computation on a Turing machine’. Church’s thesis states that this is indeed true: all algorithms (in the intuitive sense) can be carried out on a Turing machine, and all definitions of computable are equivalent to ‘computable by a Turing machine.’ Church’s thesis cannot be proved, because it is not a theorem with a precise statement, but it is generally accepted to be true. Therefore, we will only consider Turing machines, and interpret Hilbert’s tenth problem as the problem whether or not there exists a Turing machine that decides whether or not an equation has a solution.

A Turing machine can be thought of as a tape, divided into cells, in which a head can write symbols from an alphabet. The tape is infinite to the right, but has a left end and the left-most cell is blank. At each moment, the machine is in a state. There is a unique initial state, in which the machine is at the start of the computation, and there are some final states. During each step of the computation, the head reads a symbol on the tape, writes a symbol (on the same cell), and moves to the left or right. This can be formalized as follows:

Definition 3.17. A Turing machine consists of a sextuple $(Q, \Sigma, \Gamma, \delta, q_0, F)$. Q is a finite set of states, containing the *start state* q_0 , and the set of *final states* F , not containing q_0 . Γ is a finite set, called the *tape alphabet*, containing a special symbol B representing empty cells. Σ is the *input alphabet*, which is always a subset of $\Gamma \setminus \{B\}$. δ is a function from a subset of $Q \times \Gamma$ to $Q \times \Gamma \times \{L, R\}$ and is called the *transition function*.

Definition 3.18. The set of (finite, but possibly empty) strings of elements of Σ is denoted Σ^* .

At the beginning of the computation, an element of Σ^* is written on the tape, starting directly right from the left-most cell, which is empty. The remainder of the tape is also empty. If the machine is in a state $q_i \in Q$ and reads a symbol $\gamma_i \in \Gamma$, such that $\delta(q_i, \gamma_i) = (q_j, \gamma_j, D)$ is defined (with $D \in \{L, R\}$), then the state is changed to q_j , the machine writes γ_j , and the head moves to the direction specified by D .

Definition 3.19. A Turing machine *halts* if it is in a state q , reading a symbol γ , such that $\delta(q, \gamma)$ is not defined. The computation of a Turing machine *terminates abnormally* if it moves to the left from the left-most cell. A string s of elements of Σ is *accepted* by a Turing machine if the machine halts (normally) in a final state, when the input is s .

Definition 3.20. A subset S of Σ^* is *recursively enumerable* if there exists a Turing machine that accepts a string s if and only if $s \in S$. S is *recursive* if it is accepted by a Turing machine that halts on all input strings. In this case, S will also be called *decidable*.

A recursive set is also called decidable since the Turing machine that accepts it can be used to decide whether $s \in S$: on input s , the machine halts after a finite number of steps. If it halts in a final state, then $s \in S$, and if it halts in a non-final state, then $s \notin S$.

In all Turing machines we will construct below, the tape alphabet will be $\Gamma = \{B, 0, 1, 2, 3, \blacksquare\}$, and the input alphabet is $\Sigma = \{0, 1\}$. The symbol \blacksquare will be used to mark the left-most cell.

Definition 3.21. The *canonical representation* of a number $n \in \mathbb{N}$ consists of a 0, followed by n consecutive 1’s. A tuple (a_1, \dots, a_n) will be represented by the concatenation of the representations of a_1, \dots, a_n .

Definition 3.22. A subset $S \subseteq \mathbb{N}$ is *recursive* (or *recursively enumerable*) if the corresponding subset of Σ is recursive (or recursively enumerable).

In constructing Turing machines, we can use machines we have constructed already. We can combine two Turing machines M_1 and M_2 to give a new machine M in two ways:

First, we can construct a machine that first executes the action of M_1 , and then executes the action of M_2 if M_1 halted in a final state. We will denote this machine by $M_1; M_2$.

Secondly, we can construct a machine with the description **while** M_1 **do** M_2 . This machine works as follows: if M_1 halts in a final state, then M_2 is executed. If M_2 also halts in a final state, then M_1 is executed again, etcetera. If M_2 halts in a non-final state, then M halts in a non-final state, and if M_1 halts in a non-final state, then M halts in a final state. If either M_1 or M_2 does not halt, then M doesn't halt.

In table 3.2, a few simple machines are listed, together with their action. The first column gives the name of the machine, the second column gives a description of the action and the third column is the definition. When in the second column only a tuple is given, then this is the output on input (a_1, \dots, a_n) . If the movement of the head is not given, then the head ends in the initial position. Most machines halt, so if the action is 'halt in a final state if some condition is satisfied', then it will halt in a non-final state if the condition is not satisfied. Of course, NEVERSTOP doesn't halt, and there are some machines that end in a infinite loop if the symbol they search for is not on the tape. By carefully applying this machines, we will not have to be concerned with these infinite loops. If nothing is specified about the halting, then the machine always halts in a final state.

Some of these machines only work after the left-most cell has been marked by \blacksquare . This can simply be done by starting every computation with a transition from q_0 to some other state q_1 , while changing B into \blacksquare . Now q_1 must be used as the start state for the remainder of the computation. We will assume that the computation always starts with this step.

Table 3.1: Some examples of Turing machines

Name	Description of the action	Definition
LEFT RIGHT	Move the head one cell to the left Move the head one cell to the right	
WRITE(0), etc. READ(0), etc. READ(B), etc. STOP NEVERSTOP READNOT(0), etc.	Write 0, etc. Halt in a final state when reading 0, etc. Halts in a final state when reading B or scanning an empty cell Halt in a non-final state Always stay in state q_0 without halting Halt in a final state when not reading 0, etc.	while READ(0) do STOP
START VACANT JUMP FIND(k) LAST	Go to the left-most cell Go to the left-most cell containing B or empty cell Go to the first cell on the right of the head containing 0 Move the head to the cell containing the zero of a_k Move the head to the cell containing the zero of a_n	while READNOT(\blacksquare) do LEFT START; while READNOT(B) do RIGHT while READNOT(0) do RIGHT FIND(1) = START; JUMP FIND($k + 1$) = FIND(k); RIGHT; JUMP VACANT; while READNOT(0) do LEFT
NEW INC	$(a_1, \dots, a_n, 0)$ $(a_1, \dots, a_n + 1)$	VACANT; WRITE(0) VACANT; WRITE(1)

Table 3.1: Some examples of Turing machines (continued)

DELETE	(a_1, \dots, a_{n-1})	VACANT; while READNOT(0) do {WRITE(B); LEFT}; WRITE(B)
MARK(2) (or 3)	Replaces 1 by 2 (or 3)	while {RIGHT; READ(1)} do WRITE(2)
THEREIS(2) etc.	Halts in a final state if the symbol 2 or 3, respectively, occurs on the tape	START; while READNOT(2) do {READNOT(B); RIGHT}
THEREWAS(2) etc.	Replaces the symbol 2 or 3 by 1, if it occurs on the tape	THEREIS(2); WRITE(1)
RESTORE	Replaces all 2's and 3's by 1's	while THEREIS(2) do THEREWAS(2); while THEREIS(3) do THEREWAS(3)
APPEND(k)	$(a_1, \dots, a_n + a_k)$	FIND(k); MARK(2); while THEREWAS(2) do INC
COPY(k)	(a_1, \dots, a_n, a_k)	NEW; APPEND(k)
ADD(k, l)	$(a_1, \dots, a_n, a_k + a_l)$	COPY(k); APPEND
MULT(k, l)	$(a_1, \dots, a_n, a_k a_l)$	NEW; FIND(k); MARK(3); while THEREWAS(3) do APPEND(l)
NOTGREATER(k, l)	Halts in a final state if $a_k \leq a_l$	FIND(k); MARK(2); FIND(l); MARK(3); while THEREIS(2) and THEREIS(3) do {THEREWAS(2) and THEREWAS(3)}; while THEREIS(2) do {RESTORE; STOP}; RESTORE
EQUAL(k, l)	Halt in a final state if $a_k = a_l$	NOTGREATER(k, l) and NOTGREATER(l, k)
NOTEQUAL(k, l)	Halt in a final state if $a_k \neq a_l$	while EQUAL(k, l) do STOP
NEXT	$(a_1, \dots, a_{n-2}, b, c)$, where (b, c) immediately follows (a_{n-1}, a_n) in the Cantor numbering	LAST; WRITE(1) RIGHT; while READ(λ) do {WRITE(1); LAST RIGHT}; WRITE(0)
DECODE	(a_1, \dots, a_n, b, c) , where (b, c) has Cantor number a_n	LAST; MARK(2); NEW; NEW; while THEREWAS(2) do NEXT

3.3 Recursively enumerable and diophantine sets

Since Turing machines are very powerful machines that can compute everything that is computable by an algorithm, we would expect that they can recognize diophantine sets. This is indeed true. It is much more remarkable that the converse is also true:

Theorem 3.23. (DPRM-theorem) *A subset of \mathbb{N}^n is diophantine if and only if it is recursively enumerable.*

We prove this in two steps.

Theorem 3.24. *Every diophantine set is recursively enumerable.*

Proof. Let S be a diophantine set, given by the equation $D(a_1, \dots, a_n, x_1, \dots, x_{m+1}) = 0$. The statement is intuitively clear: given any tuple (a_1, \dots, a_n) , we can check whether $(a_1, \dots, a_n) \in S$ by computing the value of $D(a_1, \dots, a_n, x_1, \dots, x_{m+1})$ for all $(x_1, \dots, x_{m+1}) \in \mathbb{N}^{m+1}$. If we find (x_1, \dots, x_{m+1}) such that $D = 0$, then the machine stops. Otherwise it runs forever, trying all possible solutions.

This idea can be formalized as follows:

First we construct a Turing machine M_1 that halts in a final state on input (a_1, \dots, a_n, y_0) if and only if y_0 is the Cantor number of a tuple (x_1, \dots, x_{m+1}) that does not satisfy the equation.

The machine `DECODE` adds to the tuple (a_1, \dots, a_n, y_0) the pair (x_1, y_1) with Cantor₂-number y_0 . Repeating this m times, we get a tuple $(a_1, \dots, a_n, y_0, x_1, y_1, \dots, x_m, y_m)$ such that $y_{k-1} = \text{Cantor}_2(x_k, y_k)$. By definition of the Cantor code, this implies $y_0 = \text{Cantor}_{m+1}(x_1, \dots, x_m, y_m)$.

Split the polynomial D into the parts C_L with positive coefficients and C_R with negative coefficients, as in section 3.1. Then $D(a_1, \dots, a_n, x_1, \dots, x_{m+1}) = 0$ is solvable if and only if $C_L(a_1, \dots, a_n, x_1, \dots, x_m, y_m) = C_R(a_1, \dots, a_n, x_1, \dots, x_m, y_m)$ is solvable. The computation of $C_L(a_1, \dots, a_n, x_1, \dots, x_m, y_m)$ and $C_R(a_1, \dots, a_n, x_1, \dots, x_m, y_m)$ consists of a finite number of additions and multiplications. Hence it can be done by repeated use of the machines `LEFT`, `RIGHT`, `NEW`, `INC`, `FIND`, `ADD` and `MULT`. After applying these operations, the contents of tape will be $(a_1, \dots, a_n, y_0, x_1, y_1, \dots, x_m, y_m, 1, z_1, \dots, z_k)$, where $C_L(a_1, \dots, a_n, x_1, \dots, x_m, y_m) = z_i$ and $C_R(a_1, \dots, a_n, x_1, \dots, x_m, y_m) = z_j$ for some i and j . Hence, if we apply `NOTEQUAL`(i, j), the machine will halt in a final state if and only if y_0 is the Cantor number of a tuple (x_1, \dots, x_{m+1}) that does not satisfy $D(a_1, \dots, a_n, x_1, \dots, x_{m+1}) = 0$. Otherwise the machine halts in a non-final state.

The machine M_2 , that is a repetition of $2m + k + 1$ times the machine `DELETE`, transforms the tuple $(a_1, \dots, a_n, y_0, x_1, \dots, x_m, y_m, 1, z_1, \dots, z_k)$ back into the tuple $(a_1, \dots, a_n, 0)$. Define the machine M by $M = \text{NEW}; \text{while } M_1 \text{ do } \{M_2; \text{INC}\}$. This machine behaves as follows: the input on the tape is (a_1, \dots, a_n) . `NEW` changes this into $(a_1, \dots, a_n, 0)$. If 0 is the Cantor number of a solution, then M_1 halts in a non-final state, so M halts in a final state. Otherwise, M_1 halts in a non-final state, so M_2 is applied, giving $(a_1, \dots, a_n, 0)$ again. `INC` changes this into $(a_1, \dots, a_n, 1)$. Now M_1 is applied again, so M halts in a final state if 1 is the Cantor number of a solution. Continuing like this, it is checked for all numbers whether they are the Cantor number of a solution. M halts in a final state if it finds a solution, and otherwise it will continue in an infinite loop. Hence S is recursively enumerable. \square

Theorem 3.25. *Every recursively enumerable subset of \mathbb{N}^n is diophantine.*

Proof. Let $S \subseteq \mathbb{N}^n$ be recursively enumerable, and suppose it is accepted by the Turing machine M , with alphabet $\{\gamma_1, \dots, \gamma_w\}$. At each step in the computation, only a finite initial segment of the tape is used, so we can represent the tape by a tuple (s_1, \dots, s_l) , with $s_i \in \{1, \dots, w\}$ (so if $s_i = t$, then the i^{th} position of the tape is occupied by γ_t). The position of the head and the current state can be represented by a tuple of the same length: the tuple $(0, \dots, 0, i, 0, \dots, 0)$ with the i at the j^{th} position means that the head scans the j^{th} cell, and the machine is in state q_i . Choose a base $\beta \geq \max\{3, v, w\}$, where v is the number of states of M , and let p and t be the ciphers to the base β of the tuples $(0, \dots, 0, i, 0, \dots, 0)$ and (s_1, \dots, s_l) , respectively. These uniquely determine the configuration of the machine (i.e. the state, the position of the head and the contents of the tape). (p, t) is called the *configuration code*.

We will construct a diophantine equation $D(p, t, x_1, \dots, x_m) = 0$, that, whenever (p, t) is configuration code, is solvable if and only if M halts when beginning in configuration (p, t) .

Define the functions `NextP`(p, t) and `NextT`(p, t) as follows: if (p, t) is the configuration code of a non-final state, let $(\text{NextP}(p, t), \text{NextT}(p, t))$ be the configuration code after performing the next step. When (p, t) represents a final state, let $(\text{NextP}(p, t), \text{NextT}(p, t)) = (0, t)$. When (p, t) is not a configuration code, we do not care about the value of $(\text{NextP}(p, t), \text{NextT}(p, t))$. These functions simulate a single step of the Turing machine. We claim that `NextP` and `NextT` are diophantine.

Recall the function $\delta(q_i, \gamma_{i'}) = (\delta_1(q_i, \gamma_{i'}), \delta_2(q_i, \gamma_{i'}), \delta_3(q_i, \gamma_{i'})) = (q_j, \gamma_{j'}, D)$. Extend this function by putting $\delta(q_i, \gamma_{i'}) = (q_R, \gamma_{i'}, L)$ and $\delta(q_R, \gamma_{i'}) = (q_a, \gamma_{i'}, R)$, where q_a and q_R are new states, when q_i is a final state. This means that the machine goes to a new state, without moving or writing. Extend δ even further by putting $\delta(q_i, \gamma_{i'}) = (q_R, \gamma_{i'}, L)$ for all $(q_i, \gamma_{i'})$ such that $\delta(q_i, \gamma_{i'})$ was not defined already. Define $\Delta_2(i, j) = \delta_2(q_i, \gamma_j)$. Δ_2 can be extended to tuples. This means that $\Delta_2[\beta](p, t; l)$ is the cipher to the base β of the tuple $(\Delta_2(p_1, t_1), \dots, \Delta_2(p_l, t_l))$ when p and t are the ciphers of (p_1, \dots, p_l) and (t_1, \dots, t_l) . Since $p_j = 0$ for all j , except when the current state is q_j , the new tuple $(\Delta_2(p_1, t_1), \dots, \Delta_2(p_l, t_l))$ is the same as the old tuple represented by t , except at the position of the current state. There the symbol is changed into the next symbol. Hence $\Delta_2[\beta](p, t, l)$ represents the content of the tape, after applying one step. This gives the

following equivalence:

$$t' = \text{NextT}(p, t) \iff \exists l : t' = \Delta_2[\beta](p, t; l)$$

Since $\delta_2[\beta]$ can be given by a finite table, it can also be given by a polynomial (by giving a disjunction of all possibilities and transforming this into a polynomial). To show that NextP is diophantine, define $p^R = p\beta$, $p^L = \left\lfloor \frac{p}{\beta} \right\rfloor$, $t^R = t\beta$ and $t^L = \left\lfloor \frac{t}{\beta} \right\rfloor$. If p is the cipher of (p_1, \dots, p_l) , then p^R is the cipher of $(0, p_1, \dots, p_l)$ and p^L is the cipher of (p_2, \dots, p_l) , and similarly for t . Define the function DQ by

$$\text{DQ}(i^L, i^R, j^L, j^R) = \begin{cases} \Delta_2(i^L, j^L) & \text{if } i^L > 0, i^R = 0, \delta_3(q_{i^L}, \gamma_{j^L}) = L \\ \Delta_2(i^R, j^R) & \text{if } i^L = 0, i^R > 0, \delta_3(q_{i^R}, \gamma_{j^R}) = R \\ 0 & \text{otherwise} \end{cases}$$

Then, if $p = (p_1, \dots, p_l)$ and $t = (t_1, \dots, t_l)$, we get that $\text{DQ}[\beta] \left(\left\lfloor \frac{p}{\beta} \right\rfloor, p, p\beta, \left\lfloor \frac{t}{\beta} \right\rfloor, t, t\beta; l \right)$ is the cipher to the base β of $(\text{DQ}(p_2, p_1, 0, t_2, t_1, 0), \text{DQ}(p_3, p_2, 0, p_1, t_3, t_2, t_1), \dots, \text{DQ}(0, p_l, p_{l-1}, 0, t_l, t_{l-1}), \text{DQ}(0, 0, p_l, 0, 0, t_l))$. DQ takes the element that represents the position of the head and moves the head accordingly. For example, if the head is in position 2 and state q_i , then all elements of the above tuple are zero, except for $\text{DQ}(i, 0, 0, t_2, t_1, 0)$, $\text{DQ}(0, i, 0, t_3, t_2, t_1)$ and $\text{DQ}(0, 0, i, t_4, t_3, t_2)$. If $\delta_3(q_i, \gamma_{t_2}) = L$, then only $\text{DQ}(0, 0, i, t_4, t_3, t_2) = Q(i, t_2)$ is nonzero. Hence $\text{DQ}[\beta] \left(\left\lfloor \frac{p}{\beta} \right\rfloor, p, p\beta, \left\lfloor \frac{t}{\beta} \right\rfloor, t, t\beta; l \right)$ is the cipher to the base β of $(Q(i, t_2), 0, 0, 0, \dots, 0)$, so the tuple that represents the new position is $(Q(i, t_2), 0, 0, 0, \dots, 0)$, while the old position was $(0, i, 0, 0, \dots, 0)$. Hence the function DQ indeed represents the changes in the position and the state. This implies that NextP is diophantine (since DQ can also be given by a finite table):

$$p' = \text{NextP}(p, t) \iff \exists l : p' = \text{DQ}[\beta] \left(\left\lfloor \frac{p}{\beta} \right\rfloor, p, p\beta, \left\lfloor \frac{t}{\beta} \right\rfloor, t, t\beta; l \right)$$

Using NextT and NextP , it is easy to simulate k steps of the Turing machine:

$$\begin{aligned} \text{AfterP}(0, p, t) &= p \\ \text{AfterT}(0, p, t) &= t \\ \text{AfterP}(k+1, p, t) &= \text{NextP}(\text{AfterP}(k, p, t), \text{AfterT}(k, p, t)) \\ \text{AfterT}(k+1, p, t) &= \text{NextT}(\text{AfterP}(k, p, t), \text{AfterT}(k, p, t)) \end{aligned}$$

We claim that the functions AfterP and AfterT are also diophantine.

Suppose that the machine goes from the configuration (p, t) to the configuration (p', t') in k steps. Then $p' = \text{AfterP}(k, p, t)$ and $t' = \text{AfterT}(k, p, t)$. Consider the sequence of configurations $(p_0, t_0), (p_1, t_1), \dots, (p_k, t_k)$ with $(p_0, t_0) = (p, t)$ and $(p_{i+1}, t_{i+1}) = (\text{NextP}(p_i, t_i), \text{NextT}(p_i, t_i))$, so $(p_k, t_k) = (p', t')$. Choose l such that $p, t < \beta^{l-k-2}$. Then in the configuration (p, t) , only the first $l-k-2$ cells are nonempty. In each step, only one symbol can be written, so $p_i, t_i < \beta^{l-2}$ and in particular

$$p', t' < \beta^{l-2} \tag{3.2}$$

Define superconfigurations (p_L, t_L) and (p_R, t_R) by concatenating configurations, which is a diophantine operation (+ denotes concatenation):

$$\begin{aligned} (p_L, \beta, kl) &= (p_0, \beta, l) + \dots + (p_{k-1}, \beta, l) \\ (t_L, \beta, kl) &= (t_0, \beta, l) + \dots + (t_{k-1}, \beta, l) \\ (p_R, \beta, kl) &= (p_1, \beta, l) + \dots + (p_k, \beta, l) \\ (t_R, \beta, kl) &= (t_1, \beta, l) + \dots + (t_k, \beta, l) \end{aligned}$$

Notice that these superconfigurations are not configurations of M , since they represents tuples in which the head is in k positions and k states, and the tape is divided into k parts. However,

it does hold that NextP and NextT give the next superconfiguration, since they give the next configuration for each of the k parts of the tape:

$$p_R = \text{NextP}(p_L, t_L) \quad \text{and} \quad t_R = \text{NextT}(p_L, t_L) \quad (3.3)$$

Define the superconfiguration (p_M, t_M) by

$$\begin{aligned} (p_M, \beta, (k-1)l) &= (p_1, \beta, l) + \dots + (p_{k-1}, \beta, l) \\ (t_M, \beta, (k-1)l) &= (t_1, \beta, l) + \dots + (t_{k-1}, \beta, l) \end{aligned}$$

Then, by recalling that $(p_0, t_0) = (p, t)$, it is clear that the following system of equations holds:

$$(p_L, \beta, kl) = (p, \beta, l) + (p_M, \beta, (k-1)l) \quad (3.4a)$$

$$(t_L, \beta, kl) = (t, \beta, l) + (t_M, \beta, (k-1)l) \quad (3.4b)$$

$$(p_R, \beta, kl) = (p_M, \beta, (k-1)l) + (p', \beta, l) \quad (3.4c)$$

$$(t_R, \beta, kl) = (t_M, \beta, (k-1)l) + (t', \beta, l) \quad (3.4d)$$

We will now show that these four equations together with the equations (3.2) and (3.3) uniquely determine $p_L, p_R, p_M, t_M, p_R, t_R, p'$ and t' , given k, l, p, t .

The equations (3.4a) and (3.4b) uniquely determine the first l elements of (p_L, β, kl) and (t_L, β, kl) . Hence, by equation (3.3) and the definition of NextP and NextT, the first $l-1$ elements of (p_R, β, kl) and (t_R, β, kl) are uniquely determined. This uniquely determines the first $l-1$ elements of $(p_M, \beta, (k-1)l)$ and $(t_M, \beta, (k-1)l)$ by the equations (3.4c) and (3.4d). But now, by (3.4a) and (3.4b), the first $2l-1$ elements of (p_L, β, kl) and (t_L, β, kl) are uniquely determined. Repeating this reasoning n times, we get that the first $n(l-1)+1$ elements of (p_L, β, kl) and (t_L, β, kl) , and the first $n(l-1)$ elements of (p_R, β, kl) , (t_R, β, kl) , $(p_M, \beta, (k-1)l)$ and $(t_M, \beta, (k-1)l)$ are uniquely determined. Hence, for n large enough, we find that p_L, t_L, p_M and t_M are unique, and that all elements of (p_R, β, kl) , (t_R, β, kl) , (p', β, l) and (t', β, l) , except for the last one, are unique. But by inequality (3.2) these last elements must be zero. Hence the system of equations has a unique solution.

Therefore, this system of equations gives a diophantine definition of $p_L, p_R, p_M, t_M, p_R, t_R, p'$ and t' , so AfterP and AfterT are diophantine.

Let $\omega_1, \dots, \omega_z$ be the subscripts of the final states of M . Then M halts eventually, when starting in the configuration (p, t) , if there exists a number of steps k after which the state is $q_{\omega_1}, q_{\omega_2}, \dots$, or q_{ω_z} . This is a diophantine relation:

$$\text{Halts}(p, t) \iff \exists k, r : \text{Elem}(\text{AfterT}(k, p, t), \beta, r) = \omega_1 \vee \dots \vee \text{Elem}(\text{AfterT}(k, p, t), \beta, r) = \omega_z$$

Hence there is a corresponding diophantine equation $D = 0$, with parameters p and t .

Let $\gamma_\kappa = \blacksquare$, $\gamma_\mu = 0$ and $\gamma_\nu = 1$. Suppose that the machine starts in state q_0 and position 1, with a_1, \dots, a_n written on the tape. Then $p = 1$, and the tape looks like

$$\blacksquare, 0, \overbrace{1, \dots, 1}^{a_1}, 0, \dots, 0, \overbrace{1, \dots, 1}^{a_n}$$

so, with $a = a_1 + \dots + a_n + n + 1$,

$$(t, \beta, a) = (\kappa, \beta, 1) + (\mu, \beta, 1) + (\text{Repeat}(\nu, \beta, a_1), \beta, a_1) + \dots + (\mu, \beta, 1) + (\text{Repeat}(\nu, \beta, a_n), \beta, a_n)$$

where $\text{Repeat}(\nu, \beta, a_i)$ is the cipher of the tuple (ν, \dots, ν) of length a_i to the base β . Hence M halts on input (a_1, \dots, a_n) if and only if the equation $D = 0$ has a solution for $p = 1$ and this value of t . This shows that S is diophantine. \square

Proof of theorem 3.23. This follows from theorem 3.24 and 3.25. \square

To prove that Hilbert's tenth problem is undecidable, we need one more lemma:

Lemma 3.26. *If $S \subseteq \mathbb{N}^n$ is recursive, then \overline{S} is recursively enumerable.*

Proof. Let S be recursive, and suppose it is accepted by the machine M , which stops on all input values. Consider the machine $M_1 = \mathbf{while} \ M \ \mathbf{do} \ \mathbf{NEVERSTOP}$. This machine halts if and only if M halts in a non-final state, i.e. if and only if $(a_1, \dots, a_n) \notin S$. \square

Theorem 3.27. *The set of codes of all solvable parameter-free diophantine equations is undecidable, i.e. there is no Turing machine that decides whether a given equation has an integral solution.*

Proof. In definition 3.15, this set was denoted by \mathfrak{h}_0 . By theorem 3.16, it is a diophantine set while its complement is non-diophantine. Suppose that \mathfrak{h}_0 is decidable. Then by lemma 3.26, its complement is recursively enumerable. Hence by theorem 3.25, $\overline{\mathfrak{h}_0}$ is diophantine. Contradiction, so \mathfrak{h}_0 is not decidable. \square

Chapter 4

Languages with divisibility relations

4.1 Positive existential models

The fact that Hilbert's tenth problem is undecidable means that the diophantine theory of \mathbb{Z} in the language $(0, 1; +, \cdot)$ is undecidable. By lemma 1.14, this is equivalent to the statement that the positive existential theory of \mathbb{Z} in $(0, 1; +, \cdot)$ is undecidable.

In the remaining of this thesis, we will show that the positive existential or diophantine theory of some other sets and languages is undecidable as well.

The idea of the proofs will be to 'embed' the positive existential theory of \mathbb{Z} in $(0, 1; +, \cdot)$ in the positive existential theory of another set S in a language L , such that positive existential formulas over \mathbb{Z} can be 'translated' into positive existential formulas over S in L . It follows that the positive existential theory of S in L is also undecidable.

Of course, this can also be done for other theories than the positive existential theory of \mathbb{Z} in $(0, 1; +, \cdot)$, as long as we know that it is undecidable.

We will first make this notion of embedding more precise, and show that the existence of such an embedding indeed implies that the positive existential theory of S in L is undecidable. The remaining of this chapter will be devoted to proving that the positive existential theories of \mathbb{Z} in $(0, 1; +; |^n)$, \mathbb{Z} in $(0, 1; +; |_p)$ and \mathbb{N} in $(0, 1; +; |_p)$ are undecidable. These results are not only interesting by themselves, but will also be useful in later chapters to show that Hilbert's tenth problem over polynomial rings and some function fields is undecidable.

Definition 4.1. Let two pairs (S_1, L_1) and (S_2, L_2) be given, consisting of a set and a language as in definition 1.2. Let the set of constant of L_1 be V_1 , the set of functions symbols $\{f_1, \dots, f_l\}$ and the set of relation symbols $\{r_1, \dots, r_m\}$. A *positive existential model* of (S_1, L_1) in (S_2, L_2) is given by an injective map $\phi : S_1 \rightarrow S_2^n$, with $n \geq 1$, such that ϕ is computable by a Turing machine and the sets $\phi(S_1)$,

$$\phi(f_i) := \{(\phi(x_1), \dots, \phi(x_{k_i}), \phi(f_i(x_1, \dots, x_{k_i}))) \mid x_1, \dots, x_{k_i} \in S_1\}$$

and

$$\phi(r_j) := \{(\phi(x_1), \dots, \phi(x_{k_j})) \mid x_1, \dots, x_{k_j} \in S_1, r_j(x_1, \dots, x_{k_j}) \text{ holds}\}$$

are positive existential over S_2 in L_2 (for all i, j).

For example, if S_1 is a subset of S_2 and $L_1 = L_2$, we can take $\phi = \text{Id}_{S_1}$. Then we only have to show is that S_1 is positive existentially definable over S_2 . In all cases we will consider, we take $S_1 = \mathbb{Z}$ or \mathbb{N} . As a model, we will sometimes just use \mathbb{Z} or \mathbb{N} , but also the solutions of the Pell equation or points on an elliptic curve.

Theorem 4.2. *Let (S_1, L_1) and (S_2, L_2) be as in definition 4.1. If there exists a positive existential model of (S_1, L_1) in (S_2, L_2) and the positive existential theory of S_1 in L_1 is undecidable, then the positive existential theory of S_2 in L_2 is also undecidable.*

Proof. Suppose that there exists a positive existential model, given by $\phi : S_1 \rightarrow S_2$, and let a positive existential formula ψ over S_1 in L_1 be given. ψ is equivalent to a formula $\exists x_1, \dots, x_n \in S_1 : \chi(x_1, \dots, x_m)$ where χ is built up from conjunctions and disjunction of formulas of the form $y = x$, $y = f_i(x_1, \dots, x_{k_i})$, and $r_j(x_1, \dots, x_{k_j})$. Since ϕ is injective, the first formula is equivalent to $\phi(y) = \phi(x)$. The second formula is equivalent to $(\phi(x_1), \dots, \phi(x_{k_i}), \phi(y)) \in \phi(f_i)$, and the third formula is equivalent to $(\phi(x_1), \dots, \phi(x_{k_j})) \in \phi(r_j)$. Since ϕ is computable, these formulas are all computable, and they are positive existential. Hence we obtain a positive existential formula $\chi'(\phi(x_1), \dots, \phi(x_m))$, such that $\exists x_1, \dots, x_n \in S_1 : \chi(x_1, \dots, x_m)$ is equivalent to $\exists x_1, \dots, x_n \in S_1 : \chi'(\phi(x_1), \dots, \phi(x_m))$. Finally, since $\phi(S_1)$ is positive existential over S_2 in L_2 , the formula

$$\exists y_1, \dots, y_n \in S_2 : y_1 \in \phi(S_1) \wedge \dots \wedge y_m \in \phi(S_1) \wedge \chi'(y_1, \dots, y_m)$$

is positive existential over S_2 in L_2 , and is equivalent to ψ .

Hence for every positive existential formula over S_1 in L_1 , we can compute an equivalent positive existential formula over S_2 in L_2 . So if the positive existential theory of S_2 in L_2 is decidable, then we can decide the positive existential theory of S_1 in L_1 by deciding the equivalent formulas over S_2 . It follows that if the positive existential theory of S_1 in L_1 is undecidable, then the positive existential theory of S_2 in L_2 is also undecidable. \square

4.2 The theories of \mathbb{Z} in $(0, 1; +; |^n)$ and \mathbb{Z} in $(0, 1; +; |, |_p)$

As an example how to use positive existential models, we will prove the undecidability of the positive existential theories of \mathbb{Z} in $(0, 1; +; |^n)$ and \mathbb{Z} in $(0, 1; +; |, |_p)$. The second one will be used in the proof of the undecidability of Hilbert's tenth problem for polynomial rings of positive characteristic in sections 5.2 and 5.3. The proofs are based on [Den79].

Definition 4.3. By $|^n$ we denote divisibility in $\mathbb{Z}[\frac{1}{n}]$, i.e.

$$x|^n y \iff \exists q \in \mathbb{Z}, \exists r \in \mathbb{N} : yn^r = xq$$

Furthermore, for a prime number p , we define $|_p$ by

$$x|_p y \iff \exists r \in \mathbb{N} : y = \pm xp^r$$

Remark 4.4. Note that if $\gcd(x, n) = 1$, then $x|^n y \iff x|y$.

Theorem 4.5. *Let $n > 1$. Then the positive existential theory of \mathbb{Z} in $(0, 1; +; |^n)$ is undecidable; i.e. there is no algorithm to decide whether formulas of the form*

$$\exists x_1, \dots, x_m \in \mathbb{Z} : \bigwedge_{i=1}^s \bigvee_{j=1}^{t_i} \phi_{i,j}(x_1, \dots, x_m)$$

hold, where $\phi(i, j)$ is either of the form $F_{i,j}(x_1, \dots, x_m) = G_{i,j}(x_1, \dots, x_m)$, with $F_{i,j}$ and $G_{i,j}$ constant or linear polynomials over \mathbb{Z} , or of the form $F_{i,j}(x_1, \dots, x_m) |^n G_{i,j}(x_1, \dots, x_m)$.

Corollary 4.6. *Let p be a prime number. Then the positive existential theory of \mathbb{Z} in the language $(0, 1; +; |, |_p)$ is undecidable.*

Proof of corollary 4.6. We construct a positive existential model of $(\mathbb{Z}, (0, 1; +; |^p))$ in $(\mathbb{Z}, (0, 1; +; |, |_p))$.

The embedding is simply $\phi : \mathbb{Z} \rightarrow \mathbb{Z} : n \mapsto n$. Then $+$ corresponds to $+$, or more precisely

$$\{(\phi(x), \phi(y), \phi(x+y)) \mid x, y \in \mathbb{Z}\} = \{(x, y, z) \in \mathbb{Z}^3 \mid z = x+y\}$$

which is clearly diophantine. It remains to show that $\{(\phi(x), \phi(y)) \in \mathbb{Z}^2 \mid x|^p y\}$ is positive existential over \mathbb{Z} in $(0, 1; +; |, |_p)$. We have the following equivalences:

$$\begin{aligned} x|^p y &\iff \exists q \in \mathbb{Z}, \exists r \in \mathbb{N} : yp^r = xq \iff \exists r \in \mathbb{N} : x|yp^r \iff \\ &\iff \exists z \in \mathbb{Z}, \exists r \in \mathbb{N} : x|z \wedge z = \pm yp^r \iff \exists z \in \mathbb{Z} : x|z \wedge y|_p z \end{aligned}$$

Hence

$$\{(\phi(x), \phi(y)) \in \mathbb{Z}^2 \mid x|_p y\} = \{(x, y) \in \mathbb{Z}^2 \mid \exists z \in \mathbb{Z} : x|z \wedge y|_p z\}$$

This is a positive existential set over \mathbb{Z} in $(0, 1; +; |, |_p)$.

Now the statement follows from theorems 4.5 and 4.2. \square

In the remaining of this section, we will prove theorem 4.5. We will construct a model of $(\mathbb{Z}, (0, 1; +, \cdot))$ in $(\mathbb{Z}, (0, 1; +; |^n))$ by putting $\phi : \mathbb{Z} \rightarrow \mathbb{Z} : n \mapsto n$. The only hard part is to show that \cdot is given by a positive existential formula over \mathbb{Z} in $(0, 1; +; |^n)$.

Definition 4.7. Let n, x and y be fixed integers with $n > 1$. For prime numbers p , we define $h(p)$ by $h(p) = 0$ if $v_p(ny) = v_p(x)$, and $h(p) = 1$ otherwise.

Lemma 4.8. Suppose that $h \equiv h(p) \pmod{p}$, and let $i > 0$. If $p^i | ny - hx$, then $p^i | x$.

Proof. Suppose that $h(p) = 0$. Then $p|h$, and $v_p(x) = v_p(ny)$. This implies $v_p(hx) > v_p(ny)$, so $p^i | ny - hx$ doesn't hold for $i \geq v_p(ny) + 1$. Hence from $p^i | ny - hx$ it follows that $i \leq v_p(ny) = v_p(x)$, so $p^i | x$ holds.

On the other hand, suppose $h(p) = 1$. Then $v_p(ny) \neq v_p(x) = v_p(hx)$, so $p^i | ny - hx$ implies $i \leq \min(v_p(ny), v_p(hx)) = \min(v_p(ny), v_p(x)) \leq v_p(x)$. Hence $p^i | x$. \square

Lemma 4.9. Let $n > 1$, and let x, y be integers such that $x|^{n-1}$ and $y|^{n-1}$. Then $y = x^2$ if and only if the following conditions hold:

$$2nx + 1 | 4n^2y - 1 \tag{4.1a}$$

$$2nx - 1 | 4n^2y - 1 \tag{4.1b}$$

$$ny - kx |^{n-1} nx - k \quad \text{for all } k \text{ with } |k| < n. \tag{4.1c}$$

Proof. First we prove the ‘only if’ part. Suppose that $y = x^2$. Then $4n^2y - 1 = 4n^2x^2 - 1 = (2nx + 1)(2nx - 1)$ so $2nx + 1 | 4n^2y - 1$ and $2nx - 1 | 4n^2y - 1$. Furthermore, $ny - kx = x(nx - k)$ and $x|^{n-1}$, so $ny - kx |^{n-1} nx - k$ for all k .

Now we prove the ‘if’ part. Suppose that the conditions hold. Since $\gcd(2nx \pm 1, n) = 1$, remark 4.4 gives $2nx + 1 | 4n^2y - 1$ and $2nx - 1 | 4n^2y - 1$. From $\gcd(2nx + 1, 2nx - 1) = \gcd(2nx - 1, 2) = 1$ it follows that $(2nx + 1)(2nx - 1) | 4n^2y - 1$, i.e. $4n^2x^2 - 1 | 4n^2y - 1$. Since $4n^2y \neq 1$, this gives $4n^2x^2 - 1 \leq |4n^2y - 1| \leq 4n^2|y| + 1$, so $x^2 \leq |y| + \frac{1}{2n^2}$. Since x and y are integers, this implies that $x^2 \leq |y|$.

For all primes p dividing n , consider $h(p)$. By the Chinese remainder theorem, there exists $|h| < n$ such that $h \equiv h(p) \pmod{p}$ for all these p . By replacing h by $-h$ if necessary, we may assume $hx \geq 0$. Equation (4.1c) implies that $ny - hx |^{n-1} nx - h$, so there exist $q \in \mathbb{Z}$ and $r \in \mathbb{N}$ such that $(ny - hx)q = (nx - h)n^r$.

Let p be a prime, and let i be such that $p^i | ny - hx$. If p divides n , then we obtain $p^i | x$ from lemma 4.8, so in particular $p^i | x(nx - h)$. For primes p that don't divide n , $p^i | ny - hx$ and $(ny - hx)q = (nx - h)n^r$ together imply $p^i | nx - h$, so $p^i | x(nx - h)$. This shows that $ny - hx | x(nx - h)$, so $|ny - hx| \leq |x(nx - h)|$.

Since $|h| < n$ and $x \neq 0$, we have $x(nx - h) > 0$, and $hx \geq 0$, so

$$n|y| - hx = n|y| - |hx| \leq |ny - hx| \leq |x(nx - h)| = x(nx - h) = nx^2 - hx$$

Hence $n|y| \leq nx^2$, so $|y| \leq x^2$.

Together with $x^2 \leq |y|$, this gives $|y| = x^2$, so $y = \pm x^2$. If $y = -x^2$, then from $2nx + 1 | 4n^2y - 1$ it follows that $2nx + 1 | -4n^2x^2 - 1$, so $2nx + 1 | (-4n^2x^2 - 1) + (4n^2x^2 - 1)$, i.e. $2nx + 1 | 2$. Contradiction, so $y = x^2$. \square

Lemma 4.10. *Let $n > 1$ and $x, u, z \in \mathbb{Z}$. Suppose that the following conditions hold:*

$$nz + nx - 1 \mid n^2u - (nx - 1)^2 \quad (4.2a)$$

$$2nz + 1 \mid nx - 1 \quad (4.2b)$$

$$2nz - 1 \mid nx - 1 \quad (4.2c)$$

$$2n^2u + 1 \mid nx - 1 \quad (4.2d)$$

Then $u = z^2$.

Proof. Since $\gcd(nz + nx - 1, n) = \gcd(2nz \pm 1, n) = \gcd(2n^2u + 1, n) = 1$, it follows from remark 4.4 that $nz + nx - 1 \mid n^2u - (nx - 1)^2$, $2nz + 1 \mid nx - 1$, $2nz - 1 \mid nx - 1$ and $2n^2u + 1 \mid nx - 1$. Hence $n^2u - (nx - 1)^2 \equiv 0 \pmod{nz + nx - 1}$, and clearly $nx - 1 \equiv -nz \pmod{nz + nx - 1}$, so $n^2u - (-nz)^2 \equiv 0 \pmod{nz + nx - 1}$, so $nz + nx - 1 \mid n^2u - n^2z^2$. Suppose that $u \neq z^2$. Then

$$|nx - 1| - n|z| \leq |nz - nx - 1| \leq |n^2u - n^2z^2| \leq n^2|u| + n^2z^2$$

Since $\gcd(2nz + 1, 2nz - 1) = 1$, we have $(2nz + 1)(2nz - 1) \mid nx - 1$, i.e. $4n^2z^2 - 1 \mid nx - 1$. $nx - 1 \neq 0$ because $n \neq 1$, so $4n^2z^2 - 1 \leq |nx - 1|$. Similarly, $2n^2|u| - 1 \leq |2n^2u + 1| \leq |nx - 1|$. This gives:

$$\begin{aligned} (n|z|)^2 - n|z| - 1 &= |nx - 1| - n|z| + n^2z^2 - 1 - |nx - 1| \leq n^2|u| + 2n^2z^2 - 1 - |nx - 1| \leq \\ &\frac{1}{2}|nx - 1| + 2n^2z^2 - \frac{1}{2} - |nx - 1| = 2n^2z^2 - \frac{1}{2} - \frac{1}{2}|nx - 1| \leq \frac{1}{2}|nx - 1| - \frac{1}{2}|nx - 1| = 0 \end{aligned}$$

so $(n|z|)^2 - n|z| - 1 \leq 0$. Since the polynomial $t^2 - t - 1$ has roots $t = \frac{1 \pm \sqrt{5}}{2}$, this inequality implies

$$-1 < \frac{1 - \sqrt{5}}{2} \leq n|z| \leq \frac{1 + \sqrt{5}}{2} < 2$$

But since $n > 1$, we either have $n|z| > 2$, or $z = 0$. Hence $z = 0$.

Now $|nx - 1| - n|z| \leq n^2|u| + n^2z^2$ implies $|nx - 1| \leq n^2|u|$, and $2n^2|u| - 1 \leq |nx - 1|$ gives $n^2|u| \leq \frac{1}{2}(|nx - 1| + 1)$. It follows that $n^2|u| \leq \frac{1}{2}(n^2|u| + 1)$, so $n^2|u| \leq 1$, so $u = 0$.

So either $u = z^2$, or $u = z = 0$, in which case $u = z^2$ also holds. \square

Lemma 4.11. *Let $n > 1$ and $d \in \mathbb{Z}_{\neq 0}$. Then there exists $x \in \mathbb{Z}$ such that $x \mid^n 1$ and $d \mid^n nx - 1$.*

Proof. Factor d as $d = p_1^{k_1} \dots p_r^{k_r} q_1^{l_1} \dots q_s^{l_s}$, where all p_i and q_j are distinct primes, $k_i, l_j > 0$, $p_i \mid n$ and $q_j \nmid n$ for all i, j . Define $d_0 = p_1^{k_1} \dots p_r^{k_r}$ and $d_1 = q_1^{l_1} \dots q_s^{l_s}$. Then $d = d_0 d_1$ and $d_0 \mid^n 1$. Furthermore, define $x = n^{\varphi(d_1) - 1}$. Then clearly $x \mid^n 1$. Since $\gcd(d_1, n) = 1$, by Euler's theorem we have $n^{\varphi(d_1)} \equiv 1 \pmod{d_1}$. This implies $nx - 1 = n^{\varphi(d_1)} - 1 \equiv 0 \pmod{d_1}$ so $d_1 \mid nx - 1$. Together with $d_0 \mid^n 1$, this gives $d \mid^n nx - 1$. \square

Lemma 4.12. *$u = z^2$ if and only if there are $x, y \in \mathbb{Z}$ such that $x \mid^n 1$, $y \mid^n 1$ and the equations (4.1) and (4.2) hold.*

Proof. Suppose that $u = z^2$. Take $d = (2nz + 1)(2nz - 1)(2n^2u + 1)$ and apply lemma 4.11 to obtain $x \in \mathbb{Z}$ such that $x \mid^n 1$ and $(2nz + 1)(2nz - 1)(2n^2u + 1) \mid^n nx - 1$. Then conditions (4.2b), (4.2c) and (4.2d) follow immediately. By taking $y = x^2$, the other conditions, except for (4.2a) are satisfied by lemma 4.9. Since $n^2u - n^2y + 2nx - 1 = n^2z^2 - n^2x^2 + 2nx - 1 = (nz)^2 - (nx - 1)^2 = (nz + nx - 1)(nz - nx + 1)$, condition (4.2b) is also satisfied.

Conversely, suppose that the conditions hold. Then by lemma 4.9, we have $y = x^2$. Substituting this into $nz + nx - 1 \mid n^2u - n^2y + 2nx - 1$ gives $nz + nx - 1 \mid n^2u - (nx - 1)^2$, so the conditions of lemma 4.11 are fulfilled. Hence $u = z^2$. \square

Proof of theorem 4.5. By lemma 4.12 there exist constant or linear polynomials A_i and B_i over \mathbb{Z} such that

$$u = z^2 \iff \exists x, y : \bigwedge_{i=1}^s A_i(x, y) \mid^n B_i(x, y)$$

Furthermore, we have the equivalences

$$\begin{aligned} z = x + y &\iff 0 \mid^n x + y - z \\ z = xy &\iff 4z = (x + y)^2 - (x - y)^2 \end{aligned}$$

Hence we can also define \cdot by such polynomials, i.e. there exist constant or linear C_i and D_i such that

$$z = xy \iff \exists x_1, \dots, x_m : \bigwedge_{i=1}^s C_i(x_1, \dots, x_m) \mid^n D_i(x_1, \dots, x_m)$$

This gives the desired model of $(\mathbb{Z}, (0, 1; +, \cdot))$ in $(\mathbb{Z}, (0, 1; +; |^n))$, so the positive existential theory of \mathbb{Z} in $(0, 1; +, |^n)$ is undecidable. \square

4.3 The theory of \mathbb{N} in $(0, 1; +; |_p)$

Corollary 4.6 gives that the positive existential theory of \mathbb{Z} in $(0, 1; +; |, |_p)$ is undecidable. It is a remarkable fact that if we change \mathbb{Z} to \mathbb{N} , we don't need $|$ anymore: the positive existential theory of \mathbb{N} in $(0, 1; +, |_p)$ is already undecidable. This theorem is due to Pheidas, and is proven in [Phe87]. In chapter 7 we will use this to prove that Hilbert's tenth problem for function fields over finite fields is undecidable.

Theorem 4.13. *The positive existential theory of \mathbb{N} in $(0, 1; +; |_p)$ is undecidable.*

Notice that in \mathbb{N} we have $x|_p y \iff \exists r \in \mathbb{N} : y = xp^r$ instead of $y = \pm xp^r$.

Similar to the proof of the undecidability of the positive existential theory of \mathbb{Z} in $(0, 1; +; |^n)$ from theorem 4.5, we will prove that \cdot is definable over \mathbb{N} in $(0, 1; +; |_p)$.

Lemma 4.14. *Suppose that $i, j, k, l \geq 1$, such that $(p^i - 1)(p^j - 1) = (p^k - 1)(p^l - 1)$. Then either $i = k$ and $j = l$, or $i = l$ and $j = k$.*

Proof. Without loss of generality, we can assume $j \leq i, k, l$ and $l \leq k$. From $p^{i+j} - p^i - p^j = p^{k+l} - p^k - p^l$ we obtain $p^i - p^{i-j} - 1 = p^{k+l-j} - p^{k-j} - p^{l-j}$. Modulo p , the left-hand side is either -1 or -2 , so the right-hand side must also be -1 or -2 modulo p . Hence at least one of $k-j$ and $l-j$ is zero. Since $l \leq k$, we must have $l = j$. This implies $p^i - p^{i-j} - 1 = p^k - p^{k-j} - 1$, so $p^i - p^{i-j} = p^k - p^{k-j}$. $p^i - p^{i-j}$ is $i-j$ times divisible by p , and $p^k - p^{k-j}$ is $k-j$ times divisible by p . Hence it must hold that $i-j = k-j$ so $i = k$. \square

Lemma 4.15. *Let $m, n > 0$ and $s \in \mathbb{N}$ be given. Then*

$$m = p^s n \iff n|_p m \wedge (n+1)|_p(m+p^s) \wedge (n+p)|_p(m+p^{s+1})$$

Given m, n and p^s , the relation $m = p^s n$ is positive existential over \mathbb{N} in the language $(0, 1; +; |_p)$.

Proof. The second statement follows immediately from the first, so we only have to prove the first statement.

If $m = p^s n$, then $m + p^s = p^s(n+1)$ and $m + p^{s+1} = p^s(n+p)$, so $n|_p m, (n+1)|_p(m+p^s)$ and $(n+p)|_p(m+p^{s+1})$.

Conversely, suppose that $n|_p m, (n+1)|_p(m+p^s)$ and $(n+p)|_p(m+p^{s+1})$. Then there exist $r, q, k \in \mathbb{N}$ such that $m = p^r n, m + p^s = p^q(n+1)$ and $m + p^{s+1} = p^k(n+p)$. Substituting $m = p^r n$ in the second and third equation gives $p^r n + p^s = p^q(n+1)$ and $p^r n + p^{s+1} = p^k(n+p)$. From this it

follows that $n = \frac{p^q - p^s}{p^r - p^q}$, so $p^r \frac{p^q - p^s}{p^r - p^q} + p^{s+1} = p^k \left(\frac{p^q - p^s}{p^r - p^q} + p \right)$ and hence $p^r(p^q - p^s) + p^{s+1}(p^r - p^q) = p^k(p^q - p^s) + p^{k+1}(p^r - p^q)$. This is equivalent to

$$(p^r - p^q)(p^{k+1} - p^{s+1}) = (p^q - p^s)(p^r - p^k) \quad (4.3)$$

There are three possibilities: both sides of the equation are zero, they are both positive, or they are both negative.

We start with the case that they are both positive. Then we have four possibilities:

1. $r > q$ and $k > s$ and $q > s$ and $r > k$
Then $(p^{r-q} - 1)(p^{k-s} - 1)p^{q+s+1} = (p^{q-s} - 1)(p^{r-k} - 1)p^{s+k}$. Since both sides have to be divisible by p the same number of times, we must have $q + s + 1 = s + k$, so $q + 1 = k$. Now by lemma 4.14, we either have $r - q = q - s$ and $k - s = r - k$, or $r - q = r - k$ and $k - s = q - s$. In the first case, we obtain $2q = r + s = 2k$, so $k = q$. In the second case we also get $k = q$. Both cases contradict $q + 1 = k$.
2. $r > q$ and $k > s$ and $q < s$ and $r < k$
Then $(p^{r-q} - 1)(p^{k-s} - 1)p^{q+s+1} = (p^{s-q} - 1)(p^{k-r} - 1)p^{q+r}$, so $q + s + 1 = q + r$ and hence $s + 1 = r$. Either $r - q = s - q$ and $k - s = k - r$, or $r - q = k - r$ and $k - s = s - q$. In both cases $r = s$, so again we have a contradiction.
3. $r < q$ and $k < s$ and $q > s$ and $r > k$
Then $(p^{q-r} - 1)(p^{s-k} - 1)p^{r+k+1} = (p^{q-s} - 1)(p^{r-k} - 1)p^{s+k}$, so $r + 1 = s$. This contradicts $q - r = q - s$ and $s - k = r - k$, and it also contradicts $q - r = r - k$ and $s - k = q - s$.
4. $r < q$ and $k < s$ and $q < s$ and $r < k$
Then $(p^{q-r} - 1)(p^{s-k} - 1)p^{r+k+1} = (p^{s-q} - 1)(p^{k-r} - 1)p^{q+r}$, so $k + 1 = q$. This contradicts $q - r = s - q$ and $s - k = k - r$, and it also contradicts $q - r = k - r$ and $s - k = s - q$.

Now suppose that both sides of the equation are negative. Again we have four possibilities:

1. $r > q$ and $k < s$ and $q > s$ and $r < k$
Then $r > q > s > k > r$, which is clearly a contradiction.
2. $r > q$ and $k < s$ and $q < s$ and $r > k$
Then $(p^{r-q} - 1)(p^{s-k} - 1)p^{q+k+1} = (p^{s-q} - 1)(p^{r-k} - 1)p^{q+k}$, so $q + k + 1 = 0$, which is clearly impossible.
3. $r < q$ and $k > s$ and $q > s$ and $r < k$
Then $(p^{q-r} - 1)(p^{k-s} - 1)p^{r+s+1} = (p^{q-s} - 1)(p^{k-r} - 1)p^{s+r}$, so $r + s + 1 = s + r$, which is also impossible.
4. $r < q$ and $k > s$ and $q < s$ and $r > k$
Then $r < q < s < k < r$, which is again a contradiction.

So all these possibilities cannot occur, and hence both sides have to be zero. It follows that $r = q$ or $k = s$. If $r = q$, then $m + p^s = p^r(n + 1)$ and $m = p^r n$, so $p^s = p^r$ and hence $m = p^s n$, as was to be shown. If $k = s$, then $m + p^{s+1} = p^s(n + p)$, and again $m = p^s n$. \square

Lemma 4.16. *Let $m, n \in \mathbb{N}$. Then*

$$m = n^2 \iff \exists s, r > 0 : n < p^s - 1 \wedge m < p^{2s} - 1 \wedge p^{2s} - 1 | p^r - 1 \wedge \frac{p^r - 1}{p^{2s} - 1} \equiv n \pmod{p^{2s} - 1} \wedge \left(\frac{p^r - 1}{p^{2s} - 1} \right)^2 \equiv m \pmod{p^{2s} - 1}$$

Proof. Suppose that $m = n^2$. Choose s such that $n < p^s - 1$, and put $r = 2ns$. Then $m = n^2 < (p^s - 1)^2 < p^{2s} - 1$, and

$$p^r - 1 = p^{2ns} - 1 = (p^{2s} - 1)(p^{2s(n-1)} + p^{2s(n-2)} + \dots + p^{2s} + 1)$$

so $p^{2s} - 1 | p^r - 1$. Furthermore,

$$\frac{p^r - 1}{p^{2s} - 1} = p^{2s(n-1)} + p^{2s(n-2)} + \dots + p^{2s} + 1 \equiv n \pmod{p^{2s} - 1}$$

since $p^{2s} \equiv 1 \pmod{p^{2s} - 1}$. Finally,

$$\left(\frac{p^r - 1}{p^{2s} - 1} \right)^2 \equiv n^2 = m \pmod{p^{2s} - 1}$$

For the converse direction, suppose that such s and r exist. Then

$$m \equiv \left(\frac{p^r - 1}{p^{2s} - 1} \right)^2 \equiv n^2 \pmod{p^{2s} - 1}$$

so $m \equiv n^2 \pmod{p^{2s} - 1}$. Since $0 \leq m, n^2 < p^{2s} - 1$, it follows that $m = n^2$. \square

Proof of theorem 4.13. By lemma 4.15 and 4.16, the relation $m = n^2$ is positive existential over \mathbb{N} in the language $(0, 1; +; |_p)$. Since $m = nk$ if and only if $(n+k)^2 = n^2 + 2m + k^2$, the relation $m = nk$ is also positive existential in this language.

This gives a positive existential model of $(\mathbb{N}, (0, 1; +, \cdot))$ in $(\mathbb{N}, (0, 1; +; |_p))$. Since the positive existential theory of \mathbb{N} in $(0, 1; +, \cdot)$ is undecidable by Hilbert's original tenth problem, it follows from theorem 4.2 that the positive existential theory of \mathbb{N} in $(0, 1; +; |_p)$ is also undecidable. \square

Chapter 5

Polynomial rings

In this chapter, we will prove Hilbert's tenth problem for polynomial rings $R[t]$ over integral domains R , with coefficients in $\mathbb{Z}[t]$. In the proof, the characteristic of the ring will be important. We will make a distinction between characteristic zero and positive characteristic, and divide this even further into odd characteristic and characteristic two. The proofs are based on [Den78] for rings of characteristic zero, and [Den79] for rings of positive characteristic.

Throughout this chapter, we will use the language $(0, 1, t; +, \cdot)$. Therefore, by 'diophantine' (or 'positive existential'), we will always mean 'diophantine (or positive existential) over $R[t]$ in the language $(0, 1, t; +, \cdot)$ '. Recall that all positive existential sets are diophantine by lemma 1.15.

5.1 Rings of characteristic zero

Theorem 5.1. *Let R be an integral domain of characteristic zero. Then there is no algorithm to decide whether a diophantine equation with coefficients in $\mathbb{Z}[t]$ has a solution in $R[t]$.*

In the proof of this theorem, we will use the *Pell equation*, i.e. the equation

$$x^2 - (v^2 - 1)y^2 = 1 \quad (5.1)$$

where $v \in R[t] \setminus R$. Let $u \in \overline{R[t]}$ satisfy $u^2 = v^2 - 1$. Notice that u is not an element of $R(t)$, since if it would be, then we could write $u = \frac{p}{q}$ with $p, q \in R[t]$ and $\gcd(p, q) = 1$. Since $u^2 = v^2 - 1$, this gives $\frac{p^2}{q^2} = v^2 - 1 \in R[t]$, so $q = 1$ and $u = p$. But then $(v + u)(v - u) = v^2 - u^2 = 1$, with $v + u, v - u \in R[t]$, so $v + u, v - u \in R$. This implies $2v \in R$, which clearly doesn't hold.

This makes it possible to give the following definition:

Definition 5.2. For $n \in \mathbb{Z}$, define $x_n(v), y_n(v) \in R[t]$ by

$$x_n(v) + uy_n(v) = (v + u)^n \quad (5.2)$$

Remark 5.3. Notice that $(v + u)(v - u) = 1$, so $(v + u)^{-n} = (v - u)^n$. Hence $x_{-n}(v) = x_n(v)$ and $y_{-n}(v) = -y_n(v)$.

We will often use that for all $n \geq 0$

$$x_n(t) + uy_n(t) = (t + u)^n = \sum_{i=0}^n \binom{n}{i} t^{n-i} u^i = \sum_{i=0}^n \binom{n}{i} t^{n-i} (t^2 - 1)^{\lfloor \frac{i}{2} \rfloor} u^{i-2\lfloor \frac{i}{2} \rfloor}$$

so

$$x_n(v) = \sum_{\substack{i=0 \\ i \text{ even}}}^n \binom{n}{i} v^{n-i} (v^2 - 1)^{\frac{i}{2}} \quad (5.3)$$

and

$$y_n(t) = \sum_{\substack{i=1 \\ i \text{ odd}}}^n \binom{n}{i} t^{n-i} (t^2 - 1)^{\lfloor \frac{i}{2} \rfloor} \quad (5.4)$$

We need some lemmas about the solutions of the Pell equation:

Lemma 5.4. *The solutions of the Pell equation (5.1) in $R[t]$ are $x = \pm x_n(v)$, $y = \pm y_n(v)$ for $n = 0, 1, 2, \dots$*

Proof. First we show that $(x_n(v), y_n(v))$ is indeed a solution. Since $(v+u)(v-u) = v^2 - u^2 = 1$, we have $x_n(v) - uy_n(v) = (v-u)^n = (v+u)^{-n}$. Hence

$$\begin{aligned} x_n(v)^2 - (v^2 - 1)y_n(v)^2 &= x_n(v)^2 - u^2y_n(v)^2 = \\ &= (x_n(v) + uy_n(v))(x_n(v) - uy_n(v)) = (v+u)^n(v+u)^{-n} = 1 \end{aligned}$$

so $(x_n(v), y_n(v))$ is a solution of the Pell equation.

Now we have to show that all solutions are of the form $x = \pm x_n(v)$ and $y = \pm y_n(v)$. We will only prove this for the special case $v = t$. This is the only case that is needed if the characteristic is zero. If the characteristic is odd, we will also need this for other v . For the proof of this case, we refer to [Den79], lemma 2.1, part 2.

So suppose that $v = t$, and assume that $x^2 - (t^2 - 1)y^2 = 1$. By lemma 5.5, we can parametrize the curve $u^2 = t^2 - 1$ by $t(s) = \frac{s^2+1}{s^2-1}$, $u(s) = \frac{2s}{s^2-1}$. Then t has poles only at $s = \pm 1$, and since $x, y \in R[t]$, the same holds for x and y . u also has poles only at $s = \pm 1$. This implies that the functions $x + uy$ and $x - uy$ only have poles at $s = \pm 1$, and since $(x - uy)(x + uy) = 1$, each of them can only have a zero at a pole of the other one, of the same order. It follows that there exist $m, n \in \mathbb{Z}$ such that

$$x + uy = c(s-1)^m(s+1)^n = c \left(\frac{2}{t+u-1} \right)^m \left(\frac{2(t+u)}{t+u-1} \right)^n = c \frac{2^{m+n}(t+u)^n}{(t+u-1)^{m+n}}$$

Since $x, y \in R[t]$ and $u \notin R[t]$, this needs to be a polynomial in t and u , so $m = -n$, and $x + uy = c(t+u)^n$. This implies that $x - uy = c(t-u)^n$. Substituting this into $(x - uy)(x + uy) = 1$ shows that $c^2 = 1$ so $c = \pm 1$.

Hence if $n \geq 0$, then $x + uy = \pm(t+u)^n = \pm(x_n(t) + uy_n(t))$, so $x = \pm x_n(t)$, $y = \pm y_n(t)$ (with the same sign), and if $n < 0$, then $x + uy = \pm(t+u)^n = \pm(t-u)^{|n|} = \pm(x_{|n|}(t) - uy_{|n|}(t))$, so $x = \pm x_{|n|}(t)$, $y = \pm y_{|n|}(t)$ (with different signs). \square

Lemma 5.5. *The curve $u^2 = t^2 - 1$ can be parametrized by $t(s) = \frac{s^2+1}{s^2-1}$, $u(s) = \frac{2s}{s^2-1}$.*

Proof. For any s , we have

$$t(s)^2 - u(s)^2 = \frac{s^4 + 2s^2 + 1}{(s^2 - 1)^2} - \frac{4s^2}{(s^2 - 1)^2} = \frac{s^4 - 2s^2 + 1}{(s^2 - 1)^2} = \frac{(s^2 - 1)^2}{(s^2 - 1)^2} = 1$$

so we indeed have $u(s)^2 = t(s)^2 - 1$.

The map $\varphi : s \mapsto (t(s), u(s)) = (1 + \frac{2}{s^2-1}, \frac{2s}{s^2-1})$ is a bijection, since it has an inverse, given by $\psi : (t, u) \mapsto \frac{t+u+1}{t+u-1}$. This is indeed an inverse, since

$$\psi(\varphi(s)) = \psi \left(\frac{s^2+1}{s^2-1}, \frac{2s}{s^2-1} \right) = \frac{\frac{s^2+2s+1}{s^2-1} + 1}{\frac{s^2+2s+1}{s^2-1} - 1} = \frac{\frac{s+1}{s-1} + 1}{\frac{s+1}{s-1} - 1} - 1 = \frac{2s}{2} = s$$

and

$$\left(\frac{t+u+1}{t+u-1} \right)^2 - 1 = \left(\frac{t+u+1}{t+u-1} + 1 \right) \cdot \left(\frac{t+u+1}{t+u-1} - 1 \right) = \frac{2t+2u}{t+u-1} \cdot \frac{2}{t+u-1} = \frac{4(t+u)}{(t+u-1)^2}$$

so

$$\begin{aligned} \varphi(\psi(t, u)) &= \varphi \left(\frac{t+u+1}{t+u-1} \right) = \left(\frac{4(t+u) + 2(t+u-1)^2}{4(t+u)}, \frac{2(t+u+1)(t+u-1)}{4(t+u)} \right) = \\ &= \left(\frac{1 + 2tu + t^2 + u^2}{2(t+u)}, \frac{t^2 + u^2 - 1 + 2tu}{2(t+u)} \right) = \left(\frac{2tu + 2t^2}{2(t+u)}, \frac{2u^2 + 2tu}{2(t+u)} \right) = (t, u) \end{aligned}$$

Hence this is a parametrization. \square

Lemma 5.6. *In $R[t]$, the congruence $x_n(v) \equiv 1 \pmod{v-1}$ holds.*

Proof. First suppose that $n \geq 0$. By equation (5.3) we have

$$x_n(v) = \sum_{\substack{i=0 \\ i \text{ even}}}^n \binom{n}{i} v^{n-i} (v^2 - 1)^{\frac{i}{2}} \equiv \sum_{\substack{i=0 \\ i \text{ even}}}^n \binom{n}{i} 1^{n-i} (1 - 1)^{\frac{i}{2}} \equiv \binom{n}{0} 1^n (1 - 1)^0 \equiv 1 \pmod{v-1}$$

If $n < 0$, then $x_n(v) = x_{-n}(v) \equiv 1 \pmod{v-1}$. □

By lemma 5.4 and lemma 5.6,

$$x^2 - (v^2 - 1)y^2 = 1 \wedge x \equiv 1 \pmod{v-1} \iff \exists n \in \mathbb{Z} : x = x_n(v) \wedge y = y_n(v)$$

Hence the relation $\exists m \in \mathbb{Z} : x = x_m(v) \wedge y = y_m(v)$ is diophantine. We can extend the definition of x_n and y_n to $v \in R$ using this diophantine representation:

Definition 5.7. For $v \in R[t]$, define the relation $\exists n \in \mathbb{Z} : x = x_n(v) \wedge y = y_n(v)$ to hold if and only if the relation $x^2 - (v^2 - 1)y^2 = 1 \wedge x \equiv 1 \pmod{v-1}$ holds.

In the proof of theorem 5.1, we will only use $x_n(t)$. The next definition and lemma deal with this special case.

Definition 5.8. For $v, w \in R[t]$, define the relation $v \sim w$ by $v \sim w \iff v(1) = w(1)$, and define the predicate $\text{Imt}(y)$ by

$$\text{Imt}(y) \iff y \in R[t] \wedge \exists x \in R[t] : x^2 - (t^2 - 1)y^2 = 1$$

Lemma 5.9. (i) *The predicate $v \sim 0$ is diophantine.*

(ii) *For $n \in \mathbb{N}$ we have $y_n(t) \sim n$.*

(iii) *The predicate $\text{Imt}(y)$ is diophantine.*

(iv) *If $\text{Imt}(y)$ holds, then there is an integer n such that $y \sim n$.*

(v) *For every integer n there exists $y \in R[t]$ such that $y \sim n$ and $\text{Imt}(y)$.*

Proof. (i) $v \sim 0$ if and only if 1 is a zero of v , i.e. $v \sim 0 \iff \exists w \in R[t] : v = (t-1)w$. It is clear that this is diophantine.

(ii) By substituting $t = 1$ into equation (5.4) we obtain

$$y_n(1) = \sum_{\substack{i=1 \\ i \text{ odd}}}^n \binom{n}{i} 1^{n-i} 0^{\lfloor \frac{i}{2} \rfloor} = \binom{n}{1} = n$$

so $y_n \sim n$.

(iii) This follows immediately from the definition.

(iv) Suppose that $\text{Imt}(y)$ holds. Then by lemma 5.4, there exists $n \geq 0$ such that $y = \pm y_n(t)$. By part (ii), $y_n(t) \sim n$, so $y \sim \pm y_n(t) \sim \pm n$ by part (ii).

(v) Choose $y = \varepsilon y_{|n|}(t)$, where ε denotes the sign of n . Then $y \sim \varepsilon(n)|n| = n$ and $\text{Imt}(y)$ holds by choosing $x = x_{|n|}$. □

Using these lemmas, we can prove that Hilbert's tenth problem for polynomial rings of characteristic zero is undecidable:

Proof of theorem 5.1. Suppose that there is an algorithm to decide whether a diophantine equation with coefficients in $\mathbb{Z}[t]$ has a solution in $R[t]$, and let $P(z_1, \dots, z_n)$ be a polynomial over \mathbb{Z} . By lemma 5.9, we have

$$\begin{aligned} \exists z_1, \dots, z_n \in \mathbb{Z} : P(z_1, \dots, z_n) = 0 &\iff \\ \exists Z_1, \dots, Z_n \in R[T] : \text{Imt}(Z_1) \wedge \dots \wedge \text{Imt}(Z_n) \wedge P(Z_1, \dots, Z_n) \sim 0 & \end{aligned}$$

Indeed, for the ‘if’ part, we can take $z_i \in R$ such that $Z_i \sim z_i$. Then $z_i \in \mathbb{Z}$ by lemma 5.9 (iv), and $P(z_1, \dots, z_n) = P(Z_1(1), \dots, Z_n(1)) = P(Z_1, \dots, Z_n)(1) = 0$. For the ‘only if’ part, take $Z_i \sim z_i$ satisfying $\text{Imt}(Z_i)$. These exist by lemma 5.9 (v). Then $P(Z_1, \dots, Z_n)(1) = P(Z_1(1), \dots, Z_n(1)) = P(z_1, \dots, z_n) = 0$.

Since we have diophantine representations of \sim and Imt , we can compute a polynomial P^* such that

$$\begin{aligned} \exists Z_1, \dots, Z_n \in R[T] : P^*(Z_1, \dots, Z_n) = 0 &\iff \\ \exists Z_1, \dots, Z_n \in R[T] : \text{Imt}(Z_1) \wedge \dots \wedge \text{Imt}(Z_n) \wedge P(Z_1, \dots, Z_n) \sim 0 & \end{aligned}$$

Hence

$$\exists z_1, \dots, z_n \in \mathbb{Z} : P(z_1, \dots, z_n) = 0 \iff \exists Z_1, \dots, Z_n \in R[T] : P^*(Z_1, \dots, Z_n) = 0$$

By deciding whether P^* has a zero in $R[t]$, we can decide whether P has a zero in \mathbb{Z} . Hence this gives an algorithm that decides whether a polynomial over \mathbb{Z} has a zero in \mathbb{Z} , which is a contradiction with the undecidability of Hilbert’s tenth problem over \mathbb{Z} . \square

5.2 Rings of odd characteristic

The previous theorem only dealt with rings of characteristic zero. Now we will prove that the same theorem holds, if R has characteristic p , with p an odd prime. Of course, the coefficients of the equations must be taken modulo p .

Theorem 5.10. *Let R be an integral domain of characteristic $p \geq 3$. Then there is no algorithm to decide whether a diophantine equation with coefficients in $(\mathbb{Z}/p\mathbb{Z})[t]$ has a solution in $R[t]$.*

We have the following lemma:

Lemma 5.11. *For all $n, m \in \mathbb{Z}$ and $v \in R[t]$, the following statements hold:*

- (i) *As polynomials in v , $x_n(v)$ has degree $|n|$ and $y_n(v)$ has degree $|n| - 1$ (if $n \neq 0$). In particular, if $n \neq 0$, then $x_n(t) \notin R$, and if $n \neq 0, \pm 1$, then $y_n(t) \notin R$.*
- (ii) *$x_0(v) = 1, y_0(v) = 0, x_1(v) = v$ and $y_1(v) = 1$.*
- (iii) *If $n \geq 0$, then $x_{mp^n}(v) = x_m(v)^{p^n}$ and $x_{p^n}(v) = v^{p^n}$.*
- (iv) *$x_m(v+1) = x_m(v) + 1$ holds if and only if $m = \pm p^n$ for some $n \in \mathbb{N}$.*
- (v) *$x_{m+n}(v) = x_m(v)x_n(v) + (v^2 - 1)y_m(v)y_n(v)$ and $y_{m+n}(v) = x_m(v)y_n(v) + y_m(v)x_n(v)$.*
- (vi) *$n|m \iff y_n(v)|y_m(v)$.*

Proof. (i) Since

$$x_n(v) = \sum_{\substack{i=0 \\ i \text{ even}}}^n \binom{n}{i} v^{n-i} (v^2 - 1)^{\frac{i}{2}} = \sum_{\substack{i=0 \\ i \text{ even}}}^n \binom{n}{i} v^n + \dots$$

for all $n \geq 0$ (by equation (5.3)), $x_n(v)$ has degree n if $n \geq 0$. For $n < 0$ we have $x_n(v) = x_{-n}(v)$, so x_n has degree $-n = |n|$.

Furthermore, if $n > 0$, then

$$y_n(v) = \sum_{\substack{i=1 \\ i \text{ odd}}}^n \binom{n}{i} v^{n-i} (v^2 - 1)^{\frac{i-1}{2}}$$

which has degree $n - 1$. For $n < 0$ we have $y_n(v) = -y_{-n}(v)$, so y_n has degree $-n - 1 = |n| - 1$. The second statement follows immediately from this.

(ii) $x_0(v) + y_0(v) = (v + u)^0 = 1$ so $x_0(v) = 1$ and $y_0(v) = 0$.

Furthermore, $x_1(v) + y_1(v) = (v + u)^1 = v + u$ so $x_1(v) = v$ and $y_1(v) = 1$.

(iii) Since the characteristic of R is p , we have

$$x_{mp^n}(v) + uy_{mp^n}(v) = (v + u)^{mp^n} = ((v + u)^m)^{p^n} = (x_m(v) + uy_m(v))^{p^n} = x_m(v)^{p^n} + u^{p^n} y_m(v)^{p^n}$$

Since p^n is odd, this gives $x_{mp^n}(v) = x_m(v)^{p^n}$.

Substituting $m = 1$ and using $x_1(v) = v$ shows that $x_{p^n}(v) = v^{p^n}$.

(iv) Suppose that $m = \pm p^n$. Notice that by part (iii) and the fact that the characteristic of R is p , we have

$$x_m(v + 1) = x_{p^n}(v + 1) = (v + 1)^{p^n} = v^{p^n} + 1 = x_{p^n}(v) + 1 = x_m(v) + 1$$

Conversely, assume that $x_m(v + 1) = x_m(v) + 1$. Notice that $m \neq 0$, so we can assume $m > 0$. Write $m = qp^n$, with $p \nmid q$. Then

$$x_q(v + 1)^{p^n} = x_{qp^n}(v + 1) = x_{qp^n}(v) + 1 = x_q(v)^{p^n} + 1 = (x_q(v) + 1)^{p^n}$$

Hence $x_q(v + 1) = x_q(v) + 1$.

As a polynomial in v , x_q has degree q (by part (i)), so we can write $x_q(v) = \alpha v^q + \beta v^{q-1} + \dots$ with $\alpha \neq 0$. Hence $x_q(v + 1) = \alpha(v + 1)^q + \beta(v + 1)^{q-1} + \dots = \alpha v^q + (q\alpha + \beta)v^{q-1} + \dots$, so if $q \geq 2$, then $\beta = q\alpha + \beta$, which is clearly a contradiction with $p \nmid q$. It follows that $q = 1$ and $m = p^n$.

(v) This follows from an easy computation:

$$\begin{aligned} x_{m+n}(v) + uy_{m+n}(v) &= (v + u)^{m+n} = \\ &= (v + u)^m \cdot (v + u)^n = (x_m(v) + uy_m(v)) \cdot (x_n(v) + uy_n(v)) = \\ &= (x_m(v)x_n(v) + (v^2 - 1)y_m(v)y_n(v)) + u(x_m(v)y_n(v) + y_m(v)x_n(v)) \end{aligned}$$

(vi) Since the relations $n|m$ and $y_n(v)|y_m(v)$ are both independent of the signs of n and m (because $y_{-n}(v) = -y_n(v)$), we can assume that n and m are non-negative.

Assume that $n|m$. Then there exists $q \in \mathbb{N}$ such that $m = nq$. This gives

$$\begin{aligned} x_{nq}(v) + uy_{nq}(v) &= (v + u)^{nq} = ((v + u)^n)^q = (x_n(v) + uy_n(v))^q = \\ &= \sum_{i=0}^q \binom{q}{i} x_n(v)^{q-i} (uy_n(v))^i = \sum_{i=0}^q \binom{q}{i} x_n(v)^{q-i} (v^2 - 1)^{\lfloor \frac{i}{2} \rfloor} u^{i-2\lfloor \frac{i}{2} \rfloor} y_n(v)^i \end{aligned}$$

Hence

$$y_n(nq) = \sum_{\substack{i=1 \\ i \text{ odd}}}^q \binom{q}{i} x_n(v)^{q-i} (v^2 - 1)^{\frac{i-1}{2}} y_n(v)^i$$

This is clearly divisible by $y_n(v)$, so $y_n(v)|y_m(v)$.

Conversely, suppose that $y_n(v)|y_m(v)$. If $n = 0$, then $y_n(v) = 0$ so $y_m(v) = 0$ and $m = 0$ (by part (i) and (ii)). In this case $n|m$, so we can assume $n > 0$.

Write $m = nq + r$ with $0 \leq r < n$. By part (vi), we have $y_m(v) = x_{nq}(v)y_r(v) + y_{nq}(v)x_r(v)$. Since $y_n(v)|y_m(v)$ and $y_n(v)|y_{nq}(v)$ (by the other direction), we must have $y_n(v)|x_{nq}(v)y_r(v)$, and hence $y_n(v)|x_{nq}(v)^2 y_r(v)$. Since $(x_{nq}(v), y_{nq}(v))$ is a solution of the Pell equation, we have $x_{nq}(v)^2 + (v^2 - 1)y_{nq}(v)^2 = 1$, so $x_{nq}(v)^2 = 1 - (v^2 - 1)y_{nq}(v)^2$. So $y_n(v)|(1 - (v^2 - 1)y_{nq}(v)^2)y_r(v)$, and hence $y_n(v)|y_r(v)$. If $r \neq 0$, then $y_n(v)$ and $y_r(v)$ are polynomials of degree $n - 1$ and $r - 1$, respectively. Hence this divisibility relation gives $n \leq r$, which is a contradiction. So $r = 0$ and $m = nq$, so $n|m$. \square

Lemma 5.12. *Let R be an integral domain of characteristic $p > 2$, and let $m, q \in \mathbb{Z}$. Then*

$$m|_p q \iff (x_m(t) = x_q(t) = 1) \vee (\exists a \in R[t], \exists l, s \in \mathbb{Z} : a = x_m(t) \wedge x_q(t) = x_l(a) \wedge x_s(a+1) = x_l(a) + 1)$$

Proof. Suppose that $m|_p q$. Then there exists $n \in \mathbb{N}$ such that $q = \pm mp^n$. If $m = 0$, then $q = 0$, and hence $x_m(t) = x_q(t) = 1$ by lemma 5.11(ii), so we can assume $m \neq 0$.

Put $a = x_m(t)$. Then $a \notin R$ by lemma 5.11(i). Furthermore, if we choose $l = s = p^n$, then we have $x_s(a+1) = x_l(a) + 1$ (lemma 5.11(v)). Finally, lemma 5.11(iii) implies that

$$x_q(t) = x_{mp^n}(t) = x_m(t)^{p^n} = a^{p^n} = x_{p^n}(a) = x_l(a)$$

Conversely, if $x_m(t) = x_q(t) = 1$ then $m = q = 0$ by lemma 5.11(i), so $m|_p q$. In the other case, let a, l, s be as in the lemma. If $m = 0$, then $a = 1$. By the extension of the definition of x_l , this implies $x_l(a) = 1$, so $x_q(t) = 1$. By lemma 5.11(i), we have $q = 0$, and $m|_p q$ clearly holds.

If $m \neq 0$, then $a = x_m(t) \notin R$ by lemma 5.11(i). Since $x_s(a+1)$ and $x_l(a) + 1$ have the same degree, we get $s = \pm l$ (lemma 5.11(i), and hence $x_s(a+1) = x_s(a) + 1$. By lemma 5.11(v), this implies $l = \pm s = \pm p^n$ for some $n \in \mathbb{N}$, so

$$x_l(a) = x_{\pm p^n}(a) = x_{p^n}(a) = a^{p^n} = x_m(t)^{p^n} = x_{mp^n}(t)$$

by lemma 5.11(iii). Hence $x_q(t) = x_l(a) = x_{mp^n}(t)$, so $q = \pm mp^n$ (by lemma 5.11(i), since $x_q(t)$ and $x_{mp^n}(t)$ have the same degree). This gives $m|_p q$. \square

Proof of theorem 5.10. Consider the map $n \mapsto (x_n(t), y_n(t))$ for $n \in \mathbb{Z}$. This map is injective, since if $x_m(t) = x_n(t)$, then $m = \pm n$, and hence if also $y_m(t) = y_n(t)$, then $m = n$. Hence we have a model of \mathbb{Z} in $R[t]$, given by the set $\{(x_n(t), y_n(t)) \mid n \in \mathbb{Z}\}$. By lemma 5.11(v) and (vi), addition and divisibility in this model are diophantine over $R[t]$ in $(0, 1, t; +, \cdot)$. By lemma 5.12, the same holds for $|_p$. Hence we have a positive existential model of $(\mathbb{Z}, (0, 1; +; |, |_p))$ in $(R[t], (0, 1, t; +, \cdot))$.

Since the positive existential theory of \mathbb{Z} in $(0, 1; +; |, |_p)$ is undecidable by corollary 4.6, the same holds for the positive existential theory of $R[t]$ in $(0, 1, t; +, \cdot)$ by theorem 4.2. Since the diophantine theory of $R[t]$ in $(0, 1, t; +, \cdot)$ equals the positive existential theory, the diophantine theory is also undecidable. \square

5.3 Rings of characteristic two

In the previous proof, we used the Pell equation $x^2 - (v^2 - 1)y^2 = 1$, and defined $u^2 = v^2 - 1$. If the characteristic of R is 2, then we get $(x + (v+1)y)^2 = 1$, so $x = (v+1)y$, and the proof doesn't work anymore. Therefore, to prove the theorem for R with characteristic 2, we will use another equation. After redefining x_n and y_n , the proof will almost be the same as for characteristic greater than 2.

In the remaining of this chapter, R will be an integral domain of characteristic 2.

Theorem 5.13. *Let R be an integral domain of characteristic 2. Then there is no algorithm to decide whether a diophantine equation with coefficients in $(\mathbb{Z}/2\mathbb{Z})[t]$ has a solution in $R[t]$.*

As before, let $v \in R[t] \setminus R$, and consider the following equation

$$x^2 + vxy + y^2 = 0 \tag{5.5}$$

Let $u \in \overline{R[t]}$ be a solution of $x^2 + vx + 1 = 0$. If u is an element of $R(t)$, then we can write $u = \frac{p}{q}$ with $p, q \in R[t]$ with $\gcd(p, q) = 1$. This gives $p^2 + vpq + q^2 = 0$, so $q|p$ and hence $q = 1$. But then $p^2 + vp + 1 = 0$, so $p|1$. This implies $r \in R$, so $u \in R$, which is a contradiction.

Notice that $u(v+u) = u^2 + uv = 1$. This makes it possible to give the following definition:

Definition 5.14. For $n \in \mathbb{Z}$, define $x_n(v), y_n(v) \in R[t]$ by

$$x_n(v) + uy_n(v) = (v+u)^{-n} = u^n \tag{5.6}$$

Again we need some lemmas about the solutions of the equation:

Lemma 5.15. *The solutions of equation (5.5) in $R[t]$ are $x = x_n(v)$, $y = y_n(v)$ for $n \in \mathbb{Z}$.*

Proof. First we show that $(x_n(v), y_n(v))$ is indeed a solution. Since $u^2 + vu + 1 = 0$, we have $v = u + \frac{1}{u}$. Hence the equation (5.5) is equivalent to $x^2 + (u + \frac{1}{u})xy + y^2 = 1$. Since

$$x_n(v)^2 + \left(u + \frac{1}{u}\right)x_n(v)y_n(v) + y_n(v)^2 = (x_n(v) + uy_n(v)) \cdot \left(x_n(v) + \frac{1}{u}y_n(v)\right) = u^n \cdot \left(\frac{1}{u}\right)^n = 1$$

$(x_n(v), y_n(v))$ is a solution of this equation.

As in the proof of lemma 5.4, we again only prove the converse for the case $v = t$, and refer to [Den79] for the general case. Suppose that $x^2 + txy + y^2 = 1$. By lemma 5.16, we can parametrize the curve $u^2 + ut + 1 = 0$ by $t(s) = s + \frac{1}{s}$, $u(s) = s$. Then the functions $x + uy$ and $x + \frac{1}{u}y$ have only poles at $s = 0$, and since $(x + uy)(x + \frac{1}{u}y) = x^2 + vxy + y^2 = 1$, they also only have zeros at $s = 0$. Hence $x + uy = c \cdot s^{-n}$ for some $n \in \mathbb{Z}$. Hence

$$x + uy = c \cdot s^{-n} = c \cdot u^{-n} = x_n(v) + uy_n(v)$$

Hence $x = x_n(v)$ and $y = y_n(v)$. □

Lemma 5.16. *The curve $u^2 + ut + 1 = 0$ can be parametrized by $t(s) = s + \frac{1}{s}$, $u(s) = s$.*

Proof. For any s , we have $u(s)^2 + u(s)t(s) + 1 = s^2 + s^2 + 1 + 1 = 0$ (since the characteristic is 2).

The map $\varphi : s \mapsto (t(s), u(s)) = (s + \frac{1}{s}, s)$ is clearly injective. If $u^2 + ut + 1 = 0$, then $t = \frac{u^2+1}{u} = u + \frac{1}{u}$, so $(t, u) = \phi(u)$. Hence φ is an injection, so this is a parametrization. □

Lemma 5.17. *For all $n, m \in \mathbb{Z}$ and $v \in R[t] \setminus R$, the following statements hold:*

- (i) $x_{-n}(v) = x_n(v) + vy_n(v)$ and $y_{-n}(v) = y_n(v)$.
- (ii) $x_0(v) = 1$, $y_0(v) = 0$, $x_1(v) = 0$ and $y_1(v) = 1$.
- (iii) $x_{m+n}(v) = x_m(v)x_n(v) + y_m(v)y_n(v)$ and $y_{m+n}(v) = x_m(v)y_n(v) + y_m(v)x_n(v) + vy_m(v)y_n(v)$.
- (iv) If $n \geq 2$, then $x_n(v)$ has degree $n - 2$ and $y_n(v)$ has degree $n - 1$ (as polynomials in v).
- (v) If $n \geq 0$, then $y_{2^n m}(v) = v^{2^n - 1}y_m(v)^{2^n}$ and $y_{2^n}(v) = v^{2^n - 1}$.
- (vi) $(v + 1)y_m(v + 1) = vy_m(v) + 1$ holds if and only if $m = \pm 2^n$ for some $n \in \mathbb{N}$.
- (vii) $n|m \iff y_n(v)|y_m(v)$.

Proof. (i) Notice that $v + u = \frac{1}{u}$, so

$$\begin{aligned} (x_n(v) + uy_n(v)) \cdot (x_n(v) + vy_n(v) + uy_n(v)) &= (x_n(v) + uy_n(v)) \cdot (x_n(v) + \frac{1}{u}y_n(v)) = \\ x_n(v)^2 + (u + \frac{1}{u})x_n(v)y_n(v) + y_n(v)^2 &= x_n(v)^2 + vx_n(v)y_n(v) + y_n(v)^2 = 1 \end{aligned}$$

Hence

$$x_{-n}(v) + uy_{-n}(v) = (v + u)^n = \frac{1}{(v + u)^{-n}} = \frac{1}{x_n(v) + y_n(v)} = x_n(v) + vy_n(v) + uy_n(v)$$

This shows that $x_{-n}(v) = x_n(v) + vy_n(v)$ and $y_{-n}(v) = y_n(v)$.

(ii) $x_0(v) + uy_0(v) = u^0 = 1$ so $x_0(v) = 1$ and $y_0(v) = 0$.

Furthermore, $x_1(v) + y_1(v) = u^1 = u$ so $x_1(v) = 0$ and $y_1(v) = 1$.

(iii) This follows from an easy computation:

$$\begin{aligned} x_{m+n}(v) + uy_{m+n}(v) &= (v + u)^{m+n} = (v + u)^m \cdot (v + u)^n = (x_m(v) + uy_m(v)) \cdot (x_n(v) + uy_n(v)) = \\ x_m(v)x_n(v) + u(x_m(v)y_n(v) + y_m(v)x_n(v)) &+ u^2y_m(v)y_n(v) = \\ (x_m(v)x_n(v) + y_m(v)y_n(v)) + u(x_m(v)y_n(v) &+ y_m(v)x_n(v) + vy_m(v)y_n(v)) \end{aligned}$$

(iv) For $n = 2$, we have $x_2(v) + uy_2(v) = u^2 = uv + 1$ so $x_2(v) = 1$ and $y_2(v) = v$, of degree 0 and 1, respectively. Now let $n > 2$ and suppose that $x_n(v)$ has degree $n - 2$ and $y_n(v)$ has degree $n - 1$. Then, by part (ii) and (iii), we have $x_{n+1}(v) = x_1(v)x_n(v) + y_1(v)y_n(v) = y_n(v)$, of degree $n - 1$, and $y_{n+1}(v) = x_1(v)y_n(v) + y_1(v)x_n(v) + vy_1(v)y_n(v) = x_n(v) + vy_n(v)$. Since $x_n(v)$ has degree $n - 2$ and $vy_n(v)$ has degree n , $y_{n+1}(v)$ also has degree n . The statement now follows by induction.

(v) If $n = 0$, then the statements are clear.

Suppose that $y_{2^n m}(v) = v^{2^n - 1} y_m(v)^{2^n}$ for some $n \geq 0$. Then

$$\begin{aligned} y_{2^{n+1}m}(v) &= x_{2^n m}(v)y_{2^n m}(v) + y_{2^n m}(v)x_{2^n m}(v) + vy_{2^n m}(v)y_{2^n m}(v) = \\ &= vy_{2^n m}(v)^2 = v(v^{2^n - 1} y_m(v)^{2^n})^2 = v^{2^{n+1} - 1} y_m(v)^{2^{n+1}} \end{aligned}$$

The first statement now follows by induction.

For the second statement, notice that $y_1(v) = 1$.

(vi) For the ‘if’ part, suppose that $m = \pm 2^n$. Then $y_m(v) = y_{2^n}(v)$ for all v , so by part (v)

$$(v+1)y_m(v+1) = (v+1)y_{2^n}(v+1) = (v+1) \cdot (v+1)^{2^n - 1} = (v+1)^{2^n} = v^{2^n} + 1 = y_{2^n}(v) + 1 = y_m(v) + 1$$

in which the fourth equality follows from the fact that the characteristic of R is 2.

For the ‘only if’ part, assume that $(v+1)y_m(v+1) = vy_m(v) + 1$. Notice that $m \neq 0$, so since $y_{-m}(v) = y_m(v)$, we can assume $m > 0$. Write $m = 2^n q$, with q odd. Then

$$(v+1)y_m(v+1) = (v+1)y_{2^n q}(v+1) = (v+1)^{2^n} y_q(v+1)^{2^n}$$

so

$$((v+1)y_q(v+1))^{2^n} = (v+1)y_m(v+1) = vy_m(v) + 1 = vy_{2^n q}(v) + 1 = v^{2^n} y_q(v)^{2^n} + 1 = (vy_q(v) + 1)^{2^n}$$

Hence $(v+1)y_q(v+1) = vy_q(v) + 1$.

Now suppose that $q \geq 3$. Then $y_q(v)$ has degree $q - 1$ (by part (iv)), so we can write $y_q(v) = \alpha v^{q-1} + \beta v^{q-2} + \dots$ with $\alpha \neq 0$. But then $(v+1)y_q(v+1) = \alpha(v+1)^q + \beta(v+1)^{q-1} + \dots = \alpha v^q + (q\alpha + \beta)v^{q-1} + \dots$ and $vy_q(v) + 1 = \alpha v^q + \beta v^{q-1} + \dots$. Hence $q\alpha + \beta = \beta$, so $q\alpha = 0$, which is a contradiction with $\alpha \neq 0$. So $q = 1$ and $m = 2^n$.

(vii) Since the relations $n|m$ and $y_n(v)|y_m(v)$ are both independent of the signs of n and m (because $y_{-n}(v) = -y_n(v)$), we can assume that n and m are non-negative.

Assume that $n|m$. Then there exists $q \in \mathbb{N}$ such that $m = qn$. If $q = 1$, then it is clear that $y_n(v)|y_m(v)$. Now suppose that $y_n(v)|y_{kn}(v)$ for some $k \geq 1$. Then

$$y_{(k+1)n}(v) = x_{kn}(v)y_n(v) + y_{kn}(v)x_n(v) + vy_{kn}(v)y_n(v)$$

which is also divisible by $y_n(v)$. By induction, $y_{qn}(v)$ is also divisible by $y_n(v)$, so $y_n(v)|y_m(v)$.

Conversely, suppose that $y_n(v)|y_m(v)$. If $n = 0$, then $y_n(v) = 0$ so $y_m(v) = 0$ and $m = 0$ (by part (ii) and (iv)). In this case $n|m$, so we can assume $n > 0$.

Write $m = nq + r$ with $0 \leq r < n$. $y_m(v) = x_{nq}(v)y_r(v) + y_{nq}(v)x_r(v) + vy_{nq}(v)y_r(v)$, from $y_n(v)|y_m(v)$ and $y_n(v)|y_{nq}(v)$ it follows that $y_n(v)|x_{nq}(v)y_r(v)$, and hence $y_n(v)|x_{nq}(v)^2 y_r(v)$. Since $(x_{nq}(v), y_{nq}(v))$ is a solution of equation (5.5), we have $x_{nq}(v)^2 + vx_{nq}(v)y_{nq}(v) + y_{nq}(v)^2 = 1$, so $x_{nq}^2 = 1 + vx_{nq}(v)y_{nq}(v) + y_{nq}(v)^2$. This implies that $y_n(v)|(1 + vx_{nq}(v)y_{nq}(v) + y_{nq}(v)^2)y_r(v)$, so $y_n(v)|y_r(v)$. If $r \neq 0$, then $y_n(v)$ and $y_r(v)$ are polynomials of degree $n - 1$ and $r - 1$, respectively (this also holds if $r = 1$ or $n = 1$). Hence this divisibility relation gives $n \leq r$, which is a contradiction. So $r = 0$ and $m = nq$, so $n|m$. \square

Lemma 5.18. *Let R be an integral domain of characteristic 2, and let $m, q \in \mathbb{Z}$. Then*

$$\begin{aligned} m|_{2q} &\iff (y_m(t) = y_q(t) = 0) \vee \\ &(\exists a \in R[t], \exists l, s \in \mathbb{Z} : a = ty_m(t) \wedge ty_q(t) = ay_l(a) \wedge (a+1)y_s(a+1) = ax_l(a+1)) \end{aligned}$$

Proof. Suppose that $m|_2q$. Then there exists $n \in \mathbb{N}$ such that $q = \pm 2^n m$. If $m = 0$, then $q = 0$, and hence $y_m(t) = y_q(t) = 0$ by lemma 5.17(ii), so we can assume $m \neq 0$.

Put $a = ty_m(t)$. Then it is clear that $a \notin R$. Furthermore, if we choose $l = s = 2^n$, then we have $(a+1)y_s(a+1) = ay_l(a) + 1$ (lemma 5.17(vi)). Finally, lemma 5.17(iii) implies that

$$ty_q(t) = ty_{2^n m}(t) = t^{2^n} y_m(t)^{2^n} = a^{2^n} = ay_{2^n}(a) = ay_l(a)$$

Conversely, if $y_m(t) = y_q(t) = 0$, then $m = q = 0$ by lemma 5.17(ii) and (iv), so $m|_2q$. Otherwise, let a, l, s be as in the lemma. If $m = 0$, then $a = ty_0(t) = 0$, and hence $ty_q(t) = 0$, so $q = 0$ and $m|_2q$. If $m \neq 0$, then $a = ty_m(t) \notin R$. Since $(a+1)y_s(a+1)$ and $ay_l(a) + 1$ have the same degree, lemma 5.17(iv) gives $s = \pm l$, and hence $(a+1)y_s(a+1) = ay_s(a) + 1$. By lemma 5.17(vi), this implies $l = \pm s = \pm 2^n$ for some $n \in \mathbb{N}$, so

$$ay_l(a) = ay_{\pm 2^n}(a) = a^{2^n} = (ty_m(t))^{2^n} = t^{2^n} y_m(t)^{2^n} = ty_{2^n m}(t)$$

by lemma 5.17(v). Hence $ty_q(t) = ay_l(a) = ty_{2^n m}(t)$, so $q = \pm 2^n m$ (since $y_q(t)$ and $y_{2^n m}(t)$ have the same degree). This gives $m|_2q$. \square

Proof of theorem 5.13. This proof is exactly the same as for theorem 5.10, except that we use lemma 5.17(iii) and (vii) and lemma 5.18 instead of lemma 5.11(v) and (vi) and lemma 5.12. \square

Chapter 6

Quadratic rings

In this chapter, we will prove that Hilbert's tenth problem over $\mathbb{Z}[\pi]$ is undecidable. Here π denotes a quadratic integer, satisfying the equation $\pi^2 + a\pi + b = 0$ for some $a, b \in \mathbb{Z}$. A special case of rings of the form $\mathbb{Z}[\pi]$ are rings of integers of quadratic extensions of \mathbb{Q} . For these rings, the undecidability of Hilbert's tenth problem was proven by Denef in [Den75]. The results in this chapter are a small generalization of Denef's result, and the proofs are based on his article.

Throughout this chapter, D will denote square-free part of the discriminant $a^2 - 4b$ of π . Hence $\mathbb{Z}[\pi]$ is a subset of $\mathbb{Q}(\sqrt{D})$. In the proof, we will make a distinction between real and non-real π , i.e. positive and negative D . In both cases, we will use the solutions of the Pell equation.

In both characteristics, we will use the norm on $\mathbb{Q}(\pi)$. We have the following lemma:

Lemma 6.1. *Let π be a solution of the equation $x^2 + ax + b = 0$, where a and b are integers.*

(i) *For all $x, y \in \mathbb{Z}$, the norm of $x + y\pi$ in $\mathbb{Q}(\pi)$ is $N(x + y\pi) = x^2 - axy + by^2$, so in particular, the norm is an integer.*

(ii) *If π is non-real, then the norm equals the square of the complex norm, i.e. $N(x + y\pi) = |x + y\pi|^2$ for all $x, y \in \mathbb{Q}$.*

Proof. (i) By definition, the norm is the product of $x + y\pi$ and its conjugate $\overline{x + y\pi}$ in $\mathbb{Q}(\pi) = \mathbb{Q}(\sqrt{D})$. Since π satisfies $\pi^2 + a\pi + b = 0$, we have $\pi = \frac{-a \pm \sqrt{a^2 - 4b}}{2}$, so $\overline{x + y\pi} = x - \frac{a}{2}y \mp \frac{\sqrt{a^2 - 4b}}{2}y$. Hence the norm of $x + y\pi$ is

$$\begin{aligned} N(x + y\pi) &= \left(x - \frac{a}{2}y \pm \frac{\sqrt{a^2 - 4b}}{2}y \right) \cdot \left(x - \frac{a}{2}y \mp \frac{\sqrt{a^2 - 4b}}{2}y \right) = \\ &= \left(x - \frac{a}{2}y \right)^2 - (a^2 - 4b) \frac{y^2}{4} = x^2 - axy + by^2 \end{aligned}$$

Since a and b are integers, $N(x + y\pi)$ is also an integer.

(ii) Suppose that π is non-real, so $a^2 - 4b < 0$. Then

$$|x + \pi y|^2 = \left| x - \frac{a}{2}y \pm \frac{\sqrt{4b - a^2}}{2}iy \right|^2 = \left(x - \frac{a}{2}y \right)^2 + (4b - a^2) \frac{y^2}{4} = x^2 - axy + by^2 = N(x + \pi y)$$

□

6.1 The Pell equation

Let $d \in \mathbb{N}$ and suppose that d is not a square. The *Pell equation* is the equation

$$x^2 - dy^2 = 1$$

A standard result from number theory tells us that this equation has an integral solution:

Lemma 6.2. *If d is not a square, then the Pell equation $x^2 - dy^2 = 1$ has a solution $(x, y) \in \mathbb{Z}^2$ with $x, y > 0$.*

Proof. See [Beu00], chapter 15, theorem 15.1.1. \square

Similar to lemma 5.4, we have the following definition and lemma about the structure of the set of solutions:

Definition 6.3. Let $d > 1$, and write $\alpha = \sqrt{d^2 - 1} \in \mathbb{R}$. We define $x_n(d), y_n(d) \in \mathbb{N}$ by

$$x_n(d) + \alpha y_n(d) = (d + \alpha)^n$$

Lemma 6.4. *Let $d > 1$. Then all solutions in \mathbb{N} of the Pell equation $x^2 - (d^2 - 1)y^2 = 1$ are $(x_n(d), y_n(d))$ for $n = 0, 1, 2, \dots$*

Proof. For all $n \in \mathbb{N}$, $(x_n(d), y_n(d))$ is indeed a solution, since

$$\begin{aligned} x_n(d)^2 - (d^2 - 1)y_n(d)^2 &= x_n(d)^2 - \alpha^2 y_n(d)^2 = \\ &= (x_n(d) + \alpha y_n(d)) \cdot (x_n(d) - \alpha y_n(d)) = (d + \alpha)^n \cdot (d - \alpha)^n = (d^2 - \alpha^2)^n = 1 \end{aligned}$$

Notice that $(x_1(d), y_1(d)) = (d, 1)$. It is clear that of all solutions (x, y) with $x, y > 0$, this one has the smallest value for $x + \alpha y$. Hence by [Beu00], chapter 15, theorem 15.1.2, all solutions (x, y) are of the form $x + \alpha y = (x_1(d) + \alpha y_1(d))^n = (d + \alpha)^n$ for some $n \in \mathbb{N}$. This implies that all solutions are of the form $(x_n(d), y_n(d))$ with $n \in \mathbb{N}$. \square

Lemma 6.5. *If $d, n, k \in \mathbb{N}$ with $d > 1$, then*

$$y_{nk}(d)^2 \equiv k^2 y_n(d)^2 \pmod{y_n(d)^4}$$

Proof. As we noted already in chapter 5,

$$x_{nk}(d) + \alpha y_{nk}(d) = (x_n(d) + \alpha y_n(d))^k = \sum_{i=0}^k \binom{k}{i} \alpha^i x_n(d)^{k-i} y_n(d)^i$$

so

$$y_{nk}(d) = \sum_{\substack{i=1 \\ i \text{ odd}}}^k \binom{k}{i} (d^2 - 1)^{\frac{i-1}{2}} x_n(d)^{k-i} y_n(d)^i = k x_n(d)^{k-1} y_n(d) + z y_n(d)^3$$

for some $z \in \mathbb{N}$. Hence $\frac{y_{nk}(d)}{y_n(d)} = k x_n(d)^{k-1} + z y_n(d)^2$. By squaring this, it follows that

$$\left(\frac{y_{nk}(d)}{y_n(d)} \right)^2 = k^2 x_n(d)^{2(k-1)} \pmod{y_n(d)^2}$$

Since $(x_n(d), y_n(d))$ is a solution of the Pell equation $x^2 - (d^2 - 1)y^2 = 1$,

$$x_n(d)^{2(k-1)} = (1 + (d^2 - 1)y_n(d)^2)^{k-1} \equiv 1 \pmod{y_n(d)^2}$$

Hence $\left(\frac{y_{nk}(d)}{y_n(d)} \right)^2 \equiv k^2 \pmod{y_n(d)^2}$, so $y_{nk}(d)^2 \equiv k^2 y_n(d)^2 \pmod{y_n(d)^4}$. \square

6.2 Real rings

In this section and the next, we will give a system of equations that almost defines \mathbb{Z} in $\mathbb{Z}[\pi]$. In the last section, we will use this to show that \mathbb{Z} is positive existential over $\mathbb{Z}[\pi]$ in $(0, 1, \pi; +, \cdot)$. We will need to make a distinction between real and non-real π . Throughout this section, we assume that π is real, so $\mathbb{Z}[\pi] \subseteq \mathbb{R}$.

First we need two lemmas:

Lemma 6.6. *Let (u, v) be a solution in \mathbb{N} of the Pell equation $u^2 - Dv^2 = 1$ and suppose that $v \neq 0$. If (x, y) is a solution in $\mathbb{Z}[\pi]$ of the Pell equation $x^2 - Dv^2y^2 = 1$, then $y^2 \in \mathbb{N}$.*

Proof. Since $v \neq 0$, we have $u > 1$. Notice that $(x + \sqrt{D}vy)(x - \sqrt{D}vy) = 1$, so $x + \sqrt{D}vy$ is a unit in $\mathbb{Z}[\pi]$. Define $z = x + \sqrt{D}vy$. Then $z^{-1} = x - \sqrt{D}vy$, and

$$4Dv^2y^2 + 2 = (z - z^{-1})^2 + 2 = z^2 + z^{-2}$$

By lemma 6.1(i), the norm of z is an integer, and since z is a unit, the norm must be equal to ± 1 . This implies that the conjugate of z in $\mathbb{Q}(\sqrt{D})$ equals $\pm z^{-1}$, i.e. if we write $z = z_1 + z_2\sqrt{D}$ with $z_1, z_2 \in \mathbb{Q}$, then $z^{-1} = \pm(z_1 - z_2\sqrt{D})$. Hence

$$4Dv^2y^2 + 2 = z^2 + z^{-2} = (z_1 + z_2\sqrt{D})^2 + (z_1 - z_2\sqrt{D})^2 = 2(z_1^2 + Dz_2^2)$$

It follows that $4Dv^2y^2 + 2 \in \mathbb{Q}$, so $y^2 \in \mathbb{Q}$. Since $y^2 \in \mathbb{Z}[\pi]$, this implies that $y^2 \in \mathbb{Z}$. Because $y \in \mathbb{Z}[\pi] \subseteq \mathbb{R}$, we even have $y^2 \in \mathbb{N}$. \square

Lemma 6.7. *Suppose that $x, y, z \in \mathbb{Z}[\pi]$ satisfy $x \equiv y \pmod{z}$, $0 \leq x, y < z$ and $0 \leq \bar{x}, \bar{y} < \bar{z}$. Here \bar{x} denotes the conjugate of x in $\mathbb{Q}(\sqrt{D})$. Then $x = y$.*

Proof. Suppose that $x \neq y$. Then there exists $w \in \mathbb{Z}[\pi] \setminus \{0\}$ such that $x - y = wz$. Hence

$$|x - y| \cdot |\bar{x} - \bar{y}| = |wz| \cdot |\bar{w}\bar{z}| = |w\bar{w}| \cdot |z\bar{z}| = |N(w)| \cdot |z\bar{z}| \geq |z\bar{z}|$$

where the last inequality follows from the fact that $N(w)$ is an integer by lemma 6.1(i), and is non-zero since $w \neq 0$. But the assumptions imply that $0 \leq |x - y| < z$ and $0 \leq |\bar{x} - \bar{y}| < \bar{z}$. Contradiction, so $x = y$. \square

Now we can give the system of equations:

Lemma 6.8. *Let (u, v) be a solution in \mathbb{N} of the Pell equation $u^2 - Dv^2 = 1$ and assume that $v \neq 0$. Consider the following system of equations in the unknowns $x, y, l, m, t, z, w, h, q, r, s$:*

$$x^2 - Dv^2y^2 = 1 \tag{6.1a}$$

$$l^2 - Dv^2m^2 = 1 \tag{6.1b}$$

$$m^2 - y^2t = zy^4 \tag{6.1c}$$

$$t = w^2 \tag{6.1d}$$

$$y^2 - t = 1 + h^2 + q^2 + r^2 + s^2 \tag{6.1e}$$

The following statements hold:

- (i) If these equations have a solution in $\mathbb{Z}[\pi]$, then $t \in \mathbb{Z}$.
- (ii) For every $k \in \mathbb{N}_{>0}$, there is a solution in $\mathbb{Z}[\pi]$ with $t = k^2$.

Proof. First notice that (u, v) exists by lemma 6.2.

(i) Suppose that there is a solution $(x, y, l, m, t, z, w, h, q, r, s)$ in $\mathbb{Z}[\pi]$. By lemma 6.6, y^2 and m^2 are natural numbers. Since $\mathbb{Z}[\pi] \subseteq \mathbb{R}$, all squares are non-negative. Hence $t > 0$ by equation

(6.1d) and $t < y^2$ by (6.1e). It follows that $y \neq 0$. By equation (6.1c), $y^2|m^2$, and $\frac{m^2}{y^2} \equiv t \pmod{y^2}$. Since y^2 and m^2 are natural numbers, this implies

$$\bar{t} \equiv \frac{\overline{m^2}}{y^2} \equiv \frac{m^2}{y^2} \pmod{\overline{y^2}}$$

which is equivalent to $\bar{t} \equiv \frac{m^2}{y^2} \pmod{y^2}$. It follows that $t \equiv \frac{m^2}{y^2} \equiv \bar{t} \pmod{y^2}$. Notice that $t, \bar{t} \geq 0$ by equation (6.1d) (because $\bar{t} = \overline{w^2}$), and $t, \bar{t} < y^2 = \overline{y^2}$ by equation (6.1e). Hence we can apply lemma 6.7 to obtain $t = \bar{t}$, so $t \in \mathbb{Z}$.

(ii) Let $k \in \mathbb{N}_{>0}$. Choose $n \in \mathbb{N}$ such that $y_n(u) > k$, and put $x = x_n(u), y = y_n(u), l = x_{nk}(u)$ and $m = y_{nk}(u)$. By the choice of (u, v) , we have $Dv^2 = u^2 - 1$, so by lemma 6.4, the equations (6.1a) and (6.1b) are satisfied. z exists by lemma 6.5. Furthermore, let $t = k^2$ and $w = k$. Then equation (6.1d) is clearly satisfied. Finally, since $y^2 > k^2 = t$, Lagrange's theorem implies that h, q, r and s exist. \square

6.3 Imaginary rings

Throughout this section, suppose that π is non-real. We will prove a lemma similar to lemma 6.8. The proof will be a bit harder, and we will use a biquadratic field, i.e. a field generated over \mathbb{Q} by two quadratic numbers.

Recall Dirichlet's unit theorem: if a number field K has r real embeddings and s pairs of complex embeddings, then K has $r + s - 1$ fundamental units $\eta_1, \dots, \eta_{r+s-1}$, and every unit u in K can be written as $u = \zeta \cdot \eta_1^{k_1} \cdot \dots \cdot \eta_{r+s-1}^{k_{r+s-1}}$ with $\zeta \in K$ a root of unity and $k_i \in \mathbb{Z}$.

When we apply this theorem, it will be useful to know which roots of unity can be elements of K . Therefore, we have the following lemma:

Lemma 6.9. *Let K be a biquadratic field $\mathbb{Q}(\sqrt{F}, \sqrt{D})$ with $F > 1$ and $D \leq -1$, D and F square-free. Then the only roots of unity K can contain are $\pm 1, \pm i, \frac{\pm 1 \pm \sqrt{3}i}{2}, \frac{\pm 2 \pm \sqrt{2}i}{2}$ and $\frac{\sqrt{3} \pm i}{2}$.*

Proof. The Galois group of K over \mathbb{Q} is isomorphic to $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Hence if ζ is a root of unity contained in K , then the Galois group of $\mathbb{Q}[\zeta]$ over \mathbb{Q} is a subgroup of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. In particular, ζ has degree 1, 2 or 4.

If ζ is a n^{th} root of unity, then the minimal polynomial of ζ is

$$\prod_{\substack{1 \leq k \leq n \\ \gcd(k, n) = 1}} (X - \zeta^k)$$

This is a polynomial of degree $\phi(n)$, where ϕ denotes Euler's totient function. Hence n must satisfy $\phi(n) = 1, 2, 4$.

If $n = p_1^{k_1} \cdot \dots \cdot p_r^{k_r}$ is the prime factorization of n , then $\phi(n) = p_1^{k_1-1} \cdot \dots \cdot p_r^{k_r-1} \cdot (p_1-1) \cdot \dots \cdot (p_r-1)$. Hence the only prime factors that can occur in n are 2, 3 and 5. Notice that $\phi(2^l) = 2^{l-1}$, $\phi(3^l) = 2 \cdot 3^{l-1}$ and $\phi(5^l) = 4 \cdot 5^{l-1}$. Hence at most one factor 3 or 5 can occur. If 5 is a factor of n , then there is at most one factor 2, so $n = 5$ or $n = 10$. If 3 is a factor of n , then there are at most two factors 2, so $n = 3, 6$ or 12 . Finally, if both 3 and 5 don't divide n , then n has at most three factors 2, so $n = 1, 2, 4$ or 8 .

$n = 1$ gives $\zeta = 1$ and $n = 2$ gives $\zeta = -1$. The 4^{th} roots of unity are $\pm i$, and the 8^{th} roots of unity are $\frac{\pm 2 \pm \sqrt{2}i}{2}$. Furthermore, $n = 3$ gives $\zeta = \frac{-1 \pm \sqrt{3}i}{2}$ and $n = 6$ gives $\zeta = \frac{1 \pm \sqrt{3}i}{2}$. The 12^{th} roots of unity are $\frac{\pm 1 \pm \sqrt{3}i}{2}$. Finally, if $n = 5$ or $n = 10$, then the Galois group of ζ is cyclic, since it is generated by $\zeta \mapsto \zeta^2$ and $\zeta \mapsto \zeta^3$, respectively. Hence in this case, it is not a subgroup of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. \square

Lemma 6.10. *Let $F = 3$ if $D \neq -1, -3$, and let $F = 15$ if $D = -1$ or -3 . If $x, y \in \mathbb{Z}[\pi]$ and $x^2 - Fy^2 = 1$, then $y^2 \in \mathbb{Z}$.*

Proof. We write $K = \mathbb{Q}(\sqrt{F}, \sqrt{D})$. Notice that $(x - \sqrt{F}y)(x + \sqrt{F}y) = 1$. Hence $x + \sqrt{F}y$ is a unit in K . K has no real embeddings and 2 pairs of complex embeddings, so by Dirichlet's unit theorem, there exists one fundamental unit. Denote this fundamental unit by η . Then there is a root of unity $\zeta \in K$ and $m \in \mathbb{Z}$ such that $x + \sqrt{F}y = \zeta\eta^m$, and hence $x - \sqrt{F}y = \zeta^{-1}\eta^{-m}$. This implies that

$$4Fy^2 + 2 = ((x + \sqrt{F}y) - (x - \sqrt{F}y))^2 + 2 = (\zeta\eta^m - \zeta^{-1}\eta^{-m})^2 + 2 = \zeta^2\eta^{2m} + \zeta^{-2}\eta^{-2m} \quad (6.2)$$

Again by Dirichlet's unit theorem, $\mathbb{Q}(\sqrt{F})$ has one fundamental unit, since there is one real embedding and no complex embeddings. Denote this fundamental unit by ε . Explicit computations show that $\varepsilon = 2 + \sqrt{3}$ if $F = 3$ and $\varepsilon = 4 + \sqrt{15}$ if $F = 15$. ε is also a unit in K , so there exists a root of unity $\tau \in K$ and $k \in \mathbb{Z}$ such that

$$\varepsilon = \tau\eta^k \quad (6.3)$$

Let σ_1 and σ_2 be the automorphisms of K defined by $\sigma_1(\sqrt{F}) = -\sqrt{F}$, $\sigma_1(\sqrt{D}) = \sqrt{D}$ and $\sigma_2(\sqrt{F}) = \sqrt{F}$, $\sigma_2(\sqrt{D}) = -\sqrt{D}$. Then, since $\varepsilon \in \mathbb{Q}(\sqrt{F})$,

$$\varepsilon^2 = \varepsilon\sigma_2(\varepsilon) = \tau\eta^k\sigma_2(\tau\eta^k) = \tau\sigma_2(\tau)(\eta\sigma_2(\eta))^k$$

If we write $\tau = \tau_1 + \tau_2\sqrt{D}$ with $\tau_1, \tau_2 \in \mathbb{Q}(\sqrt{F})$, then $\tau\sigma_2(\tau) = (\tau_1 + \tau_2\sqrt{D})(\tau_1 - \tau_2\sqrt{D}) = \tau_1^2 - D\tau_2^2 \in \mathbb{Q}(\sqrt{F})$. Hence $\tau' = \tau\sigma_2(\tau)$ is a root of unity in $\mathbb{Q}(\sqrt{F})$. Similarly, $e = \eta\sigma_2(\eta)$ is a unit in $\mathbb{Q}(\sqrt{F})$. Since F is positive, $\mathbb{Q}(\sqrt{F})$ is a real field, so the only roots of unity are ± 1 . Hence $\tau' = \pm 1$, and $\varepsilon^2 = \pm e^k$.

e is a unit in $\mathbb{Q}(\sqrt{F})$, so there exists a root of unity ± 1 and some $l \in \mathbb{Z}$ such that $e = \pm\varepsilon^l$. This implies that $\varepsilon^2 = \pm\varepsilon^{kl}$, so $\varepsilon^{kl-2} = \pm 1$. Since $\mathbb{Q}(\sqrt{F})$ is a real field, the fundamental unit doesn't have absolute value 1 (since otherwise it equals ± 1 , so it is not a fundamental unit). This implies that $kl - 2 = 0$, so $kl = 2$. Recall that $\varepsilon = \tau\eta^k$ by equation (6.3). It follows that

$$\eta^2 = (\eta^k)^l = (\varepsilon\tau^{-1})^l$$

with $l = \pm 1, \pm 2$. So $\zeta_1 = \tau^{-l}$ is a root of unity in K such that $\eta^2 = \zeta_1\varepsilon^l$.

Equation (6.2) reduces to

$$4Fy^2 + 2 = \zeta^2\zeta_1\varepsilon^{lm} + \zeta^{-2}\zeta_1^{-1}\varepsilon^{-lm}$$

Since both ζ and ζ_1 are roots of unity on K , $\zeta_2 := \zeta^2\zeta_1$ is also a root of unity in K . Furthermore, $\varepsilon_l := \varepsilon^l$ is a unit in $\mathbb{Q}(\sqrt{F})$. This change of variables gives the equation $4Fy^2 + 2 = \zeta_2\varepsilon_l^m + \zeta_2\varepsilon_l^{-m}$. Notice that if $F = 3$, then $(2 + \sqrt{3})\sigma_1(2 + \sqrt{3}) = (2 + \sqrt{3})(2 - \sqrt{3}) = 1$, and if $F = 15$, then $(4 + \sqrt{15})\sigma_1(4 + \sqrt{15}) = (4 + \sqrt{15})(4 - \sqrt{15}) = 1$. Since ε_l is a power of either $2 + \sqrt{3}$ or $4 + \sqrt{15}$, this implies that $\sigma_1(\varepsilon_l) = \varepsilon_l^{-1}$, so $4Fy^2 + 2 = \zeta_2\varepsilon_l^m + \zeta_2\sigma_1(\varepsilon_l^m)$.

Now we compute the imaginary part of both sides of this equation. Since $y \in \mathbb{Z}[\pi] \subseteq \mathbb{Q}(\sqrt{D})$, there exists $q_1 \in \mathbb{Q}$ such that $\text{Im}(4Fy^2 + 2) = q_1\sqrt{|D|}$. Furthermore, ε_l is an element of the real field $\mathbb{Q}(\sqrt{F})$, its imaginary part is zero. ζ_2 is a root of unity, so ζ_2^{-1} equals the complex conjugate of ζ_2 , so $\text{Im}(\zeta_2^{-1}) = -\text{Im}(\zeta_2)$. Hence $\text{Im}(4Fy^2 + 2) = \text{Im}(\zeta_2)(\varepsilon_l^m - \sigma_1(\varepsilon_l^m))$. From the definition of σ_1 , it is clear that $\varepsilon_l^m - \sigma_1(\varepsilon_l^m) = q_2\sqrt{F}$ for some $q_2 \in \mathbb{Q}$. Since ζ_2 is a root of unity in K , lemma 6.9 implies that it is of the form $q_3\sqrt{S}$ with $q_3 \in \mathbb{Q}$ and $S \in \{0, 1, 2, 3\}$. It follows that

$$q_1\sqrt{|D|} = \text{Im}(4Fy^2 + 2) = q_2q_3\sqrt{FS}$$

Hence $\sqrt{\frac{|D|}{FS}} \in \mathbb{Q}$.

Suppose that $S = 1$. Then

$$\sqrt{\frac{|D|}{FS}} = \begin{cases} \sqrt{\frac{|D|}{3}} & \text{if } D \neq -1, -3 \\ \sqrt{\frac{|D|}{15}} & \text{if } D = -1, -3 \end{cases}$$

In the first case, we must have $|D| = 3E^2$ for some $E \in \mathbb{Z}$. But since D is square-free, $E = \pm 1$, so $D = -3$. Contradiction, so in this case $\sqrt{\frac{|D|}{FS}}$ is not a rational number. The same holds for the other case.

Now suppose that $S = 2$. Then

$$\sqrt{\frac{|D|}{FS}} = \begin{cases} \sqrt{\frac{|D|}{6}} & \text{if } D \neq -1, -3 \\ \sqrt{\frac{|D|}{30}} & \text{if } D = -1, -3 \end{cases}$$

This can only be a rational number if $D = -6$. But then $K = \mathbb{Q}(\sqrt{3}, \sqrt{-6})$ and $\zeta_2 = \frac{\pm 2 \pm \sqrt{2}i}{2}$ by lemma 6.9. It follows that $\frac{\pm 2 \pm \sqrt{2}i}{2}$ can be written as $r_1 + r_2\sqrt{3} + r_3\sqrt{-6} + r_4\sqrt{-18}$, with $r_i \in \mathbb{Q}$. Since $\sqrt{-2} = \frac{1}{3}\sqrt{-18}$, $\sqrt{2}$ must also be of this form, which is clearly not true. So this case also gives a contradiction.

Finally, suppose that $S = 3$. Then

$$\sqrt{\frac{|D|}{FS}} = \begin{cases} \sqrt{\frac{|D|}{9}} = \frac{1}{3}\sqrt{|D|} & \text{if } D \neq -1, -3 \\ \sqrt{\frac{|D|}{45}} = \frac{1}{3}\sqrt{\frac{|D|}{5}} & \text{if } D = -1, -3 \end{cases}$$

which again cannot be a rational number.

It follows that $S = 0$, so $\text{Im}(4Fy^2 + 2) = 0$. Hence $y^2 \in \mathbb{Q}$, and since also $y^2 \in \mathbb{Z}[\pi]$, we must have $y^2 \in \mathbb{Z}$. \square

Lemma 6.11. *Suppose that $t \in \mathbb{Z}[\pi]$. If there exist $r, w \in \mathbb{Z}$ such that $t \equiv r \pmod{w}$ and $|t| < \frac{|w|}{2}$, then $t \in \mathbb{Z}$.*

Proof. Since $t \equiv r \pmod{w}$, there exists $s \in \mathbb{Z}[\pi]$ such that $t = r + sw$. Hence $|t| = |r + sw|$. Write $s = s_1 + s_2\pi = s_1 + s_2 \frac{-a \pm \sqrt{a^2 - 4b}}{2}$ with $s_1, s_2 \in \mathbb{Z}$. Then

$$\begin{aligned} |t|^2 = |r + sw|^2 &= \left| r + s_1w - s_2w \frac{a}{2} + s_2w \frac{\sqrt{4b - a^2}i}{2} \right|^2 = \\ &= \left(r + s_1w - s_2w \frac{a}{2} \right)^2 + \left(s_2w \frac{\sqrt{4b - a^2}}{2} \right)^2 \geq \left(s_2w \frac{\sqrt{4b - a^2}}{2} \right)^2 = (4b - a^2) \frac{s_2^2 w^2}{4} \end{aligned}$$

Since $4b - a^2 \geq 1$ (because π is non-real), we obtain $|t|^2 \geq \frac{s_2^2 w^2}{4}$, so $|t| \geq \frac{|s_2 w|}{2}$. But we assumed $|t| < \frac{|w|}{2}$, so $\frac{|s_2 w|}{2} \leq |t| < \frac{|w|}{2}$. Hence $|s_2| < 1$, so $s_2 = 0$, and $s \in \mathbb{Z}$. This implies that $t \in \mathbb{Z}$. \square

Lemma 6.12. *Let $F = 3$ if $D \neq -1, 3$ and $F = 15$ if $D = -1, -3$. Consider the following system of equations in the unknowns $x, y, l, m, t, z, r, s, h, w$:*

$$x^2 - Fy^2 = 1 \tag{6.4a}$$

$$l^2 - Fm^2 = 1 \tag{6.4b}$$

$$m^2 - y^2 t = zy^4 \tag{6.4c}$$

$$ry + s(5h + 2) = 1 \tag{6.4d}$$

$$y = 2tw \tag{6.4e}$$

The following statements hold:

(i) If these equations have a solution in $\mathbb{Z}[\pi]$, then $t \in \mathbb{Z}$.

(ii) For every $k \in \mathbb{N}_{>0}$, there is a solution in $\mathbb{Z}[\pi]$ with $t = k^2$.

Proof. (i) Suppose that there is a solution $x, y, l, m, t, z, r, s, h, w$ in $\mathbb{Z}[\pi]$. By lemma 6.10, y^2 and m^2 are integers.

Suppose that $y = 0$. Then equation (6.4d) reduces to $s(5h + 2) = 1$, so $5h + 2$ is a unit in $\mathbb{Z}[\pi]$. Write $h = h_1 + h_2\pi = h_1 + h_2 \frac{-a + \sqrt{a^2 - 4b}}{2}$ with $h_1, h_2 \in \mathbb{Z}$. Then by lemma 6.1(ii),

$$N(5h + 2) = |5h + 2|^2 = \left(5h_1 + 2 - a \frac{5h_2}{2}\right)^2 + (4b - a^2) \left(\frac{5h_2}{2}\right)^2$$

Since the norm of an element of $\mathbb{Z}[\pi]$ is a natural number (lemma 6.1(i)), the norm of a unit must be 1. Hence both $(5h_1 + 2 - a \frac{5h_2}{2})^2$ and $(4b - a^2) (\frac{5h_2}{2})^2$ cannot be greater than 1, so in particular $\frac{5h_2}{2} \in \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$, so $h_2 \in \{-\frac{2}{5}, -\frac{1}{5}, 0, \frac{1}{5}, \frac{2}{5}\}$. Since $h_2 \in \mathbb{Z}$, this implies that $h_2 = 0$, and hence $(5h_1 + 2)^2 = 1$. This is clearly impossible since $h_1 \in \mathbb{Z}$. Hence $y \neq 0$.

By equation (6.4c), $y^2 | m^2$ and $\frac{m^2}{y^2} \equiv t \pmod{y^2}$ with $\frac{m^2}{y^2}, y^2 \in \mathbb{Z}$. Since $y \neq 0$, equation (6.4e) implies $w \neq 0$ so $|t|^2 = |\frac{y}{2w}|^2 = \frac{y^2}{4N(w)} \leq \frac{y^2}{4}$, so $|t| \leq \frac{y}{2} < \frac{y^2}{2}$. Hence we can apply lemma 6.11 to obtain $t \in \mathbb{Z}$.

(ii) Suppose that $k \in \mathbb{N}_{>0}$. Put $t = k^2$. By lemma 6.2, the Pell equation $v^2 - F(2t)^2 w^2 = 1$ has a solution (v, w) in \mathbb{N} with $w > 0$. If we define $A = 2$ if $D \neq -1, 3$ and $A = 4$ if $D = -1, -3$, then $F = A^2 - 1$, so $v^2 - (A^2 - 1)(2tw)^2 = 1$. Hence by lemma 6.4, there exists $n \in \mathbb{N}$ such that $v = x_n(A)$ and $2tw = y_n(A)$. Now put $x = x_n(A), y = y_n(A), l = x_{nk}(A)$ and $m = y_{nk}(A)$. It is clear that $y = 2tw$, and the equations (6.4a) and (6.4b) hold by lemma 6.4. By lemma 6.5, there exists $z \in \mathbb{Z}$ satisfying equation (6.4c). Write $y = 2^c h$ with h odd and $c \in \mathbb{N}$. Then $5h + 2$ is also odd, and $\gcd(2^c \cdot 5h, 5h + 2) = \gcd(2^{c+1}, 5h + 2) = 1$. Since $\gcd(2^c h, 5h + 2)$ divides $\gcd(2^c \cdot 5h, 5h + 2)$, y and $5h + 2$ are coprime. Hence there exists r and s in \mathbb{Z} satisfying equation (6.4d). \square

6.4 Proof of the negative solution to Hilbert's tenth problem

Now we will use the results from the previous sections to show that Hilbert's tenth problem over $\mathbb{Z}[\pi]$ is undecidable. We will use a model of $(\mathbb{Z}, (0, 1; +, \cdot))$ in $(\mathbb{Z}[\pi], (0, 1, \pi; +, \cdot))$, given by the identity map $n \mapsto n$. The results in this section hold for both real and non-real π .

Lemma 6.13. \mathbb{Z} is positive existential over $\mathbb{Z}[\pi]$ in $(0, 1, \pi; +, \cdot)$.

Proof. Denote the statement 'the system of equations (6.1) (or (6.4)) has a solution with $t = a$ ' by $\Sigma(a)$. Of course, the choice of the equations depends on whether π is real or non-real. It is clear that $\Sigma(a)$ is a positive existential formula.

We claim that the statement ' $k \in \mathbb{Z}$ ' is equivalent to

$$\begin{aligned} \exists a, b, c, d \in \mathbb{Z}[\pi] : (a = 0 \vee \Sigma(a^2)) \wedge (b = 0 \vee \Sigma(b^2)) \wedge (c = 0 \vee \Sigma(c^2)) \wedge \\ (d = 0 \vee \Sigma(d^2)) \wedge (k = a^2 + b^2 + c^2 + d^2 \vee k + a^2 + b^2 + c^2 + d^2 = 0) \end{aligned} \quad (6.5)$$

Suppose that $k \in \mathbb{Z}$. Then either $k \in \mathbb{N}$ or $-k \in \mathbb{N}$, so by Lagrange's theorem, there exist $a, b, c, d \in \mathbb{N}$ such that $k = a^2 + b^2 + c^2 + d^2$ or $-k = a^2 + b^2 + c^2 + d^2$. Furthermore, by lemma 6.8(ii) or 6.12(ii), if $a \neq 0$, then $\Sigma(a^2)$ holds. The same holds for b, c and d , so formula (6.5) holds.

Conversely, suppose that formula (6.5) holds. If $a \neq 0$, then $a^2 \in \mathbb{Z}$ by lemma 6.8(i) or 6.12(i). Of course, $a^2 \in \mathbb{Z}$ also holds for $a = 0$. Similarly, b^2, c^2 and d^2 are integers. Hence $k = \pm(a^2 + b^2 + c^2 + d^2)$ is also an integer.

This gives a positive existential definition of \mathbb{Z} in $\mathbb{Z}[\pi]$. \square

Theorem 6.14. Hilbert's tenth problem over $\mathbb{Z}[\pi]$ is undecidable.

Proof. Hilbert's tenth problem over \mathbb{Z} is undecidable, so the positive existential theory of \mathbb{Z} in $(0, 1; +, \cdot)$ is undecidable. \mathbb{Z} is a positive existential subset of $\mathbb{Z}[\pi]$ by lemma 6.13, and under the

map $n \mapsto n$, addition and multiplication in \mathbb{Z} correspond to addition and multiplication in $\mathbb{Z}[\pi]$. Hence this map gives a positive existential model of $(\mathbb{Z}, (0, 1; +, \cdot))$ in $(\mathbb{Z}[\pi], (0, 1, \pi; +, \cdot))$. Now it follows from theorem 4.2 that the positive existential theory of $\mathbb{Z}[\pi]$ in $(0, 1, \pi; +, \cdot)$ is undecidable.

The field of fractions of $\mathbb{Z}[\pi]$ is $\mathbb{Q}(\pi) = \mathbb{Q}(\sqrt{D})$, and this has an extension $\mathbb{Q}(\sqrt{D}, \sqrt{E})$ with $E \neq D$, E square-free. The norm map of $\mathbb{Q}(\sqrt{D}, \sqrt{E})$ over $\mathbb{Q}(\sqrt{D})$ is given by $N(x + y\sqrt{E}) = x^2 + Ey^2$, with coefficients in \mathbb{Z} . Hence by theorem 1.13, the positive existential theory of $\mathbb{Z}[\pi]$ equals the diophantine theory, so the diophantine theory of $\mathbb{Z}[\pi]$ in $(0, 1, \pi; +, \cdot)$ is undecidable. \square

Corollary 6.15. *For any quadratic extension of \mathbb{Q} , Hilbert's tenth problem over the ring of integers is undecidable.*

Proof. Notice that any quadratic extension of \mathbb{Q} is of the form $\mathbb{Q}(\sqrt{D})$, where D is a square-free integer. The ring of integers is

$$\mathcal{O}_{\mathbb{Q}(\sqrt{D})} = \begin{cases} \mathbb{Z}[\frac{1+\sqrt{D}}{2}] & \text{if } D \equiv 1 \pmod{4} \\ \mathbb{Z}[\sqrt{D}] & \text{if } D \equiv 2, 3 \pmod{4} \end{cases}$$

If $D \equiv 1 \pmod{4}$, then the minimal polynomial of $\frac{1+\sqrt{D}}{2}$ is $x^2 - x + \frac{1-D}{4} = 0$, so $\frac{1+\sqrt{D}}{2}$ is an algebraic integer. It is clear that \sqrt{D} is an algebraic integer, with minimal polynomial $x^2 - D = 0$. Hence we can apply theorem 6.14 with either $\pi = \frac{1+\sqrt{D}}{2}$ or $\pi = \sqrt{D}$. \square

Chapter 7

Function fields over finite fields

Now we will prove Hilbert's tenth theorem for rational function fields over finite fields with coefficients in $\mathbb{Z}[t]$. This was proved by Pheidas in [Phe91] for fields of odd characteristic, and by Videla in [Vid94] for fields of characteristic two. Videla's proof is a small modification of Pheidas's proof. Therefore, we will first present some parts of the proof which are the same for both characteristic two and odd characteristic, and give the parts of the proof for which the characteristic is important. The final part of the proof is again the same for all characteristics.

Throughout this chapter, K will be a finite field of characteristic p . The language is always $(0, 1, t; +, \cdot)$.

7.1 Defining the order of a rational function

We start by giving some definitions about rational functions. After that, we give some lemmas that will often be needed in the proofs below. The main result of this section is the fact that there is a nice criterion to determine whether a rational map has positive order at 0 (theorem 7.6). Later, this will be shown to be a positive existential definition.

Definition 7.1. Let $x \in K(t)$. Then there are unique $a, b \in K[t]$ such that $x = \frac{a}{b}$, $\gcd(a, b) = 1$ and b is monic. a is called the *numerator* of x and b is the *denominator*. Whenever we write $x = \frac{a}{b}$ with $a, b \in K[t]$, we will assume that a and b are coprime, and b is monic. The same holds for $\overline{K}[t]$.

Definition 7.2. A *prime* in $K(t)$ is an irreducible monic polynomial in $K[t]$. If q is a prime and $x \in K(t)$, then we can write $x = q^k \frac{y_1}{y_2}$ with $k \in \mathbb{Z}$, $y_1, y_2 \in K[t]$ and $q \nmid y_1, y_2$. k is called the *order* of x at q and is denoted by $\text{ord}_q(x)$. By convention, $\text{ord}_q(0) = \infty$ for every prime q .

The next lemmas hold for fields of any positive characteristic:

Lemma 7.3. For any $x \in K(t)$ and $n \in \mathbb{N}$, there exists $y \in K(t)$ such that $x^{p^n} - x = y^p - y$.

Proof. If $n = 0$, we can take $y = 0$.

If $n \geq 1$, put $y = \sum_{i=0}^{n-1} x^{p^i}$. Then

$$y^p - y = \left(\sum_{i=0}^{n-1} x^{p^i} \right)^p - \sum_{i=0}^{n-1} x^{p^i} = \sum_{i=1}^n x^{p^i} - \sum_{i=0}^{n-1} x^{p^i} = x^{p^n} - x$$

where we use $\text{char}(K) = p$ in the second equality. □

Lemma 7.4. Let $x \in K(t)$. Then the multiplicity of every pole of $x^p - x$ is divisible by p .

Proof. Write $x = \frac{a}{b}$ with $a, b \in K[t]$. Then $x^p - x = \frac{a^p - ab^{p-1}}{b^p}$, and $\gcd(a^p - ab^{p-1}, b^p) = 1$ (recall our convention that $\gcd(a, b) = 1$). Hence the denominator of $x^p - x$ is b^p , and the poles of x are the zeros of b^p . Any pole q has multiplicity $\text{ord}_q(b^p) = p \text{ord}_q(b)$, which is clearly divisible by p . □

Lemma 7.5. *Let $a, b \in K[t]$ such that $a^2|b$. Then $a|b'$ in $K[t]$, where $'$ denotes the derivative with respect to t . The same holds with $K[t]$ replaced by $\overline{K}[t]$.*

Proof. There exists $c \in K[t]$ or $\overline{K}[t]$ such that $b = a^2c$. Differentiate this to obtain $b' = a'ac + a^2c'$, so $a|b'$, either in $K[t]$ or $\overline{K}[t]$. \square

The next theorem is the only part of the proof in which we use that K is finite.

Theorem 7.6. *Suppose that K has r elements, and let $x \in K(t)$. Then*

$$\text{ord}_t(x) \geq 0 \iff \exists n > 0, \exists a, a_1, \dots, a_{r-1} \in K(t) : \frac{(1 - t^{p^n-1})tx^p}{1 + tx^p} = a^r - a + ta_1^r + t^2a_2^r + \dots + t^{r-1}a_{r-1}^r$$

We use two lemmas in the proof of theorem 7.6:

Lemma 7.7. *Let K be finite, and let $f \in K[t]$ only have simple roots, and suppose that f is not divisible by t . Then there exists $n \geq 1$ such that $f|t^{p^n-1} - 1$.*

Proof. Consider the ring $R = K[t]/(f)$. This ring is finite, since K is finite and every element of R can be represented by an element of $K[t]$ of degree less than $\deg(f)$. Hence the elements $1, t, t^2, \dots$ of R cannot all be distinct, so there exist $k > m$ such that $t^k = t^m$ in R , i.e. $f|t^k - t^m$. Since f is not divisible by t , we have $\gcd(f, t) = 1$, and hence from $f|t^m(t^{k-m} - 1)$ it follows that $f|t^{k-m} - 1$.

Choose $l > 0$ such that $f|t^l - 1$, and write $l = p^r s$ with $p \nmid s$. Then $t^l - 1 = t^{p^r s} - 1 = (t^s - 1)^{p^r}$. Since f has only simple roots, this implies $f|t^s - 1$. In the ring $\mathbb{Z}/s\mathbb{Z}$, p is a unit, so there exist $n \geq 1$ such that $p^n \equiv 1 \pmod{s}$. Then $s|p^n - 1$, say $p^n - 1 = qs$. Notice that

$$t^{p^n-1} - 1 = t^{qs} - 1 = (t^s - 1)(t^{(q-1)s} + t^{(q-2)s} + \dots + 1)$$

so $t^s - 1|t^{p^n-1} - 1$. Hence $f|t^{p^n-1} - 1$. \square

Lemma 7.8. *Suppose that K is a finite field with r elements, and $f \in K[t]$ is divisible by t . Then there exist $a, a_1, \dots, a_r \in K[t]$ such that*

$$f = a^r - a + ta_1^r + t^2a_2^r + \dots + t^{r-1}a_{r-1}^r$$

Proof. We prove this by induction on the degree of f .

If $f = 0$, then we can take $a = a_1 = \dots = a_{r-1} = 0$.

f cannot have degree 0, since $t|f$, so then we would have $f = 0$.

If $\deg(f) = 1$, then $f = ct$ for some $c \in K$. Then we take $a = a_2 = \dots = a_{r-1} = 0$ and $a_1 = c$, since $c^r = c$.

Now suppose that $\deg(f) \geq 2$, and assume that the lemma has already been proven for all polynomials of lower degree. $K[t]$ is a free module of dimension r over $K[t^r]$, and a basis is given by $\{1, t, \dots, t^{r-1}\}$. Hence there exist $g_0, \dots, g_{r-1} \in K[t^r]$ such that $f = g_0 + tg_1 + \dots + t^{r-1}g_{r-1}$. Write $g_i = \sum_{j=0}^d b_{ij}t^{jr}$ with $b_{ij} \in K$. Since K has r elements, we have

$$g_i = \sum_{j=0}^d b_{ij}t^{jr} = \left(\sum_{j=0}^d b_{ij}t^j \right)^r = f_i^r$$

for $f_i = \sum_{j=0}^d b_{ij}t^j \in K[t]$. Hence $f = f_0^r + tf_1^r + \dots + t^{r-1}f_{r-1}^r$. All terms in this sum have different degrees, since the degree of the terms of f_i are i modulo r . Therefore, there is no cancellation of terms, and $\deg(f) \geq \deg(t^i f_i^r) = i + r \deg(f_i)$ for all i such that $f_i \neq 0$. Hence either $f_0 = 0$, or $\deg(f_0) > 0$, and $\deg(f) \geq r \deg(f_0) > \deg(f_0)$. So in any case, we have $\deg(f_0) < \deg(f)$, so we can apply the induction hypothesis to f_0 . This gives that there exist $b, b_1, \dots, b_{r-1} \in K[t]$ such that

$$f_0 = b^r - b + tb_1^r + t^2b_2^r + \dots + t^{r-1}b_{r-1}^r$$

Put $a = f_0 + b$ and $a_i = f_i + b_i$ for all $1 \leq i \leq r-1$. Then

$$\begin{aligned} a^r - a + \sum_{i=1}^{r-1} t^i a_i^r &= (f_0 + b)^r - (f_0 + b) + \sum_{i=1}^{r-1} t^i (f_i + b_i)^r = \\ &= f_0^r + b^r - f_0 - b + \sum_{i=1}^{r-1} t^i f_i^r + \sum_{i=1}^{r-1} t^i b_i^r = \\ &= \sum_{i=0}^{r-1} t^i f_i^r + (b^r - b + \sum_{i=1}^{r-1} t^i b_i^r) - f_0 = f + f_0 - f_0 = f \end{aligned}$$

□

Proof of theorem 7.6. Suppose there exist $n > 0$ and $a, a_1, \dots, a_{r-1} \in K(t)$ such that

$$\frac{(1 - t^{p^n-1})tx^p}{1 + tx^p} = a^r - a + ta_1^r + t^2 a_2^r + \dots + t^{r-1} a_{r-1}^r$$

Every element of $K(t)$ can be viewed as a Laurent series in t . Write $a = \sum_{i=-k}^l b_i t^i$ for $b_i \in K$. Then

$$a^r = \left(\sum_{i=-k}^l b_i t^i \right)^r = \sum_{i=-k}^l b_i t^{ir}$$

since $b_i^r = b_i$ for all $b_i \in K$. Similarly, if $a_j = \sum_{i=-k}^l b_{ij} t^i$, then $a_j^r = \sum_{i=-k}^l b_{ij} t^{ir}$. Hence the constant term of a^r is b_0 , which is also the constant term of a , and the constant term of $t^i a_i$ is zero. It follows that the constant term of $a^r - a + ta_1^r + t^2 a_2^r + \dots + t^{r-1} a_{r-1}^r$ is zero.

Assume that $\text{ord}_t(x) < 0$, say $\text{ord}_t(x) = -l$ for some $l \geq 1$. Then $\text{ord}_t(tx^p) = 1 - pk < 0$ so $\text{ord}_t(1 + tx^p) = \text{ord}_t(tx^p) = 1 - pk$. Since $\text{ord}_t(1 - t^{p^n-1}) = 0$, we have $\text{ord}_t\left(\frac{1-t^{p^n-1}tx^p}{1+tx^p}\right) = 0$, so when we view this as a Laurent series, the constant term is non-zero. Contradiction, so $\text{ord}_t(x) \geq 0$.

For the converse direction, suppose that $\text{ord}_t(x) \geq 0$. Consider $\frac{tx^p}{1+tx^p}$. Write $x = \frac{a}{b}$ with $a, b \in K[t]$. Then $\frac{tx^p}{1+tx^p} = \frac{ta^p}{b^p+ta^p}$. The numerator $b^p + ta^p$ only has simple zeros, since if it has a double zero, then there exists a prime q such that $q^2 | b^p + ta^p$. But then $q | a^p$ by lemma 7.5, but then also $q | b^p$, which contradicts $\text{gcd}(a, b) = 1$.

Apply lemma 7.7 to $b^p + ta^p$. It follows that there exists $n \geq 1$ such that $b^p + ta^p | t^{p^n-1} - 1$. Hence $\frac{1-t^{p^n-1}}{b^p+ta^p}$ is a polynomial, so $\frac{(1-t^{p^n-1})tx^p}{1+tx^p} = \frac{(1-t^{p^n-1})ta^p}{b^p+ta^p}$ is a polynomial, which is divisible by t , since $t \nmid b$ because $\text{ord}_t(x) \geq 0$. Now it follows from lemma 7.8 that a, a_1, \dots, a_{r-1} exist. □

7.2 Finite fields of odd characteristic

In this section, assume that the characteristic p is odd. We will now show that the set $\{t^{p^n} \mid n \in \mathbb{N}\}$ and the relation $\exists n \geq 1 : y = x^{p^n}$ are positive existential over $K(t)$.

Lemma 7.9. *Let $x \in K(t)$. Then*

$$\exists n \in \mathbb{N} : x = t^{p^n} \iff \exists u, v \in K(t) : x - t = u^p - u \wedge x^{-1} - t^{-1} = v^p - v$$

Proof. If $x = t^{p^n}$, then the existence of u and v follows from lemma 7.3, applied to t and t^{-1} , respectively.

Now suppose that u and v exist. First assume x is not a p -th power in $\overline{K}(t)$. By lemma 7.4, the multiplicity of every pole of $u^p - u$ is divisible by p . Hence all poles of $x - t$ has a multiplicity divisible by p . Again by lemma 7.4, all poles of $v^p - v$ also have a multiplicity divisible by p . These poles, except for 0, are exactly the zeros of x . Hence the multiplicities of all zeros of x , except for 0, are divisible by p . It follows that there exists $z \in \overline{K}(t)$ and $0 \leq k \leq p-1$ such that

$x = z^p t^k$. Since we assumed that x is not a p -th power, $k = 0$ cannot occur. Substitute $x = z^p t^k$ into $x^{-1} - t^{-1} = v^p - v$ to obtain

$$z^{-p} t^{-k} - t^{-1} = v^p - v$$

Suppose that $k > 1$. Then $\text{ord}_t(z^{-p} t^{-k}) = -p \text{ord}_t(z) - k \equiv k \pmod{p}$, and $\text{ord}_t(t^{-1}) = -1$, so $\text{ord}_t(z^{-p} t^{-k}) \neq \text{ord}_t(t^{-1})$. Hence

$$\text{ord}_t(z^{-p} t^{-k} - t^{-1}) = \min(\text{ord}_t(z^{-p} t^{-k}), \text{ord}_t(t^{-1})) = \min(-p \text{ord}_t(z) - k, -1) < 0$$

It follows that t is a pole of $z^{-p} t^{-k}$, so it is also a pole of $v^p - v$. From lemma 7.4 we get $p | \text{ord}_t(v^p - v)$. But $\text{ord}_t(z^{-p} t^{-k} - t^{-1}) \equiv -k$ or $-1 \pmod{p}$. This gives a contradiction, since we assumed that $1 < k \leq p - 1$.

Hence $k = 1$, and $x = tz^p$. It follows that

$$u^p - u = x - t = t(z^p - 1) = t(z - 1)^p$$

Write $z - 1 = \frac{a}{b}$ and $u = \frac{c}{d}$ with $a, b, c, d \in \overline{K}[t]$. If $t|b$, then t is a pole of $z - 1$, so it is also a pole of $x - t$ and hence of $u^p - u$. All poles of $u^p - u$ has a multiplicity divisible by p (lemma 7.4), so $p | \text{ord}_t(x)$. But this is impossible, since $x = tz^p$. It follows that $t \nmid b$.

Since $t(z - 1)^p = u^p - u$, we have

$$t \frac{a^p}{b^p} = \frac{c^p}{d^p} - \frac{c}{d} = \frac{c^p - cd^{p-1}}{d^p}$$

Since b is not divisible by t , $\text{gcd}(ta^p, b^p) = \text{gcd}(c^p - cd^{p-1}, d^p) = 1$ and hence $b^p = d^p$ so $b = d$ and $ta^p = c^p - cb^{p-1}$. This shows that $b^2 | ta^p - c^p$, and hence $b|a^p$ by lemma 7.5. But since $\text{gcd}(b, a^p) = 1$, b must be a unit, and since b is monic, we have $b = 1$. It follows that $ta^p = c^p - c$.

Suppose that both sides of the equality are non-zero. Then the degree of the left-hand side is 1 modulo p , but the degree of the right-hand side is divisible by p . Contradiction, so $a = 0$, so $z = 1$ and $x = t$.

Now suppose that x is a p -th power, say $x = x_1^p$ for some $x_1 \in \overline{K}(t)$. Put $u_1 = u - x_1$ and $v_1 = v - x_1^{-1}$. Then u_1 and v_1 satisfy $u_1^p - u_1 = x_1 - t$ and $v_1^p - v_1 = x_1^{-1} - t^{-1}$. If x_1 is a p -th power, then we can apply this again, by putting $x_k = x_{k+1}^p$, $u_{k+1} = u_k - x_{k+1}$ and $v_{k+1} = v_k - x_{k+1}^{-1}$ until we end up with x_n, u_n and v_n , such that x_n is not a p -th power and $u_n^p - u_n = x_n - t$ and $v_n^p - v_n = x_n^{-1} - t^{-1}$. Then it follows from the previous part of the proof that $x_n = t$, and hence $x = t^{p^n}$. \square

Theorem 7.10. *Let $x, y \in K(t)^\times$. Furthermore, let $u = \frac{x^p + t}{x^p - t}$ and $v = \frac{y + t^{p^n}}{y - t^{p^n}}$ for some $n \in \mathbb{N}$. Then $y = x^{p^{n+1}}$ if and only if there exist $\sigma, \tau, \mu, \theta, \delta \in K(t)$ such that the following system of equations holds:*

$$v^2 - u^2 = \sigma^p - \sigma \tag{7.1a}$$

$$v^{-2} - u^{-2} = \tau^p - \tau \tag{7.1b}$$

$$v^2 t^{p^n} - u^2 t = \mu^p - \mu \tag{7.1c}$$

$$v^{-2} t^{-p^n} - u^{-2} t^{-1} = \theta^p - \theta \tag{7.1d}$$

$$v - u = \delta^p - \delta \tag{7.1e}$$

To prove this theorem, we need two lemmas:

Lemma 7.11. *Let $x \in K(t)$. Define $u = \frac{x^p + t}{x^p - t}$. Then all zeros and poles of u are simple.*

Proof. Write $x = \frac{a}{b}$ with $a, b \in K[t]$. Then $u = \frac{a^p + tb^p}{a^p - tb^p}$. Then for all double zeros c of u , the prime $q = t - c$ satisfies $q^2 | a^p + tb^p$ (notice that this also holds if $a^p + tb^p$ and $a^p - tb^p$ have factors in common). Lemma 7.5 shows that $q|b^p$. But $q|a^p + tb^p$, so $q|a^p$ and hence $q|a$. This is clearly a contradiction with $\text{gcd}(a, b) = 1$. Hence all zeros of u are simple.

By considering $a^p - tb^p$ instead of $a^p + tb^p$, the same result follows for the poles. \square

Lemma 7.12. *Let $n \in \mathbb{N}$ and $x, y, u, v \in K(t)$ be as in theorem 7.10. Suppose that $\sigma, \tau, \mu, \theta, \delta \in K(t)$ satisfy the system of equations (7.1). If $n \geq 1$, then there exists $z \in \overline{K}(t)$ such that $y = z^p$.*

Proof. Suppose that z doesn't exist. We first show that v is also not a p -th power in $\overline{K}(t)$. Suppose that v is a p -th power, and let $v = h^p$ for some $h \in \overline{K}(t)$. Then $\frac{y+t^{p^n}}{y-t^{p^n}} = h^p$, so $y(h^p - 1) = t^{p^n}(1 + h^p)$. Notice that $h^p \neq 1$, since $h = v = 1$ would imply that $t^{p^n} = -t^{p^n}$, which is clearly not true in characteristic $p > 2$. Hence we can divide by $h^p - 1$ to obtain

$$y = \left(\frac{t^{p^n-1}(1+h)}{h-1} \right)^p$$

But we assumed that y is not a p -th power, so v is also not a p -th power.

Next, we show that all poles of v different from zero have multiplicity divisible by p . Let $c \in \overline{K}^\times$ be a pole of v , and let $q = t - c$ be the corresponding prime. Then $q \neq t$ and $\text{ord}_q(v) < 0$. Suppose that $p \nmid \text{ord}_q(v)$. If $\text{ord}_q(v) \neq \text{ord}_q(u)$, then

$$\text{ord}_q(\sigma^p - \sigma) = \text{ord}_q(v^2 - u^2) = 2 \min(\text{ord}_q(u), \text{ord}_q(v)) = 2 \text{ord}_q(v) < 0$$

But if $\text{ord}_q(\sigma^p - \sigma) < 0$, then $\text{ord}_q(\sigma^p - \sigma) = p \text{ord}_q(\sigma)$, which is clearly divisible by p , while $\text{ord}_q(v)$ is not. Hence $\text{ord}_q(v) = \text{ord}_q(u)$. Since u only has simple poles and zeros by lemma 7.16, we have $\text{ord}_q(u) \in \{-1, 0, 1\}$, so $\text{ord}_q(\sigma^p - \sigma) \geq \text{ord}_q(u) \geq -1$. As we showed above, $\text{ord}_q(\sigma^p - \sigma)$ is divisible by p if it is negative. Hence $\text{ord}_q(\sigma^p - \sigma) \geq 0$. Similarly, equation (7.1c) implies that $\text{ord}_q(\mu^p - \mu) \geq 0$. But then

$$\text{ord}_q(v^2(t^{p^n} - t)) = \text{ord}_q((v^2t^{p^n} - u^2t) - t(v^2 - u^2)) = \text{ord}_q((\mu^p - \mu) - t(\sigma^p - \sigma)) \geq 0$$

Since $\text{ord}_q(v^2) = 2 \text{ord}_q(v) \leq -2$, we must have $\text{ord}_q(t^{p^n} - t) \geq 2$, so $q^2 | t^{p^n} - t$. By lemma 7.5, this implies that $q | -1$, which is clearly impossible. Hence such q does not exist, and all poles of v , different from zero, have multiplicities divisible by p .

By using equations (7.1b) and (7.1d) instead of (7.1a) and (7.1c), it follows that also all zeros of v , except for 0, have multiplicities divisible by p . Hence there exists $r \in \overline{K}(t)$ and $0 \leq j < p$ such that $v = t^j r^p$. Since v is not a p -th power, we have $j > 0$. Notice that $\text{ord}_t(u) = \text{ord}_t(x^p + t) - \text{ord}_t(x^p - t) = 0$. By equation (7.1b), we have

$$\text{ord}_t(\tau^p - \tau) = \text{ord}_t(v^{-2} - u^{-2}) = \text{ord}_t(v^{-2}) \equiv -2j \pmod{p}$$

But by lemma 7.4, all poles of $\tau^p - \tau$ has a multiplicity divisible by p . Hence $p | 2j$, which is clearly impossible.

It follows that y is a p -th power in $\overline{K}(t)$. □

Proof of theorem 7.10. Suppose that $y = x^{p^{n+1}}$. Then

$$u^{p^n} = \left(\frac{x^p + t}{x^p - t} \right)^{p^n} = \frac{x^{p^{n+1}} + t^{p^n}}{x^{p^{n+1}} - t^{p^n}} = \frac{y + t^{p^n}}{y - t^{p^n}} = v$$

The existence of $\sigma, \tau, \mu, \theta$ and δ satisfying (7.1) now follows from lemma 7.3.

For the converse, suppose that there exist $\sigma, \tau, \mu, \theta$ and δ satisfying (7.1). First we consider the case $n = 0$. Then equation (7.1c) reduces to $t(v^2 - u^2) = \mu^p - \mu$. Combined with equation (7.1a), this gives $t(\sigma^p - \sigma) = \mu^p - \mu$. Write $\sigma = t^i \frac{a}{b}$ and $\mu = t^j \frac{c}{d}$, with $a, b, c, d \in K[t]$, b and d monic and $\text{gcd}(a, b) = \text{gcd}(c, d) = \text{gcd}(abcd, t) = 1$. Then we have

$$\frac{t(t^{pi} a^p - t^i ab^{p-1})}{b^p} = \frac{t^{pj} c^p - t^j cd^{p-1}}{d^p}$$

In both fractions, the numerator is coprime to the denominator, and hence we have $b^p = d^p$, which implies $b = p$ since b and p are both monic. This gives

$$t(t^{pi} a^p - t^i ab^{p-1}) = t^{pj} c^p - t^j cb^{p-1} \tag{7.2}$$

Suppose that $i < 0$. Then $\text{ord}_t(t(t^{pi}a^p - t^i ab^{p-1})) = pi + 1 < 0$, so the right-hand side of (7.2) also has negative order at t . Hence $j < 0$, and the order is $\text{ord}_t(t^{pj}c^p - t^j cb^{p-1}) = pj$. But then $pi + 1 = pj$, which is clearly not possible. It follows that $i \geq 0$, so $j \geq 0$.

Notice that $b^{p-1}(t^j c - t^{i+1}a) = t^{pj}c^p - t^{pi+1}a^p$ by equation (7.2), so $b^{p-1} | t^{pj}c^p - t^{pi+1}a^p$. By lemma 7.5, this implies $b | a^p$. But b is coprime to a , so b must be a unit in $K[t]$, and since b is monic, we have $b = 1$ and equation (7.1) reduces to $t(t^{pi}a^p - t^i a) = t^{pj}c^p - t^j c$. If both sides of this equality are non-zero, then the degree of the left-hand side is $pi + 1 + p \deg(a) \equiv 1 \pmod{p}$, while the right-hand side has degree $pj + p \deg(c) \equiv 0 \pmod{p}$. Hence $v^2 - u^2 = \sigma^p - \sigma = t^{pi}a^p - t^i a = 0$, so $u^2 = v^2$ and $u = \pm v$.

Suppose that $u = -v$. Then by equation (7.1e), $-2u = v - u = \delta^p - \delta$. By lemma 7.11, u has only simple poles, but the multiplicities of the poles of $\delta^p - \delta$ are divisible by p by lemma 7.4. Hence u has no poles, i.e. $u \in K[t]$. Now write $x = \frac{e}{f}$ with $e, f \in K[t]$. Then

$$u = \frac{x^p + t}{x^p - t} = \frac{e^p + tf^p}{e^p - tf^p}$$

Let $q = \gcd(e^p + tf^p, e^p - tf^p)$. Then q divides both e^p and tf^p . Since e and f are coprime, this implies $q | t$. But u has no poles, so $e^p - tf^p | e^p + tf^p$, and hence $q = e^p - tf^p$. It follows that $e^p - tf^p | t$. By considering the degree of both sides, we obtain that e and f must both have degree zero, so $e, f \in K$. Since f is monic, $f = 1$, and hence $t - e^p | t$. This is only possible if $e = 0$, so $x = 0$, which we assumed not to be the case. Hence the case $u = -v$ cannot occur, so we always have $u = v$. This gives

$$\frac{x^p + t}{x^p - t} = \frac{y + t}{y - t}$$

Solving this equation for y gives $y = x^p = x^{p^{n+1}}$. This proves the lemma in the case $n = 0$.

Now suppose $n > 0$. Then, by lemma 7.12, y is a p -th power, say $y = z_1^p$. Put $v_1 = \frac{z_1 + t^{p^{n-1}}}{z_1 - t^{p^{n-1}}}$, $\sigma_1 = \sigma - v_1^2$, $\tau_1 = \tau - v_1^{-2}$, $\mu_1 = \mu - v_1^2 t^{p^{n-1}}$, $\theta_1 = \theta - v_1^{-2} t^{-p^{n-1}}$ and $\delta_1 = \delta - v_1$. An easy calculation shows that $(\sigma_1, \tau_1, \mu_1, \theta_1, \delta_1)$ satisfy the system of equations (7.1) with v replaced by v_1 and n by $n-1$. If $n-1 = 0$, then it follows from the previous argument that $z_1 = x^p$, and hence $y = x^{p^2}$. Otherwise, z_1 has to be a p -th power by lemma 7.12, so there exists z_2 such that $z = z_2^p$. By applying this change of variables n times, we will find $v_n, \sigma_n, \tau_n, \mu_n, \theta_n$ and δ_n satisfying this system of equations for $n = 0$, with $v_i = v_{i+1}^p$ for all i . But then by the previous proof $v_n = u$, so $v_{n-1} = u^p, \dots, v_1 = u^{p^{n-1}}$ and $v = u^{p^n}$. Hence

$$\frac{x^{p^{n+1}} + t^{p^n}}{x^{p^{n+1}} - t^{p^n}} = \left(\frac{x^p + t}{x^p - t} \right)^{p^n} = \frac{y + t^{p^n}}{y - t^{p^n}}$$

so $y = x^{p^{n+1}}$. □

7.3 Finite fields of characteristic two

In this part, we will prove a lemma and a theorem similar to lemma 7.9 and theorem 7.10. Throughout this section, K is a finite field of characteristic two.

From lemma 7.9 it follows that, if the characteristic is odd, the set $\{t^{2^n} | n \in \mathbb{N}\}$ can be defined by two equations: $x - t = u^p - u$ and $x^{-1} - t^{-1} = v^p - v$. Therefore, we could expect that in characteristic two, the same set is defined by the equations $x + t = u^2 + u$ and $x^{-1} + t^{-1} = v^2 + v$. It turns out that this is not the case. For example, take $x = \frac{(1+t+t^3)t}{(1+t^2+t^3)^2}$, $u = \frac{t^3}{1+t^2+t^3}$ and $v = \frac{1}{1+t+t^3}$. An easy calculation shows that $x + t = u^2 + u$ and $x^{-1} + t^{-1} = v^2 + v$, but it is clear that $x \notin \{t^{2^n} | n \in \mathbb{N}\}$. Fortunately, there exists another characterization of this set:

Lemma 7.13. *Let $x \in K(t)$. There exists $n \geq 1$ such that $x = t^{2^n}$ if and only if there exist $u, v, w, s \in K(t)$ satisfying the following system of equations:*

$$x + t = u^2 + u \quad (7.3a)$$

$$u = w^2 + t \quad (7.3b)$$

$$x^{-1} + t^{-1} = v^2 + v \quad (7.3c)$$

$$v = s^2 + t^{-1} \quad (7.3d)$$

The proof of this lemma is more difficult than the proof of lemma 7.9, and we need another lemma:

Lemma 7.14. *Let $x \in K(t)$. If there exist $a, b \in K[t]$ with $x = t \frac{a^2}{b^2}$ and $\gcd(at, b) = 1$, then the system of equations (7.3) has no solutions in $K(t)$.*

Proof. Let $x = t \frac{a^2}{b^2}$, and suppose that there exists a solution (u, v, w, s) . Write $u = \frac{c}{d}$, $w = \frac{k}{l}$, $v = \frac{e}{f}$ and $s = \frac{g}{h}$ with $c, d, e, f, g, h, k, l \in K[t]$. Equation (7.3a) gives

$$\frac{t(a^2 + b^2)}{b^2} = t \frac{a^2}{b^2} + t = x + t = u^2 + u = \frac{c^2 + cd}{d^2}$$

Since both fractions have coprime numerators and denominators, this implies $b^2 = d^2$ and hence $b = d$.

By equation (7.3a), $\frac{c}{d} = u = w^2 + t = \frac{k^2 + l^2}{l^2}$, so $b = d = l^2$. Equation (7.3c) gives

$$\frac{a^2 + b^2}{ta^2} = \frac{b^2}{ta^2} + \frac{1}{t} = \frac{1}{x} + \frac{1}{t} = v^2 + v = \frac{e^2 + ef}{f^2}$$

so $(a^2 + b^2)f^2 = t(e^2 + ef)a^2$. Hence $t|(a^2 + b^2)f^2$. Suppose that $t \nmid a^2 + b^2$. Then $t|f^2$, so $\text{ord}_t((a^2 + b^2)f^2) = 2 \text{ord}_t(f)$. Since $t|f$ and $\gcd(e, f) = 1$, we have $t \nmid e$ and $t \nmid e^2 + ef$, and hence $\text{ord}_t(t(e^2 + ef)a^2) = 1 + 2 \text{ord}_t(a)$. But $2 \text{ord}_t(f) = 1 + 2 \text{ord}_t(a)$ clearly gives a contradiction, so it must hold that $t|a^2 + b^2$. Write $a(t) = a_1(t)t + a_0$ and $b(t) = b_1(t)t + b_0$, with $a_1, b_1 \in K[t]$ and $a_0, b_0 \in K$. Then $a_0^2 + b_0^2 \equiv a^2 + b^2 \equiv 0 \pmod{t}$, so $t|a_0^2 + b_0^2$. Hence $a_0^2 + b_0^2 = 0$, so $a_0^2 = b_0^2$ and hence $a_0 = b_0$. Notice that $\gcd(\frac{a^2 + b^2}{t}, a^2) | \gcd(a^2 + b^2, a^2) = 1$, so in the identity $\frac{a^2 + b^2}{a^2} = \frac{e^2 + ef}{f^2}$, both fractions have coprime numerators and denominators. It follows that $a = f$.

Now equation (7.3d) gives

$$\frac{e}{a} = \frac{e}{f} = v = s^2 + \frac{1}{t} = \frac{g^2}{h^2} + \frac{1}{t} = \frac{g^2t + h^2}{h^2t}$$

Suppose that $\gcd(g^2t + h^2, h^2t) = 1$. Then $a = h^2t$ so $a_0 = 0$, and hence $b_0 = 0$, and $t|b$. But this contradicts $\gcd(at, b) = 1$. Hence $\gcd(g^2t + h^2, h^2t) \neq 1$. Let $q = \gcd(g^2t + h^2, h^2t)$. If s is a prime factor of q , then $s|h$ or $s = t$. If $s|h$, then we must also have $s|g^2t$, and since $\gcd(g, h) = 1$, this implies $s = t$. Hence q is a power of t , say $q = t^r$ with $r \geq 1$. Since $q|g^2t + h^2$ and $r \geq 1$, we have $t|h^2$, so $t|h$ and $t \nmid g$. Hence $t^2 \nmid g^2t + h^2$. This implies that $q = t$.

Write $h = th'$. Then we have

$$\frac{e}{a} = \frac{g^2t + h^2}{h^2t} = \frac{g^2t + t^2(h')^2}{(h')^2t^3} = \frac{g^2 + t(h')^2}{t^2(h')^2}$$

with $\gcd(g^2 + t(h')^2, t^2(h')^2) = 1$. But now it follows that $a = t^2(h')^2$, so $t|a$ and again $b_0 = a_0 = 0$. This again gives a contradiction, so the solution (u, w, v, s) doesn't exist. \square

Proof of lemma 7.13. Suppose that $x = t^{2^n}$. If $n = 1$, then we can take $u = t, v = t^{-1}$ and $w = s = 0$. If $n \geq 2$, then put $w = t^{2^{n-2}} + t^{2^{n-3}} + \dots + t$, $v = t^{-2^{n-2}} + t^{-2^{n-3}} + \dots + t^{-1}$ and $u = w^2 + t, v = s^2 + t$. An easy calculation, as in the proof of lemma 7.3, shows that $x + t = u^2 + u$ and $x^{-1} + t^{-1} = v^2 + v$.

For the converse direction, suppose that the system of equations has a solution (u, v, w, s) . First we show that x must be a square. Suppose that x is not a square. Since $x + t = u^2 + u$, all poles of $x + t$ has even multiplicity (lemma 7.4), so the same holds for all poles of x . Similarly, by $x^{-1} + t^{-1} = v^2 + v$, all zeros of x , except for 0, have even multiplicity. It follows that $x = \frac{a^2}{b^2}t^k$ for some $a, b \in K[t]$ with $\gcd(at, b) = 1$. We can assume $k \in \{0, 1\}$, and since x is not a square, we have $k = 1$ so $x = t\frac{a^2}{b^2}$. Suppose $t|b$. Then $t^2|b^2$, so 0 is a pole of x . But then it is also a pole of $u^2 + u$, and hence it has multiplicity divisible by two (see lemma 7.4). But then x would be a square. Hence $t \nmid b$, and it follows from lemma 7.14 that (u, v, w, s) cannot exist. This shows that x is a square, say $x = z^2$.

Suppose that z is not a square. Notice that

$$(u + z)^2 + (u + z) = u^2 + u + z^2 + z = x + t + x + z = z + t \quad (7.4)$$

and $z^{-2} = x^{-1} = v^2 + v + t^{-1} = (s^2 + t^{-1})^2 + (s^2 + t^{-1}) + t^{-1} = (s^2 + s + t^{-1})^2$, so $z^{-1} + t^{-1} = s^2 + s$. In the same way as before, we get $z = t\frac{a^2}{b^2}$ for some $a, b \in K[t]$ with $\gcd(at, b) = 1$ and b monic. Write $u + z = \frac{c}{d}$, $w = \frac{k}{l}$ with $c, d, k, l \in K[t]$. Then $\frac{c^2 + cd}{d^2} = (u + z)^2 + (u + z) = z + t = \frac{(a^2 + b^2)t}{b^2}$, so $b = d$ (since both fractions have coprime numerators and denominators). Hence

$$\frac{c}{b} = \frac{c}{d} = u + z = w^2 + t + z = \frac{k^2}{l^2} + t + t\frac{a^2}{b^2} = \frac{(kb)^2 + t(lb)^2 + t(la)^2}{(lb)^2}$$

Suppose that $q := \gcd((kb)^2 + t(lb)^2 + t(la)^2, (lb)^2) > 1$, and assume $q \nmid l^2$. Define $h = \gcd(q, n^2)$, and $g = \frac{q}{h}$. Then g is non-constant, and $\gcd(g, n^2) = 1$. Since $g|q$ and $q|(lb)^2$, this implies $g|b^2$. We also have $g|(kb)^2 + t(lb)^2 + t(la)^2$, and hence $g|t(la)^2$ so $g|ta^2$. Since $g|b^2$ and $\gcd(a, b) = 1$, we must have $\gcd(g, a^2) = 1$ and hence $g|t$, so $g = t$. But then $t|b$, which we assumed not to be the case. Therefore, if $q > 1$ then $q|l^2$. Hence there exists l' such that $ql' = l^2$. This clearly also holds if $q = 1$. We have

$$\frac{c}{b} = \frac{(kb)^2 + t(lb)^2 + t(la)^2}{(lb)^2} = \frac{(kb)^2 + t(kb)^2 + t(la)^2}{l'b^2}$$

so $b = l'b^2$. This implies $l' = b = 1$, so $z = ta^2$ and hence

$$t(a^2 + 1) = z + t = (u + z)^2 + (u + z) = \frac{c^2 + cd}{d^2} = \frac{c^2 + cb}{b^2} = c^2 + c$$

by equation (7.4). If both sides of this equality are non-zero, then the degree of the left-hand side is odd, but the degree of the right-hand side is even. Hence $a^2 + 1 = 0$ so $a = 1$ and $z = t$, so $x = t^2$.

Now suppose z is a square, say $z = z_1^2$. Put $u_1 = u + z$, $w_1 = w + z_1$, $v_1 = v + z^{-1}$ and $s_1 = s + z_1^{-1}$. Then (u_1, w_1, v_1, s_1) is a solution of the system of equations (7.3), with x replaced by z . If z_1 is not a square, it follows from the above that $z_1 = t$, so $z = t^2$ and $x = t^4$. Otherwise, we can repeat this until we end up with (u_n, w_n, v_n, s_n) satisfying the system of equations for $z_n = \sqrt[n+1]{x}$, such that z_n is not a square. Then $z_n = t$ so $x = t^{2^{n+1}}$ for some $n \in \mathbb{N}$. \square

In theorem 7.10, we used $u = \frac{x^p + t}{x^p - t}$ and $v = \frac{y + t^{p^n}}{y - t^{p^n}}$. In characteristic 2, this gives $u = \frac{x^2 + t}{x^2 + t} = 1$ and $v = \frac{y + t^{2^n}}{y + t^{2^n}} = 1$. Therefore, we will use $u = \frac{x^2 + t^2 + t}{x^2 + t}$ and $v = \frac{y + t^{2^{n+1}} + t^{2^n}}{y + t^{2^n}}$. If we transform the system of equations in theorem 7.10 to characteristic two, we get

$$\begin{aligned} v^2 + u^2 &= \sigma^2 + \sigma \\ v^{-2} + u^{-2} &= \tau^2 + \tau \\ v^2 t^{2^n} + u^2 t &= \mu^2 + \mu \\ v^{-2} t^{-2^n} + u^{-2} t^{-1} &= \theta^2 + \theta \\ v + u &= \delta^2 + \delta \end{aligned}$$

Notice that $v + u = \delta^2 + \delta$ implies $v^2 + u^2 = (v + u)^2 = (\delta^2 + \delta)^2 = \delta^4 + \delta^2$, so we can take $\sigma = \delta^2$. Therefore, we will leave out the first equation. To keep the symmetry between the equations, we will replace the second equation by $v^{-1} + u^{-1} = \tau^2 + \tau$. This suggests the following theorem:

Theorem 7.15. *Let $x, y \in K(t)^\times$. Furthermore, let $u = \frac{x^2+t^2+t}{x^2+t}$ and $v = \frac{y+t^{2n+1}+t^{2n}}{y+t^{2n}}$ for some $n \in \mathbb{N}$. Then $y = x^{2^{n+1}}$ if and only if there exist $\sigma, \tau, \mu, \theta \in K(t)$ such that the following system of equations holds:*

$$v + u = \sigma^2 + \sigma \tag{7.5a}$$

$$v^{-1} + u^{-1} = \tau^2 + \tau \tag{7.5b}$$

$$v^2 t^{2n} + u^2 t = \mu^2 + \mu \tag{7.5c}$$

$$v^{-2} t^{-2n} + u^{-2} t^{-1} = \theta^2 + \theta \tag{7.5d}$$

In the proof, we need two lemmas similar to lemmas 7.11 and 7.12:

Lemma 7.16. *Let $x \in K(t)$. Define $u = \frac{x^2+t^2+t}{x^2+t}$. Then all zeros and poles of u are simple.*

Proof. The proof is similar to the proof of lemma 7.11. □

Lemma 7.17. *Let $n \in \mathbb{N}$ and $x, y, u, v \in K(t)$ be as in theorem 7.15. Suppose that $\sigma, \tau, \mu, \theta \in K(t)$ satisfy the system of equations (7.5). If $n \geq 1$, then there exists $z \in \overline{K}(t)$ such that $y = z^2$.*

Proof. This proof is similar to the proof of lemma 7.12.

Suppose that z does not exist. We first show that v is also not a square in \overline{K} . Suppose that v is a square, and let $v = h^2$ for some $h \in \overline{K}$. Then $\frac{y+t^{2n+1}+t^{2n}}{y+t^{2n}} = h^2$, so $y(h^2+1) = h^2 t^{2n} + t^{2n+1} + t^{2n}$. Notice that $h^2 \neq 1$, since $h = v = 1$ would imply $t^{2n+1} = 0$. Hence we can divide by $h^2 + 1$ to obtain

$$y = \left(\frac{ht^{2n-1} + t^{2n} + t^{2n-1}}{h+1} \right)^2$$

But we assumed y is not a square, so v is also not a square.

Next, we show that all poles of v different from zero have even multiplicity. Let $c \in \overline{K}^\times$ be a pole of v , and let $q = t - c$ be the corresponding prime. Then $q \neq t$ and $\text{ord}_q(v) < 0$. Assume that $\text{ord}_q(v)$ is odd. If $\text{ord}_q(v) \neq \text{ord}_q(u)$, then

$$\text{ord}_q(\sigma^2 + \sigma) = \text{ord}_q(u + v) = \min(\text{ord}_q(u), \text{ord}_q(v)) = \text{ord}_q(v) < 0$$

But if $\text{ord}_q(\sigma^2 + \sigma) < 0$, then $\text{ord}_q(\sigma^2 + \sigma) = 2\text{ord}_q(\sigma)$, which is clearly even, while $\text{ord}_q(v)$ is odd. Hence $\text{ord}_q(v) = \text{ord}_q(u)$. Since u only has simple poles and zeros by lemma 7.16, we have $\text{ord}_q(u) \in \{-1, 0, 1\}$, so $\text{ord}_q(\sigma^2 + \sigma) \geq \text{ord}_q(u) \geq -1$. As we showed above, $\text{ord}_q(\sigma^2 + \sigma)$ is even if it is negative. Hence $\text{ord}_q(\sigma^2 + \sigma) \geq 0$, and hence also $\text{ord}_q(\sigma^4 + \sigma^2) \geq 0$. Similarly, equation (7.5c) implies that $\text{ord}_q(\mu^p - \mu) \geq 0$. But then

$$\text{ord}_q(v^2(t^{2n} + t)) = \text{ord}_q((v^2 t^{2n} + u^2 t) + t(v^2 + u^2)) = \text{ord}_q((\mu^2 + \mu) + t(\sigma^4 + \sigma^2)) \geq 0$$

Since $\text{ord}_q(v^2) = 2\text{ord}_q(v) \leq -2$, we must have $\text{ord}_q(t^{2n} + t) \geq 2$, so $q^2 | t^{2n} + t$. By lemma 7.5, this implies that $q|1$, which is clearly impossible. Hence such q doesn't exist, and all poles of v , different from zero, have even multiplicity.

By using equations (7.5b) and (7.5d) instead of (7.5a) and (7.5c), it follows that also all zeros of v , except for 0, have even multiplicities. The rest of the proof is the same as in lemma 7.12. Notice that we even get $j = 1$ instead of $0 < j < p$.

It follows that y is a square in $\overline{K}(t)$. □

Proof of theorem 7.15. This proof is similar to the proof of theorem 7.10. Therefore, we only mention the steps that need to be changed.

The proof of the ‘only if’ part is the same as in theorem 7.10.

For the converse, suppose that there exist σ, τ, μ and θ satisfying (7.5). First we consider the case $n = 0$. Since $v + u = \sigma^2 + \sigma$, we have $v^2 + u^2 = \sigma^4 + \sigma^2$ and $\mu^2 + \mu = t(v^2 + u^2)$. Again write $\sigma = t^i \frac{a}{b}$ and $\mu = t^j \frac{c}{d}$, with $a, b, c, d \in K[t]$, b and d monic and $\gcd(a, b) = \gcd(c, d) = \gcd(abcd, t) = 1$. Then we have

$$\frac{t(t^{4i}a^4 + t^{2i}a^2b^2)}{b^4} = \frac{t^{2j}c^2 + t^jcd}{d^2}$$

so $b^2 = d$, and

$$t(t^{4i}a^4 + t^{2i}a^2b^2) = t^{2j}c^2 + t^jcb^2$$

In the same way as in theorem 7.10, it follows that $i, j \geq 0$. Notice that $b^2(t^{2i+1}a^2 + t^j c) = t^{4i+1}a^4 + t^{2j}c^2$. By lemma 7.5, this implies $b|a^4t^{4i}$. But b is coprime to a and t , so again we have $b = 1$. Hence $t(t^{4i}a^4 + t^{2i}a^2) = t^{2j}c^2 + t^j c$. Considering the parity of the degrees of both sides of this equation, we find that both sides are zero, and hence $v^2 + u^2 = \sigma^4 + \sigma^2 = t(t^{4i}a^4 + t^{2i}a^2) = 0$, so $u^2 = v^2$ and $u = v$ (notice that this proof is easier than the proof of theorem 7.10, because we don’t have to consider the case $u = -v$.) It follows that $y = x^2$. This proves the lemma in the case $n = 0$.

Now suppose $n > 0$. Then, by lemma 7.17, y is a square, say $y = z_1^2$. The same argument applies as in the proof of theorem 7.10, using $v_1 = \frac{z_1 + t^{2^n} + t^{2^{n-1}}}{z_1 + t^{2^{n-1}}}$, $\sigma_1 = \sigma + v_1$, $\tau_1 = \tau + v_1^{-1}$, $\mu_1 = \mu + v_1^2 t^{2^{n-1}}$ and $\theta_1 = \theta + v_1^{-2} t^{-2^{n-1}}$. \square

7.4 The undecidability of the diophantine theory of $(K(t); +, \cdot)$

Lemma 7.18. *In both odd characteristic and characteristic two, the relation $\exists n \in \mathbb{N} : y = x^{p^n}$ is positive existential.*

Proof. The set $\{t^{p^n} \mid n \in \mathbb{N}\}$ is positive existential by lemma 7.9 or lemma 7.13. By theorems 7.10 and 7.15, the relation $\exists n \in \mathbb{N} : y = x^{p^{n+1}}$ is positive existential for $xy \neq 0$. Since

$$\exists n \in \mathbb{N} : y = x^{p^n} \iff (x = y \vee \exists n \in \mathbb{N} : y = x^{p^{n+1}})$$

the relation $\exists n \in \mathbb{N} : y = x^{p^{n+1}}$ is also positive existential. \square

Now we can finally prove Hilbert’s tenth problem for rational function fields over finite fields:

Theorem 7.19. *If K is a finite field, then there is no algorithm to decide whether a diophantine equation with coefficients in $\mathbb{Z}/p\mathbb{Z}[t]$ has a solution in $K(t)$.*

Proof. Define the equivalence relation \sim on $K(t)$ by

$$x \sim y \iff \text{ord}_t(x) = \text{ord}_t(y)$$

Then multiplication in $K(t)/\sim$ is well-defined.

Denote the relation $\exists n \in \mathbb{N} : y = x^{p^n}$ by $x \wr y$. Notice that \wr induces a relation on $K(t)/\sim$, that we shall also denote \wr .

Define the predicate P on $K(t)$ by

$$P(x) \iff \exists n \in \mathbb{N} : x = t^{p^n}$$

Finally, define the predicate O on $K(t)$ by

$$O(x) \iff \text{ord}_t(x) \geq 0$$

This induces a predicate on $K(t)/\sim$.

Consider the language $L = (0, 1; \cdot; \wr, O)$. We claim that the positive existential theory of $K(t)/\sim$ in L is undecidable. Recall that the positive existential theory of \mathbb{N} in $(0, 1; +; |_p)$ is undecidable. Define an injective map by $\phi : \mathbb{N} \rightarrow K(t)/\sim : n \mapsto [t^n]$, where $[t^n]$ denote the equivalence class of t^n under the relation \sim . Then

$$\phi(\mathbb{N}) = \{x \in K(t)/\sim \mid \text{ord}_t(x) \geq 0\} = \{x \in K(t)/\sim \mid O(x) \text{ holds}\}$$

Futhermore, addition of elements of \mathbb{N} corresponds to multiplication of elements of $K(t)/\sim$, and

$$\begin{aligned} k|_p l &\iff \exists n \in \mathbb{N} : l = kp^n \iff \\ &\iff \exists n \in \mathbb{N} : \phi(l) = \phi(kp^n) = [t^{kp^n}] = [t^k]^{p^n} = \phi(k)^{p^n} \iff \phi(k) \wr \phi(l) \end{aligned}$$

Hence this is a positive existential model of $(\mathbb{N}, (0, 1; +; |_p))$ in $(K(t)/\sim, L)$. It follows by theorem 4.2 that the positive existential theory of $K(t)/\sim$ in L is undecidable.

Now we show that the positive existential theory of $K(t)$ in L is also undecidable. Take a positive existential formula over $K(t)/\sim$ in L . Then we can ‘translate’ this into a positive existential formula over $K(t)$ in L as follows: in $K(t)$, we get formulas of the form $[x] = [y] \cdot [z]$, $[x] \wr [y]$ and $O([x])$. The first time a variable $[x]$ appears, replace it by x_1 . The second time it appears, replace it by x_2 , and add the relation $x_1 \sim x_2$. Continue like this, until we have a formula over $K(t)$, containing the relation \sim , that is equivalent to the formula in $K(t)/\sim$ we started with. Now notice that \sim is positive existential over $K(t)$ in L , since

$$x \sim y \iff \text{ord}_t(x) = \text{ord}_t(y) \iff \text{ord}_t\left(\frac{x}{y}\right) = 0 \iff \text{ord}_t\left(\frac{x}{y}\right) \geq 0 \wedge \text{ord}_t\left(\frac{x}{y}\right) \leq 0$$

Hence every positive existential formula over $K(t)/\sim$ in L can be ‘translated’ to an equivalent formula over $K(t)$ in L . Hence if the positive existential theory of $K(t)$ in L is undecidable, then so is the positive existential theory of $K(t)/\sim$ in L . It follows that the positive existential theory of $K(t)$ in L is undecidable.

Finally, we show that the predicate O on $K(t)$ is positive existential over $K(t)$ in $(0, 1, t; +, \cdot)$. First notice that the predicate P is positive existential by lemma 7.9 or lemma 7.13. It follows that O is positive existential, since by theorem 7.6

$$\begin{aligned} O(x) &\iff \text{ord}_t(x) \geq 0 \iff \\ \exists n > 0, \exists a, a_1, \dots, a_{r-1} \in K(t) : &\frac{(1 - t^{p^n-1})tx^p}{1 + tx^p} = a^r - a + ta_1^r + t^2 a_2^r + \dots + t^{r-1} a_{r-1}^r \iff \\ \exists z, w, a, a_1, \dots, a_{r-1} \in K(t) : &P(z) \wedge w = z^p \wedge \frac{(1 - \frac{w}{t})tx^p}{1 + tx^p} = a^r - a + ta_1^r + t^2 a_2^r + \dots + t^{r-1} a_{r-1}^r \end{aligned}$$

This shows that O is positive existential.

By lemma 7.18, \wr is also positive existential. It follows that the positive existential theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable. By lemma 1.15, the diophantine theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable. \square

Chapter 8

Formal groups

Up to now, we have proven that the diophantine or positive existential theories of several sets and languages are undecidable, by using elementary methods. In the next chapters, we will use methods based on the theory of elliptic curves. Using points on an elliptic curve as a model for \mathbb{Z} , in chapter 9 we will prove the undecidability of Hilbert's tenth problem over $K(t)$ in the language $(0, 1, t; +, \cdot)$ for certain infinite fields of characteristic zero. After that, in chapter 10 we try give another proof for the fact that exponentiation in \mathbb{N} is diophantine in $(0, 1; +, \cdot)$, by using sequences associated to an elliptic curve. Finally, in chapter 11, we will try to prove the undecidability of Hilbert's tenth problem over infinite fields of positive characteristic in $(0, 1, t; +, \cdot)$, again using points on an elliptic curve as a model for \mathbb{Z} .

In these chapters, we will need some properties of elliptic curves, that are most easily proven by using the formal group of the elliptic curve. Therefore, this chapter will be devoted to formal groups. We will first introduce formal groups and give some basic facts about these. After that, we will consider multiplication by integers, and consider formal groups over valuation rings. These sections handle formal groups in general. In the last section, we will finally define the formal group of an elliptic curve.

8.1 Definition and basic properties

Definition 8.1. Let R be a ring. A (*one-parameter commutative*) formal group \mathcal{F} defined over R is a power series $F(X, Y) \in R[[X, Y]]$ satisfying

(i) $F(X, Y) = X + Y + G(X, Y)$, where all terms of $G(X, Y)$ have degree at least 2.

(ii) Associativity: $F(X, F(Y, Z)) = F(F(X, Y), Z)$.

(iii) Commutativity: $F(X, Y) = F(Y, X)$.

(iv) Inverse: there exists a unique power series $i(T) \in R[[T]]$ satisfying $F(T, i(T)) = 0$.

(v) $F(X, 0) = X$ and $F(0, Y) = Y$.

$F(X, Y)$ is the *formal group law* of \mathcal{F} . We will also write (\mathcal{F}, F) to denote this formal group.

Example 8.2. An important example of a formal group is the *formal additive group* $\widehat{\mathbb{G}}_a$ with formal group law $F(X, Y) = X + Y$.

Remark 8.3. Write $F(X, Y) = X + Y + G(X, Y)$, and suppose that G has a term X^k with $k \geq 2$. Then $F(X, 0) = X + X^k + \dots$, contradicting 8.1(v). Hence $Y|G(X, Y)$, and similarly, $X|G(X, Y)$. This implies that $XY|G(X, Y)$.

Definition 8.4. Let (\mathcal{F}, F) and (\mathcal{G}, G) be formal groups over R . A *homomorphism* from \mathcal{F} to \mathcal{G} is a power series $f(T) \in R[[T]]$ satisfying

$$f(F(X, Y)) = G(f(X), f(Y))$$

\mathcal{F} and \mathcal{G} are *isomorphic* if there exist homomorphisms $f : \mathcal{F} \rightarrow \mathcal{G}$ and $g : \mathcal{G} \rightarrow \mathcal{F}$ such that

$$f(g(T)) = g(f(T)) = T$$

Remark 8.5. Notice that a homomorphism f has no constant term: since G has no constant term, the constant term of $G(f(X), g(Y))$ is also zero, so $f(F(X, Y))$ has no constant term. This implies that f has no constant term. Hence for all homomorphisms f , it holds that $f(0) = 0$.

Lemma 8.6. *Let (\mathcal{F}, F) be a formal group. Then the inverse i is an isomorphism of (\mathcal{F}, F) , so it has no constant term.*

Proof. Notice that

$$\begin{aligned} F(F(X, Y), F(i(X), i(Y))) &= F(F(Y, X), F(i(X), i(Y))) = F(Y, F(X, F(i(X), i(Y)))) = \\ &= F(Y, F(F(X, i(X)), i(Y))) = F(Y, F(0, i(Y))) = F(Y, i(Y)) = 0 \end{aligned}$$

i is the unique power series satisfying $F(T, i(T)) = 0$, so this implies $i(F(X, Y)) = F(i(X), i(Y))$ (by taking $T = F(X, Y)$). i has no constant term by remark 8.5. \square

Lemma 8.7. (i) *Let $f(T) = aT + \dots \in R[[T]]$ with $a \in R^\times$. Then there exists a unique power series $g(T) \in R[[T]]$ such that $f(g(T)) = T$. g also satisfies $g(f(T)) = T$.*

(ii) *If f is a homomorphism $(\mathcal{F}, F) \rightarrow (\mathcal{G}, G)$, then g is a homomorphism $(\mathcal{G}, G) \rightarrow (\mathcal{F}, F)$.*

Proof. (i) We define a sequence of polynomials $g_n(T) \in R[T]$ of degree n such that

$$f(g_n(T)) \equiv T \pmod{T^{n+1}} \quad \text{and} \quad g_{n+1}(T) \equiv g_n(T) \pmod{T^{n+1}}$$

Let $g_1(T) = a^{-1}T$. Then it is clear that $f(g_1(T)) = T + \dots$, so $f(g_1(T)) \equiv T \pmod{T^2}$.

Now we define g_{n+1} recursively, by $g_{n+1}(T) = g_n(T) + \lambda T^{n+1}$ for some $\lambda \in R$ such that the first condition holds. Notice that the second condition holds for all λ . By definition of g_n , $f(g_n(T)) \equiv T \pmod{T^{n+1}}$, so $f(g_n(T)) = T + b_n T^{n+1} + h_n(T) T^{n+2}$ for some $b_n \in R$ and $h_n \in R[[T]]$. Take $\lambda = \frac{-b_n}{a}$. Then, if we write $f(T) = \sum_{k=1}^{\infty} a_k T^k$,

$$\begin{aligned} f(g_{n+1}(T)) &= f(g_n(T) + \lambda T^{n+1}) = \sum_{k=1}^{\infty} a_k (g_n(T) + \lambda T^{n+1})^k = \\ &= \sum_{k=1}^{\infty} a_k \sum_{i=0}^k \binom{k}{i} g_n(T)^{k-i} \lambda^i T^{i(n+1)} \equiv \sum_{k=1}^{\infty} a_k (g_n(T)^k + k\lambda g_n(T)^{k-1} T^{n+1}) = \\ &= f(g_n(T)) + \sum_{k=1}^{\infty} k a_k \lambda g_n(T)^{k-1} T^{n+1} \pmod{T^{n+2}} \end{aligned}$$

Since $f(g_n(T)) \equiv T + b_n T^{n+1} \pmod{T^{n+2}}$ and g_n has no constant term (by definition), it follows that

$$f(g_{n+1}(T)) \equiv T + b_n T^{n+1} + a_1 \lambda T^{n+1} \pmod{T^{n+2}}$$

But $a_1 = a$, and $\lambda = \frac{-b_n}{a}$, so $f(g_{n+1}(T)) \equiv T \pmod{T^{n+2}}$. This shows that this sequence of polynomials indeed satisfies the conditions.

Now define $g(T) = \lim_{n \rightarrow \infty} g_n(T)$. This limit exists in $R[[T]]$ and satisfies $f(g(T)) = T$. This shows that g exists.

Next, we show that this g also satisfies $g(f(T)) = T$. Apply g to $f(g(T)) = T$ to obtain $g(f(g(T))) = g(T)$. Considering this as an identity in the ring $R[[g(T)]]$, we get $g(f(T)) = T$.

Finally, g is unique. For suppose $h \in R[[T]]$ also satisfies $f(h(T)) = T$. Then

$$g(T) = g(f(h(T))) = (g \circ f)(h(T)) = h(T)$$

so $g = h$.

(ii) From $f(g(T)) = T$ it follows that $G(X, Y) = G(f(g(X)), f(g(Y)))$. f is a homomorphism $(\mathcal{F}, F) \rightarrow (\mathcal{G}, g)$, so $f(F(X, Y)) = G(f(X), f(Y))$. Applying this with $g(X)$ and $g(Y)$ instead of X and Y , we get $G(X, Y) = G(f(g(X)), f(g(Y))) = f(F(g(X), g(Y)))$. Since $g(f(T)) = T$, it follows that $g(G(X, Y)) = g(f(F(g(X), g(Y)))) = F(g(X), g(Y))$. Hence g is a homomorphism $(\mathcal{G}, G) \rightarrow (\mathcal{F}, F)$. \square

Next, we define the logarithm and exponential of a formal group. For this, we need the invariant differential.

Definition 8.8. A *differential form* is an expression of the form $P(T)dT$ with $P(T) \in R[[T]]$. A differential $\omega(T) = P(T)dT$ is *invariant* if it satisfies $(\omega \circ F)(T, S) = \omega(T)$. An invariant differential is *normalized* if $P(0) = 1$.

Remark 8.9. The condition $(\omega \circ F)(T, S) = \omega(T)$ can be rewritten as follows: $\omega(T) = P(T)dT$, so

$$(\omega \circ F)(T, S) = P(F(T, S)) \frac{\partial F(T, S)}{\partial T} dT$$

Hence the condition is equivalent to

$$P(F(T, S))F_X(T, S) = P(T)$$

where $F_X(X, Y)$ denotes the partial derivative of F with respect to the first variable.

Lemma 8.10. *For every formal group \mathcal{F} , there exists a unique normalized invariant differential, given by*

$$\omega(T) = F_X(0, T)^{-1} dT$$

All invariant differentials are of the form $a\omega$ for some $a \in R$.

Proof. First we show that ω is indeed a normalized invariant differential.

By the associative law, we have

$$F(U, F(T, S)) = F(F(U, T), S)$$

Differentiating this with respect to U gives

$$F_X(U, F(T, S)) = F_X(F(U, T), S)F_X(U, T)$$

Now put $U = 0$. Since $F(0, T) = T$, this gives

$$F_X(0, F(T, S)) = F_X(F(0, T), S)F_X(0, T) = F_X(T, S)F_X(0, T)$$

so, writing $\omega = Q(T)dT$ with $Q(T) = F_X(0, T)^{-1}$,

$$Q(F(T, S))F_X(T, S) = F_X(0, F(T, S))^{-1}F_X(T, S) = F_X(0, T)^{-1} = Q(T)$$

By remark 8.9, this implies that ω is indeed an invariant differential.

Furthermore, $F(X, Y) = X + Y + XYH(X, Y)$ for some $H \in R[[X, Y]]$ by remark 8.3, so $F_X(X, Y) = 1 + YH(X, Y) + XYH_X(X, Y)$. Hence $F_X(0, 0) = 1$, so ω is normalized.

Now let $P(T)dT$ be an invariant differential. Then

$$P(F(T, S))F_X(T, S) = P(T)$$

so by putting $S = 0$ and using $F(0, S) = S$, we get

$$P(S)F_X(0, S) = P(F(0, S))F_X(0, S) = P(0)$$

$F(T, S) = T + S + (\text{higher order terms})$, so $F_X(T, S) = 1 + (\text{h.o.t.})$, and $F_X(0, S) = 1 + (\text{h.o.t.})$. Hence $F(0, S)$ is invertible, and

$$P(S) = F_X(0, S)^{-1}P(0) = P(0)Q(S)$$

so $P(T)dT = P(0)\omega$. □

Now we can define the formal logarithm and formal exponential, and prove that the logarithm behaves as expected.

Definition 8.11. Let \mathcal{F} be a formal group over R with normalized invariant differential be $\omega(T) = (\sum_{n=0}^{\infty} c_n T^n) dT$ (with $c_0 = 1$). The *formal logarithm* is defined by

$$\log_{\mathcal{F}}(T) = \int \omega(T) = \left(\sum_{n=1}^{\infty} \frac{c_{n-1}}{n} T^n \right) dT$$

By lemma 8.7, there exists a unique power series $\exp_{\mathcal{F}}$ satisfying

$$(\log_{\mathcal{F}} \circ \exp_{\mathcal{F}})(T) = (\exp_{\mathcal{F}} \circ \log_{\mathcal{F}})(T) = T$$

This power series is called the *formal exponential*. Notice that $\log_{\mathcal{F}}$ and $\exp_{\mathcal{F}}$ are power series over $R \otimes \mathbb{Q}$, but not necessarily over R .

Lemma 8.12. $\log_{\mathcal{F}} : \mathcal{F} \rightarrow \widehat{\mathbb{G}}_a$ is an isomorphism of formal groups over $R \otimes \mathbb{Q}$.

Proof. Let ω be the normalized invariant differential on \mathcal{F} . Then by definition $\omega(F(T, S)) = \omega(T)$. Integrate this with respect to T , and put $T = 0$:

$$\log_{\mathcal{F}} F(T, S) = \log_{\mathcal{F}}(T) + f(S)$$

for some $f(S) \in R[[S]]$, so

$$\log_{\mathcal{F}}(S) = \log_{\mathcal{F}} F(0, S) = \log_{\mathcal{F}}(0) + f(S) = f(S)$$

This shows that

$$\log_{\mathcal{F}} F(T, S) = \log_{\mathcal{F}}(T) + \log_{\mathcal{F}}(S)$$

so $\log_{\mathcal{F}}$ is a homomorphism. Its inverse is $\exp_{\mathcal{F}}$, so $\log_{\mathcal{F}}$ is an isomorphism. \square

8.2 Multiplication in formal groups

In a formal group, we can multiply by integers:

Definition 8.13. Let (\mathcal{F}, F) be a formal group, and let $n \in \mathbb{Z}$. Then the *multiplication-by- n map* $[n] : \mathcal{F} \rightarrow \mathcal{F}$ is defined inductively by

$$[n](T) = \begin{cases} 0 & \text{if } n = 0 \\ F([n-1](T), T) & \text{if } n > 0 \\ F([n+1](T), i(T)) & \text{if } n < 0 \end{cases}$$

Lemma 8.14. $[n]$ is a homomorphism of formal groups.

Proof. $[n]$ is a homomorphism if $[n](F(X, Y)) = F([n]X, [n]Y)$. We prove this by induction.

For $n = 0$, we have $[0](F(X, Y)) = 0$ and $F([0]X, [0]Y) = F(0, 0) = 0$.

Let $n > 0$ and suppose that $[n-1]$ is a homomorphism. Then $[n-1](F(X, Y)) = F([n-1]X, [n-1]Y)$ so

$$\begin{aligned} [n](F(X, Y)) &= F([n-1](F(X, Y)), F(X, Y)) = F(F([n-1]X, [n-1]Y), F(X, Y)) = \\ &F(F(F([n-1]X, [n-1]Y), X), Y) = F(F(F([n-1]Y, [n-1]X), X), Y) = \\ &F(F([n-1]Y, F([n-1]X, X)), Y) = F(F([n-1]Y, [n](X)), Y) = \\ &F(F([n]X, [n-1](Y)), Y) = F([n]X, F([n-1](Y), Y)) = F([n]X, [n]Y) \end{aligned}$$

so $[n]$ is a homomorphism.

Now let $n < 0$ and suppose that $[n+1]$ is a homomorphism. Then $[n+1](F(X, Y)) = F([n+1]X, [n+1]Y)$ so

$$[n](F(X, Y)) = F([n+1](F(X, Y)), i(F(X, Y))) = F(F([n+1](X), [n+1]Y), i(F(X, Y)))$$

By lemma 8.6, i is a homomorphism, so $i(F(X, Y)) = F(i(X), i(Y))$. Hence

$$\begin{aligned} [n](F(X, Y)) &= F(F([n+1](X), [n+1]Y), F(i(X), i(Y))) = \\ &= F([n+1](X), F([n+1]Y, F(i(X), i(Y)))) = F([n+1](X), F(F([n+1]Y, i(X)), i(Y))) = \\ &= F([n+1](X), F(F(i(X), [n+1]Y), i(Y))) = F([n+1](X), F(i(X), F([n+1]Y, i(Y)))) = \\ &= F(F([n+1](X), i(X)), F([n+1](Y), i(Y))) = F([n](X), [n](Y)) \end{aligned}$$

so $[n]$ is a homomorphism.

By induction, $[n]$ is a homomorphism for all $n \in \mathbb{Z}$. \square

Lemma 8.15. *Let (\mathcal{F}, F) be a formal group over a ring R , and let $n \in \mathbb{Z}$. Then*

- (i) $[n]T = nT + (\text{higher order terms})$
- (ii) If $n \in R^\times$, then $[n] : \mathcal{F} \rightarrow \mathcal{F}$ is an isomorphism.

Proof. (i) Again we prove this by induction.

For $n = 0$, the statement clearly holds.

Suppose that it holds for some $n > 0$. By definition, $F(X, Y) = X + Y + (\text{higher order terms})$. Hence

$$[n+1](T) = F([n](T), T) = [n](T) + T + (\text{higher order terms}) = nT + T + (\text{h.o.t.}) = (n+1)T + (\text{h.o.t.})$$

so the statement also holds for $n + 1$.

For $n < 0$, notice that by remark 8.3, $F(X, Y) = X + Y + XYG'(X, Y)$ for some $G' \in R[[X, Y]]$. Furthermore, i is a power series without a constant term by lemma 8.6, so $i(T) = TH(t)$ for some $H \in R[[X, Y]]$. This gives:

$$0 = F(T, i(T)) = T + i(T) + Ti(T)G'(T, i(T)) = T + i(T) + T^2H(T)G'(T, TH(T))$$

so $i(T) = -T + (\text{h.o.t.})$. Hence, if $[n]T = nT + (\text{h.o.t.})$, then

$$[n-1](T) = F([n](T), i(T)) = [n](T) + i(T) + (\text{h.o.t.}) = nT + (\text{h.o.t.}) - T + (\text{h.o.t.}) = (n-1)T + (\text{h.o.t.})$$

so the statement also holds for $n - 1$.

By induction, it holds for all $n \in \mathbb{Z}$.

(ii) Apply lemma 8.7 to $[n]$. This gives a homomorphism g satisfying $[n](g(T)) = g([n](T)) = T$. Hence $[n]$ is an isomorphism. \square

8.3 Formal groups over valuation rings

Up to now, we have considered formal groups as consisting of a group law, without an actual group. But in some cases, we can associate a group to a formal group. To do this, we need a ring in which the power series converges. For this, we use the ring \mathbb{Q}_p of p -adic numbers, which consists of Laurent series in p , for some prime p :

$$\mathbb{Q}_p = \left\{ \sum_{i=k}^{\infty} a_i p^i \mid k \in \mathbb{Z}, a_i \in \{0, \dots, p-1\} \right\}$$

This ring is equipped with the p -adic valuation.

Definition 8.16. A (*non-archimedean*) valuation on a ring R is a map $v : R \rightarrow \mathbb{Z} \cup \{\infty\}$, such that the following properties hold:

- (i) $v(x) = \infty$ if and only if $x = 0$.
- (ii) $v(xy) = v(x) + v(y)$ for all $x, y \in R$.
- (iii) $v(x + y) \geq \min\{v(x), v(y)\}$ for all $x, y \in R$.

Definition 8.17. For every prime p , we define the p -adic valuation on \mathbb{Q}_p by v_p by $v_p(0) = \infty$ and $v_p(p^a/c) = a$ if p does not divide b and c .

Remark 8.18. (i) Notice that v_p is indeed a valuation.

(ii) Let $v(x) < v(y)$, and suppose $v(x+y) > \min\{v(x), v(y)\} = v(x)$. Then

$$v(x) = v(x+y-y) \geq \min\{v(x+y), v(-y)\} = \min\{v(x+y), v(y)\}$$

Since we assumed $v(x) < v(y)$, this minimum can not be $v(y)$, so

$$v(x) \geq v(x+y) > \min\{v(x), v(y)\} = v(x)$$

Contradiction, so if $v(x) \neq v(y)$, then $v(x+y) = \min\{v(x), v(y)\}$.

There is an absolute value $\|\cdot\|$ associated to v_p , given by $\|x\| = p^{-v_p(x)}$. Note that this is indeed an absolute value. Since this absolute value is non-archimedean, a sequence converges if and only if the valuation of the terms tends to infinity.

With respect to the p -adic valuation, the Laurent series $\sum_{i=n}^{\infty} a_i p^i$ (with $a_i \in \{0, \dots, p-1\}$ and $a_n \neq 0$) has valuation $v_p(p^n) = n$. The valuation ring of the ring of p -adic numbers is the ring of p -adic integers \mathbb{Z}_p , consisting of all elements with non-negative valuation:

$$\mathbb{Z}_p = \left\{ \sum_{i=0}^{\infty} a_i p^i \mid a_i \in \{0, \dots, p-1\} \right\}$$

\mathbb{Z}_p is complete with respect to the absolute value associated to v_p , i.e. all Cauchy sequences in \mathbb{Z}_p converge to a limit in \mathbb{Z}_p .

\mathbb{Z}_p has a unique maximal ideal

$$\mathcal{M}_p = \left\{ \sum_{i=1}^{\infty} a_i p^i \mid a_i \in \{0, \dots, p-1\} \right\}$$

consisting of all p -adic numbers with strictly positive valuation.

Definition 8.19. Let (\mathcal{F}, F) be a formal group defined over \mathbb{Q}_p , and let \mathcal{M}_p be the maximal ideal of \mathbb{Z}_p . The *group associated to \mathcal{F}* , denoted $\mathcal{F}(\mathcal{M}_p)$, is the set \mathcal{M}_p , with the operations $x+y = F(x, y)$ and $-x = i(x)$. For $n \geq 1$, $\mathcal{F}(\mathcal{M}_p^n)$ is the subgroup of $\mathcal{F}(\mathcal{M}_p)$ consisting of the set \mathcal{M}_p^n with the induced operations.

Remark 8.20. The power series $F(x, y)$ and $i(x)$ converges, since $v_p(x^k y^l) = kv_p(x) + lv_p(y)$, which tends to infinity when $k+l \rightarrow \infty$, since $v_p(x), v_p(y) \geq 1$. Hence $\mathcal{F}(\mathcal{M}_p)$ is indeed a group, with subgroups $\mathcal{F}(\mathcal{M}_p^n)$ for all n .

To prove that the formal group is often isomorphic to the additive group, we need some lemmas.

Lemma 8.21. For all integers n and all primes p

$$v_p(n!) \leq \frac{n-1}{p-1}$$

Proof. The proof is an easy computation:

$$\begin{aligned} v_p(n!) &= \sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor = \sum_{i=1}^{\lfloor \log_p(n) \rfloor} \left\lfloor \frac{n}{p^i} \right\rfloor \leq \sum_{i=1}^{\lfloor \log_p(n) \rfloor} \frac{n}{p^i} = \\ &= \frac{n}{p} \sum_{i=0}^{\lfloor \log_p(n) \rfloor - 1} \frac{1}{p^i} = \frac{n}{p} \cdot \frac{1-p^{-\lfloor \log_p(n) \rfloor}}{1-p^{-1}} = \frac{n-np^{-\lfloor \log_p(n) \rfloor}}{p-1} \end{aligned}$$

Since $n = p^{\log_p(n)}$, we have $np^{-\lfloor \log_p(n) \rfloor} = p^{\log_p(n) - \lfloor \log_p(n) \rfloor} \geq 1$ so

$$v_p(n!) = \frac{n - np^{-\lfloor \log_p(n) \rfloor}}{p-1} \leq \frac{n-1}{p-1}$$

□

Lemma 8.22. *Let $f(T) = \sum_{n=1}^{\infty} \frac{a_n}{n!} T^n$ be a power series with $a_n \in \mathbb{Z}_p$ and $v_p(a_1) = 0$. If $x \in \mathbb{Z}_p$ satisfies $v_p(x) > \frac{1}{p-1}$, then the series $f(x)$ converges in \mathbb{Z}_p , and $v_p(f(x)) = v_p(x)$.*

Proof. For the terms in the series of $f(x)$, we have

$$v_p\left(\frac{a_n x^n}{n!}\right) = v_p(a_n) + nv_p(x) - v_p(n!) \geq nv_p(x) - v_p(n!)$$

By lemma 8.21, $v_p(n!) \leq \frac{n-1}{p-1}$, so

$$v_p\left(\frac{a_n x^n}{n!}\right) \geq nv_p(x) - \frac{n-1}{p-1} = v_p(x) + (n-1)\left(v_p(x) - \frac{1}{p-1}\right)$$

Since $v_p(x) > \frac{1}{p-1}$, this tends to infinity (for $n \rightarrow \infty$), so the series converges in \mathbb{Z}_p . Furthermore, for $n \geq 2$,

$$v_p\left(\frac{a_n x^n}{n!}\right) \geq v_p(x) + (n-1)\left(v_p(x) - \frac{1}{p-1}\right) > v_p(x)$$

so $v_p(f(x)) = v_p(a_1 x) = v_p(x)$ since $v_p(a_1) = 0$.

□

Lemma 8.23. *Let \mathcal{F} be a formal group over \mathbb{Z}_p . Then there exists $a_n, b_n \in \mathbb{Z}_p$ with $a_1 = b_1 = 1$ such that*

$$\log_{\mathcal{F}}(T) = \sum_{n=1}^{\infty} \frac{a_n}{n} T^n \quad \text{and} \quad \exp_{\mathcal{F}}(T) = \sum_{n=1}^{\infty} \frac{b_n}{n!} T^n$$

Proof. For $\log_{\mathcal{F}}$, this follows immediately from definition 8.11. Here $a_n = c_{n-1}$ and $a_1 = c_0 = 1$.

We can write $\exp_{\mathcal{F}}(T) = \sum_{n=1}^{\infty} \frac{b_n}{n!} T^n$, but then we still have to show that $b_n \in \mathbb{Z}_p$ (instead of $\mathbb{Z}_p \otimes \mathbb{Q}$) and $b_1 = 1$. Differentiate $\log_{\mathcal{F}}(\exp_{\mathcal{F}}(T)) = T$ to obtain

$$\omega(\exp_{\mathcal{F}}(T)) \exp'_{\mathcal{F}}(T) = 1$$

Evaluating this at $T = 0$ gives $\omega(0) \cdot b_1 = 1$, so $b_1 = 1$ since ω is normalized.

Differentiating again will give a polynomial expression for $\omega(\exp_{\mathcal{F}}(T)) \exp_{\mathcal{F}}^{(n)}(T)$ in terms of $\omega^{(i)}(\exp_{\mathcal{F}}(T))$ and $\exp_{\mathcal{F}}^{(j)}(T)$ with $0 \leq i \leq n-1$ and $1 \leq j \leq n-1$. Evaluating this at $T = 0$ gives a polynomial expression for $\omega(0)b_n$ in terms of $a_2, \dots, a_n, b_1, \dots, b_{n-1}$. Since $\omega(0) = 1$, this shows that b_n is a polynomial in $a_2, \dots, a_n, b_1, \dots, b_{n-1}$, so by induction, it is an element of \mathbb{Z}_p . □

Theorem 8.24. *Let \mathcal{F} be a formal group over \mathbb{Z}_p and let $r > \frac{1}{p-1}$ be an integer. Then the formal logarithm induces an isomorphism $\log_{\mathcal{F}} : \mathcal{F}(\mathcal{M}_p^r) \rightarrow \widehat{\mathbb{G}}_a(\mathcal{M}_p^r)$, and this isomorphism preserves valuations.*

Proof. By lemma 8.23, $\log_{\mathcal{F}}$ and $\exp_{\mathcal{F}}$ are both given by power series of the form $\sum_{n=1}^{\infty} \frac{a_n}{n!} T^n$ with $a_n \in \mathbb{Z}_p$ and $a_1 = 1$. Since $x \in \mathcal{M}_p^r$ is equivalent to $v_p(x) \geq r$, and $r > \frac{1}{p-1}$, we can apply lemma 8.22 to get that $\log_{\mathcal{F}}(x)$ and $\exp_{\mathcal{F}}(x)$ converge to a limit in \mathbb{Z}_p . Hence $\log_{\mathcal{F}}$ and $\exp_{\mathcal{F}}$ are well-defined on \mathcal{M}_p^r .

By lemma 8.12, $\log_{\mathcal{F}}$ is an isomorphism $\mathcal{F} \rightarrow \widehat{\mathbb{G}}_a$, so it induces an isomorphism $\mathcal{F}(\mathcal{M}_p^r) \rightarrow \widehat{\mathbb{G}}_a(\mathcal{M}_p^r)$.

Now suppose that $v_p(x) = s_1$. Then $x \in \mathcal{M}_p^{s_1}$, so $\log_{\mathcal{F}}(x) \in \mathcal{M}_p^{s_1}$ and hence $v_p(\log_{\mathcal{F}}(x)) \geq s_1$. If $v_p(\log_{\mathcal{F}}(x)) = s_2 > s_1$, then $\log_{\mathcal{F}}(x)$ lies in the image of $\log_{\mathcal{F}} : \mathcal{F}(\mathcal{M}_p^{s_2}) \rightarrow \widehat{\mathbb{G}}_a(\mathcal{M}_p^{s_2})$. Since the isomorphism is the same for all s , this implies that, under the isomorphism $\log_{\mathcal{F}} : \mathcal{F}(\mathcal{M}_p^r) \rightarrow \widehat{\mathbb{G}}_a(\mathcal{M}_p^r)$, the pre-image of $\log_{\mathcal{F}}(x)$ lies in $\mathcal{M}_p^{s_2}$, so $x \in \mathcal{M}_p^{s_2}$. Contradiction, so $v_p(\log_{\mathcal{F}}(x)) = s_1$, and hence the isomorphism preserves valuations. □

8.4 The formal group of an elliptic curve

Finally we come to the application to elliptic curves. Let E be an elliptic curve, given by a Weierstrass equation $y^2 = x^3 + ax^2 + bx + c$. We can make a change of variables $z = -\frac{x}{y}$ and $w = -\frac{1}{y}$. Since $x = \frac{z}{w}$ and $y = -\frac{1}{w}$, the Weierstrass equation becomes

$$\frac{1}{w^2} = \frac{z^3}{w^3} + a\frac{z^2}{w^2} + b\frac{z}{w} + c$$

so

$$w = z^3 + az^2w + b zw^2 + cw^3 =: f(z, w)$$

Now we can substitute this equation into itself:

$$w = z^3 + az^2f(z, w) + bzf(z, w)^2 + cf(z, w)^3$$

We can repeat this, to get a power series $w \in \mathbb{Z}[a, b, c][[z]]$ that satisfies $w(z) = f(z, w(z))$. By the following lemma, this procedure indeed gives a power series:

Lemma 8.25. *Define $f_1(z, w) = f(z, w)$ and recursively $f_{m+1}(z, w) = f_m(z, f(z, w))$. Then the limit $w(z) = \lim_{m \rightarrow \infty} f_m(z, 0)$ exists in $\mathbb{Z}[a, b, c][[z]]$ and it is the unique power series satisfying $w(z) = f(z, w(z))$. Furthermore, the first term of w is $A_0 z^3$ for some $A_0 \in \mathbb{Z}[a, b, c]$.*

The proof uses *Hensel's lemma*:

Lemma 8.26. (Hensel) *Let R be an integral domain, and let $I \neq R$ be an ideal, such that R is complete with respect to I (i.e. every Cauchy sequence converges to a limit $u \in R$). A Cauchy sequence is a sequence $\{u_n \in R\}_{n \in \mathbb{N}}$ such that for all $n \in \mathbb{N}$, there exists $N \in \mathbb{N}$ such that $u_m - u_{m'} \in I^n$ for all $m, m' \geq N$). Furthermore, let $F(t) \in R[t]$ be a polynomial such that $F(0) \in I^n$ for some $n \geq 1$ and $F'(0) \in R^\times$. Suppose that $F'(0) \equiv -1 \pmod{I}$. Then the sequence*

$$w_0 = 0 \quad \text{and} \quad w_{m+1} = w_m + F(w_m)$$

converges to $w \in R$ with

$$F(w) = 0 \quad \text{and} \quad w \in I^n$$

and w is the unique element of R satisfying this conditions.

Proof. Since $F(0) \in I^n$, the constant term of F must be an element of I^n . Hence if $w_m \in I^n$, then $F(w_m) \in I^n$, and also $w_{m+1} = w_m + F(w_m) \in I^n$. Since $w_0 = 0 \in I^n$, it follows that $w_m \in I^n$ for all m .

Next, we prove that $w_{m+1} - w_m \in I^{m+n}$ for all $m \in \mathbb{N}$. For $m = 0$, this says that $F(0) = w_1 \equiv w_0 = 0 \pmod{I^n}$, which is true by assumption. Now assume that the claim holds for all natural numbers less than m . Write $F(X) = \sum_{k=0}^n a_k X^k$. Then

$$\begin{aligned} F(X) - F(Y) &= \sum_{k=1}^n a_k (X^k - Y^k) = (X - Y) \sum_{k=1}^n a_k (X^{k-1} + X^{k-2}Y + \dots + XY^{k-2} + Y^{k-1}) = \\ &= (X - Y)(F'(0) + X \sum_{k=2}^n a_k (X^{k-2} + X^{k-3}Y + \dots + Y^{k-2}) + Y \sum_{k=2}^n a_k Y^{k-2} \end{aligned}$$

Define $G(X, Y) = \sum_{k=2}^n a_k (X^{k-2} + X^{k-3}Y + \dots + Y^{k-2})$ and $H(X, Y) = \sum_{k=2}^n a_k Y^{k-2}$; then $F(X) - F(Y) = (X - Y)(F'(0) + XG(X, Y) + YH(X, Y))$.

Applying this with $X = w_m$ and $Y = w_{m-1}$ gives

$$\begin{aligned} w_{m+1} - w_m &= (w_m + F(w_m)) - (w_{m-1} + F(w_{m-1})) = \\ &= (w_m - w_{m-1}) + (F(w_m) - F(w_{m-1})) = \\ &= (w_m - w_{m-1})(1 + F'(0) + w_m G(w_m, w_{m-1}) + w_{m-1} H(w_m, w_{m-1})) \end{aligned}$$

By the induction hypothesis, $w_m - w_{m-1} \in I^{m+n-1}$. Since $1 + F'(0) \in I$ and $w_m, w_{m-1} \in I^n$, we also have $1 + F'(0) + w_m G(w_m, w_{m-1}) + w_{m-1} H(w_m, w_{m-1}) \in I$, so $w_{m+1} - w_m \in I^{m+n}$. This proves the claim.

It follows that $(w_m)_{m \in \mathbb{N}}$ is a Cauchy sequence, and since R is complete, it converges to a limit w . Since all w_m are elements of I^n , w is also an element of I^n . Furthermore, by taking limits in the relation $w_{m+1} = w_m + F(w_m)$, we get $w = w + F(w)$, so $F(w) = 0$.

It only remains to prove that w is unique. Suppose that v also satisfies $F(v) = 0$ and $v \in I^n$. Then

$$0 = F(w) - F(v) = (w - v)(F'(0) + wG(w, v) + vH(w, v))$$

so either $w = v$ or $F'(0) + wG(w, v) + vH(w, v) = 0$. But $w, v \in I$ so the latter would imply that $F'(0) \in I$, so $I = R$ since $F'(0) \in R^\times$. Contradiction, so $w = v$ and w is unique. \square

Proof of lemma 8.25. First we claim that $f_m(z, f(w)) = f(z, f_m(w))$ for all $m \geq 1$. For $m = 1$, this is clear. Suppose that it holds for some m . Then

$$f_{m+1}(z, f(w)) = f_m(z, f(z, f(w))) = f(z, f_m(z, f(w))) = f(z, f_{m+1}(z, w))$$

Now the claim follows by induction.

It follows that $f_{m+1}(z, w) = f_m(z, f(z, w)) = f(z, f_m(z, w))$. Hence the recursion is equivalent to $f_0(z, w) = w$ and $f_{m+1}(z, w) = f(z, f_m(z, w))$. Now define $w_m(z) = f_m(z, 0)$. Then w_m satisfies $w_0(z) = 0$ and $w_m(z) = f(z, f_m(z, 0))$. Furthermore, let $F(t) = f(z, w) - w$. Then

$$w_{m+1}(z) = f_{m+1}(z, 0) = f(z, f_m(z, 0)) = f(z, w_m(z)) = w_m(z) + F(w_m(z))$$

Apply Hensel's lemma to $R = \mathbb{Z}[a, b, c][[z]]$, $I = (z)$ and $F(w) = f(z, w) - w$. Then the conditions are satisfied since $F(w) = z^3 + az^2w + bw^2 + cw^3 - w$, so $F(0) = z^3 \in I^3$ and $F'(0) = az^2 - 1 \in R^\times$. It follows that there exists a unique $w(z) \in \mathbb{Z}[a, b, c][[z]]$ such that $F(w(z)) = 0$, i.e. $w(z) = f(z, w(z))$, and $w(z) \in (z)^3$. \square

w is given by a power series in $\mathbb{Z}[a, b, c][[z]]$, which starts with Az^3 for some A . We write

$$w(z) = z^3(A_0 + A_1z + A_2z^2 + \dots) = \sum_{n=3}^{\infty} A_{n-3}z^n$$

Using the power series $w(z)$ and the change of coordinates $x = \frac{z}{w}$ and $y = -\frac{1}{w}$, for every z we get a formal solution $(x(z), y(z))$ of the Weierstrass equation.

Now we want to define an addition on points (z, w) that corresponds to the addition of points in the (x, y) -plane. Take $z_1 \neq z_2$, and let $w_i = w(z_i)$, $x_i = \frac{z_i}{w_i}$ and $y_i = -\frac{1}{w_i}$. To compute the sum $P_3 = (x_3, y_3)$ of $P_1 = (x_1, y_1)$ and $P_2 = (x_2, y_2)$, we first compute the line through P_1 and P_2 . This line has an equation of the form $y = \lambda x + \nu$. Then the third intersection point of the line with the curve is the inverse of (x_3, y_3) .

When we change the coordinates to z and w , the line $y = \lambda x + \nu$ changes to $-\frac{1}{w} = \lambda \frac{z}{w} + \nu$, so $\lambda z + \nu w = -1$. Hence the line in the (x, y) -plane through P_1 and P_2 corresponds to the line in the (z, w) -plane through (z_1, w_1) and (z_2, w_2) . Since the inverse of (x, y) is $(x, -y)$, the z -coordinate of the inverse of a point is $\frac{x}{-y} = -z$. Hence $-(z, w(z))$ is $(-z, w(-z))$. Now we can give a formula for the z -coordinate of the sum of two points:

Let the line through (z_1, w_1) and (z_2, w_2) be $w = \alpha z + \beta$. Then

$$\alpha = \alpha(z_1, z_2) = \frac{w_1 - w_2}{z_1 - z_2} = \sum_{n=3}^{\infty} A_{n-3} \frac{z_1^n - z_2^n}{z_1 - z_2} = \sum_{n=3}^{\infty} A_{n-3} \sum_{k=0}^{n-1} z_1^k z_2^{n-1-k} \in \mathbb{Z}[a, b, c][[z]]$$

and

$$\begin{aligned}\beta = \beta(z_1, z_2) &= w_1 - \alpha z_1 = \left(\sum_{n=3}^{\infty} A_{n-3} z_1^n \right) - \left(\sum_{n=3}^{\infty} A_{n-3} \sum_{k=0}^{n-1} z_1^{k+1} z_2^{n-1-k} \right) = \\ &= \sum_{n=3}^{\infty} A_{n-3} \left(z_1^n - \sum_{k=0}^{n-1} z_1^{k+1} z_2^{n-1-k} \right) = - \sum_{n=3}^{\infty} A_{n-3} \left(\sum_{k=0}^{n-2} z_1^{k+1} z_2^{n-1-k} \right) \in \mathbb{Z}[a, b, c][[z]]\end{aligned}$$

The intersection of this line with the curve consists of three points, namely (z_1, w_1) , (z_2, w_2) and $(-z_3, -w_3)$. The Weierstrass equation is $w = z^3 + az^2w + bz w^2 + cw^3$. Substitute $w = \alpha z + \beta$ to get

$$\alpha z + \beta = z^3 + az^2(\alpha z + \beta) + bz(\alpha z + \beta)^2 + c(\alpha z + \beta)^3$$

so

$$z^3 + az^2(\alpha z + \beta) + bz(\alpha z + \beta)^2 + c(\alpha z + \beta)^3 - \alpha z - \beta = (z - z_1)(z - z_2)(z + z_3)$$

If we compare the quadratic terms, then we get

$$a\beta + 2b\alpha\beta + 3c\alpha^2\beta = -(z_1 + z_2 - z_3)$$

so

$$z_3 = z_1 + z_2 + a\beta + 2b\alpha\beta + 3c\alpha^2\beta \in \mathbb{Z}[a, b, c][[z]]$$

We define a power series F by this formula:

$$F(z_1, z_2) = z_3 = z_1 + z_2 + a\beta + 2b\alpha\beta + 3c\alpha^2\beta \in \mathbb{Z}[a, b, c][[z]]$$

Lemma 8.27. F is a formal group law, and hence gives a formal group (\widehat{E}, F) associated to the elliptic curve E .

Proof. We prove the five properties of definition 8.1. First we prove (ii)-(v), and then (i).

(ii) The addition law on elliptic curves is associative. Hence F is also associative.

(iii) This follows from the fact that the addition law on elliptic curves is commutative.

(iv) First we show that i is unique. If $F(z, i(z)) = 0$ for all z , then the sum of $(z, w(z))$ and $(i(z), w(i(z)))$ is $(0, w(0))$. In the definition of w , $f(0, 0) = 0$ so $f_m(0, 0) = 0$ for all m . This implies that $w(0) = 0$. It follows that the sum of $(z, w(z))$ and $(i(z), w(i(z)))$ is $(0, 0)$, which corresponds to the point \mathcal{O} in (x, y) -coordinates. Hence $(x(z), y(z))$ and $(x(i(z)), y(i(z)))$ add up to \mathcal{O} , so $(x(i(z)), y(i(z))) = -(x(z), y(z))$, and hence $(i(z), w(i(z))) = -(z, w(z)) = (-z, w(-z))$. It follows that $i(z) = -z$. Now notice that $F(z, -z)$ is by definition the point that corresponds to \mathcal{O} , so $F(z, -z) = 0$ for all z .

(v) $F(z, 0)$ is the z -coordinate of the sum of the points $(x(z), y(z))$ and $(x(0), y(0)) = \mathcal{O}$. This sum is $(x(z), y(z))$, so $F(z, 0) = z$ for all z . Hence $F(X, 0) = X$. Similarly, $F(0, Y) = Y$.

(i) Write $F(X, Y) = \sum_{i,j=0}^{\infty} a_{ij} X^i Y^j$. By substituting $Y = 0$, we get $X = F(X, 0) = \sum_{i=0}^{\infty} a_{i0} X^i$. Hence $a_{10} = 1$ and $a_{i0} = 0$ for all other i . In the same way, $a_{01} = 1$ and $a_{0j} = 0$ for all other j . This gives:

$$F(X, Y) = \sum_{i,j=0}^{\infty} a_{ij} X^i Y^j = X + Y + \sum_{i,j=1}^{\infty} a_{ij} X^i Y^j$$

□

Let (x_i, y_i) ($i \in \{1, 2, 3\}$) be points on an elliptic curve, and let the corresponding points in (z, w) -coordinates be (z_i, w_i) . Then by construction of the formal group law, we have the following properties:

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) = (x_3, y_3) &\iff F(z_1, z_2) = z_3 \\(x_1, y_1) = -(x_2, y_2) &\iff z_1 = i(z_2)\end{aligned}$$

Chapter 9

Function fields over formally real fields

In this chapter, we return to Hilbert's tenth problem. In chapter 7, we proved that there is no algorithm to decide whether an equation with coefficients in $\mathbb{Z}/p\mathbb{Z}[t]$ has a solution in $K(t)$ if K is a finite field of characteristic p . Now we will prove that there also doesn't exist an algorithm to determine the solvability in K of equations over \mathbb{Z} if K is a formally real field. A *formally real field* is a field in which -1 is not the sum of squares. This is equivalent to the statement that zero is not the sum of squares distinct from zero (i.e. if $a_1^2 + \dots + a_n^2 = 0$, then $a_1 = \dots = a_n = 0$): if $a_1^2 + \dots + a_n^2 = -1$, then $1^2 + a_1^2 + \dots + a_n^2 = 0$, and conversely, if $a_1^2 + \dots + a_n^2 = 0$ with $a_1 \neq 0$, then $\frac{a_2^2}{a_1^2} + \dots + \frac{a_n^2}{a_1^2} = -1$. Notice that the characteristic of a formally real field is always zero, because if the characteristic is p , then $-1 = 1^2 + \dots + 1^2$ ($p-1$ times 1^2).

The proof we will give here is based on [Den78], and uses endomorphisms of an elliptic curve as a model for \mathbb{Z} . We will associate the curve $f(t)y^2 = f(x)$ to the curve $y^2 = f(x)$. The points on the associated curve correspond to endomorphisms of the original curve. The proof will consist of two parts: first we will compute the points of the associated elliptic curve. Then we define some predicates, similar to the predicate `Imt` in section 5.1. The final part of the proof will be similar to the proof of theorem 5.1.

Throughout this chapter, we use the language $(0, 1, t; +, \cdot)$. Notice that by lemma 1.15, a set is diophantine if and only if it is positive existential.

9.1 The rational points on the twist of the elliptic curve

Fix an elliptic curve E over \mathbb{Q} , given by the equation $s^2 = f(t) = t^3 + at^2 + bt + c$, without complex multiplication. To E , we associate the curve $E_t : f(t)y^2 = f(x)$. This curve is called the *twist*. By construction of E_t , it contains the point $P = (t, 1)$.

Notice that E is also defined over K , since for every field K of characteristic zero, there exists a unique injection $\mathbb{Q} \rightarrow K$. To make it possible that this curve E exists, we need the assumption that $\text{char}(K) = 0$, since otherwise the Frobenius map is always an element of the endomorphism ring, different from multiplication by an integer.

Lemma 9.1. *There exists a curve E without complex multiplication.*

Proof. To every curve, there is an associated number, called the *j -invariant*. For curves in Legendre form, i.e. with equation $E : y^2 = x(x-1)(x-\lambda)$, this j -invariant is $j(E) = 2^8 \frac{\lambda^2 - \lambda + 1}{\lambda^2(\lambda-1)^2}$. If an elliptic curve has complex multiplication, then its j -invariant is an algebraic integer (see [Deu41]). There are many curves for which j is not an algebraic integer. For example, take $\lambda = 3$ to obtain $j(E) = \frac{448}{9}$, which is not an algebraic integer. \square

Theorem 9.2. *$P = (t, 1)$ has infinite order, and generates $E_t(K(t))$ modulo the points of order two.*

In the proof of this theorem, we need some facts about endomorphisms of elliptic curves:

Lemma 9.3. *Let C be an elliptic curve, and let $\psi : C \rightarrow C$ be a rational map satisfying $\psi(\mathcal{O}) = \mathcal{O}$. Then ψ is an endomorphism of C .*

Proof. Since C is a smooth curve, ψ is a morphism of curves by [Sil86], proposition II.2.1. Hence ψ is an isogeny, so it is a homomorphism of groups by [Sil86], theorem III.4.8. This implies that ψ is an endomorphism of C .

Notice that Silverman assumes that K is a perfect field, i.e. a field of which every algebraic extension is separable. We don't need this assumption: if K is not a perfect field, then it has an extension L that is perfect. Now ψ is an endomorphism over L , and since it is a K -rational map, it is also an endomorphism over K . \square

Lemma 9.4. *Every nonconstant endomorphism e over K of E is of the form $e(t, s) = (x(t), sy(t))$ with $x, y \in \overline{K}(t)$ and $(x, y) \in E_t(\overline{K}(t, s))$.*

Proof. Endomorphisms are rational maps, so we can write $e(t, s) = (x(t, s), z(t, s))$ where x and z are rational maps over \overline{K} . Furthermore, write $x(t, s) = \frac{x_1(t, s)}{x_2(t, s)}$ and $z(t, s) = \frac{y_1(t, s)}{y_2(t, s)}$ with x_1, x_2, y_1 and y_2 polynomials in t and s . Since $s^2 = t^3 + at^2 + bt + c$ for all points (t, s) on E , we can assume that the degree of these polynomials in s is at most 1, so we can write $x_1(t, s) = x_{11}(t) + sx_{12}(t)$ etcetera. Since

$$x(t, s) = \frac{x_{11}(t) + sx_{12}(t)}{x_{21}(t) + sx_{22}(t)} = \frac{x_{11}(t) + sx_{12}(t)}{x_{21}(t) + sx_{22}(t)} \cdot \frac{x_{21}(t) - sx_{22}(t)}{x_{21}(t) - sx_{22}(t)} = \frac{(x_{11}(t)x_{21}(t) - (t^3 + at^2 + bt + c)x_{12}(t)x_{22}(t)) + s(x_{12}(t)x_{21}(t) - x_{11}(t)x_{22}(t))}{x_{21}(t)^2 - (t^3 + at^2 + bt + c)x_{22}(t)^2}$$

we can assume that $x_{22}(t) = 0$. Similarly, we can assume that $y_{22}(t) = 0$.

Since e is an endomorphism, for any point P we have $e(-P) = -e(P)$, i.e.

$$(x(t, -s), z(t, -s)) = (x(t, s), -z(t, s))$$

This implies

$$\frac{x_{11}(t) - sx_{12}(t)}{x_{21}(t)} = x(t, -s) = x(t, s) = \frac{x_{11}(t) + sx_{12}(t)}{x_{21}(t)}$$

so $x_{12}(t) = 0$. Hence $x(t, s) = \frac{x_{11}(t)}{x_{21}(t)}$, so $x \in \overline{K}(t)$. Similarly,

$$\frac{y_{11}(t) - sy_{12}(t)}{y_{21}(t)} = z(t, -s) = -z(t, s) = -\frac{y_{11}(t) + sy_{12}(t)}{y_{21}(t)}$$

so $y_{11}(t) = 0$. Hence $z(t, s) = s\frac{y_{12}(t)}{y_{21}(t)}$, so there exists $y \in \overline{K}(t)$ such that $z(t, s) = sy(t)$.

Finally, since e maps $(t, s) \in E$ to a point on E , we must have

$$f(t)y(t)^2 = (t^3 + at^2 + bt + c)y(t)^2 = (sy(t))^2 = x(t)^3 + ax(t)^2 + bx(t) + c = f(x(t))$$

for any (t, s) such that $s^2 = t^3 + at^2 + bt + c$. Hence $(x, y) \in E_t(\overline{K}(t, s))$. \square

Lemma 9.5. *The map*

$$\Phi : E_t(K(t)) \rightarrow \text{Rat}_K(E(K), E(K)) : (z(t), w(t)) \mapsto ((t, s) \mapsto (z(t), sw(t)))$$

is an injective homomorphism.

Proof. Consider the maps

$$\Phi_1 : E_t(K(t)) \rightarrow E(K(t, s)) : (z(t), w(t)) \mapsto (z(t), sw(t))$$

and

$$\Phi_2 : E(K(t, s)) \rightarrow \text{Rat}_K(E(K), E(K)) : (z(t, s), w(t, s)) \mapsto ((t, s) \mapsto (z(t, s), w(t, s)))$$

It can easily be checked that Φ_1 indeed maps $E_t(K(t))$ to $E(K(t, s))$, and $\Phi = \Phi_2 \circ \Phi_1$. Φ_1 is a homomorphism by lemma 9.3, since it is a rational map satisfying $\mathcal{O} \mapsto \mathcal{O}$, and it is clearly injective. It is also clear that Φ_2 is an injective homomorphism. Hence Φ is also an injective homomorphism. \square

Lemma 9.6. *If $(x, y) \in E_t(K(t))$, then there exist an endomorphism $e : E \rightarrow E$ and a point $R \in E(K)$ of order dividing two such that $(x, sy) = e + R$.*

Proof. Suppose that $(x, y) \in E_t(K(t))$. Then (x, sy) maps E to E . The map $\phi : E(K) \rightarrow E(K) : (t, s) \mapsto (x(t), sy(t))$ is rational, so by lemma 9.3, $\psi := \phi - \phi(\mathcal{O})$ is an endomorphism of E . Hence there exist an endomorphism e of E and a point $R = \phi(\mathcal{O}) \in E(K)$ such that $\phi = e + R$. We will now prove that the order of R divides 2. This is clear if $R = \mathcal{O}$, so we can assume $R \neq \mathcal{O}$.

First suppose that e is constant. Then $e(Q) = \mathcal{O}$ for all $Q \in E(K)$, so $\phi(Q) = R$. Hence for all $(t, s) \in E(K)$, we have $(x(t), sy(t)) = R$. Since $(t, -s) \in E(K)$ for all $(t, s) \in E(K)$, this implies $y(t) = 0$. Hence the order of $R = (x, 0)$ divides two.

Now suppose that e is non-constant. By lemma 9.4, all non-constant endomorphisms of E are of the form $e(t, s) = (z(t), sw(t))$ with $z, w \in \overline{K}(t)$. Since $e = \phi - R$, $z, w \in K(t)$. Hence $R = \phi - e = (x(t), sy(t)) - (z(t), sw(t))$. We compute R from this: the line through ϕ and $-e$ is given by $Y(t, s) = \lambda(t, s)X(t) + \nu(t, s)$ with $\lambda(t, s) = s \frac{y(t) + w(t)}{x(t) - z(t)}$ and $\nu(t, s) = sy(t) - \lambda(t, s)x(t) = s \left(y(t) - \frac{y(t) + w(t)}{x(t) - z(t)} x(t) \right)$. By the addition law, the x -coordinate of $\phi - e$ is $x_R(t, s) = \lambda^2(t, s) - a - x(t) - z(t) \in K(t, s)$. Since $s^2 = t^3 + at^2 + bt + c$, λ^2 only depends on t , so $x_R \in K(t)$. The y -coordinate is

$$y_R(t, s) = -(\lambda(t, s)x_R(t) + \nu(t, s)) = -s \cdot \left(\frac{y(t) + w(t)}{x(t) - z(t)} \cdot x_R(t) + y(t) - \frac{y(t) + w(t)}{x(t) - z(t)} \right) \in sK(t)$$

so $R = (x_R(t), y_R(t, s))$ with $y_R(t, s) \in sK(t)$. But R is a constant point in $E(K)$, not depending on t and s . In particular, it should not change if we replace s by $-s$. Hence $R = (x_R(t), 0)$, so R has order 2. \square

Proof of theorem 9.2. First notice that $\Phi(P) = \Phi(t, 1) = ((t, s) \mapsto (t, s)) = \text{Id}_{E(K)}$. If P has finite order, then there exists $n \in \mathbb{Z}_{\neq 0}$ such that $nP = \mathcal{O}$. By applying the map Φ to this equality, we get

$$nS = n \text{Id}_{E(K)}(S) = n\Phi(P)(S) = \Phi(nP)(S) = \Phi(\mathcal{O})(S) = \mathcal{O}$$

for all $S \in E(K)$, so multiplication by n is the same as multiplication by 0. But since E has no complex multiplication, the endomorphism ring is isomorphic to \mathbb{Z} . Under this isomorphism, multiplication by n corresponds to $n \in \mathbb{Z}$. Since $n \neq 0$, multiplication by n cannot be the same as multiplication by 0, i.e. the map $S \mapsto \mathcal{O}$. This shows that P has infinite order.

Suppose that $Q \in E_t(K(t)) \setminus \{\mathcal{O}\}$. Then there exist an endomorphism e and a point $R \in E(K)$ of order dividing two such that $\Phi(Q) = e + R$ (lemma 9.6). Since E has no complex multiplication, all endomorphisms are given by multiplication by a fixed number. Hence there exists some $n \in \mathbb{Z}$ such that $e(t, s) = n(t, s) = n\Phi(P)(t, s)$. Since Φ is a homomorphism, this is equal to $\Phi(nP)(t, s)$.

If $R = \mathcal{O}$, then $\Phi(Q) = e = \Phi(nP)$. Since Φ is injective by lemma 9.5, it follows that $Q = nP$, so Q is an element of the group of points generated by P .

If $R \neq \mathcal{O}$, then $R = (x_R, 0)$ with $f(x_R) = 0$. Hence R is also an element of $E_t(K(t))$, and $\Phi(R)(t, s) = R$ for all $(t, s) \in E(K)$. It follows that

$$\Phi(Q)(t, s) = e(t, s) + R(t, s) = \Phi(nP)(t, s) + \Phi(R)(t, s) = \Phi(nP + R)(t, s)$$

This implies that $Q = nP + R$. R has order dividing two, so Q is an element of the group generated by P and the points of order two.

This shows that P generates $E_t(K(t))$ modulo the points of order two. \square

Definition 9.7. For $n \neq 0$, define x_n and y_n by $nP = (x_n(t), y_n(t))$. By theorem 9.2, P has infinite order, so x_n and y_n are well-defined.

Definition 9.8. Let $f, g \in K(t)$, and let \tilde{f} and \tilde{g} be the homogenizations, i.e. \tilde{f} and \tilde{g} are quotients of two homogeneous polynomials of the same degree, that have no common factors. Then \tilde{f} and \tilde{g} are functions on the projective line. We define the relation $f \sim g$ by:

$$f \sim g \iff (\tilde{f} - \tilde{g})([1 : 0]) = 0$$

Lemma 9.9. $\frac{x_n}{ty_n} \sim n$ for all $n \in \mathbb{Z}_{\neq 0}$.

Proof. In the theory of formal groups, we only considered curves with equations of the form $y^2 = x^3 + ax^2 + bx + c$, while E_t has the equation $f(t)y^2 = f(x)$. Therefore, we define C_t by $y^2 = x^3 + af(t)x^2 + bf(t)^2x + cf(t)^3$. Then we have a bijection $\chi : E_t(K(t)) \rightarrow C_t(K(t)) : (x, y) \mapsto (f(t)x, f(t)^2y)$. Define $P' = \chi(P)$, and write $P' = (x'_n, y'_n)$ and $z'_i = -\frac{x'_i}{ty'_i}$ for all i . By definition of the multiplication-by- n map in the formal group of C_t , we have $z'_n = [n]z'_1$. Since $P' = (f(t)t, f(t)^2)$, we have $z'_1 = -\frac{1}{f(t)}$ so $z'_n = [n]\left(-\frac{1}{f(t)}\right)$. By lemma 8.15, $[n](T) = nT + (\text{higher order terms})$. Write $[n](T) = nT + T^2G(T)$, where $G \in K[[T]]$. Then

$$z'_n = [n](z'_1) = [n]\left(-\frac{1}{f(t)}\right) = -\frac{n}{f(t)} + \frac{1}{f(t)^2} \cdot G\left(\frac{1}{f(t)}\right)$$

Notice that

$$\frac{x_n}{ty_n} = \frac{\frac{x'_n}{f(t)}}{t\frac{y'_n}{f(t)^2}} = f(t)\frac{x'_n}{ty'_n}$$

Hence

$$\widetilde{\frac{x_n}{ty_n}}([1 : 0]) = -\widetilde{f(t)z'_n}([1 : 0]) = \left(n - \frac{1}{f(t)} \cdot G\left(\frac{1}{f(t)}\right)\right)([1 : 0])$$

Since $f(t) = t^3 + at^2 + bt + c$, we have $\tilde{f}(T, S) = T^3 + aT^2 + bTS^2 + cS^3$, so

$$\widetilde{\frac{1}{f(T, S)}}([1 : 0]) = \frac{S^3}{T^3 + aT^2S + bTS^2 + cS^3} \Big|_{(T, S) = ([1 : 0])} = \frac{0}{1} = 0$$

This implies that $\tilde{G}\left(\frac{1}{f(t)}\right) = 0$, so

$$\widetilde{\frac{x_n}{ty_n}}([1 : 0]) = n([1 : 0]) = n$$

Hence $\frac{x_n}{ty_n} \sim n$. \square

9.2 Proof of the negative solution to Hilbert's tenth problem

Definition 9.10. Define the predicate Imt on $K(t)$ by

$$\text{Imt}(z) \iff z \in K(t) \wedge (z = 0 \vee \exists x, y \in K(t) : (x, y) \in 2E_t(K(t)) \wedge 2tyz = x)$$

Lemma 9.11. (i) Imt is diophantine.

(ii) If $\text{Imt}(z)$ holds, then there is an integer n such that $z \sim n$.

(iii) For every integer n , there exists $z \in \mathbb{Q}(t)$ satisfying $\text{Imt}(z)$ and $z \sim n$.

Proof. (i) Notice that

$$(x, y) \in 2E_t(K(t)) \iff (x, y) \in E_t(K(t)) \wedge (\exists u, v : (u, v) \in E_t(K(t)) \wedge (x, y) = 2(u, v))$$

The expression $(x, y) \in E_t(K(t))$ is diophantine, since it is equivalent to $(t^3 + at^2 + bt + c)y^2 = x^3 + ax^2 + bx + c$. Similarly, the expression $(u, v) \in E_t(K(t))$ is diophantine. Finally, duplication is given by a rational formula, so $(x, y) = 2(u, v)$ is also diophantine.

(ii) Suppose that $\text{Imt}(z)$ holds. We can assume that $z \neq 0$. Then there exist $x, y \in 2E_t(K(t))$ such that $z = \frac{x}{2iy}$. Hence there also exists $(u, v) \in E_t(K(t))$ such that $(x, y) = 2(u, v)$. Since P generates $E_t(K(t))$ modulo points of order two (theorem 9.2), there exists $n \in \mathbb{Z}$ and a point Q of order at most 2 such that $(u, v) = nP + Q$. This implies $(x, y) = 2nP = (x_{2n}, y_{2n})$. By lemma 9.9, we have $\frac{x_{2n}}{iy_{2n}} \sim 2n$, so $z \sim n$.

(iii) If $n = 0$, take $z = 0$. For $n \neq 0$, put $z = \frac{x_{2n}}{2iy_{2n}}$. Then $\text{Imt}(z)$ holds with $x = x_{2n}, y = y_{2n}$, and $z \sim n$ by lemma 9.9. \square

Definition 9.12. Denote the elliptic curve with Weierstrass equation $y^2 = x^3 - 4$ by C , and define the predicate Com by

$$\text{Com}(y) \iff y \in K(t) \wedge \exists x \in K(t) : (x, y) \in C(K(t))$$

Lemma 9.13. (i) $\text{Com}(y)$ is diophantine.

(ii) If $\text{Com}(y)$ holds, then $y \in K$.

(iii) For all $z \in \mathbb{Q}$, there exists $y \in \mathbb{Q}$ such that $\text{Com}(y)$ holds and $y > z$.

In the proof of this lemma, we need the fact that the group of real points of C $y^2 = x^3 - 4$ is, as a Lie group, isomorphic to the unit circle. This follows from the following theorem:

Theorem 9.14. (*Theorem 3.6 in [BtD85]*) A connected abelian Lie group G is a product of a torus and a vector space: $G \cong \mathbb{T}^k \times \mathbb{R}^s$. Here \mathbb{T} denotes the torus \mathbb{R}/\mathbb{Z} .

Corollary 9.15. As a Lie group, $C(\mathbb{R})$ is isomorphic to the unit circle, i.e. $C(\mathbb{R}) \cong \mathbb{T}$.

Proof. It is well known that elliptic curves over \mathbb{C} are isomorphic to \mathbb{C}/Λ , for some lattice $\Lambda \subseteq \mathbb{C}$ (for example, see [Sil86] proposition VI.5.2b). This is clearly a compact group, so $C(\mathbb{C})$ is compact. $C(\mathbb{R})$ is a closed subgroup of $C(\mathbb{C})$, so it is also compact. Furthermore, C has only one real two-torsion point, namely $(\sqrt[3]{4}, 0)$. Hence $C(\mathbb{R})$ is connected as a topological space. Now it follows from theorem 9.14 that $C(\mathbb{R}) \cong \mathbb{T}^k$ for some $k \in \mathbb{N}$ (since \mathbb{R}^s is not compact for $s > 0$). Since C is one-dimensional, it has to be equal to \mathbb{T} . \square

Lemma 9.16. Every infinite subgroup of \mathbb{T} is dense in \mathbb{T} .

Proof. Let S be an infinite subgroup of \mathbb{T} . Take all elements of \mathbb{T} to be in the interval $[0, 1)$. There are two possibilities for S : either it has an element of infinite order, or all elements have finite order. In both cases, we prove that there exists an element α such that $0 < \alpha < \varepsilon$ for all $\varepsilon > 0$.

If all elements have finite order, then they are all rational. For $\beta \in S$, write $\beta = \frac{p\beta}{q\beta}$. Then every interval of length at least $\frac{1}{q\beta}$ contains a multiple of β . Hence if α does not exist, then $\varepsilon < \frac{1}{q\beta}$ for all β , so $q\beta < \frac{1}{\varepsilon}$. But this gives only finitely many possibilities for $q\beta$, and since $0 \leq p\beta < q\beta$, there are also only finitely many possibilities for β . This gives a contradiction with the assumption that S is infinite. So in this case, α exists.

Now suppose there exists β of infinite order, i.e. $\beta \notin \mathbb{Q}$. For all $x \in \mathbb{R}$, denote the fractional part of x by $\{x\}$, i.e. $\{x\} = x - \lfloor x \rfloor$. Define

$$V = \{\{k\beta\} \mid k \in \mathbb{Z}_{\neq 0}\}$$

Since β is irrational, there is no $m \neq 0$ such that $\{m\beta\} = 0$, so there are also no m, n such that $\{m\beta\} = \{n\beta\}$ (since then $\{(m-n)\beta\} = 0$). Hence V has infinitely many elements, and they all

lie in the interval $[0, 1)$. But this implies that there are two elements with distance smaller than ε , i.e. there exist $m, n \in \mathbb{Z}$ such that $0 < \{n\beta\} - \{m\beta\} < \varepsilon$. Hence $0 < \{(n - m)\beta\} < \varepsilon$, so $\alpha = (n - m)\beta$ exists.

Now it is almost clear that S is dense in \mathbb{T} : let $a, b \in S$, and choose $\varepsilon = b - a$. Then there exists $\alpha \in S$ such that $0 < \alpha < \varepsilon$. Hence every interval of length at least ε contains a multiple of α , so in particular, there exists $k \in \mathbb{Z}$ such that $a < k\alpha < b$. This shows that S is dense in \mathbb{T} . \square

Proof of lemma 9.13. (i) This is clear from the definition.

(ii) In this proof, we will use some concepts of algebraic geometry that we only need here. Therefore, we will not define them, but they can be found in any book about algebraic geometry (for example, [Har06]).

Suppose that $y \in K(t) \setminus K$ satisfies $\text{Com}(y)$. Then there exists $x \in K(t)$ such that $y(t)^2 = x(t)^3 - 4$. Denote the projective line over K by \mathbb{P}^1 , and define the map φ by

$$\varphi : \mathbb{P}^1 \rightarrow C : [t, s] \mapsto \begin{cases} (x(t'), y(t')) = [x(t') : y(t') : 1] & \text{if } [t : s] = [t' : 1] \\ \mathcal{O} = [0 : 1 : 0] & \text{if } [t : s] = [1 : 0] \end{cases}$$

Then φ is a morphism of curves, and $\varphi(\mathbb{P}^1)$ contains more than one point. Proposition II.6.8 from [Har06] implies that $\varphi(\mathbb{P}^1) = C$, and φ is a finite morphism. By [Har06], IV.2.5.4 this implies that $g(\mathbb{P}^1) \geq g(C)$, where g denotes the genus. But $g(\mathbb{P}^1) = 0$ (for example, see [Sil86], example II.5.6), and $g(C) = 1$ since C is an elliptic curve (see [Sil86], section III.3, definition of elliptic curve and proposition 3.1(c)). Contradiction, so y does not exist. Hence if y satisfies $\text{Com}(y)$, then $y \in K$.

(iii) Notice that $Q := (2, 2) \in C(\mathbb{Q})$. By the duplication formula, $2Q = (5, -11)$, and the x -coordinate of $4Q$ is $\frac{785}{484}$. Hence by Lutz-Nagell's theorem, Q has infinite order, so $C(\mathbb{Q})$ contains infinitely many points. By corollary 9.15, $C(\mathbb{R})$ is (as a Lie group) isomorphic to \mathbb{T} . By lemma 9.16, every infinite subgroup of \mathbb{T} is dense, so $C(\mathbb{Q})$ is dense in $C(\mathbb{R})$. For every $z \in \mathbb{Q}$, there exists $y' \in \mathbb{R}$ such that $y' > z$ and $\text{Com}(y')$ holds. For example, take $y' = z + 1$ and $x = \sqrt[3]{y'^2 + 4}$; then $(x, y') \in C(\mathbb{R})$. Since $C(\mathbb{Q})$ is dense in $C(\mathbb{R})$, there exists $y \in \mathbb{Q} \cap (y', y' + 1)$ such that $(x, y) \in C(\mathbb{Q})$ for some $x \in \mathbb{Q}$. This y clearly satisfies the conditions of the lemma. \square

Definition 9.17. Define the predicate ~ 0 on $K(t)$ by

$$z \sim 0 \iff z \in K(t) \wedge (\exists x_1, \dots, x_5, y \in K(t) : \text{Com}(y) \wedge (y - t)z^2 + 1 = x_1^2 + \dots + x_5^2)$$

Lemma 9.18. (i) The predicate $z \sim 0$ is diophantine.

(ii) If $z \sim 0$, then $z \sim 0$.

(iii) If $z \in \mathbb{Q}(t)$ and $z \sim 0$, then $z \sim 0$.

In the proof of this lemma, we need the following theorem of Pourchet:

Theorem 9.19. (Théorème 2 in [Pou71]) Let K be a number field, and let $a, b, c \in K^\times$. Let $f \in K[X] \setminus \{0\}$ be of even degree $2n$. Suppose the following condition holds: for every ordering on K such that $a > 0, b > 0$ and $c > 0$, f is positive definite, and for every ordering on K such that $a > 0, b > 0$ and $c < 0$, f is not positive definite and n is odd. Then there exist $f_1, \dots, f_5 \in K[X]$ such that

$$f = f_1^2 + af_2^2 + bf_3^2 + abf_4^2 + cf_5^2 \quad \text{and} \quad \deg(f - cf_5^2) = 2n$$

Definition 9.20. A total ordering on a field is a relation \leq that is reflexive, transitive and anti-symmetric, satisfying the following conditions for all $a, b, c \in K$:

$$\begin{aligned} a &\leq b \text{ or } b \leq a \\ \text{if } a &\leq b \text{ then } a + c \leq b + c \\ \text{if } 0 &\leq a \text{ and } 0 \leq b \text{ then } 0 \leq ab \end{aligned}$$

Definition 9.21. A function $f \in K[X]$ is *positive definite* for an ordering \leq if $0 \leq f(x)$ for all $x \in K$.

Lemma 9.22. *The only total ordering on \mathbb{Q} is the standard ordering.*

Proof. Let \leq be a total ordering on \mathbb{Q} . Suppose that $1 \leq 0$. Add -1 to both sides: $0 \leq -1$. But now also $0 \leq -1 \cdot -1 = 1$, which gives a contradiction with the anti-symmetry. Since \leq is a total ordering, it follows that $0 \leq 1$. Now repeatedly add 1 to both sides to get $1 \leq 2$, $2 \leq 3$, etc. By adding -1 we obtain $-1 \leq 0$, $-2 \leq -1$, etcetera. Hence the ordering on \mathbb{Z} is unique.

Now suppose that $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$ with $a, c, \in \mathbb{Z}$, $b, d \in \mathbb{Z}_{>0}$ and $\frac{a}{b} \neq \frac{c}{d}$. Then there are two possibilities: either $ad \leq bc$ or $bc < ad$. In the former case, we must also have $\frac{a}{b} \leq \frac{c}{d}$, since if $\frac{c}{d} \leq \frac{a}{b}$, then multiplying by bc gives $bc \leq ad$. In the same way, $bc \leq ad$ implies $\frac{c}{d} \leq \frac{a}{b}$. Hence the ordering on \mathbb{Q} is also unique. \square

Corollary 9.23. *Let $f \in \mathbb{Q}[t]$ be positive definite for the standard ordering. Then f can be written as the sum of five squares.*

Proof. Take $K = \mathbb{Q}$ and $a = b = c = 1$ in Pourchet's theorem, and note that the only total ordering on \mathbb{Q} is the standard ordering, so $1 > 0$ and f is positive definite for all orderings. \square

Proof of lemma 9.18. (i) This follows immediately from the definition and lemma 9.13(i).

(ii) Suppose that $z \sim 0$. Then

$$(y-t)z^2 + 1 = x_1^2 + \dots + x_5^2$$

for some $y, x_1, \dots, x_5 \in K(t)$ with y satisfying $\text{Com}(y)$. Each of the terms x_i^2 has even degree. Since zero is not a sum of squares in K , the coefficient of the term of highest degree is non-zero, so $x_1^2 + \dots + x_5^2$ also has even degree. Hence $(y-t)z^2 + 1$ also has even degree. But z^2 has even degree, and by lemma 9.13(ii), $y \in K$, so $(y-t)z^2$ has odd degree. Hence we must have $\deg((y-t)z^2) < \deg(1) = 0$, so $\deg(tz^2) < 0$, so $\deg(z) = \frac{1}{2} \deg(z^2) = \frac{1}{2}(\deg(z^2) - 1) < \frac{1}{2} \cdot -1 < 0$. But this implies that $z \sim 0$.

(iii) Suppose that $z \sim 0$ for some $z \in \mathbb{Q}(t)$. Then $\deg(z) \leq -1$, so $\deg(tz^2) = 1 + 2 \deg(z) \leq -1$, so $tz^2 \sim 0$. Hence there exists $N \in \mathbb{N}$ such that $|tz(t)^2| \leq \frac{1}{2}$ for all $|t| > N$. By lemma 9.13(iii), there exists $y \in \mathbb{Q}$ such that $\text{Com}(y)$ holds, and $y > N \geq 0$. Now consider the map $\mathbb{Q} \rightarrow \mathbb{Q} : t \mapsto (y-t)z(t)^2 + 1$. If $|t| \leq N$, then $|t| < y$, so $y-t > 0$ and hence $(y-t)z(t)^2 + 1 \geq 1 > 0$. On the other hand, if $|t| > N$, then $tz(t)^2 \leq \frac{1}{2}$ so $(y-t)z(t)^2 + 1 \geq yz(t)^2 + \frac{1}{2} \geq \frac{1}{2} > 0$. Hence $(y-t)z(t)^2 + 1$ is a positive definite rational map.

Write $z(t) = \frac{z_1(t)}{z_2(t)}$ with $z_1, z_2 \in \mathbb{Q}[t]$. Then $(y-t)z_1(t)^2 + z_2(t)^2$ is a positive-definite element of $\mathbb{Q}[t]$, so by corollary 9.23, there exist $x_1, \dots, x_5 \in \mathbb{Q}[t]$ such that $(y-t)z_1(t)^2 + z_2(t)^2 = x_1^2 + \dots + x_5^2$. It follows that

$$(y-t)z(t)^2 + 1 = \left(\frac{x_1(t)}{z_2(t)} \right)^2 + \dots + \left(\frac{x_5(t)}{z_2(t)} \right)^2$$

\square

Theorem 9.24. *If K is a formally real field of characteristic zero, then there is no algorithm to decide whether a diophantine equation with coefficients in $\mathbb{Z}[t]$ has a solution in $K(t)$.*

Proof. This proof is similar to the proof of theorem 5.1.

Suppose that there is an algorithm to decide whether a diophantine equation with coefficients in $\mathbb{Z}[t]$ has a solution in $K(t)$, and let $P(z_1, \dots, z_n)$ be a polynomial over \mathbb{Z} . By lemmas 9.11 and 9.18, we have

$$\begin{aligned} \exists z_1, \dots, z_n \in \mathbb{Z} : P(z_1, \dots, z_n) = 0 & \iff \\ \exists Z_1, \dots, Z_n \in K(t) : \text{Imt}(Z_1) \wedge \dots \wedge \text{Imt}(Z_n) \wedge P(Z_1, \dots, Z_n) \sim 0 \end{aligned}$$

Indeed, for the ‘if’ part, we can take $z_i \in K$ such that $Z_i \sim z_i$. Then $z_i \in \mathbb{Z}$ by lemma 9.11(ii), and

$$P(z_1, \dots, z_n) = P(Z_1([1:0]), \dots, Z_n([1:0])) = P(Z_1, \dots, Z_n)([1:0]) = 0$$

The last equality follows from the fact that $P(Z_1, \dots, Z_n) \sim 0$ by lemma 9.18(ii).

For the ‘only if’ part, take $Z_i \sim z_i$ with $Z_i \in \mathbb{Q}(t)$, satisfying $\text{Imt}(Z_i)$. These exists by lemma 9.11(iii). Then

$$P(Z_1, \dots, Z_n)([1:0]) = P(Z_1([1:0]), \dots, Z_n([1:0])) = P(z_1, \dots, z_n) = 0$$

so $P(Z_1, \dots, Z_n) \sim 0$. Since P is a polynomial over \mathbb{Z} and $Z_i \in \mathbb{Q}(t)$, we have $P(Z_1, \dots, Z_n) \in \mathbb{Q}(t)$. Hence $P(Z_1, \dots, Z_n) \sim 0$ by lemma 9.18(iii).

Since we have diophantine representations of \sim and Imt , we can compute a polynomial P^* over $\mathbb{Z}[t]$ such that

$$\begin{aligned} \exists Z_1, \dots, Z_n \in K(t) : \text{Imt}(Z_1) \wedge \dots \wedge \text{Imt}(Z_n) \wedge P(Z_1, \dots, Z_n) \sim 0 & \iff \\ \exists Z_1, \dots, Z_m \in K(t) : P^*(Z_1, \dots, Z_m) = 0 & \end{aligned}$$

Hence

$$\exists z_1, \dots, z_n \in \mathbb{Z} : P(z_1, \dots, z_n) = 0 \iff \exists Z_1, \dots, Z_m \in K(t) : P^*(Z_1, \dots, Z_m) = 0$$

By deciding whether P^* has a zero in $K(t)$, we can decide whether P has a zero in \mathbb{Z} . This gives an algorithm that decides whether a polynomial over \mathbb{Z} has a zero in \mathbb{Z} , which is a contradiction with the undecidability of Hilbert’s tenth problem over \mathbb{Z} . \square

Remark 9.25. (i) We have to use the predicate ~ 0 instead of ~ 0 , since ~ 0 is not a diophantine predicate.

(ii) The assumption that K is formally real is only necessary for the proof of lemma 9.18(ii). If we could remove this assumption, the proof would hold for all fields of characteristic zero. Unfortunately, for other fields, it is indeed possible that $z \sim 0$ holds for $z \in K(t)$, while $z \sim 0$ doesn’t hold. For example, take $K = \mathbb{C}$ and $z = 1$. Then clearly $z \not\sim 0$. Take $y = 0$, then $\text{Com}(y)$ holds by putting $x = \sqrt[3]{4}$. Furthermore, by putting $x_1 = t - \frac{5}{4}$, $x_2 = it - \frac{3i}{4}$ and $x_3 = x_4 = x_5 = 0$, we have $(y-t)z^2 + 1 = -t + 1 = x_1^2 + \dots + x_5^2$. So for these fields, we really cannot use the predicate ~ 0 .

(iii) Although we cannot use the predicate ~ 0 for arbitrary fields of characteristic zero, it might be possible to use another predicate to alter the proof. The only properties of ~ 0 we used are that it is diophantine, it implies ~ 0 (for formally real fields), and is implied by ~ 0 on $\mathbb{Q}(t)$. Unfortunately I do not know whether there exists a predicate with these properties for all fields of characteristic zero. In fact, no proof is known for the (un)decidability of Hilbert’s tenth problem for $\mathbb{C}(t)$.

Chapter 10

Divisibility sequences and exponentiation

In this chapter, we return to the original problem for \mathbb{Z} in the language $(0, 1; +, \cdot)$. The proof of the undecidability of the diophantine theory of \mathbb{Z} in $(0, 1; +, \cdot)$ is based on the fact that exponentiation is diophantine. Matiyasevich proved this in 1970 using elementary methods, as we did in section 2.2. Historically, this was the last step of the proof.

In this chapter, we will try to give another proof using sequences associated to elliptic curves. These sequences turn out to have interesting divisibility properties. First we will introduce these sequences and prove that they satisfy divisibility relations. Then we will try to replace Matiyasevich's proof by a proof using elliptic divisibility sequences. Unfortunately, the proof is incomplete.

10.1 Elliptic divisibility sequences

Let E be an elliptic curve over \mathbb{Q} , given by a Weierstrass equation $y^2 = x^3 + ax^2 + bx + c$ with $a, b, c \in \mathbb{Z}$. By the Mordell-Weil theorem, the group $E(\mathbb{Q})$ of rational points of E is isomorphic to $T \oplus \mathbb{Z}^r$, where T is a finite group of torsion points, and r is the rank of the curve. We will only consider curves with non-zero rank, so we assume that there exists a point $P \in E(\mathbb{Q})$ of infinite order.

We can write $nP = (x_n, y_n) = (\frac{a_n}{b_n}, \frac{c_n}{d_n})$ with $a_n, b_n, c_n, d_n \in \mathbb{Q}$ and $\gcd(a_n, b_n) = \gcd(c_n, d_n) = 1$. Since $nP \in E$, we have

$$\frac{c_n^2}{d_n^2} = \frac{a_n^3 + aa_n^2b_n + ba_nb_n^2 + cb_n^3}{b_n^3}$$

In both fractions, the numerator is coprime to the denominator, so $d_n^2 = b_n^3$. Hence there exists $B_n \in \mathbb{Z}$ such that $b_n = B_n^2$ and $d_n = B_n^3$, and $nP = (x_n, y_n) = (\frac{a_n}{B_n^2}, \frac{c_n}{B_n^3})$, with $a_n, B_n, c_n \in \mathbb{Z}$, and $\gcd(a_n, B_n) = \gcd(c_n, B_n) = 1$. Notice that B_n and c_n are defined up to sign.

Definition 10.1. A sequence $(u_n)_{n \in \mathbb{N}}$ is a *divisibility sequence* if the following holds: if $n|m$, then $u_n|u_m$. The sequence is a *strong divisibility sequence* if $\gcd(u_n, u_m) = u_{\gcd(n, m)}$. A sequence $(u_n)_{n \in \mathbb{N}}$ is an *odd divisibility sequence* if the following holds: if $n|m$ and $\frac{m}{n}$ is odd, then $u_n|u_m$.

Remark 10.2. Notice that a strong divisibility sequence is indeed a divisibility sequence: if $n|m$, then $\gcd(n, m) = n$, so $\gcd(u_n, u_m) = u_{\gcd(n, m)} = u_n$. Hence $u_n|u_m$.

Theorem 10.3. (i) If p is a prime such that $p|B_n$, then for every $t \in \mathbb{Z}$

$$v_p(B_{tn}) = v_p(B_n) + v_p(t)$$

(ii) $(B_n)_{n \in \mathbb{N}}$ is a strong divisibility sequence.

Proof. (i) Let (\widehat{E}, F) be the formal group of E , and let p be a prime number, with corresponding valuation v_p . To (\widehat{E}, F) , we associated the group $\widehat{E}(\mathcal{M}_p)$, where \mathcal{M}_p is the unique maximal ideal

of \mathbb{Z}_p . Since the Weierstrass equation has coefficients in \mathbb{Z} , (\widehat{E}, F) is a formal group over \mathbb{Z} , so in particular, it is a formal group over \mathbb{Z}_p , and we can apply the theory from chapter 8.

Let

$$E_1 = \{Q \in E(\mathbb{Q}_p) \mid Q \text{ is non-singular, } Q = (x, y), v_p(x) < 0\}$$

Consider the map $z : E_1 \rightarrow \widehat{E}(\mathcal{M}_p) : Q = (x, y) \mapsto z(Q) = -\frac{x}{y}$. This is the same map as we used in constructing the formal group of an elliptic curve. Notice that $v_p(-\frac{x}{y}) = v_p(x) - v_p(y)$. Let $Q = (x, y) \in E_1$, so $v_p(x) < 0$. Since $y^2 = x^3 + ax^2 + bx + c$, and $a, b, c \in \mathbb{Z}$, we have $v_p(x^3) = 3v_p(x)$, $v_p(ax^2) = v_p(a) + 2v_p(x) \geq 2v_p(x) > 3v_p(x)$ etc. so $2v_p(y) = v_p(y^2) = v_p(x^3) = 3v_p(x)$. Hence $v_p(z(Q)) = -\frac{1}{2}v_p(x)$, so z indeed maps E_1 to $\widehat{E}(\mathcal{M}_p)$. Furthermore, it is a homomorphism by definition of the formal addition law. Since $(x, y) \mapsto (z, w)$ is a bijection, and w is given by a power series in z , the map $(x, y) \mapsto z$ is also a bijection. Hence this map is an isomorphism, satisfying $v_p(z) = -\frac{1}{2}v_p(x)$, and E_1 is a group.

For any integer $r > \frac{1}{p-1}$, theorem 8.24 gives an isomorphism $\log_{\widehat{E}} : \widehat{E}(\mathcal{M}_p^r) \rightarrow \widehat{\mathbb{G}}_a(\mathcal{M}_p^r)$. This isomorphism preserves valuations, so for all $Q \in \mathcal{M}_p^r$, we have $v_p(Q) = v_p(\log_{\widehat{E}}(Q))$. By lemma 8.12, $\log_{\widehat{E}}(z(2Q)) = \log_{\widehat{E}}(F(z(Q), z(Q))) = 2\log_{\widehat{E}}(z(Q))$, so by induction $\log_{\widehat{E}}(z(nQ)) = n\log_{\widehat{E}}(z(Q))$. Hence

$$v_p(z(nQ)) = v_p(\log_{\widehat{E}}(z(nQ))) = v_p(n\log_{\widehat{E}}(z(Q))) = v_p(n) + v_p(\log_{\widehat{E}}(z(Q))) = v_p(n) + v_p(z(Q))$$

Now write $nP = (x_n, y_n) = (\frac{a_n}{B_n^2}, \frac{c_n}{B_n^3})$, with $a_n, B_n, c_n \in \mathbb{Z}$, and $\gcd(a_n, B_n) = \gcd(c_n, B_n) = 1$, and assume that $p \mid B_n$. Then $v_p(B_n) \geq 1$.

If $p \neq 2$ or $v_p(z(nP)) > 1$, then we have an isomorphism as above with $r = v_p(z(nP))$. Since $nP \in E_1$, we can apply the above to $Q = nP$. It follows that $v_p(z(tnP)) = v_p(t) + v_p(z(nP))$, i.e. $v_p(-\frac{atn}{c_{tn}}) = v_p(t) + v_p(-\frac{a_n B_n}{c_n})$ for all $t \in \mathbb{Z}$. Since $\gcd(a_n, B_n) = \gcd(c_n, B_n) = 1$, we have $v_p(a_n) = v_p(c_n) = 1$. Because we also have $tnP \in E_1$, $p \mid B_{tn}$, so $v_p(a_{tn}) = v_p(c_{tn}) = 1$. Hence $v_p(B_{tn}) = v_p(t) + v_p(B_n)$.

The only case left is $p = 2$ and $v_2(z(nP)) \leq 1$. Since $v_2(z(nP)) > 0$, we have $v_2(z(nP)) = 1$. Write $2nP = (x_{2n}, y_{2n})$. Then by the duplication formula,

$$x_{2n} = \frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4(x^3 + ax^2 + bx + c)} = \frac{x}{4} \cdot \frac{1 - 2bx^{-2} - 8cx^{-3} + (b^2 - 4ac)x^{-4}}{1 + ax^{-1} + bx^{-2} + cx^{-3}}$$

where $nP = (x, y)$. Since $v_2(x) < 0$, the second factor in this product has valuation $v_2(1) - v_2(1) = 0$, so $v_2(x_{2n}) = v_2(x) - v_2(4) = v_2(x) - 2$. Hence $v_2(B_{2n}) = -\frac{1}{2}v_2(x_{2n}) = -\frac{1}{2}v_2(x) + 1 = v_2(B_n) + 1 = v_2(B_n) + v_2(2)$, so the claim follows for $t = 2$.

Now suppose the claim holds for $t = 2^l$. Then by applying the above reasoning to $2^l P$, we get $v_2(B_{2^{l+1}n}) = v_2(B_{2^l n}) + 1$. By the assumption, $v_2(B_{2^l n}) = v_2(B_n) + l$, so $v_2(B_{2^{l+1}n}) = v_2(B_n) + l + 1 = v_2(B_n) + v_2(2^{l+1})$. By induction, the claim holds for all $t = 2^l$.

Now let t be arbitrary, and write $t = 2^l \cdot s$ with s odd. Then, as we proved above, $v_2(B_{2^l n}) = v_2(B_n) + l$. Write $r = v_2(B_n) + l$. Under the isomorphism $E_1 \rightarrow \widehat{E}(\mathcal{M}_2)$, the pre-image of $\widehat{E}(\mathcal{M}_2^r)$ is the set

$$\{Q \in E(\mathbb{Q}_2) \mid Q \text{ is non-singular, } Q = (x, y), v_2(x) \leq -2r\}$$

so this is the set of points with $v_2(B) \geq r$. Since this corresponds to the group $\widehat{E}(\mathcal{M}_2^r)$, this set is also a group. Since $2^l nP$ belongs to this group, $2^l s nP$ also belongs to it, so $v_2(B_{2^l s n}) \geq r$. Suppose that $v_2(B_{2^l s n}) > r$. Since s is odd, we can write $s = 2s' + 1$ with $s' \in \mathbb{Z}$. Then $2^l n = 2^l s n - 2^{l+1} s' n$. By the same reasoning as for $2^l s nP$, $v_2(2^{l+1} s' nP) \geq r + 1$. Since the points with $v_2(B) \geq r + 1$ form a group, this implies that $v_2(B_{2^l n}) \geq r + 1$. This contradicts the definition of r . Hence $v_2(B_{2^l s n}) = r = v_2(B_{2^l n})$. It follows that

$$v_2(B_{tn}) = v_2(B_{2^l s n}) = v_2(B_{2^l n}) = v_2(B_n) + l = v_2(B_n) + v_2(2^l) = v_2(B_n) + v_2(t)$$

(ii) Let $d = \gcd(m, n)$. Since $d|m$, we can write $m = kd$ for some integer k . Let p be a prime that divides B_d . Then by (i), we get $v_p(B_m) = v_p(B_d) + v_p(k) \geq v_p(B_d)$. Hence if $p^r|B_d$, then also $p^r|B_m$, so $B_d|B_m$. By the same reasoning, $B_d|B_n$. Hence $B_d|\gcd(B_m, B_n)$.

For the converse, choose integers x, y such that $xm + yn = d$, and let $r = v_p(\gcd(B_m, B_n))$. If $p \nmid B_m$, then clearly $v_p(B_{xm}) \geq 0 = v_p(B_m)$, and if $p|B_m$, then $v_p(B_{xm}) = v_p(B_m) + v_p(x) \geq v_p(B_m)$ by (i). In both cases, $v_p(B_{xm}) \geq v_p(B_m) \geq v_p(\gcd(B_m, B_n)) = r$. Similarly, $v_p(B_{yn}) \geq r$. Under the isomorphism $E_1 \rightarrow \widehat{E}(\mathcal{M}_p)$, the pre-image of $\widehat{E}(\mathcal{M}_p^r)$ is the set of points with $v_p(B) \geq r$. Since this corresponds to the group $\widehat{E}(\mathcal{M}_p^r)$, this set itself is also a group. Since xmP and ynP belong to this group, $dP = xmP + ynP$ also belongs to it. Hence $v_p(B_d) \geq r = v_p(\gcd(B_m, B_n))$. This holds for all p , so $\gcd(B_m, B_n)|B_d$.

It follows that $B_d = \gcd(B_m, B_n)$, and hence $(B_n)_{n \in \mathbb{N}}$ is a strong divisibility sequence. \square

Although we will only use theorem 10.3 in the next section, we will prove a few more facts about a_n , B_n and c_n . The reason to do this is that we didn't managed to complete the proof of the next section. Therefore it might be interesting to try to redo this proof, using another sequence instead of B_n . As we will prove below, $(\sqrt{a_n}B_n)_{n \in \mathbb{N}}$ is a strong divisibility sequence, and $(\sqrt{a_n})_{n \in \mathbb{N}}$ and $(\frac{c_n}{\sqrt{a_n}})_{n \in \mathbb{N}}$ are odd divisibility sequences. Maybe it can be proved that exponentiation is diophantine by using one of these sequences.

In the remaining of this chapter, we will focus on elliptic curves with a rational 2-torsion point. By translation, we can assume that this point is $(0, 0)$. Then the Weierstrass equation is $y^2 = x^3 + ax^2 + bx$.

Lemma 10.4. *Suppose that $(0, 0)$ is a rational point of E , and let P be a point of infinite order in $2E(\mathbb{Q})$. Then a_n is a square for all n , and $\sqrt{a_n}|c_n$.*

Proof. Write $P = 2Q$, with $Q \in E(\mathbb{Q})$, and let $nP = (x_n, y_n) = (\frac{a_n}{B_n^2}, \frac{c_n}{B_n^3})$ and $nQ = (x, y)$. The duplication formula gives that

$$x_n = \frac{x^4 - 2bx^2 + b^2}{4y^2} = \left(\frac{x^2 - b}{2y} \right)^2$$

Hence $\frac{a_n}{B_n^2}$ is a rational square, so a_n is the square of an integer.

Now write $a_n = A_n^2$, and substitute $x_n = \frac{A_n^2}{B_n^2}$, $y_n = \frac{c_n}{B_n^3}$ into the equation $y^2 = x^3 + ax^2 + bx$ to obtain

$$c_n^2 = A_n^6 + aA_n^4B_n^2 + bA_n^2B_n^4 = A_n^2(A_n^4 + aA_n^2B_n^2 + bB_n^4) \quad (10.1)$$

This implies $A_n^2|c_n^2$, so $A_n|c_n$. \square

Definition 10.5. We write $A_n = \sqrt{a_n}$ and $C_n = \frac{c_n}{A_n}$, so $nP = ((\frac{A_n}{B_n})^2, \frac{A_n C_n}{B_n^3})$.

By lemma 10.4, both A_n and C_n are integers. Furthermore, it is clear that $\gcd(A_n, B_n) = \gcd(C_n, B_n) = 1$.

Lemma 10.6. (i) $\gcd(A_n, C_n)^2$ divides b , so if b is square-free, then $\gcd(A_n, C_n) = 1$.

(ii) B_{2n} divides $2A_n B_n C_n$.

(iii) Suppose that P is non-singular modulo all primes dividing b . Then A_n divides B_{2n} .

Proof. (i) By equation (10.1), we have the identity $A_n^2 C_n^2 = c_n^2 = A_n^2(A_n^4 + aA_n^2B_n^2 + bB_n^4)$, so $C_n^2 = A_n^4 + aA_n^2B_n^2 + bB_n^4$. Write $d = \gcd(A_n, C_n)^2$. Then $d|C_n^2$, and $d|A_n^4 + aA_n^2B_n^2$, so $d|bB_n^4$. Since $\gcd(A_n, B_n) = 1$, this implies $d|b$, so $\gcd(A_n, C_n)^2|b$. If b is square-free, this immediately implies $\gcd(A_n, C_n) = 1$.

(ii) We have $nP = ((\frac{A_n}{B_n})^2, \frac{A_n C_n}{B_n^3})$ and $2nP = ((\frac{A_{2n}}{B_{2n}})^2, \frac{A_{2n} C_{2n}}{B_{2n}^3})$. By the duplication formula (see the proof of lemma 10.4),

$$\left(\frac{A_{2n}}{B_{2n}} \right)^2 = \left(\frac{(\frac{A_n}{B_n})^4 - b}{2 \frac{A_n C_n}{B_n^3}} \right)^2$$

so

$$\pm \frac{A_{2n}}{B_{2n}} = \frac{\left(\frac{A_n}{B_n}\right)^4 - b}{2\frac{A_n C_n}{B_n}} = \frac{A_n^4 - bB_n^4}{2A_n B_n C_n} \quad (10.2)$$

In the right-most fraction, there can be common divisors in the numerator and the denominator, but in any case, the denominator B_{2n} of x_{2n} will divide $2A_n B_n C_n$.

(iii) Suppose that P is non-singular modulo all primes dividing b . Since the non-singular points of an elliptic curve form a group under addition, nP is also non-singular for all n . The Weierstrass equation of E is $y^2 = x^3 + ax^2 + bx$, so a point (x, y) is singular modulo p if and only if $p|y$ and $p|3x^2 + 2ax + b$. Since $nP = \left(\left(\frac{A_n}{B_n}\right)^2, \frac{A_n C_n}{B_n^3}\right)$, any p dividing b cannot divide A_n , so $\gcd(A_n, b) = 1$. Since $\gcd(A_n, B_n) = \gcd(A_n, b) = 1$, there is no cancellation of factors of A_n in the numerator and denominator of $\frac{A_n^4 - bB_n^4}{2A_n B_n C_n}$. Hence A_n is a factor of the denominator, which equals B_{2n} by equation (10.2), so $A_n|B_{2n}$. \square

Lemma 10.7. *Assume that b and $a^2 - 4b$ are square-free, and P is non-singular modulo all primes dividing b or $a^2 - 4b$. Then the following holds:*

- (i) $(A_n B_n)_{n \in \mathbb{N}}$ is a strong divisibility sequence.
- (ii) If t is odd and $p|A_n$, then $v_p(A_{tn}) = v_p(A_n) + v_p(t)$ for all n . The same holds for $(C_n)_{n \in \mathbb{N}}$.
- (iii) If t is even, then $\gcd(A_n, A_{tn}) = 1$ for all n . The same holds for $(C_n)_{n \in \mathbb{N}}$.
- (iv) $(A_n)_{n \in \mathbb{N}}$ and $(C_n)_{n \in \mathbb{N}}$ are odd divisibility sequences.

Proof. The curves $E : y^2 = x^3 + ax^2 + bx$ and $E' : y^2 = x^3 - 2ax^2 + (a^2 - 4b)x$ are isomorphic, with an isomorphism given by

$$\phi : E \rightarrow E' : P \mapsto \begin{cases} \mathcal{O} & \text{if } P = \mathcal{O}, (0, 0) \\ \left(\frac{y^2}{x^2}, \frac{y(x^2 - b)}{x^2}\right) & \text{if } P = (x, y) \neq \mathcal{O}, (0, 0) \end{cases}$$

(for example, see [Sil86], example III.4.5) Write $Q = nP = \left(\left(\frac{A_n}{B_n}\right)^2, \frac{A_n C_n}{B_n^3}\right)$, and $Q' = \phi(Q) = (x(Q'), y(Q')) = \left(\left(\frac{A'_n}{B'_n}\right)^2, \frac{A'_n C'_n}{B'_n{}^3}\right)$ with $\gcd(A'_n, B'_n) = \gcd(C'_n, B'_n) = 1$. Then

$$\left(\frac{A'_n}{B'_n}\right)^2 = x(Q') = \left(\frac{\frac{A_n C_n}{B_n^3}}{\left(\frac{A_n}{B_n}\right)^2}\right)^2 = \left(\frac{C_n}{A_n B_n}\right)^2$$

Hence $\frac{A'_n}{B'_n} = \pm \frac{C_n}{A_n B_n}$. It holds that $\gcd(C_n, B_n) = 1$, and since b was assumed to be square-free, lemma 10.6(i) implies that $\gcd(A_n, C_n) = 1$. Hence $\frac{C_n}{A_n B_n}$ has no cancellation in the numerator and the denominator, and neither has $\frac{A'_n}{B'_n}$, so $A'_n = C_n$ and $B'_n = \pm A_n B_n$.

It follows that the sequences $(A_n B_n)_{n \in \mathbb{N}}$ and $(C_n)_{n \in \mathbb{N}}$ equal the sequences $(B'_n)_{n \in \mathbb{N}}$ and $(A'_n)_{n \in \mathbb{N}}$ associated to E' .

(i) By theorem 10.3, $(B'_n)_{n \in \mathbb{N}}$ is a strong divisibility sequence, so $(A_n B_n)_{n \in \mathbb{N}}$ is also a strong divisibility sequence.

(ii) Notice that we only need to prove this for $(A_n)_{n \in \mathbb{N}}$, since $(C_n)_{n \in \mathbb{N}} = (A'_n)_{n \in \mathbb{N}}$, and $a^2 - 4b$ is also square-free.

Let t be odd, and let p be a prime dividing A_n . Since $\gcd(A_n, B_n) = 1$, p cannot divide B_n . Since $B'_n = A_n B_n$, p divides B'_n , so by theorem 10.3, we have

$$\begin{aligned} v_p(A_{tn}) + v_p(B_{tn}) &= v_p(A_{tn} B_{tn}) = v_p(B'_{tn}) = v_p(B'_n) + v_p(t) = \\ &v_p(A_n B_n) + v_p(t) = v_p(A_n) + v_p(B_n) + v_p(t) = v_p(A_n) + v_p(t) \end{aligned}$$

so $v_p(A_{tn}) + v_p(B_{tn}) = v_p(A_n) + v_p(t)$.

By lemma 10.6(iii), $A_n|B_{2n}$, so $p|B_{2n}$. From theorem 10.3 and the fact that t is odd, it follows

that $\gcd(B_{tn}, B_{2n}) = B_{\gcd(tn, 2n)} = B_{\gcd(t, 2)n} = B_n$. Now $p \nmid B_n$ implies $p \nmid B_{tn}$, so $v_p(B_{tn}) = 0$. Hence $v_p(A_{tn}) = v_p(A_{tn}) + v_p(B_{tn}) = v_p(A_n) + v_p(t)$.

(iii) Again we only have to prove this for A_n . Let t be even, and let p be a prime dividing A_n . Again $p \mid B_{2n}$, and since $2n \mid tn$, theorem 10.3 implies that $B_{2n} \mid B_{tn}$, so $p \mid B_{tn}$. Since $\gcd(A_{tn}, B_{tn}) = 1$, p cannot divide A_{tn} .

It follows that $\gcd(A_n, A_{tn}) = 1$.

(iv) This follows immediately from (ii), by noting that $v_p(t) \geq 0$ so $v_p(A_{tn}) \geq v_p(A_n)$ for all p dividing A_n . \square

Definition 10.8. The sequence $(C_n)_{n \in \mathbb{N}}$ is called the odd divisibility sequence associated to E and P .

10.2 Exponentiation revisited (incomplete proof)

In this section, we try to prove that exponentiation is diophantine, by using an elliptic divisibility sequence. The idea of the proof is to take an elliptic curve, such that every twist has rank 1. As we saw in chapter 9, this is always true if we take the twist E_t . Then the group of rational points is generated by $(t, 1)$ and the two-torsion points. However, in this chapter, we will consider E_d for every natural number d . Then it is not known whether such a curve exist.

For the first part of this section, we assume that such a curve exists. After considering to which extend Matiyasevich' proof can be replaced by a proof using this elliptic curve, we will say more about the possibility of finding such a curve.

Throughout this section, let E be an elliptic curve given by a Weierstrass equation $y^2 = f(x)$ where $f(x) = x^3 + ax^2 + bx$. Let $d \in \mathbb{N}$ with $f(d) \neq 0$, and let C_d be the curve $f(d)y^2 = f(x)$. By multiplying this equation by $f(d)^3$, we get $(f(d)^2y)^2 = (f(d)x)^3 + af(d)(f(d)x) + bf(d)^2$. This gives the elliptic curve $E_d : y^2 = x^3 + af(d)x^2 + bf(d)^2x$, where the point (x, y) on C_d corresponds to the point $(f(d)x, f(d)^2y)$ on E_d . Let $g_d(x) = x^3 + af(d)x^2 + bf(d)^2x$. By construction of C_d , it contains the point $(d, 1)$. Hence on E_d we have the point $P_d = (df(d), f(d)^2)$.

We require the curve E to have the following properties:

Assumption 10.9. E has rank 0, and for all d , E_d has rank 1 and the group of rational points is generated by P_d and the torsion points.

Using an elliptic curve that satisfies assumption 10.9, we try to mimic the proof of theorem 2.2. Instead of the sequence $\alpha_b(n)$, we use for every d the sequence B_n .

By Mazur's theorem (see [Sil86], chapter VIII, theorem 7.5), all torsion points of E_d have order at most 12, and there are no torsion points of order 11. Since $3 \cdot 5 \cdot 7 \cdot 8 = 840$, the group $840E_d(\mathbb{Q})$ is generated by $840P_d$. We will denote the sequence $(B_n)_{n \in \mathbb{N}}$ corresponding to $840P_d$ on E_d by $(B_d(n))_{n \in \mathbb{N}}$ (so $840nP_d = ((\frac{A_d(n)}{B_d(n)})^2, \frac{A_d(n)C_d(n)}{B_d(n)^3})$).

To replace lemmas 2.8 and 2.10, we have the following lemmas:

Lemma 10.10. *If E satisfies assumption 10.9, then*

$$\begin{aligned} \exists n : c = B_d(n) &\iff \exists f, u, v, w, x, y \in \mathbb{Z} : f = d^3 + ad^2 + d \wedge \\ &\gcd(x, c) = 1 \wedge \gcd(y, c) = 1 \wedge \gcd(u, v) = 1 \wedge \gcd(w, v) = 1 \wedge \\ &y^2 = x^3 + ac^2fx^2 + bc^4f^2x \wedge w^2 = u^3 + av^2fu^2 + bv^4f^2u \wedge \\ &\left(\frac{x}{c^2}, \frac{y}{c^3}\right) = 840 \left(\frac{u}{v^2}, \frac{w}{v^3}\right) \end{aligned}$$

Hence the relation $\exists n : c = B_d(n)$ is positive existential.

Proof. Since $840E_d(\mathbb{Q})$ is generated by $840P_d$, all points in $840E_d(\mathbb{Q})$ are of the form $840nP_d$ for some $n \in \mathbb{Z}$. Hence there exists n such that $c = B_d(n)$ if and only if $840E_d(\mathbb{Q})$ has an element of the form $(\frac{x}{c^2}, \frac{y}{c^3})$. This means that there exists a point $(\frac{u}{v^2}, \frac{w}{v^3}) \in E_d(\mathbb{Q})$, such that $(\frac{x}{c^2}, \frac{y}{c^3}) = 840(\frac{u}{v^2}, \frac{w}{v^3})$. Hence there exists $n \in \mathbb{N}$ such that $c = B_d(n)$ if and only if there exist $u, v, w, x, y \in \mathbb{Z}$ such that $(\frac{x}{c^2}, \frac{y}{c^3}) \in E_d(\mathbb{Q})$, $(\frac{u}{v^2}, \frac{w}{v^3}) \in E_d(\mathbb{Q})$, where all numerators and denominators of the fractions are coprime, satisfying $(\frac{x}{c^2}, \frac{y}{c^3}) = 840(\frac{u}{v^2}, \frac{w}{v^3})$. Since addition is given by a positive existential formula, the same holds for multiplication by 840, so this gives a positive existential definition of the relation $\exists n : c = B_d(n)$. \square

Lemma 10.11. *Suppose that $d_1 \equiv d_2 \pmod{n}$. Then for all k , $B_{d_1}(k) \equiv B_{d_2}(k) \pmod{n}$.*

Proof. Since $d_1 \equiv d_2 \pmod{n}$, we also have $f(d_1) \equiv f(d_2)$ and $g_{d_1} \equiv g_{d_2}$. Hence the curves E_{d_1} and E_{d_2} are given by the same Weierstrass equation modulo n , and the points P_{d_1} and P_{d_2} also are the same. This implies that the points $840kP_{d_1}$ and $840kP_{d_2}$ are the same for all k , so $B_{d_1}(k) \equiv B_{d_2}(k) \pmod{n}$. \square

To find an analogon of lemmas 2.11 and 2.12, we need another assumption:

Assumption 10.12. For all d , if $B_d(n) = B_d(m)$, then $n = m$.

This might seem to be a strong assumption, but it is not. When calculating multiples of a point, one easily noted that the denominators grow very quickly and hence are different. This is indeed true:

Lemma 10.13. *There is a constant c such that $\log(|B_d(n)|) \sim cn^2$.*

Proof. See [Eve]. Since $B_d(n)$ corresponds to $840P_d$ instead of P_d , we get a constant c that is 840^2 times the constant Everest has. \square

Lemma 10.14. *Suppose that $B_d(k)|B_d(m)$. Then $k|m$.*

Proof. By theorem 10.3, B_d is a strong divisibility sequence, i.e.

$$B_d(\gcd(k, m)) = \gcd(B_d(k), B_d(m))$$

Since $B_d(k)|B_d(m)$, we have $\gcd(B_d(k), B_d(m)) = B_d(k)$, so $B_d(\gcd(k, m)) = B_d(k)$. Now it follows from assumption 10.12 that $\gcd(k, m) = k$, so $k|m$. \square

Lemma 10.15. *Suppose that $B_d(k)^2|B_d(m)$. Then $B_d(k)|\frac{m}{k}$, so in particular $B_d(k)|m$.*

Proof. By lemma 10.14, $k|m$. Write $m = lk$. Let p be a prime dividing $B_d(k)$. Then, by theorem 10.3(i),

$$v_p(B_d(m)) = v_p(B_d(kl)) = v_p(B_d(k)) + v_p(l)$$

It is given that $B_d(k)^2|B_d(m)$, so $2v_p(B_d(k)) \leq v_p(B_d(m))$. Hence

$$2v_p(B_d(k)) \leq v_p(B_d(m)) = v_p(B_d(k)) + v_p(l)$$

so

$$v_p(B_d(k)) \leq v_p(l)$$

This holds for all p dividing $B_d(k)$ so $B_d(k)|l$. In particular, $B_d(k)|m$. \square

Up to now, we have shown only that there are lemmas corresponding to the first part of lemma 2.8 and lemmas 2.10, 2.11 and 2.12, but we still have to find an analogon for the second part of lemma 2.8 and for theorem 2.13. In lemma 2.8 we proved that there is a diophantine relation between $\alpha_b(n)$ and $\alpha_b(n+1)$, namely the equation $\alpha_b(n)^2 - b\alpha_b(n-1)\alpha_b(n) + \alpha_b(n)^2 = 0$, and that this equation defines the set S_1 . This gives the following interesting result: the equation that defines the set is the same as the relation between consecutive terms of the sequence. For the sequences $(B_d(n))_{n \in \mathbb{N}}$, the situation is different. In lemma 10.10, we have given an equation that defines the set of terms of the sequence, but this is not a relation between terms of the sequence.

Such relations do exist, but they are not as nice as the relation for α_b . For example, if we change the notation to let $B_d(n)$ correspond to nP_d instead of $840nP_d$, we have the relations

$$\begin{aligned} B_d(2n+1) &= B_d(n+2)B_d(n)^3 - B_d(n-1)B_d(n+1)^3 \\ B_d(2n)B_d(2) &= B_d(n)(B_d(n+2)B_d(n-1)^2 - B_d(n-2)B_d(n+1)^2) \end{aligned}$$

but these are not relations between $B_d(n)$ and $B_d(n+1)$, and they don't define the terms of the sequence.

Furthermore, I don't know how to replace theorem 2.13, since we use the defining relation from lemma 2.8 in this.

In this attempt to prove theorem 2.2, we use a curve E with some special properties, as stated in assumptions 10.9 and 10.12. We mentioned already that we don't know whether this curve exists.

When searching for such a curve, there seem to be no curves such that all E_d have rank one. Therefore, I tried to find a curve such that E_d has rank one for all but finitely many d . Then we should have to prove that exponentiation is diophantine for the other d separately, but this might be doable. The other assumption can be weakened in the same way: if $B_d(n) = B_d(m)$ for some small n and m , then we can give a diophantine definition for d^n and d^m separately, and give a proof as we tried above for all large n . Then we can use lemma 10.13 instead of assumption 10.12.

To find curves satisfying a weakened version of assumption 10.9, I used a computer. Of course, it is not really possible to find curves with rank one for all but finitely many d using a computer, but it seems reasonable to look for a curve such that E_d has rank one for many small d . For this, I used SAGE to try all curves $E : y^2 = x^3 + ax^2 + bx$ with $|a|, |b| \leq 25$. For these E , I have computed the rank of E_d for $d = 1, 2, 3, \dots$, until there were 10 d such that E_d has rank at least 2. Let

$$N := \min\{n \in \mathbb{N} \mid \#\{d \leq n \mid E_d \text{ has rank } 1\} \geq 10\}$$

There are 9 curves for which $N \geq 40$. For these curves, I computed the number of $d \leq 199$ such that E_d has rank 1. Let this number be N_{199} . The results are shown in table 10.2. The SAGE-code can be found in appendix A.

(a, b)	N	N_{199}
(+4, -2)	44	153
(-4, +6)	66	158
(+8, -8)	48	151
(-8, +3)	40	136
(-8, +24)	54	164
(+10, +20)	46	155
(+12, +18)	44	154
(+12, -18)	53	159
(-16, +12)	48	147

Table 10.1: The number of curves with rank 1

From this table, we see that for all these curves, there are many $d \leq 199$ such that E_d has rank at least 2. We assumed that E_d had rank 1 for all but finitely many d , so this is still possible for these E . Of course it is also possible that there are large a, b such that E_d has rank 1 for all but finitely many d . But although this table does not prove anything, it seems unreasonable to expect that we can find such a curve E . Together with the problems concerning the analogues of the lemmas, as sketched above, I decided to give up this approach.

Chapter 11

Function fields over fields of odd characteristic

In chapters 7 and 9 we have proven that Hilbert's tenth problem over $K(t)$ is undecidable when K is a formally real field of characteristic zero or a finite field. In this chapter, we will try to prove the same result for fields K of positive characteristic. If necessary, we can assume that K is infinite. The idea of this attempt towards a proof is based on [Phe04]. In this article, Pheidas also tries to prove that the diophantine theory of $K(t)$ is undecidable, but he can only prove that the first-order theory is undecidable. We will prove the same result, but give another proof. After that, we will consider some possibilities to extend this proof to a complete proof that the diophantine theory of $K(t)$ is undecidable.

Similar to the case in which K is a formally real field, we will pick an elliptic curve. We will consider a subring of the endomorphism ring. This ring is isomorphic to an order in a quadratic imaginary field or a quaternion algebra. We will embed some structure containing \mathbb{N} or \mathbb{Z} into this module, and show that the relations on \mathbb{N} or \mathbb{Z} correspond to positive existential relations on $K(t)$. If we can embed a structure with undecidable positive existential theory, then it will follow that the positive existential theory of $K(t)$ is undecidable as well. Unfortunately, we will also have to use a relation that is given by a first-order formula instead of a positive existential formula.

Throughout this chapter, K will be a field of characteristic $p \geq 5$. All formulas over K are in the language $(0, 1, t; +, \cdot)$.

11.1 Endomorphisms of an ordinary elliptic curve

Let $E : s^2 = t^3 + at^2 + bt + c$ be an elliptic curve, defined over \mathbb{F}_p (notice that we use \mathbb{F}_p here instead of K). In positive characteristic, the endomorphism ring is always larger than \mathbb{Z} . There is also the *Frobenius endomorphism*, given by $\text{Frob} : E \rightarrow E : (t, s) \mapsto (t^p, s^p)$. This is indeed an endomorphism, since for all $a, b, c \in \mathbb{F}_p$, we have $a^p = a, b^p = b$ and $c^p = c$. Hence, if $s^2 = t^3 + at^2 + bt + c$, then

$$s^{2p} = (t^3 + at^2 + bt + c)^p = t^{3p} + a^p t^{2p} + b^p t^p + c^p = t^{3p} + at^{2p} + bt^p + c$$

so (t^p, s^p) also lies on E . It follows from lemma 9.3 that Frob is indeed an endomorphism of E .

Theorem 11.1. (Hasse) *Frob satisfies a quadratic equation of the form $x^2 - tx + p = 0$ for some integer t such that $|t| \leq 2\sqrt{p}$, and $\#E(\mathbb{F}_p) = p + 1 - t$.*

Proof. Both statements are proven in [Hus04], chapter 13. The first statement follows from definition 1.5 (in [Hus04], chapter 13), and the second follows from the text between definition 1.1 and theorem 1.2. \square

The next theorem describes all possibilities for the endomorphism ring of E .

Theorem 11.2. *The endomorphism ring of an elliptic curve is either isomorphic to \mathbb{Z} , or to a maximal order in a quadratic imaginary field, or to an (not necessarily maximal) order in a quaternion algebra. The first case can only occur in characteristic zero. If the characteristic of the ground field is positive and the endomorphism ring is isomorphic to an order in a quadratic imaginary field, then this field is generated over \mathbb{Q} by Frob.*

Definition 11.3. Let K be a finitely generated algebra over \mathbb{Q} . An order R of K is a subring of K which is finitely generated as a \mathbb{Z} -module and satisfies $R \otimes \mathbb{Q} = K$.

Definition 11.4. A quaternion algebra is an algebra of the form $\mathbb{Q} + \alpha\mathbb{Q} + \beta\mathbb{Q} + \alpha\beta\mathbb{Q}$. Multiplication is generated by the rules $\alpha^2, \beta^2 \in \mathbb{Q}_{<0}$ and $\beta\alpha = -\alpha\beta$.

Proof of theorem 11.2. In [Sil86], chapter III, corollary 9.4, it is shown that the endomorphism ring is isomorphic to either \mathbb{Z} , an order in a quadratic imaginary field or an order in a quaternion algebra. Since Frob is quadratic imaginary by theorem 11.1, the endomorphism ring is strictly larger than \mathbb{Z} if the characteristic is positive. If the characteristic of the ground field is positive and the endomorphism ring is an order in a quadratic imaginary field, then this is an order in a quadratic imaginary field containing Frob. Hence this quadratic imaginary field is generated over \mathbb{Q} by Frob. By [Hus04], chapter 13, table 2 in section 7, if the endomorphism ring is an order in a quadratic imaginary field, then it is a maximal order. \square

Definition 11.5. An elliptic curve over \mathbb{F}_p is *supersingular* if its endomorphism ring is isomorphic to an order in a quaternion algebra. Otherwise it is *ordinary*.

Theorem 11.6. (*Theorem 4.1(c) in [Sil86]*) *Up to isomorphism, there are $\lfloor \frac{p}{12} \rfloor + \varepsilon_p$ supersingular curves, with $\varepsilon_2 = 0$, $\varepsilon_3 = 1$, and for $p \geq 5$,*

$$\varepsilon_p = \begin{cases} 0 & \text{if } p \equiv 1 \pmod{12} \\ 1 & \text{if } p \equiv 5, 7 \pmod{12} \\ 2 & \text{if } p \equiv 11 \pmod{12} \end{cases}$$

Lemma 11.7. *There exists an ordinary curve E given by an equation of the form $s^2 = t^3 + \delta t^2 + t$ for some $\delta \in \mathbb{F}_p$.*

Proof. The curve given by $s^2 = t^3 + \delta t^2 + t$ is singular if $t^2 + \delta t + 1$ has a double zero. This happens exactly if $\delta^2 - 4 = 0$, i.e. if $\delta = \pm 2$. Hence there are $p - 2$ possibilities for δ such that E is non-singular.

The j -invariant of E is $j(\delta) = \frac{4^4(\delta^2 - 3)^3}{\delta^2 - 4}$. This implies that given a value j , there are at most 6 values for δ with $j(\delta) = j$. Since two curves are isomorphic if and only if they have the same j -invariant, the number of isomorphism classes of non-singular elliptic curves is at least $\frac{p-2}{6}$. Hence there exists an ordinary curve if the number of isomorphism classes of supersingular curves is strictly smaller than $\frac{p-2}{6}$. By theorem 11.6, there are $\lfloor \frac{p}{12} \rfloor + \varepsilon_p$ isomorphism classes of supersingular curves. If $p \equiv 1 \pmod{12}$, then $\lfloor \frac{p}{12} \rfloor + \varepsilon_p = \frac{p-1}{12}$, which is smaller than $\frac{p-2}{6}$ for all $p > 3$. If $p \equiv 5 \pmod{12}$, then $\lfloor \frac{p}{12} \rfloor + \varepsilon_p = \frac{p-5}{12} + 1$, which is smaller than $\frac{p-2}{6}$ for all $p > 11$. If $p \equiv 7 \pmod{12}$, then $\lfloor \frac{p}{12} \rfloor + \varepsilon_p = \frac{p-7}{12} + 1$, which is smaller than $\frac{p-2}{6}$ for all $p > 9$. If $p \equiv 11 \pmod{12}$, then $\lfloor \frac{p}{12} \rfloor + \varepsilon_p = \frac{p-11}{12} + 2$, which is smaller than $\frac{p-2}{6}$ for all $p > 15$. Hence there exists an ordinary curve for all $p \neq 2, 3, 5, 7, 11$.

For $p = 2, 3$, there indeed doesn't exist an ordinary curve, which is the reason why we only consider $p \geq 5$. For $p = 5$, an easy computation shows that $j(0) = 3$ and $j(\pm 1) = 1$, so there are two non-isomorphic curves, one of which is supersingular. Hence the other one is ordinary. Similarly, for $p = 7$ we have $j(0) = 6, j(\pm 1) = 4$ and $j(\pm 3) = 2$. Of the three non-isomorphic curves, one is supersingular, so there are two ordinary curves. Finally, if $p = 11$ then $j(0) = 1, j(\pm 1) = 2, j(\pm 3) = 0, j(\pm 4) = 6$ and $j(\pm 5) = 0$, so there are 4 non-isomorphic curves. Two of these are ordinary.

Hence in all cases with $p \geq 5$, there exists a $\delta \in \mathbb{F}_p$ such that E is ordinary. \square

Lemma 11.8. $\#E(\mathbb{F}_p) = p + 1$ if and only if E is supersingular.

Proof. See [Hus04], chapter 13, section 7, table 2. \square

From now on, we fix an ordinary curve $E : s^2 = t^3 + \delta t^2 + t$. If the characteristic is 5, then we take $E : s^2 = t^3 + t$. This curve is indeed ordinary, since $E(\mathbb{F}_5) = \{\mathcal{O}, (0, 0), (2, 0), (3, 0)\}$, which has 4 elements. In this case, π satisfies the equation $\pi^2 - 2\pi + 5 = 0$ by theorem 11.1. In other characteristics, E can be any ordinary curve.

The endomorphism ring of E is isomorphic to an order R in the imaginary quadratic field generated over \mathbb{Q} by the Frobenius endomorphism. Fix an isomorphism $\text{End}(E) \rightarrow R$. Under this isomorphism, Frob corresponds to an element $\pi \in R$ satisfying $\pi^2 - t\pi + p = 0$.

As in the case with characteristic zero, all endomorphisms are of the form $(x(t), sy(t))$ with $x, y \in \overline{K}(t)$ and $(x, y) \in E_t(\overline{K}(t, s))$, where $E_t(\overline{K}(t, s))$ is given by $(t^3 + \delta t^2 + t)y^2 = x^3 + \delta x^2 + x$ (see lemma 9.4). Denote the endomorphism corresponding to $n \in R \setminus \{0\}$ by $e_n = (x_n, sy_n)$. Then in particular

$$(x_\pi(t), sy_\pi(t)) = \text{Frob}(t, s) = (t^p, s^p) = (t^p, s(t^3 + \delta t^2 + t)^{\frac{p-1}{2}})$$

Furthermore, if $n \in \mathbb{Z}$, then $e_n(P) = nP$ for all $P \in E(K)$, and $0 \in R$ corresponds to $\mathcal{O} \in E_t(K(t))$.

Define $\mathbb{I} = \mathbb{Z} + \pi\mathbb{Z}$. Since π is quadratic, this is a subring of R .

11.2 A positive existential model of $(\mathbb{I}, L_{\mathbb{I}})$ in $(K(t), (0, 1, t; +, \cdot))$

In this section, we will define several relations on \mathbb{I} , and show that they correspond to positive existential relations on $K(t)$. Hence we get a positive existential model of \mathbb{I} in some language $L_{\mathbb{I}}$ into $(K(t), (0, 1, t; +, \cdot))$. In the next sections, we will consider whether it is possible to prove that the positive existential theory of \mathbb{I} in $L_{\mathbb{I}}$ is undecidable. If we would succeed in this, it would follow that the positive existential theory of $K(t)$ is also undecidable, and hence we would have solved Hilbert's tenth problem for $K(t)$ with coefficients in $\mathbb{Z}[t]$.

We start by defining the language $L_{\mathbb{I}}$.

Definition 11.9. Define the language $L_{\mathbb{I}} = (\{0, 1, \pi\}; +, \pi*; |_{\pi})$, where $\pi*$ and $|_{\pi}$ are defined by

$$\begin{aligned} x = \pi * y &\iff x = \pi y \\ x |_{\pi} y &\iff \exists n \in \mathbb{N} : y = \pm x \pi^n \end{aligned}$$

Recall that we fixed an isomorphism $\text{End}(E) \rightarrow R$, such that $(x_n, sy_n) \mapsto n$ for all $n \neq 0$. The map we will use in constructing the model of $(\mathbb{I}, L_{\mathbb{I}})$ in $(K(t), (0, 1, t; +, \cdot))$ is

$$\phi_{\mathbb{I}} : \mathbb{I} \rightarrow E_t(K(t)) : n \mapsto \begin{cases} \mathcal{O} & \text{if } n = 0 \\ (x_n, y_n) & \text{if } n \neq 0 \end{cases}$$

Of course, $\phi_{\mathbb{I}}$ can easily be extended to a map $R \rightarrow E_t(K(t))$.

Now we first prove that $\phi_{\mathbb{I}}$ indeed defines a positive existential model. Let $\mathbb{P}^2(K(t))$ denote the projective plane over $K(t)$. As we also did in the proof of theorem 7.19, where we used an equivalence relation on $K(t)$, we can 'translate' positive existential formulas over $\mathbb{P}^2(K(t))$ into positive existential formulas over $K(t)$, with the extra relation \sim . It is clear that \sim is a positive existential relation over $K(t)$, so relations that are positive existential over $\mathbb{P}^2(K(t))$ correspond to relations that are positive existential over $K(t)$. Hence to prove that a relation is positive existential over $K(t)$, it suffices to show that it is positive existential over $\mathbb{P}^2(K(t))$.

Lemma 11.10. *The set $\phi_{\mathbb{I}}(2R)$ is equal to*

$$\begin{aligned} \{[x : y : 1] \in \mathbb{P}^2(K(t)) \mid \exists z, w \in K(t) : (t^3 + at^2 + bt + c)w^2 = z^3 + az^2 + bz + c \wedge (x, sy) = 2(z, sw)\} \cup \\ \{[0 : 1 : 0]\} \end{aligned}$$

and hence is positive existential over $K(t)$ in $(0, 1, t; +, \cdot)$.

Proof. To simplify the notation, we will write (x, y) instead of $[x : y : 1]$.

Notice that all non-constant endomorphisms of E are of the form (x, sy) with $x, y \in K(t)$ and $(x, y) \in E_t(K(t))$. Every endomorphism corresponds to some $n \in R$, and if $n \neq 0$, then it is mapped to (x, y) by $\phi_{\mathbb{I}}$.

Suppose that $(x, y) \in \phi_{\mathbb{I}}(2R)$. Then $e = (x, sy)$ is an endomorphism in $2\text{End}(E) \setminus \{\text{Id}\}$. Hence there exists an endomorphism $f \in \text{End}(E) \setminus \{\text{Id}\}$ such that $e = 2f$. By lemma 9.4, there exist $z, w \in K(t)$ such that $e = (x, sy)$, $f = (z, sw)$ and $(z, w) \in E_t(K(t))$, i.e. $(x, sy) = 2(z, sw)$ and $(t^3 + at^2 + bt + c)w^2 = z^3 + az^2 + bz + c$.

Conversely, suppose that $x, y, z, w \in K(t)$ satisfy $(t^3 + at^2 + bt + c)w^2 = z^3 + az^2 + bz + c$ and $(x, sy) = 2(z, sw)$. Then $(z, w) \in E_t(K(t))$, so by lemma 9.6, there exists an endomorphism e and a point $P \in E(K)$ of order dividing two such that $(z, sw) = e + P$. Hence

$$(x, sy) = 2(z, sw) = 2\phi = 2(e + P) = 2e$$

so $(x, sy) \in 2\text{End}(E)$. Hence by the isomorphism $\text{End}(E) \rightarrow R$, (x, sy) is mapped to an element of $2R$, so $\phi_{\mathbb{I}}$ maps it back to (x, y) .

By replacing the condition $(x, sy) = 2(z, sw)$ by the duplication formula, we get a positive existential definition of $\phi_{\mathbb{I}}(2R \setminus \{0\})$.

Finally, $0 \in 2R$ corresponds to $[0 : 1 : 0]$. \square

Lemma 11.11. $\phi_{\mathbb{I}}(\mathbb{I})$ is positive existentially definable over $K(t)$ in $(0, 1, t; +, \cdot)$.

Proof. R is an order in the quadratic imaginary field $\mathbb{Q}(\pi)$, so there exist $\alpha, \beta \in \mathbb{Q}(\pi)$ such that $R = \alpha\mathbb{Z} + \beta\mathbb{Z}$. Choose $n \in \mathbb{N}_{>0}$ such that n is even and $n\alpha, n\beta \in \mathbb{Z}[\pi]$. Since $\mathbb{I} = \mathbb{Z}[\pi]$, we have $nR \subseteq \mathbb{I}$. Since R is a ring, it is closed under multiplication. \mathbb{I} is a subring of R , so in particular, R is closed under multiplication by \mathbb{I} . Hence nR is also closed under multiplication by \mathbb{I} , so it is an ideal of \mathbb{I} .

Write $\alpha' = n\alpha = \alpha_1 + \alpha_2\pi$ and $\beta' = n\beta = \beta_1 + \beta_2\pi$ with $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{Z}$. Define $\gamma_1 = \gcd(\alpha_1, \beta_1)$. Then there exist $k, l \in \mathbb{Z}$ such that $k\alpha_1 + l\beta_1 = \gamma_1$. Define $\gamma = \gamma_1 + (k\alpha_2 + l\beta_2)\pi$. As an ideal, nR is generated by (α', β') , so it is also generated by $(\alpha', l\beta' + k\alpha')$ $= (\alpha', \gamma)$. Since $\gamma_1 | \alpha'_1$, there exists m such that $\alpha'_1 = m\gamma_1$. Define $\delta = \alpha - m\gamma$. Then there exists $\delta_2 \in \mathbb{Z}$ such that $\delta = \delta_2\pi$. It follows that nR is also generated by γ and $\delta = \delta_2\pi$.

Now take any $\varepsilon = \varepsilon_1 + \varepsilon_2\pi \in \mathbb{I}$. By repeatedly adding or subtracting γ , we can find $\varepsilon' \in \mathbb{I}$ with $0 \leq \varepsilon'_1 < |\gamma_1|$ and $\varepsilon \equiv \varepsilon' \pmod{nR}$. By adding or subtracting δ , we can find ε'' with $\varepsilon'_1 = \varepsilon''_1$ and $0 \leq \varepsilon''_2 < |\delta_2|$ and $\varepsilon \equiv \varepsilon'' \pmod{nR}$. Hence modulo nR , every element of \mathbb{I} is equal to some ε'' with $0 \leq \varepsilon''_1 < |\gamma_1|$ and $0 \leq \varepsilon''_2 < |\delta_2|$. It follows that \mathbb{I}/nR is finite.

Choose a finite set $\{k_1, \dots, k_r\} \subseteq \mathbb{I}$ that represents all elements of \mathbb{I}/nR . Then $\mathbb{I} = \bigcup_{i=1}^r k_i + nR$. It follows that

$$\phi_{\mathbb{I}}(\mathbb{I}) = \bigcup_{i=1}^r \phi_{\mathbb{I}}(k_i) + \phi_{\mathbb{I}}(nR)$$

Hence it suffices to show that $\phi_{\mathbb{I}}(k_i) + \phi_{\mathbb{I}}(nR)$ is positive existential (for all i). By lemma 11.10, the set $2R$ is positive existential over $K(t)$ in $(0, 1, t; +, \cdot)$. Since n is even, we have the following equivalences:

$$(x, y) \in \phi_{\mathbb{I}}(nR) \iff \exists m \in nR : (x, sy) = e_m \iff \exists m' \in 2R : (x, sy) = e_{\frac{n}{2}} \circ e_{m'}$$

Notice that

$$(e_{\frac{n}{2}} \circ e_{m'})(t, s) = e_{\frac{n}{2}}(x_{m'}(t), sy_{m'}(t)) = (x_{\frac{n}{2}}(x_{m'}(t)), sy_{m'}(t)y_{\frac{n}{2}}(x_{m'}(t)))$$

Hence

$$(x, y) \in \phi_{\mathbb{I}}(nR) \iff \exists (z, w) \in \phi_{\mathbb{I}}(nR) : x = x_{\frac{n}{2}} \circ z \wedge y = w \cdot (y_{\frac{n}{2}} \circ x_{m'})$$

Since we can compute $x_{\frac{n}{2}}$ and $y_{\frac{n}{2}}$, this gives a positive existential definition of $\phi_{\mathbb{I}}(nR)$. Finally,

$$(x, y) \in \phi_{\mathbb{I}}(k_i) + \phi_{\mathbb{I}}(nR) \iff \exists (z, w) \in \phi_{\mathbb{I}}(nR) : (x, y) = (x_{k_i}, y_{k_i}) + (z, w)$$

where $+$ denotes addition on $E_t(K(t))$. Replacing the $+$ by the addition formula gives a positive existential definition of $\phi_{\mathbb{I}}(k_i) + \phi_{\mathbb{I}}(nR)$, and hence of $\phi_{\mathbb{I}}(\mathbb{I})$. \square

Lemma 11.12. $\phi_{\mathbb{I}}$ gives a positive existential model of $(\mathbb{I}, L_{\mathbb{I}})$ in $(K(t), (0, 1, t; +, \cdot))$.

Proof. $\phi_{\mathbb{I}}(\mathbb{I})$ is positive existential by lemma 11.11.

+ in \mathbb{I} corresponds to addition on $E_t(K(t))$, which is positive existentially definable over $K(t)$.

Furthermore, the Frobenius endomorphism is given by $\text{Frob} : (t, s) \mapsto (t^p, s^p)$. Hence

$$(x_{\pi n}(t), sy_{\pi n}(t)) = \text{Frob}(x_n(t), sy_n(t)) = (x_n(t)^p, s^p y_n(t)^p) = (x_n(t)^p, s(t^3 + \delta t^2 + t)^{\frac{p-1}{2}} y_n(t)^p)$$

This implies that

$$\begin{aligned} \{((x_n, y_n), (x_m, y_m)) \mid n, m \in \mathbb{I}, m = \pi n\} = \\ \{((x, y), (z, w)) \mid (x, y), (z, w) \in \phi_{\mathbb{I}}(\mathbb{I}) \wedge z = x^p \wedge w = (t^3 + \delta t^2 + t)^{\frac{p-1}{2}} y^p\} \end{aligned}$$

which is a positive existential set.

Finally, notice that

$$(x_{n\pi^r}(t), sy_{n\pi^r}(t)) = e_{n\pi^r}(t, s) = e_{\pi^r}(e_n(t, s)) = \text{Frob}^r(x_n(t, s), y_n(t, s)) = (x_n(t, s)^{p^r}, y_n(t, s)^{p^r})$$

Hence

$$\begin{aligned} \{((x_n, y_n), (x_m, y_m)) \mid n, m \in \mathbb{I}, n \mid_{\pi} m\} = \\ \{((x, y), (z, w)) \mid (x, y), (z, w) \in \phi_{\mathbb{I}}(\mathbb{I}) \wedge \exists r \in \mathbb{N} : z = x^{p^r}\} \end{aligned}$$

By lemma 7.18, the relation $\exists r \in \mathbb{N} : z = x^{p^r}$ is positive existentially definable over $K(t)$ in $(0, 1, t; +, \cdot)$, so this set is positive existential.

This shows that $\phi_{\mathbb{I}}$ gives a positive existential model of $(\mathbb{I}, L_{\mathbb{I}})$ in $(K(t), (0, 1, t; +, \cdot))$. \square

11.3 A positive existential model of $(\mathbb{N}, (0, 1; +; |))$ in $(\mathbb{I}, L_{\mathbb{I}})$

$\phi_{\mathbb{I}}$ gives a positive existential model of $(\mathbb{I}, L_{\mathbb{I}})$ in $(K(t), (0, 1, t; +, \cdot))$. Hence to prove that the diophantine theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable, it suffices to show that the positive existential theory of \mathbb{I} is undecidable in $L_{\mathbb{I}}$. We will try to do this by giving a model of $(\mathbb{N}, (0, 1; +, |))$ in $(\mathbb{I}, L_{\mathbb{I}})$.

In this section, we embed $(\mathbb{N}, (0, 1; +; |))$ in $(\mathbb{I}, L_{\mathbb{I}})$ and show that this gives a positive existential model. In the next section, we will define multiplication over \mathbb{N} in $(0, 1; +; |)$ by a first-order formula, and show that the first-order theory of \mathbb{I} is undecidable. Then we will consider whether $(\mathbb{N}, (0, 1; +; |))$ can be extended to a pair of a set and a language such that the positive existential theory is undecidable.

Lemma 11.13. 2 is not divisible by $\pi - 1$ in \mathbb{I} .

Proof. Suppose that $\pi - 1$ divides 2 . Then there exist $a, b \in \mathbb{Z}$ such that $(a + b\pi)(\pi - 1) = 2$. Since $\pi^2 - t\pi + p = 0$, this gives $-(a + b\pi) + (a + b\pi - b)\pi = 2$, so $a + b\pi = -2$ and $a + b\pi - b = 0$. Solving this gives $a = -b\pi - 2$ and $b(-p + t - 1) = 2$. Hence $-p + t - 1$ divides 2 in \mathbb{Z} , so $-p + t - 1$ equals ± 1 or ± 2 . It follows that $t \in \{p - 1, p, p + 2, p + 3\}$. Since $|t| \leq 2\sqrt{p} < p$ for $p \geq 5$, this is only possible if $t = p - 1$ and $p - 1 \leq 2\sqrt{p}$. But this gives $0 \leq p \leq 3 + 2\sqrt{2} < 6$, and for $p = 5$ we had chosen $t = 2$, which clearly doesn't satisfy $t = p - 1$. Hence a and b don't exist, so $\pi - 1$ doesn't divide 2 in \mathbb{I} . \square

Lemma 11.14. For every fixed $x \in \mathbb{I}$, the relation $y \equiv z \pmod{x}$ is positive existential over \mathbb{I} in $L_{\mathbb{I}}$.

Proof. Write $x = x_1 + x_2\pi$ with $x_1, x_2 \in \mathbb{Z}$. Then multiplication by x is definable by

$$a = xb \iff a = x_1b + x_2\pi b \iff a = \overbrace{(b + \dots + b)}^{x_1} + \overbrace{\pi * b + \dots + \pi * b}^{x_2}$$

Now $y \equiv z \pmod{x}$ is definable by $\exists w \in \mathbb{I} : y - z = xw$. \square

Now we can define the model of $(\mathbb{N}, (0, 1; +; |))$ in $(\mathbb{I}, L_{\mathbb{I}})$:

Definition 11.15. Define $\phi_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{I} : n \mapsto \pi^n$.

Lemma 11.16. $\phi_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{I}$ is injective.

Proof. If $\phi_{\mathbb{N}}$ is not injective, then there exist $k > l$ such that $\phi_{\mathbb{N}}(k) = \phi_{\mathbb{N}}(l)$. Hence to prove that $\phi_{\mathbb{N}}$ is injective, it suffices to show that there exists no $n \neq 0$ such that $\pi^n = 1$. Notice that the norm of π is p . Hence $N(\pi^n) = p^n$, which can only equal 1 if $n = 0$. \square

Lemma 11.17. $\phi_{\mathbb{N}}(\mathbb{N})$ is a positive existential set over \mathbb{I} in $L_{\mathbb{I}}$.

Proof. Notice that $1|_{\pi}x$ if and only if there exists $n \in \mathbb{N}$ such that $x = \pm\pi^n$. By lemma 11.13, $\pi - 1$ is not a divisor of 2. Hence $1 \not\equiv -1 \pmod{\pi - 1}$. Since $\pi^n \equiv 1 \pmod{\pi - 1}$ and $-\pi^n \equiv -1 \pmod{\pi - 1}$, we have the equivalences

$$\phi_{\mathbb{N}}(\mathbb{N}) = \{x \in \mathbb{I} \mid \exists n \in \mathbb{N} : x = \pi^n\} = \{x \in \mathbb{I} \mid 1|_{\pi}x \wedge x \equiv 1 \pmod{\pi - 1}\}$$

This is a positive existential set by lemma 11.14. \square

To show that addition in \mathbb{N} corresponds to a positive existential relation in \mathbb{I} , we first show that multiplication by powers of π is given by a positive existential formula.

Definition 11.18. Let $x, y, z \in \mathbb{I}$. Define the formula $\psi(x, y, z)$ by

$$\begin{aligned} \psi(x, y, z) \equiv & ((\pi - 1)\pi x + 1) |_{\pi} ((\pi - 1)\pi z + y) \wedge \\ & ((\pi - 1)((\pi - 1)\pi x + 1) - 1) |_{\pi} ((\pi - 1)((\pi - 1)\pi z + y) - y) \end{aligned}$$

Lemma 11.19. If there exists $n \in \mathbb{N}$ such that $y = \pi^n$, then $\psi(x, y, z)$ is equivalent to $z = xy$.

Proof. Assume that $y = \pi^n$. Write

$$\begin{aligned} X &= (\pi - 1)\pi x + 1 \\ Z &= (\pi - 1)\pi z + y \\ \tilde{X} &= (\pi - 1)((\pi - 1)\pi x + 1) - 1 = (\pi - 1)X - 1 \\ \tilde{Z} &= (\pi - 1)((\pi - 1)\pi z + y) - y = (\pi - 1)Z - y \end{aligned}$$

Then $\psi(x, y, z) \equiv X|_{\pi}Z \wedge \tilde{X}|_{\pi}\tilde{Z}$.

First suppose that $z = xy = x\pi^n$. Then

$$Z = (\pi - 1)\pi z + y = \pi^n((\pi - 1)\pi x + 1) = \pi^n X$$

so $X|_{\pi}Z$, and

$$\tilde{Z} = (\pi - 1)Z - y = \pi^n((\pi - 1)X - 1) = \pi^n \tilde{X}$$

Hence $\tilde{X}|_{\pi}\tilde{Z}$, so $\psi(x, y, z)$ holds.

For the converse direction, suppose that $\psi(x, y, z)$ holds. Then there exist $k, l \in \mathbb{N}$ such that $Z = \pm\pi^k X$ and $\tilde{Z} = \pm\pi^l \tilde{X}$. Notice that modulo $\pi - 1$, $X \equiv 1$, $Z \equiv y \equiv \pi^n \equiv 1^n \equiv 1$, $\tilde{X} \equiv -1$ and $\tilde{Z} \equiv -y \equiv -1$. Substituting this into $Z = \pm\pi^k X$ and $\tilde{Z} = \pm\pi^l \tilde{X}$ gives that $1 \equiv \pm\pi^k \equiv \pm 1$ and $-1 \equiv \mp\pi^l \equiv \mp 1 \pmod{\pi - 1}$. Since $\pi - 1$ doesn't divide 2 (lemma 11.13), this implies that $Z = \pi^k X$ and $\tilde{Z} = \pi^l \tilde{X}$.

From

$$\pi^l \tilde{X} = (\pi - 1)\pi^l X - \pi^l \tag{11.1a}$$

$$\tilde{Z} = (\pi - 1)Z - \pi^n \tag{11.1b}$$

it follows that

$$\begin{aligned} (\pi - 1)(\pi^l - \pi^k)X &= (\pi - 1)\pi^l X - (\pi - 1)\pi^k X \stackrel{(11.1a)}{=} \\ &(\pi^l \tilde{X} + \pi^l) - (\pi - 1)Z \stackrel{(11.1b)}{=} (\tilde{Z} + \pi^l) - (\tilde{Z} + \pi^n) = \pi^l - \pi^n \end{aligned}$$

so

$$(\pi - 1)(\pi^l - \pi^k)X = \pi^l - \pi^n$$

Suppose that k, l and n are all different. Since $\pi - 1 \equiv -1$ and $X \equiv 1 \pmod{\pi}$, and π doesn't divide 1 (since $N(\pi) = p$), the left-hand side is $\min(k, l)$ times divisible by π , while the right-hand side has $\min(l, n)$ factors π . It follows that $\min(k, l) = \min(l, n)$. Since we assumed k, l and n to be different, this implies that $l < k, n$. Hence

$$(\pi - 1)(1 - \pi^{k-l})X = 1 - \pi^{n-l}$$

Modulo π , this equation reduces to $-1 \equiv 1$, which is impossible by lemma 11.13.

Hence at least two of k, l and n are equal. If $k = l$, then $\pi^l - \pi^n = 0$, which implies $k = l = n$ by lemma 11.16. If $l = n$, then $(\pi - 1)X(\pi^l - \pi^k) = 0$. Since $\pi \neq 1$ and $X \equiv 1 \not\equiv 0 \pmod{\pi}$, this implies that $\pi^l - \pi^k = 0$, so again $k = l = n$. So in all cases, we have $k = n$, and hence $Z = \pi^k X = \pi^n X = yX$. Hence

$$z = \frac{Z - y}{(\pi - 1)\pi} = \frac{X - 1}{(\pi - 1)\pi}y = xy$$

□

Lemma 11.20. *In the embedding given by $\phi_{\mathbb{N}}$, addition in \mathbb{N} is positive existentially definable in $(\mathbb{I}, L_{\mathbb{I}})$.*

Proof. This follows from lemma 11.19, since

$$m = k + l \iff \pi^m = \pi^k \cdot \pi^l \iff \psi(\pi^k, \pi^l, \pi^m) \iff \psi(\phi_{\mathbb{N}}(k), \phi_{\mathbb{N}}(l), \phi_{\mathbb{N}}(m))$$

□

Lemma 11.21. *Let $n, m \in \mathbb{N}$. Then n divides m in \mathbb{N} if and only if $\pi^n - 1$ divides $\pi^m - 1$ in \mathbb{I} .*

Proof. Suppose that $m = kn$. Then

$$\pi^m - 1 = \pi^{kn} - 1 = (\pi^n - 1)(\pi^{(k-1)n} + \pi^{(k-2)n} + \dots + \pi^n + 1)$$

so $\pi^n - 1$ divides $\pi^m - 1$.

For the converse direction, suppose that $\pi^n - 1$ divides $\pi^m - 1$. Let $q = \gcd(n, m)$. Then there exist $k, l \in \mathbb{Z}$ such that $km - ln = q$.

Suppose that $k, l \geq 0$. By the converse direction, we have $\pi^n - 1 \mid \pi^{ln} - 1$ and $\pi^m - 1 \mid \pi^{km} - 1$. From $\pi^n - 1 \mid \pi^m - 1$ it follows that $\pi^n - 1 \mid \pi^{km} - 1$ and hence

$$\pi^n - 1 \mid ((\pi^{km} - 1) - (\pi^{ln} - 1) - (\pi^{ln} - 1)(\pi^q - 1))$$

Notice that

$$\begin{aligned} (\pi^{km} - 1) - (\pi^{ln} - 1) - (\pi^{ln} - 1)(\pi^q - 1) &= \\ (\pi^{km} - 1) - (\pi^{ln} - 1) - (\pi^{q+ln} - \pi^{ln} - \pi^q + 1) &= \\ \pi^{km} - \pi^{km} + \pi^q - 1 &= \pi^q - 1 \end{aligned}$$

Hence $\pi^n - 1 \mid \pi^q - 1$, so there exists $\alpha \in \mathbb{I}$ such that $\pi^q - 1 = \alpha(\pi^n - 1)$. Since $q \mid n$, we also have $\pi^q - 1 \mid \pi^n - 1$, so there exists $\beta \in \mathbb{I}$ such that $\pi^n - 1 = \beta(\pi^q - 1)$. It follows that $\pi^n - 1 = \beta\alpha(\pi^n - 1)$,

so either $\pi^n - 1 = 0$ or $\beta\alpha = 1$. If $\pi^n - 1 = 0$, then $\pi^m - 1 = 0$, and hence $n = m = 0$ so $n|m$. If $\beta\alpha = 1$, then α is a unit in \mathbb{I} , so $N(\alpha)$ is a unit in \mathbb{Z} . Since the norm equals the square of the complex norm by lemma 6.1(ii), the norms of all elements of \mathbb{I} are positive. Hence $N(\alpha) = 1$, so

$$|\pi^n - 1|^2 = N(\pi^n - 1) = N(\pi^q - 1) = |\pi^q - 1|^2$$

Write $n = dq$ with $d \geq 1$. Since all norms are positive,

$$p^{dq} - 1 = p^n - 1 = |\pi^n| - 1 \leq |\pi^n - 1| = |\pi^q - 1| \leq |\pi^q| + 1 = p^q + 1$$

so $p^{dq} - 1 \geq p^q + 1$. But if $d > 2$, then $p^{dq} - 1 \geq (p^q)^2 - 1 \geq 5(p^q) - 1 > p^q + 1$ because $p \geq 5$ and $q > 0$. Contradiction, so $d = 1$. Hence $q = n$, so $n|m$.

Now suppose that at least one of k and l is negative. Then the other one is non-positive, since otherwise $|km - ln| \geq m + n > q$. This implies that $(-k)m - (-l)n = -q$, with $-k, -l \geq 0$. Using $(\pi^{-ln} - 1) - (\pi^{-km} - 1) - (\pi^{-km} - 1)(\pi^q - 1)$ instead of $(\pi^{km} - 1) - (\pi^{ln} - 1) - (\pi^{ln} - 1)(\pi^q - 1)$, we again get $\pi^n - 1 | \pi^q - 1$, and the same proof applies. \square

Lemma 11.22. *In the embedding given by $\phi_{\mathbb{N}}$, divisibility in \mathbb{N} is positive existentially definable in $(\mathbb{I}, L_{\mathbb{I}})$.*

Proof. By lemma 11.21, $n|m$ in \mathbb{N} if and only if $\phi_{\mathbb{N}}(n) - 1 | \phi_{\mathbb{N}}(m) - 1$ in \mathbb{I} . Hence, using lemma 11.19,

$$\begin{aligned} n|m \text{ (in } \mathbb{N}) &\iff \phi_{\mathbb{N}}(n) - 1 | \phi_{\mathbb{N}}(m) - 1 \iff \\ &\exists x \in \mathbb{I} : \phi_{\mathbb{N}}(m) - 1 = x(\phi_{\mathbb{N}}(n) - 1) \iff \\ &\exists x, y \in \mathbb{I} : \phi_{\mathbb{N}}(m) + x = y + 1 \wedge \psi(x, \phi_{\mathbb{N}}(n), y) \end{aligned}$$

\square

Corollary 11.23. *$\phi_{\mathbb{N}}$ gives a positive existential model of $(\mathbb{N}, (0, 1; +; |))$ in $(\mathbb{I}, L_{\mathbb{I}})$.*

Proof. This follows immediately from lemmas 11.16, 11.17, 11.20 and 11.22. \square

11.4 Defining multiplication in \mathbb{N} in $(0, 1; +; |)$

We have shown that there exist positive existential models of $(\mathbb{N}, (0, 1; +; |))$ in $(\mathbb{I}, L_{\mathbb{I}})$ and of $(\mathbb{I}, L_{\mathbb{I}})$ in $(K(t), (0, 1, t; +, \cdot))$. Unfortunately, this doesn't show that the diophantine theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable. In this section, we will define multiplication in \mathbb{N} by a formula in the language $(0, 1; +; |)$ containing a universal quantifier. Although it doesn't prove that the diophantine theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable, this allows us to show that the first order theory is undecidable.

If we could define multiplication in \mathbb{N} by a positive existential formula in $(0, 1; +; |)$, then it would follow that the diophantine theory of $K(t)$ is undecidable. Unfortunately, as far as known, no such formula exists. Therefore, we will allow the use of universal quantifiers, but we will try to give a formula that is 'as positive existentially as possible'. More precisely, we will still only allow the connectives \wedge and \vee , and use only one universal quantifier. This means that the formula has a low complexity. The complexity of a formula can formally be measured by its place in the positive arithmetical hierarchy:

Definition 11.24. The *positive arithmetical hierarchy* (Σ^+, Π^+) of a language L is defined as follows: $\Sigma_0^+ = \Pi_0^+$ consists of the formulas without quantifiers, with the connectives \wedge and \vee (and without \neg and \Rightarrow). Inductively, Σ_{n+1}^+ contains all formulas of the form $\exists x_1, \dots, x_n : \phi(x_1, \dots, x_n)$, with $\phi(x_1, \dots, x_n) \in \Pi_n^+$, and Π_{n+1}^+ contains all formulas of the form $\forall x_1, \dots, x_n : \phi(x_1, \dots, x_n)$, with $\phi(x_1, \dots, x_n) \in \Sigma_n^+$.

Notice that Σ_1^+ is precisely the set of positive existential formulas. If a formula is at a low level of the positive arithmetical hierarchy, the number of quantifier changes is low. The formula we will define is at level Σ_3^+ , so it is of the form $\exists \vec{x} \forall \vec{y} \exists \vec{z} : \phi(\vec{x}, \vec{y}, \vec{z})$ where ϕ contains no quantifiers. It is even possible to take a single variable for \vec{y} , but of course this doesn't reduce the number of quantifier changes.

Definition 11.25. Define the formula χ by

$$\chi(x, y, t) \equiv x|y \wedge x+1|y+x \wedge ((x|t \wedge x+1|t+x) \Rightarrow y+x|t+x)$$

Lemma 11.26. For all $x, y \in \mathbb{N}$, the formula $\forall t \in \mathbb{N} : \chi(x, y, t)$ holds if and only if $y = x^2$.

Proof. Suppose that $y = x^2$. Then $x|y$ and $x+1|y+x$ clearly hold. Suppose that t satisfies $x|t$ and $x+1|t+x$. Then also $x|t+x$, and since $\gcd(x, x+1) = 1$, this implies $x(x+1)|t+x$. Since $x(x+1) = y+x$, $\chi(x, y, t)$ holds for all $t \in \mathbb{N}$.

Conversely, suppose that $\chi(x, y, t)$ holds for all t . Then $x|y$, so there exists $z \in \mathbb{N}$ such that $y = xz$. $x+1|y+x$ so $x+1|(z+1)x$. Since $\gcd(x, x+1) = 1$, we either have $x = 0$, in which case $y = 0$, so $x = y^2$, or $x+1|z+1$. Now take $t = x^2$. Then $x+t$ and $x+1|t+x$, so $y+x|t+x$, i.e. $(z+1)x|(x+1)x$. If $x \neq 0$, this implies $z+1|x+1$. Hence $z+1 = x+1$ and $z = x$, so $y = x^2$. \square

Lemma 11.27. In the language $(0, 1; +; |)$, there exists a formula $\Xi(k, l, m) \in \Sigma_3^+$ of $(0, 1; +; |)$ that is over \mathbb{N} equivalent to $m = kl$.

Proof. First we give a positive existential definition of \dagger . Notice that the relation \neq is positive existential, since $a \neq b$ is equivalent to $\exists x \in \mathbb{N} : a = b + x + 1 \vee b = a + x + 1$. In chapter 2, we gave a definition of the greatest common divisor. Although we used the language $(0, 1; +, \cdot)$ there, the definition can easily be changed into a definition in $(0, 1; +; |)$:

$$a = \gcd(b, c) \iff a|b \wedge a|c \wedge \exists x, y : a|x \wedge b|y \wedge (a = x - y \vee a = y - x)$$

Now we can simply define $a \dagger b$ by $\exists g : g = \gcd(a, b) \wedge g \neq a$.

Using this definition of \dagger , we define a formula $\Phi(x, y) \in \Pi_2^+$ that is equivalent to $y = x^2$. We cannot use $\forall t : \chi(x, y, t)$ for this, since it contains the connective \Rightarrow . However, we can change $(x|t \wedge x+1|t+x) \Rightarrow y+x|t+x$ into $x \dagger t \vee x+1 \dagger t+x \vee y+x|t+x$. Since $x \dagger t$ and $x+1 \dagger t+x$ are given by a positive existential formula, $\chi(x, y, t)$ is equivalent to a positive existential formula $\chi'(x, y, t)$. Hence $\Phi(x, y) \equiv \forall t : \chi'(x, y, t)$ is a formula in Π_2^+ that is equivalent to $y = x^2$.

Finally, we can define multiplication by

$$m = kl \iff \exists a, b, c, d : \Phi(k, a) \wedge \Phi(l, b) \wedge c = k + l \wedge \Phi(c, d) \wedge d = a + d + m + m$$

so multiplication is definable by a formula in Σ_3^+ . \square

Definition 11.28. Let S and L be a set and a language as in definition 1.2. A *first-order formula* over S in L is inductively defined to be an equality between terms, a relation symbol applied to terms, a negation of a first-order formula (by using the connective \neg), a conjunction, disjunction or implication of two first-order formulas (by using the connectives \wedge , \vee and \Rightarrow , respectively), or a formula of the form $\exists x : \phi(x)$ or $\forall x : \phi(x)$, where ϕ is a first-order formula. The definitions of *first-order sentence* and the *first-order theory* are similar to definition 1.2, with 'diophantine' replaced by 'first-order'.

Theorem 11.29. The first order theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable.

Proof. Let a positive existential sentence over \mathbb{N} in $(0, 1; +, \cdot)$ be given. Using the above definition of multiplication in $(0, 1; +; |)$, there exists an equivalent Σ_3^+ -sentence over \mathbb{N} in $(0, 1; +; |)$. This formula is of the form

$$\exists x_1, \dots, x_k \in \mathbb{N}, \forall y_1, \dots, y_l \in \mathbb{N} : \theta(x_1, \dots, x_k, y_1, \dots, y_l)$$

where θ is a positive existential formula in $(0, 1; +; |)$. Now use the embedding $\phi_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{I}$ to obtain an equivalent formula of the form

$$\exists x'_1, \dots, x'_k \in \phi_{\mathbb{N}}(\mathbb{N}), \forall y'_1, \dots, y'_l \in \phi_{\mathbb{N}}(\mathbb{N}) : \theta'(x'_1, \dots, x'_k, y'_1, \dots, y'_l)$$

θ' is a positive existential formula over \mathbb{I} in $L_{\mathbb{I}}$, but we still have to take care of the sets over which the quantifiers run. As in the case with positive existential models, we can replace $\exists x'_i \in \phi_{\mathbb{N}}(\mathbb{N})$ by $\exists x'_i \in \mathbb{I} : x'_i \in \phi_{\mathbb{N}}(\mathbb{N})$, and this is a positive existential statement. Furthermore, we can replace $\forall y'_i \in \phi_{\mathbb{N}}(\mathbb{N}) : \theta'$ by $\forall y'_i \in \mathbb{I} : y'_i \in \phi_{\mathbb{N}}(\mathbb{N}) \Rightarrow \theta'$. Hence the above formula is equivalent to a first-order formula over \mathbb{I} in $L_{\mathbb{I}}$. Repeating this reasoning, we obtain an equivalent first-order formula over $K(t)$ in $(0, 1, t; +, \cdot)$.

Hence if there exists an algorithm to decide the first-order theory of $K(t)$ in $(0, 1, t; +, \cdot)$, then we can use it to decide the positive existential theory of \mathbb{N} in $(0, 1; +, \cdot)$. But this is undecidable, so the first-order theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is also undecidable. \square

11.5 Other approaches

Using the approach explained above, we were only able to prove that the first-order theory of $K(t)$ is undecidable. It is not known whether the diophantine theory is undecidable. We will now consider how we can modify the proof in such a way that we might be able to prove that the diophantine theory is undecidable. There are at least three ways to do this: instead of \mathbb{N} and the language $(0, 1; +; |)$, use another set or language such that the positive existential theory is undecidable, use another embedding $\mathbb{N} \rightarrow \mathbb{I}$ or use a supersingular curve. For each of these approaches, we will now consider whether they will possibly give a proof of the undecidability of the diophantine theory.

We start by considering other languages, still using the embedding $n \mapsto \pi^n$. In chapters 2 and 3, we showed that the diophantine theory of both \mathbb{N} and \mathbb{Z} in $(0, 1; +, \cdot)$ is undecidable. Furthermore, by the results in chapter 4, the positive existential theories of \mathbb{Z} in $(0, 1; +; |, |_p)$ and $(0, 1; +; |^p)$ and of \mathbb{N} in $(0, 1; +; |_p)$ are undecidable.

First consider the sets \mathbb{N} and \mathbb{Z} . We have shown already that $\{\pi^n \mid n \in \mathbb{N}\}$ is positive existential over \mathbb{I} in $L_{\mathbb{I}}$, but $\{\pi^n \mid n \in \mathbb{Z}\}$ is not even a subset of \mathbb{I} .

Now we consider the relations $+, \cdot, |, |_p$ and $|^p$. By lemmas 11.20 and 11.21, addition and divisibility in \mathbb{N} are definable in \mathbb{I} .

The relation $|_p$ on \mathbb{N} is defined by $n|_p m \iff \exists k \in \mathbb{N} : m = \pm np^k$. Although on \mathbb{N} , this is equivalent to $m = \exists k \in \mathbb{N} : m = np^k$, we prefer to include the possibility of having a minus sign, since this will be easier in the next part (because $x_n = x_m$ if and only if $n = \pm m$). In \mathbb{I} , $|_p$ gives the set

$$\{(\pi^n, \pi^m) \mid \exists k \in \mathbb{N} : \pi^m = \pi^n p^k\} = \{(\pi^n, \pi^m) \mid \exists k \in \mathbb{N} : \pi^m = (\pi^n)^{p^k}\}$$

The easiest way to show that this set is positive existential is probably to show that the relation $a \lambda_p b \equiv \exists k \in \mathbb{N} : b = a^{p^k}$ is positive existential for all $a, b \in \mathbb{I}$ (so not only for $a, b \in \phi_{\mathbb{N}}(\mathbb{N})$). Unfortunately, we don't know how to define this relation.

Instead of showing that this relation is positive existential in \mathbb{I} , we can also add it to $L_{\mathbb{I}}$. Then we have to show that $\phi_{\mathbb{I}}(\lambda_p)$ is a positive existential set over $K(t)$ in $(0, 1, t; +, \cdot)$. But we have $a \lambda_p b$ if and only if there exists $k \in \mathbb{N}$ such that $x_b = x_{a^{p^k}}$ and $y_b = y_{a^{p^k}}$. I don't know how to express $x_{a^{p^k}}$ and $y_{a^{p^k}}$ in terms of x_a and y_a .

If we restrict λ_p to $\phi_{\mathbb{N}}(\mathbb{N})$, then

$$\begin{aligned} \phi_{\mathbb{I}}(\lambda_p) &= \{((x_{\pi^n}, y_{\pi^n}), (x_{\pi^m}, y_{\pi^m})) \mid n, m \in \mathbb{N} \wedge n|_p m\} = \\ &= \{((x_{\pi^n}, y_{\pi^n}), (x_{\pi^m}, y_{\pi^m})) \mid n, m \in \mathbb{N} \wedge \exists k \in \mathbb{N} : m = \pm np^k\} = \\ &= \{((x_{\pi^n}, y_{\pi^n}), (x_{\pi^m}, y_{\pi^m})) \mid n, m \in \mathbb{N} \wedge \exists k \in \mathbb{N} : x_{\pi^m} = x_{\pi^n p^k}\} \end{aligned}$$

Since $x_{\pi^a}(t) = t^{p^a}$ for all $a \in \mathbb{N}$, the above set is equal to

$$\{(x, y), (z, w) \mid (x, y), (z, w) \in \phi_{\mathbb{I}}(\phi_{\mathbb{N}}(\mathbb{N})) \wedge \exists k, n, m \in \mathbb{N} : x = t^{p^n} \wedge z = t^{p^m} \wedge t^{p^m} = t^{p^{np^k}}\}$$

Although the set $\{t^{p^n} \mid n \in \mathbb{N}\}$ is positive existential over $K(t)$ (by lemma 7.18), the above set is probably not positive existential since we will have to find an expression for $t^{p^m} = t^{p^{np^k}}$ in terms of t^{p^n} and t^{p^m} .

The relation $|^p$ on \mathbb{N} is defined by $n|^p m \iff \exists q \in \mathbb{Z}, \exists r \in \mathbb{N} : mp^r = nq$. In \mathbb{I} this gives the set

$$\{(\pi^n, \pi^m) \mid \exists q, r \in \mathbb{N} : \pi^{mp^r} = \pi^{nq}\} = \{(\pi^n, \pi^m) \mid \exists q, r \in \mathbb{N} : (\pi^m)^{p^r} = (\pi^n)^q\}$$

Since we again don't know how to define this relation in \mathbb{I} , we can again add it to $L_{\mathbb{I}}$ and try to show that the induced relation on $K(t)$ is positive existential. So denote the relation $\exists q, r \in \mathbb{N} : b^{p^r} = a^q$ by $a|^p b$. After applying the map $\phi_{\mathbb{I}}$, this gives $a|^p b$ if and only if $\exists q, r \in \mathbb{N} : x_{b^{p^r}} = x_{a^q}$, which we again don't know how to define.

If we restrict $|^p$ to $\phi_{\mathbb{N}}(\mathbb{N})$, then over $K(t)$ we get the set

$$\begin{aligned} \phi_{\mathbb{I}}(|^p) &= \{((x_{\pi^n}, y_{\pi^n}), (x_{\pi^m}, y_{\pi^m})) \mid n, m \in \mathbb{N} \wedge \exists q, r \in \mathbb{N} : mp^r = \pm nq\} = \\ &\quad \{((x_{\pi^n}, y_{\pi^n}), (x_{\pi^m}, y_{\pi^m})) \mid n, m \in \mathbb{N} \wedge \exists q, r \in \mathbb{N} : x_{\pi^{mp^r}} = x_{\pi^{nq}}\} = \\ \{(x, y), (z, w) \mid (x, y), (z, w) \in \phi_{\mathbb{I}}(\phi_{\mathbb{N}}(\mathbb{N})) \wedge \exists n, m, q, r \in \mathbb{N} : x = t^{p^n} \wedge z = t^{p^m} \wedge t^{p^{mp^r}} = t^{p^{nq}}\} \end{aligned}$$

Here we have the same problems as with $|_p$.

Now we will investigate what happens if we try to use another embedding $\mathbb{N} \rightarrow \mathbb{I}$. Instead of $\phi_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{I} : n \mapsto \pi^n$, we simply use $n \rightarrow n$. We will consider the same languages as we did above.

In [Phe04], Pheidas tries to embed $(\mathbb{Z}; (0, 1; +, \cdot))$ into $(\mathbb{I}, L_{\mathbb{I}})$. First he shows that multiplication in \mathbb{I} can be defined by a first-order formula in $L_{\mathbb{I}}$. Then he refers to [Den75] for a proof that \mathbb{Z} is positive existential over \mathbb{I} in $(0, 1; +, \cdot)$. Hence \mathbb{Z} is definable by a first-order formula over \mathbb{I} in $L_{\mathbb{I}}$. We proved this already in chapter 6 (theorem 6.14). Since \mathbb{N} is positive existential over \mathbb{Z} in $(0, 1; +, \cdot)$ (by using $n \in \mathbb{N} \iff \exists a, b, c, d \in \mathbb{Z} : n = a^2 + b^2 + c^2 + d^2$), \mathbb{N} can also be defined by a first-order formula over \mathbb{I} in $L_{\mathbb{I}}$. No results are known about definability with positive existential formulas.

For the sake of completeness, we include Pheidas' proof that multiplication can be defined by a first-order formula:

Lemma 11.30. *Let $\chi_{\mathbb{Z}}(x, y)$ be the formula*

$$\forall r \in \mathbb{N} \exists q \in \mathbb{N} : r|q \wedge x \equiv \frac{\pi^q - 1}{\pi^r - 1} \pmod{\pi^r - 1} \wedge y \equiv \left(\frac{\pi^q - 1}{\pi^r - 1}\right)^2 \pmod{\pi^r - 1}$$

- (i) *If $x, y \in \mathbb{I}$ and $\chi_{\mathbb{Z}}(x, y)$ holds, then $y = x^2$.*
- (ii) *If $x, y \in \mathbb{Z}$ and $y = x^2$, then $\chi_{\mathbb{Z}}(x, y)$ holds.*

Proof. (i) Suppose that $x, y \in \mathbb{I}$ and $\chi_{\mathbb{Z}}(x, y)$ holds. Since $|\pi^r - 1| \geq p^r - 1$, the norm of $\pi^r - 1$ tends to infinity when r tends to infinity. Hence there exists $r \in \mathbb{N}$ such that $|\pi^r - 1| > |y - x^2|$. Since $\chi_{\mathbb{Z}}(x, y)$ holds, there exists $q \in \mathbb{N}$ such that

$$r|q \wedge x \equiv \frac{\pi^q - 1}{\pi^r - 1} \pmod{\pi^r - 1} \wedge y \equiv \left(\frac{\pi^q - 1}{\pi^r - 1}\right)^2 \pmod{\pi^r - 1}$$

This implies that $y \equiv x^2 \pmod{\pi^r - 1}$, so $\pi^r - 1|y - x^2$ in \mathbb{I} . By taking norms, it follows that $N(\pi^r - 1)|N(y - x^2)$ in \mathbb{Z} . Suppose that $y \neq x^2$. Then $N(y - x^2) > 0$, so $N(\pi^r - 1) \geq N(y - x^2)$. However, by lemma 6.1(ii),

$$|\pi^r - 1|^2 = N(\pi^r - 1) \geq N(y - x^2) = |y - x^2|^2$$

This clearly contradicts the choice of r . Hence $y = x^2$.

(ii) Suppose that $x, y \in \mathbb{Z}$ with $y = x^2$. For every $r \in \mathbb{N}$, let $q = rx$. Then clearly $r|q$, and

$$\pi^q - 1 = (\pi^r - 1)(\pi^{(x-1)r} + \dots + \pi^r + 1)$$

Hence $\frac{\pi^q - 1}{\pi^r - 1}$ exists, and equals $\pi^{(x-1)r} + \dots + \pi^r + 1$. Modulo $\pi^r - 1$, all terms are 1, so $x \equiv \frac{\pi^q - 1}{\pi^r - 1} \pmod{\pi^r - 1}$. Since $y = x^2$, $y \equiv \left(\frac{\pi^q - 1}{\pi^r - 1}\right)^2 \pmod{\pi^r - 1}$. \square

Corollary 11.31. *On the subset \mathbb{Z} of \mathbb{I} , multiplication is definable by a first-order formula in $L_{\mathbb{I}}$.*

Proof. In the formula $\chi_{\mathbb{Z}}$, $r|q$ can be replaced by $\pi^r - 1 | \pi^q - 1$ by lemma 11.21. Since the set $\{\pi^n \mid n \in \mathbb{N}\}$ is positive existential, there exists a first-order formula $\chi'_{\mathbb{Z}}(x, y)$ that holds if and only if $y = x^2$ (for all $x, y \in \mathbb{Z}$). Now we can define multiplication on \mathbb{Z} by

$$z = xy \iff \exists u, v, w \in \mathbb{I} : \chi'_{\mathbb{Z}}(x, u) \wedge \chi'_{\mathbb{Z}}(y, v) \wedge \chi'_{\mathbb{Z}}(x + y, w) \wedge z + z + u + v = w$$

because $\chi'_{\mathbb{Z}}(a, b)$ implies $b = a^2$ for all $a, b \in \mathbb{I}$ (not only in \mathbb{Z}). \square

We even have the stronger result that multiplication is definable on \mathbb{I} instead of $\mathbb{Z} \subseteq \mathbb{I}$:

Lemma 11.32. *Let $\Xi_{\mathbb{I}}(x, y)$ be the formula*

$$\exists a, b, c, d, e \in \mathbb{I} : \chi_{\mathbb{Z}}(a, c) \wedge \chi_{\mathbb{Z}}(b, d) \wedge \chi_{\mathbb{Z}}(a + b, e) \wedge x = a + \pi b \wedge y = c - pd + \pi(e - c - d + td)$$

Then for all $x, y \in \mathbb{I}$, $\Xi_{\mathbb{I}}(x, y)$ holds if and only if $y = x^2$.

Proof. Let $x, y \in \mathbb{I}$ and suppose that $\Xi_{\mathbb{I}}(x, y)$ holds. Then there exist $a, b, c, d, e \in \mathbb{I}$ such that $x = a + \pi b$, $c = a^2$, $d = b^2$ and $e = (a + b)^2$. We can compute y from this:

$$y = c - pd + \pi(e - c - d + td) = a^2 - pb^2 + \pi((a + b)^2 - a^2 - b^2 + tb^2) = a^2 - pb^2 + 2ab\pi + tb^2\pi$$

Since $\pi^2 = t\pi - p$, this equals $a^2 + 2ab\pi + b^2\pi^2 = (a + b\pi)^2$, so $y = x^2$.

Conversely, suppose that $y = x^2$. Choose $a, b, c, d, e \in \mathbb{Z}$ such that $x = a + \pi b$, $c = a^2$, $d = b^2$ and $e = (a + b)^2$. Then by lemma 11.30, $\chi_{\mathbb{Z}}(a, c)$, $\chi_{\mathbb{Z}}(b, d)$ and $\chi_{\mathbb{Z}}(a + b, e)$ hold. The same computation as above shows that $y = c - pd + \pi(e - c - d + td)$. \square

Corollary 11.33. *Multiplication in \mathbb{I} is definable by a first-order formula in $L_{\mathbb{I}}$.*

Proof. We can again define multiplication by

$$z = xy \iff \exists u, v, w \in \mathbb{I} : \Xi_{\mathbb{I}}(x, u) \wedge \Xi_{\mathbb{I}}(y, v) \wedge \Xi_{\mathbb{I}}(x + y, w) \wedge z + z + u + v = w$$

\square

Since both \mathbb{Z} and \mathbb{N} are definable by a first-order formula, but not by a positive existential formula (as far as known), in what follows \mathbb{Z} can be replaced by \mathbb{N} everywhere.

It is clear that addition on \mathbb{Z} is definable, since $L_{\mathbb{I}}$ contains $+$.

As proven above, multiplication is definable by a first-order formula, but probably it is not definable by a positive existential formula. We can try to add multiplication to $L_{\mathbb{I}}$ and try to prove that this gives a positive existential relation on $K(t)$. Notice that, if $m = kl$, then

$$e_m(t, s) = e_k(e_l(t, s)) = e_k(x_l(t), sy_l(t)) = (x_k(x_l(t)), sy_l(t)y_k(x_k(t)))$$

Hence we have to prove that the set

$$\begin{aligned} & \{((x_k, y_k), (x_l, y_l), (x_m, y_m)) \mid k, l, m \in \mathbb{Z} \wedge m = kl\} = \\ & \{((x, y), (z, w), (u, v)) \mid (x, y), (z, w), (u, v) \in \phi_{\mathbb{I}}(\phi_{\mathbb{N}}(\mathbb{Z})) \wedge u = x \circ y \wedge v = w \cdot (y \circ x)\} \end{aligned}$$

is positive existential. If this make it easier to prove, then $\phi_{\mathbb{I}}(\phi_{\mathbb{N}}(\mathbb{Z}))$ can be replaced by $\phi_{\mathbb{I}}(\mathbb{I})$. Of both sets, it seems difficult, or even impossible, to prove that they are positive existential.

In \mathbb{I} , the relation $|$ is probably not definable. We can again add it to $L_{\mathbb{I}}$ and try to prove that the corresponding relation on $K(t)$ is positive existential, but then we will need to avoid multiplication: we cannot define $n|m$ by something like $\exists k : m = nk$ or its counterpart in $K(t)$. I have no idea how to do this.

I also don't know how to define the relations $|_p$ and $|^p$ over \mathbb{I} in $L_{\mathbb{I}}$. If we add them to $L_{\mathbb{I}}$, we get the sets

$$\begin{aligned} \{((x_n, y_n), (x_m, y_m)) \mid n, m \in \mathbb{Z} \wedge \exists k \in \mathbb{N} : m = \pm np^k\} = \\ \{((x, y), (z, w)) \mid (x, y), (z, w) \in \phi_{\mathbb{I}}(\phi_{\mathbb{N}}(\mathbb{Z})) \wedge \exists k \in \mathbb{N} : z = x_{p^k} \circ x_n\} \end{aligned}$$

and

$$\begin{aligned} \{((x_n, y_n), (x_m, y_m)) \mid n, m \in \mathbb{Z} \wedge \exists q \in \mathbb{Z}, \exists r \in \mathbb{N} : mp^r = nq\} = \\ \{((x, y), (z, w)) \mid (x, y), (z, w) \in \phi_{\mathbb{I}}(\phi_{\mathbb{N}}(\mathbb{Z})) \wedge \exists q \in \mathbb{Z}, \exists r \in \mathbb{N} : x_{p^r} \circ x_m = x_q \circ x_n\} \end{aligned}$$

respectively. For both sets, I don't know how to show whether they are positive existential over $K(t)$ in $(0, 1, t; +, \cdot)$.

In conclusion, it might be possible to prove that the diophantine theory of $K(t)$ is undecidable by using one of these other languages or embeddings, but I have no idea how to do this.

11.6 Supersingular curves

Now we consider supersingular elliptic curves. By theorem 11.6, these exist in any characteristic. The endomorphism ring is isomorphic to an order in a quaternion algebra. Denote this order again by R . Let Frob again denote the Frobenius endomorphism, and choose another endomorphism I that doesn't commute with Frob.

For example, if the characteristic is $p \equiv 3 \pmod{4}$, then we can take the curve $E : s^2 = t^3 + t$. There is an element $i \in \mathbb{F}_{p^2} \setminus \mathbb{F}_p$ satisfying $i^2 = -1$, and the endomorphism $I : (t, s) \mapsto (-t, is)$ doesn't commute with Frob.

Under the isomorphism $\text{End}(E) \rightarrow R$, Frob corresponds to some element $\pi \in R$, and since E is supersingular, $\pi^2 = -p$. I corresponds to some $i \in R$ satisfying a quadratic equation. Consider the subring $\mathbb{O} = \mathbb{Z} + \pi\mathbb{Z} + i\mathbb{Z} + \pi i\mathbb{Z}$. We will use this ring instead of \mathbb{I} . Notice that \mathbb{O} is not commutative, so we have to be careful when using multiplication and divisibilities. We extend the language $L_{\mathbb{I}}$ by a constant i , and denote it by $L_{\mathbb{O}}$. The map $\phi_{\mathbb{O}} : R \rightarrow E_t(K(t))$ is defined similar to $\phi_{\mathbb{I}}$. We will now prove that $\phi_{\mathbb{O}}$ defines a positive existential model of $(\mathbb{O}, L_{\mathbb{O}})$ in $(K(t), (0, 1, t; +, \cdot))$.

Lemma 11.34. $\phi_{\mathbb{O}}(\mathbb{O})$ is positive existentially definable over $K(t)$ in $(0, 1, t; +, \cdot)$.

Proof. The proof of lemma 11.11 almost remains true. Therefore, we only mention the steps that have to be changed and use the same notation.

R is an order in a quaternion algebra containing π and i . Since quaternion algebras are generated by two quadratic numbers, this algebra has to be equal to $\mathbb{Q}(\pi, i)$. Orders in quaternion algebras are four-dimensional \mathbb{Z} -modules, so there exist $\alpha, \beta, \gamma, \delta \in \mathbb{Q}(\pi, i)$ such that $R = \alpha\mathbb{Z} + \beta\mathbb{Z} + \gamma\mathbb{Z} + \delta\mathbb{Z}$. Define n as in lemma 11.11. By the same reasoning as in the proof of lemma 11.11, nR is an ideal of \mathbb{O} . By repeatedly adding or subtracting generators of nR , we can assume that $\alpha' = \alpha_4\pi i$, $\beta' = \beta_3\pi + \beta_4\pi i$, $\gamma' = \gamma_2\pi + \gamma_3i + \gamma_4\pi i$ and $\delta' = \delta_1 + \delta_2\pi + \delta_3i + \delta_4\pi i$. Hence all

elements of \mathbb{O}/nR can be represented by some ε with $0 \leq \varepsilon_1 < |\delta_1|$, $0 \leq \varepsilon_2 < |\gamma_2|$, $0 \leq \varepsilon_3 < |\beta_3|$ and $0 \leq \varepsilon_4 < |\alpha_4|$. It follows that \mathbb{O}/nR is finite.

Notice that lemma 11.10 still holds, with the same proof. In the same way as in the proof of lemma 11.11, this gives a positive existential definition of $\phi_{\mathbb{O}}(\mathbb{O})$. \square

Lemma 11.35. $\phi_{\mathbb{O}}$ gives a positive existential model of $(\mathbb{O}, L_{\mathbb{O}})$ in $(K(t), (0, 1, t; +, \cdot))$.

Proof. This proof is exactly the same as the proof of lemma 11.12. \square

Now we still have to prove that $\phi_{\mathbb{N}}$ gives an positive existential model of $(\mathbb{N}, (0, 1; +; |))$ in $(\mathbb{O}, L_{\mathbb{O}})$. We again take the embedding $\phi_{\mathbb{N}} : n \mapsto \pi^n$, but now it maps to \mathbb{O} instead of \mathbb{I} . Then we will the same result as we obtained using an ordinary elliptic curve: the first-order theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable.

Lemma 11.36. $\pi - 1$ is not a left-divisor of 2, i.e. there doesn't exists $x \in \mathbb{O}$ such that $(\pi - 1)x = 2$.

Proof. Suppose that x exists, and write $x = x_1 + x_2\pi + x_3i + x_4\pi i$ with $x_1, x_2, x_3, x_4 \in \mathbb{Z}$. Then $2 = (\pi - 1)x = (\pi - 1)(x_1 + x_2\pi + x_3i + x_4\pi i) = -(x_1 + px_2) + (x_1 - x_2)\pi - (x_3 + px_4)i + (x_3 - x_4)\pi i$ so $x_1 + px_2 = -2$ and $x_1 - x_2 = 0$. Hence $x_1 = x_2$ and $(p + 1)x_2 = 2$. Since $p + 1 > 2$ and $x_2 \in \mathbb{Z}$, this is impossible. \square

Lemma 11.37. (i) $\phi_{\mathbb{N}} : \mathbb{N} \rightarrow \mathbb{O}$ is injective.

(ii) $\phi_{\mathbb{N}}(\mathbb{N})$ is a positive existential set over \mathbb{O} in $L_{\mathbb{O}}$.

(iii) Addition in \mathbb{N} corresponds to a positive existential set over \mathbb{O} in $L_{\mathbb{O}}$.

(iv) Division in \mathbb{N} corresponds to a positive existential set over \mathbb{O} in $L_{\mathbb{O}}$.

(v) $\phi_{\mathbb{N}}$ gives a positive existential model of $(\mathbb{N}, (0, 1; +; |))$ in $(\mathbb{O}, L_{\mathbb{O}})$.

Proof. (i) Since $\pi^2 = -p$, π^n equals $(-p)^{\frac{n}{2}}$ if n is even, and $(-p)^{\frac{n-1}{2}}\pi$ if n is odd. Hence $\pi^n = 0$ implies $n = 0$, so $\phi_{\mathbb{N}}$ is injective.

(ii) By lemma 11.36, the proof of lemma 11.17 still holds if we use that convention that $y \equiv z \pmod{x}$ means $\exists w \in \mathbb{O} : y - z = xw$.

(iii) It suffices to show that lemma 11.19 remains true. The only thing that has to be checked is that π is not a left-divisor of 1. But since $\pi(x_1 + x_2\pi + x_3i + x_4\pi i) = -px_2 + \pi x_1 - px_4i + x_3\pi i$, $\pi x = 1$ would imply $px_2 = -1$, which is clearly impossible.

(iv) Now we only have to check that lemma 11.21 remains true. Then the statement follows in the same way as in the proof of lemma 11.22.

The proof of lemma 11.21 has to be adapted a little bit. We have to redo the part of the proof in which we showed that $\pi^n - 1 | \pi^q - 1$ and $\pi^q - 1 | \pi^n - 1$ imply $q = n$. So suppose that $\pi^n - 1 | \pi^q - 1$ and $\pi^q - 1 | \pi^n - 1$. Then there exist $\alpha, \beta \in \mathbb{O}$ such that $\pi^q - 1 = \alpha(\pi^n - 1)$ and $\pi^n - 1 = \beta(\pi^q - 1)$, so $\alpha\beta = 1$. Since $\pi^q - 1$ and $\pi^n - 1$ are elements of $\mathbb{Q}(\pi)$, their quotients α and β are also elements of $\mathbb{Q}(\pi)$, and hence of $\mathbb{Z} + \pi\mathbb{Z}$. The norm of $a + b\pi$ is $a^2 + b^2p$ by lemma 6.1. Hence if $\alpha\beta = 1$, then $N(\alpha) = N(\beta) = 1$. But this implies $\alpha = \beta = \pm 1$, and hence $\pi^q - 1 = \pm(\pi^n - 1)$. If $\pi^q - 1 = -\pi^n + 1$, then $\pi^n + \pi^q = 2$, so $\pi^q | 2$. But by taking norms, this implies $p^q | 4$, which is clearly impossible since $q > 0$. Hence $\pi^q - 1 = \pi^n - 1$ so $q = n$ and $n | m$.

(v) This follows immediately from (i)-(iv). \square

Corollary 11.38. The first-order theory of $K(t)$ in $(0, 1, t; +, \cdot)$ is undecidable.

Proof. The proof is the same as the proof of theorem 11.29. \square

So far, we seem to have won nothing. We still haven't proven that the diophantine theory of $K(t)$ is undecidable. The reason to introduce these supersingular curves is that it might be useful to use such a curve when we consider another embedding and other languages.

First we look what happens if we still use the embedding $n \mapsto \pi^n$, but consider other languages. Similar to the case with an ordinary curve, the image of \mathbb{Z} under this map is not a subset of \mathbb{O} , and as we showed above, addition and divisibility are positive existential in $L_{\mathbb{O}}$, whereas multiplication is given by a first-order formula. For the other relations $|_p$ and $|^p$, we get the same sets of which we have to show that they are positive existential as we got when using an ordinary curve. Hence here we also cannot give a complete proof that the diophantine theory of $K(t)$ is undecidable.

Now we change the embedding into $n \mapsto n$. The first problem is to define \mathbb{Z} by a positive existential formula. Up to now, the language $L_{\mathbb{O}}$ only contained a function and relation symbol corresponding to π . We first extend it with the function symbol i^* and the relation symbol $|_i$, with i^* and $|_i$ defined similar to π^* and $|\pi$:

Definition 11.39. Define the function symbol i^* and the relation symbol $|_i$ by

$$\begin{aligned} x = i^* y &\iff x = iy \\ x|_i y &\iff \exists n \in \mathbb{N} : y = \pm xi^n \end{aligned}$$

The language $L_{\mathbb{O}}$ is $L_{\mathbb{O}} = (0, 1, \pi, i; +, \pi^*, i^*; |\pi, |_i)$.

Assumption 11.40. Assume that the following holds: under the embedding $\phi_{\mathbb{O}} : \mathbb{O} \rightarrow E_t(K(t))$, the sets

$$\{(\phi_{\mathbb{O}}(x), \phi_{\mathbb{O}}(y)) \mid x, y \in \mathbb{O} \wedge x = i^* y\}$$

and

$$\{(\phi_{\mathbb{O}}(x), \phi_{\mathbb{O}}(y)) \mid x, y \in \mathbb{O} \wedge x|_i y\}$$

are positive existential over $K(t)$ in $(0, 1, t; +, \cdot)$.

We already gave the example of the curve $E : s^2 = t^3 + t$ in characteristic $p \equiv 3 \pmod{4}$, with $I : E \rightarrow E : (t, s) \mapsto (-t, is)$ where $i^2 = -1$. In this example, we have

$$\{(\phi_{\mathbb{O}}(x), \phi_{\mathbb{O}}(y)) \mid x, y \in \mathbb{O} \wedge x = i^* y\} = \{((x, y), (z, w)) \mid (x, y), (z, w) \in \phi_{\mathbb{O}}(\mathbb{O}) \wedge z = -w \wedge w = is\}$$

This is clearly positive existential over $K(i)(t)$, and since $K(i)$ is positive existential over K (because $K(i)$ is isomorphic to K^2 , with the relations $(a, b) + (c, d) = (a+c, b+d)$ and $(a, b) \cdot (c, d) = (ac-bd, ad+bc)$), this gives a positive existential definition over $K(t)$. Furthermore, since $i^2 = -1$, we have $i^4 = 1$ and hence $x|_i y$ if and only if $x = y, x = iy, x = -y$ or $x = -iy$. This is clearly positive existential. Hence it is possible to fulfill assumption 11.40, at least in characteristic $p \equiv 3 \pmod{4}$.

Lemma 11.41. *If right-multiplication by π and i are positive existentially definable over \mathbb{O} in $L_{\mathbb{O}}$, then \mathbb{Z} is a positive existential subset of \mathbb{O} .*

Proof. Let $x = x_1 + x_2\pi + x_3i + x_4\pi i \in \mathbb{O}$ with $x_i \in \mathbb{Z}$. Notice that

$$\begin{aligned} \pi x &= -px_2 + x_1\pi - px_4i + x_3\pi i \\ x\pi &= -px_2 + x_1\pi + px_4i - x_3\pi i \\ ix &= -x_3 + x_4\pi + x_1i - x_2\pi i \\ xi &= -x_3 - x_4\pi + x_1i + x_2\pi i \end{aligned}$$

Hence

$$x \in \mathbb{Z} \iff x_2 = x_3 = x_4 = 0 \iff \pi x = x\pi \wedge ix = xi$$

This gives a positive existential definition of \mathbb{Z} in \mathbb{O} . □

Although I don't know how to do it, it might be possible to define right-multiplication by π and i . Of all relations we have seen so far that we couldn't define, I think this is the easiest.

Now we consider the relations $+$, \cdot , $|$ and $|_p$ and $|^p$. Of course, $+$ is already contained in \mathbb{I} , so we don't have to define it anymore. For multiplication, division and $|^p$, the same holds as in the case with an ordinary curve: I don't know how to define these in $L_{\mathbb{O}}$, and their counterpart in $K(t)$ seems to be even harder to define. It might again be possible to give a first-order formula defining multiplication, but a positive existential formula is not known. An extra complication is that \mathbb{O} is not commutative, so it doesn't suffice to define squaring.

The relation $|_p$ remains. This relation is the main reason to use a supersingular curve. We can add $|_p$ to $L_{\mathbb{O}}$, and consider the set

$$\{(\phi_{\mathbb{O}}(x), \phi_{\mathbb{O}}(y)) \mid x, y \in \mathbb{O} \wedge x|_p y\} = \{(\phi_{\mathbb{O}}(x), \phi_{\mathbb{O}}(y)) \mid x, y \in \mathbb{O} \wedge \exists n \in \mathbb{N} : y = \pm x p^n\}$$

Notice that $p = -\pi^2$. Hence this set equals

$$\begin{aligned} \{(\phi_{\mathbb{O}}(x), \phi_{\mathbb{O}}(y)) \mid x, y \in \mathbb{O} \wedge \exists n \in \mathbb{N} : y = \pm x \pi^{2n}\} = \\ \{((x, y), (z, w)) \mid (x, y), (z, w) \in \phi_{\mathbb{O}}(\mathbb{O}) \wedge \exists n \in \mathbb{N} : z = x p^{2n}\} \end{aligned}$$

By replacing p by p^2 in the proof of lemma 7.18 (and all lemmas used in the proof of this lemma), we obtain that this set is diophantine. Hence in this embedding, the relation $|_p$ is positive existential. Although this is not enough to prove that the diophantine theory of $K(t)$ is diophantine, it might be possible to use this to give a complete proof.

In conclusion, none of the approaches sketched above give a complete proof. Using an ordinary curve, depending on the embedding into \mathbb{I} , either \mathbb{N} , $+$ and $|$ are positive existential, or $+$ is positive existential and \mathbb{Z} and \cdot can be defined by a first-order formula. If we use a supersingular curve, either \mathbb{N} , $+$ and $|$ are positive existential, or the relations $+$ and $|_p$ are positive existential, but the sets \mathbb{Z} and \mathbb{N} are not (as far as known). Since we need to define $(\mathbb{Z}; +, \cdot)$, $(\mathbb{N}; +, \cdot)$, $(\mathbb{Z}; +, |)$, $(\mathbb{Z}; +, |^p)$ or $(\mathbb{N}; +, |_p)$, there is quite much to prove left.

If we use the embedding $n \mapsto \pi^n$, the image of \mathbb{Z} is not a subset of \mathbb{I} or \mathbb{O} , so with this embedding, we will have to use $(\mathbb{N}; +, \cdot)$ or $(\mathbb{N}; +, |_p)$. I think that for both relations \cdot and $|_p$ it is very difficult to show that they correspond to positive existential relations on \mathbb{I} or \mathbb{O} .

Therefore, I think it might be better to use the embedding $n \mapsto n$. However, then in both cases we still have to show that \mathbb{Z} is definable by a positive existential formula. For ordinary curves, the only known way to do this using multiplication, which is not known to be positive existential. For supersingular curves, it suffices to define right-multiplication by π and i , and I think that this might be doable. Then we still have to prove that division is a positive existential relation on \mathbb{O} . This will probably be difficult, but I expect this to be the most promising approach.

Appendix A

Computing ranks using SAGE

In chapter 10, we described an attempt to prove that exponentiation is diophantine, using elliptic divisibility sequences. To find a curve satisfying assumption 10.9, we calculated the rank of some curves, using SAGE. The computation consisted of two parts.

For the first part, we took $|a|, |b| \leq 25$ and computed the rank of E_d for $E : y^2 = x^3 + ax^2 + bx$ for $d = 1, 2, 3, \dots$, until there where 10 d such that E_d has rank at least 2. For most curves, SAGE cannot compute the rank exactly. Then it gives an answer, which it is not guaranteed to be correct. In counting the number of curves with rank 1, I assumed that all answers are correct. In the output, an ‘?’ is added whenever the answer is not certainly correct.

The code used in this computation is:

```
from sage.libs.mwrank.all import mwrank_EllipticCurve
import sys

class AlternatingRange:
    def __init__(self, start, end, incr=1):
        self.start = start
        self.end = end
        self.incr = incr

    def __iter__(self):
        val = self.start

        while val < self.end:
            yield val
            yield -val
            val += self.incr

for a in AlternatingRange(1, 25):
    for b in AlternatingRange(1, 25):
        print "%+3d, %+3d:" % (a, b),

        if sys.stdout.isatty():
            sys.stdout.flush()

    results = []
    not1 = 0
    for d in xrange(1, 100):
        f = d**3 + a * d**2 + b * d
```

```

try:
    el = mwrank_EllipticCurve([0, a * f, 0, b * f**2, 0])
    r = el.rank()
except ArithmeticError:
    r = None

if r == None:
    results.append("Err")
elif el.certain():
    results.append("%d" % r)
else:
    results.append("%d?" % r)

if r != 1:
    not1 += 1

if d >= 20 and not1 >= 10:
    break

results = ' '.join(results)
if d != 20:
    print "%2d/%2d (%s)" % (d - not1, d, results)
else:
    print "%2d (%s)" % (d - not1, results)

sys.stdout.flush()

```

After this computation, I took the curves such that

$$N := \min\{n \in \mathbb{N} \mid \#\{d \leq n \mid E_d \text{ has rank } 1\} \geq 10\}$$

is at least 40. There are nine curves satisfying this, namely the curves with

$$(a, b) \in \{(4, -2), (-4, 6), (8, -8), (-8, 3), (-8, 24), (10, 20), (12, 18), (12, -18), (-16, 12)\}$$

For these curves, the following program computes the number of curves with rank 1 for $1 \leq d < 200$.

```

from sage.libs.mwrank.all import mwrank_EllipticCurve
import sys

for (a, b) in ((4, -2), (-4, 6), (8, -8), (-8, 3), (-8, 24), (10, 20), (12, 18),
(12, -18), (-16, 12)):
    print "%+3d, %+3d:" % (a, b),

    if sys.stdout.isatty():
        sys.stdout.flush()

results = []
not1 = 0
for d in xrange(1, 200):
    f = d**3 + a * d**2 + b * d
    el = mwrank_EllipticCurve([0, a * f, 0, b * f**2, 0])
    r = el.rank()

    if el.certain():

```

```
        results.append("%d" % r)
    else:
        results.append("%d?" % r)

    if r != 1:
        not1 += 1

results = ' '.join(results)
print "%2d/%2d (%s)" % (d - not1, d, results)

sys.stdout.flush()
```

Bibliography

- [Aya92] M. Ayad. Points S-entiers des courbes elliptiques. *Manuscripta Mathematica*, 76 (3-4):305–324, 1992.
- [Beu00] F. Beukers. *Getaltheorie voor beginners*, volume 42. Epsilon Uitgaven, 2000.
- [BtD85] T. Bröcker and T. tom Dieck. *Representations of compact Lie groups*, volume 98 of *Graduate Texts in Mathematics*. Springer-Verlag, 1985.
- [CZ07] G. Cornelissen and K. Zahidi. Elliptic divisibility sequences and undecidable problems about rational points. *Journal für die reine und angewandte Mathematik*, 613:1–33, 2007.
- [Dem07a] J. Demeyer. Recursively enumerable sets of polynomials over a finite field are diophantine. *Inventiones mathematicae*, 170:655–670, 2007.
- [Dem07b] J. Demeyer. Recursively enumerable sets of polynomials over a finite field. *Journal of Algebra*, 310:801–828, 2007.
- [Den75] J. Denef. Hilbert’s tenth problem for quadratic rings. *Proceedings of the American Mathematical Society*, 48(1):214–220, 3 1975.
- [Den78] J. Denef. The diophantine problem for polynomial rings and fields of rational functions. *Transactions of the American Mathematical Society*, 242:391–399, 1978.
- [Den79] J. Denef. The diophantine problem for polynomial rings of positive characteristic. In *Logic Colloquium 78*, volume 97 of *Studies in logic and the foundations of mathematics*, pages 131–145. North-Holland publishing company, 1979.
- [Deu41] M. Deuring. Die Typen der Multiplikatorenringe elliptischer Funktionenkörper. *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 14:197–272, 1941.
- [DMR76] M. Davis, Y. Matiyasevich, and J. Robinson. Hilbert’s tenth problem. Diophantine equations: positive aspects of a negative solution. In *Mathematical developments arising from Hilbert problems*, volume 28 of *Proceeding of the Symposium in pure mathematics of the American Mathematical Society*, pages 323–378. American Mathematical Society, 1976.
- [Eve] G.R. Everest. Elliptic divisibility sequences. www.mth.uea.ac.uk/~h090/EDS.html. School of mathematics, UEA Norwich.
- [Har06] R. Hartshorne. *Algebraic Geometry*, volume 52 of *Graduate Texts in Mathematics*. Springer-Verlag, 2006.
- [Hus04] D. Husemöller. *Elliptic curves*, volume 111 of *Graduate Texts in Mathematics*. Springer-Verlag, 2004.

- [Kna06] A.W. Knapp. *Basic Algebra*. Birkhäuser, 2006.
- [Mat93] Y.V. Matiyasevich. *Hilbert's Tenth Problem*. Foundations of Computing Series. MIT Press, 1993.
- [Mat00] Y.V. Matiyasevich. Hilbert's tenth problem: What was done and what is to be done. In *Hilbert's tenth problem: relations with arithmetic and algebraic geometry*, volume 270 of *Contemporary mathematics*, pages 1–48. American Mathematical Society, 2000.
- [Phe87] T. Pheidas. An undecidability result for power series rings of positive characteristic. ii. *Proceedings of the American Mathematical Society*, 100(3):526–530, July 1987.
- [Phe91] T. Pheidas. Hilbert's tenth problem for fields of rational functions over finite fields. *Inventiones mathematicae*, 103:1–8, 1991.
- [Phe04] T. Pheidas. Endomorphisms of elliptic curves and undecidability in function fields of positive characteristic. *Journal of algebra*, 273:395–411, 2004.
- [Poo03] B. Poonen. Hilbert's tenth problem over rings of number-theoretic interest. In *Arizona Winter School on "Number theory and logic"*, 2003.
- [Pou71] Y. Pourchet. Sur la représentation en somme de carrés des polynômes à une indéterminée sur un corps de nombres algébriques. *Acta Arithmetica*, XIX:89–104, 1971.
- [Sha74] I.R. Shafarevich. *Basic Algebraic Geometry*, volume 213 of *Die Grundlehren der Mathematischen Wissenschaften Band*. Springer-Verlag, 1974.
- [Sil86] J.H. Silverman. *The Arithmetic of Elliptic Curves*, volume 106 of *Graduate Texts in Mathematics*. Springer-Verlag, 1986.
- [ST92] J.H. Silverman and J. Tate. *Rational points on elliptic curves*. Undergraduate texts in Mathematics. Springer-Verlag, 1992.
- [Sud98] T.A. Sudkamp. *Languages and machines*. Addison-Wesley, 1998.
- [Vid94] C.R. Videla. Hilbert's tenth problem for rational function fields in characteristic 2. *Proceedings of the American mathematical society*, 120(1):249–253, 1994.

List of tables

2.1	Some examples of diophantine relations	7
3.1	Some examples of Turing machines	25
10.1	The number of curves with rank 1	91

List of notation

Notation	Description	Page
~ 0	Certain predicate on $K(t)$	Page 82
$ ^n$	Divisibility in $\mathbb{Z}[\frac{1}{n}]$	Page 32
$ _p$	Equality up to powers of p	Page 32
\sim	Equality in 1	Page 40
\sim	Equality in the point at infinity	Page 80
$\dot{\sim}$	Homogenization	Page 80
$A_b(n)$	Matrix defining the recurrence relation α_b	Page 7
α_b	Certain recurrence relation	Page 7
A_n	$\sqrt{a_n}$	Page 87
a_n	Numerator of x -coordinate of point on elliptic curve	Page 85
arem	The absolute value of the remainder on division	Page 7
B	Symbol to represent empty cells	Page 24
B_d	Elliptic divisibility sequence	Page 89
B_n	Denominator of coordinates of point on elliptic curve	Page 85
C	The elliptic curve given by $y^2 = x^3 - 4$	Page 81
C_n	$\frac{c_n}{\sqrt{a_n}}$	Page 87
c_n	Numerator of y -coordinate of point on elliptic curve	Page 85
Code	True if a triple is a positional code	Page 12
Com	True for the y -coordinates of $y^2 = x^3 - 4$	Page 81
Concat	Gives the concatenation of two tuples	Page 12
D	Square-free part of discriminant	Page 47
δ	Transition function of Turing machine	Page 24
ECode	Extended code of a polynomial	Page 20
Elem	Gives an element of a tuple with given code	Page 12
Eq	Implies that two triples are codes for the same tuple	Page 14
Equal	True if two codes are codes for the same tuple	Page 14
E_t	Associated elliptic curve	Page 77
F	Final states of Turing machine	Page 24

Notation	Description	Page
$F[b]$	Function F applied to tuples on base b	Page 16
Frob	Frobenius endomorphism	Page 92
Γ	Tape alphabet	Page 24
h	Function used in proof of theorem 4.5	Page 33
\mathfrak{h}_0	Set of solvable equations	Page 23
\mathbb{I}	Subring of endomorphism ring	Page 94
Imt	Certain predicate on $K(t)$	Page 80
Imt	True for the y -coordinates of the solutions of the Pell equation	Page 40
NotGreater	True if a tuple is element-wise not greater than another tuple	Page 14
\mathbb{O}	Subring of the endomorphism ring	Page 104
π	Frobenius endomorphism	Page 94
π	Quadratic algebraic integer	Page 47
Π^+	Positive arithmetical hierachy	Page 99
PNotGreater	True if the base is prime and a tuple is element-wise not greater than another tuple	Page 14
Prime	True if a number is prime	Page 13
PSmall	True if the base is prime and a tuple is element-wise not greater than a number	Page 14
Q	State of Turing machine	Page 24
q_0	Start state of Turing machine	Page 24
R	Endomorphism ring of an elliptic curve	Page 94
rem	The (positive) remainder on division	Page 7
Σ	Input alphabet	Page 24
Σ^*	Set of strings of elements of Σ	Page 24
\mathcal{S}_1	Set of pairs defined by α_b	Page 7
\mathcal{S}_2	Set of pairs defined by α_b	Page 7
\mathcal{S}_3	Set of triples defined by α_b	Page 8
SCode	Code of a tuple w.r.t a polynomial	Page 20
Σ^+	Positive arithmetical hierachy	Page 99
Small	True if a tuple is element-wise smaller than a number	Page 14
Solution	True if a tuple is a solution of an equation	Page 22
Ξ_b	Matrix defining the recurrence relation α_b	Page 7
x_n	x -coordinate of $n(t, 1)$	Page 80
x_n	Solution of the Pell equation	Page 38, 43, 48
x_n		Page 40
y_n	y -coordinate of $n(t, 1)$	Page 80
y_n	Solution of the Pell equation	Page 38, 43, 48
y_n		Page 40

List of diophantine relations

Relation	Description	Page
$<$		Page 7
(\cdot)		Page 13
~ 0		Page 82
$[\cdot]$		Page 7
\equiv		Page 7
$!$		Page 13
\leq		Page 7
$ $		Page 7
\dagger		Page 7
\sim	Equality in 1	Page 40
$b \geq 2 \wedge \exists n : a = \alpha_b(n)$		Page 7
$b \geq 4 \wedge a = \alpha_b(c)$		Page 8
arem	The absolute value of the remainder on division	Page 7
B_d	Elliptic divisibility sequence	Page 89
Code	True if a triple is a positional code	Page 12
Com	True for the y -coordinates of $y^2 = x^3 - 4$	Page 81
Concat	Gives the concatenation of two tuples	Page 16
Elem	Gives an element of a tuple with given code	Page 12
Eq	Implies that two triples are codes for the same tuple	Page 14
Equal exponentiation	True if two codes are codes for the same tuple	Page 14 Page 7
$F[b]$	Function F applied to tuples on base b	Page 16
gcd		Page 7
Imt	Certain predicate on $K(t)$	Page 80
Imt	True for the y -coordinates of the solutions of the Pell equation	Page 40
max		Page 7

Relation	Description	Page
NotGreater	True if a tuple is element-wise not greater than another tuple	Page 14
PNotGreater	True if the base is prime and a tuple is element-wise not greater than another tuple	Page 14
Prime	True if a number is prime	Page 13
PSmall	True if the base is prime and a tuple is element-wise not greater than a number	Page 14
rem	The (positive) remainder on division	Page 7
SCode	Code of a tuple w.r.t. a polynomial	Page 20
Small	True if a tuple is element-wise smaller than a number	Page 14
Solution	True if a tuple is a solution of an equation	Page 22

Index

- $M_1; M_2$, 25
- \mathfrak{h}_0 , 23
- \mathfrak{h}_1 , 23
- while** M_1 **do** M_2 , 25
- a_n , 85
 - divides c_n^2 , 87
 - is a square, 87
- associated curve, 77
- base, 12
- biquadratic field, 50
- canonical representation, 24
- Cantor numbering, 11
- cipher, 12
- code
 - Cantor numbering, 11
 - of a polynomial, 20
 - of a tuple, 20
 - positional, 12
 - is diophantine, 12
- code parameters, 18
- decidable set, 24
- decidable theory, 2
- differential form, 68
 - invariant, 68
 - is unique, 68
 - normalized, 68
- diophantine
 - equation, 2
 - formula, 2
 - function, 3
 - on tuples, 16
 - relation, 3
 - set
 - \mathfrak{h}_0 , 23
 - \mathfrak{h}_1 , 23
 - is equivalent to positive existential set, 4, 5
 - is given by a polynomial, 3
 - is recursively enumerable, 26
 - with non-diophantine complement, 23
 - subset of S^n , 3
 - theory, 2
- Dirichlet's unit theorem, 50
- divisibility in $\mathbb{Z}[\frac{1}{n}]$, 32
- divisibility sequence, 85
 - odd, 85
 - A_n , 88
 - C_n , 88
 - associated to E and P , 89
 - strong, 85
 - $A_n B_n$, 88
 - B_n , 85
- element parameters, 18
- elliptic curve
 - endomorphism, 78
 - rational map, 78
 - is endomorphism, 78
- endomorphism of elliptic curves, 78
- exponentiation, 7
- extended code, 20
 - determines polynomial, 20
- final state, 24
- first-order
 - formula, 100
 - sentence, 100
 - theory, 100
 - of $K(t)$, 100, 105
- formal additive group, 66
- formal exponential, 69
- formal group, 66
 - additive, 66
 - associated to elliptic curve, 75
 - group associated to, 71
 - homomorphism of, 66
 - has no constant term, 67
 - inverse, 67
 - isomorphism of, 66
- formal group law, 66
- formal logarithm, 69
 - is an isomorphism, 69, 72

-
- formally real field, 77
 - format, 19
 - Frobenius endomorphism, 92
 - function on tuples, 16
 - generator of $E_t(K(t))$, 77
 - Hensel's lemma, 73
 - Hilbert's tenth problem, 1, 2
 - is undecidable, 30
 - over $K(t)$, 64, 83
 - over R , 2
 - with coefficients in S , 2
 - over polynomial rings
 - of characteristic 2, 43
 - of characteristic zero, 38
 - of odd characteristic, 41
 - over quadratic rings, 53
 - over rings of integers, 54
 - over the natural numbers, 6
 - imaginary quadratic rings, 50
 - input alphabet, 24
 - language, 2
 - length, 12
 - model, positive existential, 31
 - norm, 47
 - equals complex norm, 47
 - is an integer, 47
 - odd divisibility sequence, 85
 - A_n , 88
 - C_n , 88
 - associated to E and P , 89
 - order, 93
 - order of a function, 55
 - ordering on a field, 82
 - on \mathbb{Q} , 83
 - ordinary elliptic curve, 93
 - Pell equation, 38, 47
 - has solutions, 48
 - solutions of, 39, 48
 - positive arithmetical hierarchy, 99
 - positive definite, 83
 - positive definite function, 83
 - positive existential
 - formula, 3
 - model, 31
 - set, 4
 - is equivalent to diophantine set, 4
 - union and intersection of, 4
 - subset of S^n , 4
 - theory, 3
 - of \mathbb{N} in $(0, 1; +; |_p)$, 35
 - of \mathbb{Z} in $(0, 1; +; |^n)$, 32
 - of \mathbb{Z} in $(0, 1; +; |, |_p)$, 32
 - prime in $K(t)$, 55
 - prime producing polynomial, 14
 - quaternion algebra, 93
 - rational map of elliptic curves, 78
 - is endomorphism, 78
 - real quadratic rings, 49
 - recurrence relation, 7
 - recursive, 24
 - recursively enumerable, 24
 - set is diophantine, 26, 27
 - right-multiplication, 106
 - roots of unity, 50
 - start state, 24
 - strong divisibility sequence, 85
 - $A_n B_n$, 88
 - B_n , 85
 - supersingular elliptic curve, 93
 - tape alphabet, 24
 - transition function, 24
 - Turing machine, 24
 - accepted by, 24
 - examples of, 25
 - halts, 24
 - twist, 77
 - universal diophantine set, 18
 - universal equation, 18
 - exists, 18
 - with n element parameters, 19
 - with one code parameter, 18
 - valuation, 70
 - p -adic, 71

