Hilbert’s tenth problem

Massop, E.

Erik Massop

Hilbert's tenth problem

Bachelor thesis, June 12, 2012

Thesis supervisors:
dr. H.J. Hoogeboom
dr. R.M. van Luijk

Leiden Institute of Advanced Computer Science (LIACS)
Mathematisch Instituut

Universiteit Leiden

Contents

1 Introduction

In the year 1900 the famous German mathematician David Hilbert proposed a list of 23 problems. The tenth problem on this list is a computability problem dealing with the solvability of Diophantine equations (equalities of two polynomials with integral coefficients). Hilbert stated his challenge thus:

> Eine Diophantische Gleichung mit irgend welchen Unbekannten und mit ganzen rationalen Zahlencoefficienten sei vorgelegt: *man soll ein Verfahren angeben, nach welchem sich mittelst einer endlichen Anzahl von Operationen entscheiden läßt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist.* [5, page 276]

That is, the challenge is to design an algorithm that given a Diophantine equation with arbitrary unknowns and integral coefficients determines whether it has an integral solution. It has been shown in 1970 that such an algorithm does not exist. In this bachelor thesis we will present a proof of this.

For the proof we will first show that Hilbert's tenth problem (over the integers) is equivalent to the same problem over the *non-negative* integers (Theorem 3.5). Then we will define so-called *Diophantine sets* (Definition 4.10), which are those subsets of $\mathbb{N}^n$ that are defined by an expression of the form

$$\begin{aligned}&(\exists Y_1 \in \mathbb{N}) \cdots (\exists Y_m \in \mathbb{N})\\ &\quad f(X_1, \ldots, X_n, Y_1, \ldots, Y_m) = g(X_1, \ldots, X_n, Y_1, \ldots, Y_m),\end{aligned} \tag{$\dagger$}$$

with f and g polynomials with non-negative integral coefficients. It is not hard to see that if Hilbert's tenth problem is solvable, we can compute in finitely many steps whether or not a tuple $(x_1, \ldots, x_n) \in \mathbb{N}^n$ is a member of a particular Diophantine set $S \subseteq \mathbb{N}^n$ (Proposition 4.27). That is, if Hilbert's tenth problem is solvable, then any Diophantine set is a so-called *recursive set* (Definition 2.1). However, we will see that there are Diophantine sets that are not recursive (Theorem 7.1). More precisely, we will see, with some intermediate steps, that the class of Diophantine sets coincides with the class of so-called *recursively enumerable sets* (Definition 2.2), which is a strict superclass of the class of recursive sets (Theorem 2.5).

Figure 1.1 shows some classes of sets and relations between them. The sets that we have not yet come across are defined using expressions of the form (†), with some modifications. The difference between the expressions for Diophantine sets and exponential Diophantine sets is that for exponential Diophantine sets exponentiation is allowed in f and g in addition to the sums and products with which polynomials with non-negative integral coefficients are constructed. For polynomial sets, D-sets, and Davis sets the difference is in the quantifiers. For polynomial sets the quantifier sequence is empty, while for D-sets some of the quantifiers are allowed to be bounded universal quantifiers $(\forall \cdot \leq \cdot)$. Davis sets, finally, have a quantifier sequence of the form $(\exists \cdot)(\forall \cdot \leq \cdot)(\exists \cdot) \cdots (\exists \cdot)$.

Note that Appendix D lists some of the used notation. In particular $\mathbb{N}$ is the set of non-negative integers $\mathbb{N} = \{0, 1, 2, \ldots\}$, while $\mathbb{Z}^+$ is the set of positive integers $\mathbb{Z}^+ = \{1, 2, 3, \ldots\}$. We use the notation "$x$ rem y" for the non-negative remainder $x - \lfloor x/y \rfloor y$ of the division x/y.

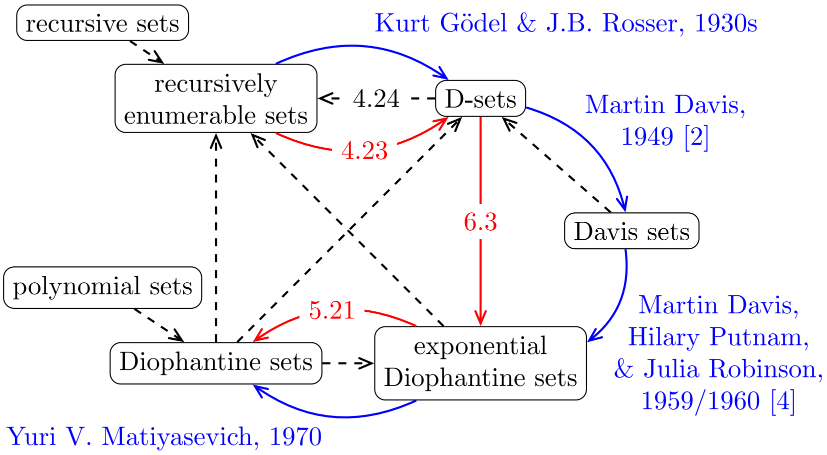
recursive sets
recursively enumerable sets
D-sets
Davis sets
polynomial sets
Diophantine sets
exponential Diophantine sets
4.24
4.23
6.3
5.21
Kurt Gödel & J.B. Rosser, 1930s
Martin Davis, 1949 [2]
Martin Davis, Hilary Putnam, & Julia Robinson, 1959/1960 [4]
Yuri V. Matiyasevich, 1970

Figure 1.1: Inclusions of certain classes. The numbers on the arrows refer to where the corresponding inclusions are proven. The inclusions displayed using dashed arrows are easily proven from the definitions. Of the inclusions displayed with solid arrows, the ones on the outside (in blue, with textual annotations) were originally used to prove that every recursively enumerable set is Diophantine. The ones on the inside (in red) are the ones that we will use. Notice that we prove the inclusions in a different order than was done historically.

2 Recursive and recursively enumerable sets

In this section we will briefly look at recursive sets and recursively enumerable sets. We will give informal definitions (Definitions 2.1 and 2.2) and state the well-known theorem that there are recursively enumerable sets that are not recursive (Theorem 2.5).

These informal definitions depend on some intuitive notion of what an algorithm is. It is important that any algorithm can be written down, that it consists of ‘elementary’ steps, and that there is an algorithm that can execute algorithms given their ‘source code’. (A non-elementary step would for instance be “determine whether this Diophantine equation has a solution”.) It is the content of the Church Thesis that any reasonable formalization of an algorithm yields the same recursive and recursively enumerable sets. (For more on the Church Thesis see for instance John Martin’s book [7, Section 9.7] or Yuri V. Matiyasevich’s book [11, Section 5.7].)

Examples of formalizations of algorithms are Turing Machines and partial recursive functions. The former formalization is used by Yuri V. Matiyasevich, who calls the resulting sets Turing decidable and Turing semidecidable respectively [11, Sections 5.6 and 5.5]. The latter formalization is used by Yu. I. Manin, who calls the resulting sets decidable and recursively enumerable respectively [6, Definitions V.4.13 and V.4.1].

Definition 2.1 (Recursive set)**.** Let $n \in \mathbb{N}$ be any non-negative integer. Then a subset S of $\mathbb{N}^n$ is called *recursive* if there is an algorithm taking a vector x from $\mathbb{N}^n$ as input that runs for a finite number of steps and then returns whether or not x is in S.

Definition 2.2 (Recursively enumerable set)**.** Let $n \in \mathbb{N}$ be any non-negative integer. Then a subset S of $\mathbb{N}^n$ is called *recursively enumerable* if there is an algorithm taking a vector x from $\mathbb{N}^n$ as input that returns TRUE after running for a finite number of steps if and only if x is in S. If x is not in S the algorithm is allowed to run indefinitely or to return anything but TRUE.

Corollary 2.3. *Any recursive set is recursively enumerable.*

Corollary 2.4. *Let $n \in \mathbb{N}$ be any non-negative integer. If $S \subseteq \mathbb{N}^n$ is a recursively enumerable set with recursively enumerable complement, then S is recursive.*

Proof. Let A and B be algorithms as in Definition 2.2 for S and $\mathbb{N}^n \setminus S$ respectively. Consider the following algorithm:

```
read a vector $x \in \mathbb{N}^n$
simulate $A$ and $B$ in parallel with input $x$, until either one terminates
if $A$ returned TRUE or $B$ terminated without returning TRUE then
    return TRUE
else
    return FALSE
```

This algorithm terminates after finitely many steps and returns whether or not x is in S. Hence S is recursive (Definition 2.1). □

Theorem 2.5. *There is a recursively enumerable set that is not a recursive set.*

Proof. We use Theorem 11.2 from John Martin's book [7]. This theorem states that there is a recursively enumerable *language* over the alphabet $\{0, 1\}$ that is not a recursive language. A language over a finite alphabet Σ is any subset of Σ^*, where Σ^* is the set of strings over Σ, which includes the empty string. (That is, Σ^* is the free monoid generated by Σ.) Recursive and recursively enumerable languages are defined analogously to recursive and recursively enumerable sets.

To carry the result over to sets in $\mathbb{N}^n$ for some non-negative integer $n \in \mathbb{N}$, notice the following: Let $\phi : X \to Y$ be any 'algorithmically computable' function with X and Y elements of $\{\Sigma^* : \Sigma \text{ finite}\} \cup \{\mathbb{N}^n : n \in \mathbb{N}\}$. Then

$$\begin{aligned} \phi^{-1} : \mathcal{P}(Y) &\to \mathcal{P}(X), \\ S &\mapsto \phi^{-1}(S) = \{x \in X : \phi(x) \in S\} \end{aligned}$$

preserves being recursive and being recursively enumerable.

We see that it is sufficient to give an algorithmically computable bijection $\{0, 1\}^* \to \mathbb{N}$ with algorithmically computable inverse. We give such a function and its inverse by the following procedures: Given a string over $\{0, 1\}$, prepend 1, interpret it as binary number, substract 1. Given a non-negative integer, add 1, write it in binary notation, remove the leading 1. (This bijection corresponds with ordering the strings first by length and then by binary value.) □

Corollary 2.6. *There is a recursively enumerable set of which the complement is not recursively enumerable.*

Proof. This is immediate from Corollary 2.4 and Theorem 2.5. □

3 From $\mathbb{N}$ to $\mathbb{Z}$ and vice versa

In this section we will look at Diophantine equations with integral and non-negative integral coefficients and establish some relationships between the solutions of these two types of equations. Moreover we prove that Hilbert's tenth problem is solvable over the integers if and only if it is solvable over the non-negative integers (Theorem 3.5).

Proposition 3.1. *Let $n \in \mathbb{N}$ be any non-negative integer and let $X_1, \dots, X_n$ different variable symbols. Then, if a polynomial $f \in \mathbb{Z}[X_1, \dots, X_n]$ with integral coefficients is given, there are polynomials $f^+, f^- \in \mathbb{N}[X_1, \dots, X_n]$ with* non-negative *integral coefficients such that we have $f = f^+ - f^-$. Moreover finding these polynomials f^+ and f^- can be done algorithmically (if f is in such a representation that it can be algorithmically converted to a finite sum of monomials).*

Proof. We obtain appropriate f^+ and f^- by writing f as a sum of monomials $f = \sum_M c_M M$ and taking

$$f^+ = \sum_{M|c_M>0} c_M M \qquad \text{and} \qquad f^- = \sum_{M|c_M<0} (-c_M)M. \qquad \square$$

Corollary 3.2. *Let $n \in \mathbb{N}$ be any non-negative integer and let $X_1, \dots, X_n$ be different variable symbols. Let R be a commutative ring (for instance $\mathbb{Z}$) and let S be any subset of R^n. Given polynomials $f, g \in \mathbb{Z}[X_1, \dots, X_n]$ with integral coefficients, there are polynomials $k, \ell \in \mathbb{N}[X_1, \dots, X_n]$ with* non-negative *integral coefficients such that the equations $f(X_1, \dots, X_n) = g(X_1, \dots, X_n)$ and $k(X_1, \dots, X_n) = \ell(X_1, \dots, X_n)$ have the same solution set in S.*

Proof. Take $h = f - g$. Now take $k = h^+$ and $\ell = h^-$ using Proposition 3.1. □

Proposition 3.3. *Let $n \in \mathbb{N}$ be a non-negative integer and let X_i, A_i, B_i, C_i, D_i be variable symbols for i in $\{1, \dots, n\}$. Consider the polynomials $f(X_1, \dots, X_n)$ and $f(A_1^2 + B_1^2 + C_1^2 + D_1^2, \dots, A_n^2 + B_n^2 + C_n^2 + D_n^2)$ with integral coefficients. Now the first polynomial has a zero in $\mathbb{N}^n$ if and only the latter one has a zero in $\mathbb{Z}^{4n}$.*

Proof. Every sum of four squares of integers is a non-negative integer. The converse, namely that every non-negative integer is the sum of four squares of integers, is Lagrange's four squares theorem. The statement follows. □

Proposition 3.4. *Let $n \in \mathbb{N}$ be any non-negative integer, let X_i, Y_i^+, Y_i^-, Z_i for i in $\{1, \dots, n\}$ be different variable symbols and let $f \in \mathbb{Z}[X_1, \dots, X_n]$ be a polynomial with integral coefficients. Then the polynomial*

$$g = f(Y_1^+ - Y_1^-, \dots, Y_n^+ - Y_n^-) \in \mathbb{Z}[Y_1^+, Y_1^-, \dots, Y_n^+, Y_n^-]$$

has a zero in $\mathbb{N}^{2n}$ if and only if f has a zero in $\mathbb{Z}^n$. Similarly,

$$h = \prod_{v \in \{-1,0,1\}^n} f(v_1 Z_1, \dots, v_n Z_n) \in \mathbb{Z}[Z_1, \dots, Z_n]$$

has a zero in $\mathbb{N}^n$ if and only if f has a zero in $\mathbb{Z}^n$.

Proof. The first assertion follows from the observation that every integer is the difference of two non-negative integers and that conversely every difference of non-negative integers is an integer. For the latter assertion let $x \in \mathbb{Z}^n$ be a zero of f. Then $(|x_1|, \ldots, |x_n|) \in \mathbb{N}^n$ is a zero of $f(\operatorname{sgn}(x_1)Z_1, \ldots, \operatorname{sgn}(x_n)Z_n)$ and hence of h. Conversely let $z \in \mathbb{N}^n$ be a zero of h, then by definition of h there is $v \in \{-1, 0, 1\}$ such that $(v_1 z_1, \ldots, v_n z_n) \in \mathbb{Z}^n$ is a zero of f. □

Theorem 3.5. *Hilbert's tenth problem over $\mathbb{Z}$ is solvable if and only if it is solvable over $\mathbb{N}$. That is, there exists an algorithm to compute in finitely many steps whether a Diophantine equation with integral coefficients has an integral solution if and only if there exists such an algorithm determining whether a Diophantine equation with* non-negative *integral coefficients has a* non-negative *integral solution.*

Proof. For the proof from left to right suppose that we have an algorithm for Hilbert's tenth problem over $\mathbb{Z}$. That is, we have an algorithm A that, when given a Diophantine equation with integral coefficients, runs for a finite number of steps and then returns correctly whether the equation has an integral solution. Now consider the following algorithm B:

```
1 read Diophantine equation $f = g$ with unknowns $X_1, \ldots, X_n$ and non-
  negative integral coefficients
2 take polynomial $h_0 = f - g \in \mathbb{Z}[X_1, \ldots, X_n]$ with integral coefficients
3 for $i = 1$ to $n$ do
4     select new variable symbols $A_i, B_i, C_i, D_i$
5     let $h_i$ be $h_{i-1}$ after substitution of $X_i$ by $(A_i^2 + B_i^2 + C_i^2 + D_i^2)$
6 execute algorithm $A$ on Diophantine equation $h_n = 0$ with integral coeffi-
  cients
7 return what algorithm $A$ returned.
```

Notice that algorithm B terminates after a finite number of steps. By Proposition 3.3 algorithm A returns in line 6 whether polynomial h_0 has zeroes in $\mathbb{N}^n$. Since the zeroes of h_0 are the solutions of Diophantine equation $f = g$, we see that the algorithm returns whether $f = g$ has a solution in $\mathbb{N}^n$. Now algorithm B is a solution to Hilbert's tenth problem over $\mathbb{N}$.

For the proof from right to left suppose that algorithm B is a solution to Hilbert's tenth problem over $\mathbb{N}$. Consider the following algorithm A:

```
1 read Diophantine equation $f = g$ with unknowns $X_1, \ldots, X_n$ and integral
  coefficients
2 take polynomial $h_0 = f - g \in \mathbb{Z}[X_1, \ldots, X_n]$ with integral coefficients
3 for $i = 1$ to $n$ do
4     select new variable symbols $Y_i^+, Y_i^-$
5     let $h_i$ be $h_{i-1}$ after substitution of $X_i$ by $(Y_i^+ - Y_i^-)$
6 compute $h_n^+$ and $h_n^-$ as in Proposition 3.1
7 execute algorithm $B$ on Diophantine equation $h_n^+ = h_n^-$ with non-negative
  integral coefficients
8 return what algorithm $B$ returned.
```

By Proposition 3.1 algorithm B returns in line 7 whether polynomial h_n has a zero in $\mathbb{N}^{2n}$. By Proposition 3.4 this corresponds to polynomial h_0 having a zero in $\mathbb{Z}^n$, which in turn corresponds to Diophantine equation $f = g$ having an integral solution. Hence algorithm A returns after finitely many operations whether Diophantine equation $f = g$ with integral coefficients has an integral solution. Consequently it is a solution to Hilbert's tenth problem over $\mathbb{Z}$. □

4 Logical expressions and sets in $\mathbb{N}^n$

In this section we will define polynomial sets, Diophantine sets and D-sets (Definition 4.10). The latter two kinds of sets are essential to our proof that Hilbert's tenth problem is unsolvable (Proposition 4.27). We will define these sets using some languages of logical expressions, which will be studied in Subsections 4.1, 4.2, and 4.5. Then in Subsections 4.3 and 4.4 the related sets and functions will be defined and explored. Finally we look at whether the defined sets are recursive or recursively enumerable (Subsections 4.4 and 4.6).

4.1 Definition of expressions

The expressions that we will formally define shortly, will roughly be of the form

$$M\ f(X_1, \ldots, X_n) = g(X_1, \ldots, X_n),$$

with M a finite sequence of quantifiers, X_1 through X_n variable symbols and f and g polynomials with non-negative integral coefficients. For polynomial equations the sequence M will be empty, so that only the equation of polynomial expressions remains. For Diophantine expressions sequence M will consist of existential quantifiers $(\exists\,\cdot\,)$, while for D-expressions bounded universal quantifiers $(\forall\,\cdot \leq \cdot\,)$ will also be allowed. For example the expression $2 = x^2 + y^2$ will be a polynomial equation, while $(\exists x)\ z = x^2$ will turn out to be a Diophantine expression.

All our variable symbols represent non-negative integers. Consequently, by "$(\exists X)\ (\ldots)$" with X some variable symbol, we mean "there exists a non-negative integer $X \in \mathbb{N}$ such that ..." and by "$(\forall X)$" we mean "for any non-negative integer $X \in \mathbb{N}$ we have ...". Now the first example expression $2 = x^2 + y^2$ is only true if x and y are both 1, since $(1, 1)$ is the only point in $\mathbb{N}^2$ at an Euclidean distance of $\sqrt{2}$ from the origin. The second expression $(\exists x)\ z = x^2$ is true precisely when z is a square (of a non-negative integer).

By the bounded universal quantification "$(\forall X \leq Y)\ (\ldots)$" with X a variable symbol and Y a non-negative integer or a variable symbol different from X, we mean "$(\forall X)\ [(X \leq Y) \longrightarrow (\ldots)]$", or in words, "for any non-negative integer $X \in \mathbb{N}$ not exceeding Y, we have ...". In any concrete expression (except in this paragraph) we will avoid the case $(\forall X \leq X)\ (\ldots)$, whose interpretation is more tricky: Let Φ be an expression and X' some variable symbol which does not occur in Φ. Then the expression $(\forall X \leq X)\ (\Phi)$ is to be understood as $(\forall X' \leq X)\ (\Phi')$ where Φ' is the expression Φ with every free occurrence of X substituted by X'. In particular $(\forall X \leq X)\ (\ldots)$ does *not* have the same meaning as $(\forall X)\ (\ldots)$!

We now turn to the formal definition of the three announced types of expressions in Definitions 4.1 through 4.3. Table 4.1 contains some examples. For those acquainted with Backus-Naur form (or one of its many variants), Alternative Definition A.1 in Appendix A gives a more concise characterisation.

Definition 4.1 (Polynomial equation)**.** Let $n \in \mathbb{N}$ be any non-negative integer, let $X_1, \dots, X_n$ variable symbols and let $f, g \in \mathbb{N}[X_1, \dots, X_n]$ be polynomials in unknowns $X_1, \dots, X_n$ with coefficients in the non-negative integers. Then we call the equation

$$f(X_1, \dots, X_n) = g(X_1, \dots, X_n)$$

a polynomial equation. We consider $X_1, \dots, X_n$ to be the free variables in this expression.

A possible notation for a polynomial with non-negative coefficients is as a sum of monomials, with the monomials written simply as strings of variable symbols prefixed by `1`. In this notation 0 is represented by the empty string, 1 by `1`, XY^2 by `1XYY`, `1YXY`, or `1YYX`, and X^2+2Y^3 by `1XX+1YYY+1YYY`, `1YYY+1XX+1YYY`, or `1YYY+1YYY+1XX`.

Alternatively we can take `0`, `1`, and the variable symbols as initial representations of polynomials and specify that if P and Q are representations of polynomials, then so are $(P+Q)$ and $(P \cdot Q)$. Using this notation we can represent $X^2 + 2Y^3$ by $((\mathtt{X} \cdot \mathtt{X}) + ((\mathtt{1}+\mathtt{1}) \cdot (\mathtt{Y} \cdot \mathtt{Y} \cdot \mathtt{Y})))$, or by a plethora of other strings (if only since we can add $\mathtt{0} \cdot P$ with P any polynomial, multiply by 1, or place brackets differently). This alternative representation has the advantage that it is easily extended to allow polynomials with coefficients in different finitely generated semirings. For instance to let the coefficients come from $\mathbb{Z}$, $\mathbb{N}[\frac{1}{2}]$, or $\mathbb{N}[\pi]$, we can simply add initial representations `-1`, $\frac{1}{2}$, or π respectively.

Moreover, we can easily introduce additional operators. For instance to add an exponentiation operator ^, we can add the following rule: if P and Q are representations of 'polynomials', then so is $(P$^$Q)$. Historically this alternative definition of 'polynomial' is significant: It was shown around 1960 by Martin Davis, Hilary Putnam and Julia Robinson that Hilbert's tenth problem is unsolvable *if we allow exponentiation* [4]. Then in 1970 Yuri V. Matiyasevich resolved Hilbert's normal tenth problem by showing that exponentiation can be expressed without introducing this additional operator (Proposition 5.21).

Definition 4.2 (Diophantine expression)**.** We define Diophantine expressions inductively by the following rules:

1. Any polynomial equation is a Diophantine expression.
2. If Φ is any Diophantine expression and if X is any variable symbol, then $(\exists X)\ \Phi$ is also a Diophantine expression.

Definition 4.3 (D-expression)**.** We define D-expressions analogously to Diophantine expressions with the alternative rule below instead of rule 2:

2′. If Φ is any D-expression and if X and Y are any two variable symbols, then $(\exists X)\ \Phi$ and $(\forall X \leq Y)\ \Phi$ are also D-expressions.

4.2 Conjunction and disjunction

Table 4.1 lists some statements that can be expressed using polynomial equations and Diophantine expressions. We will see in Propositions 4.6 and 4.8 that if two statements can be thus expressed, we can also express the conjunction ($\wedge$, and) and disjunction ($\vee$, or) of these two statements.

Expression	Type	Meaning
$a = b^2$	polynomial	a is the square of b
$ab = 0$	polynomial	a or b is zero
$a + b = 0$	polynomial	a and b are zero
$ab = c$	polynomial	$a \mid c$ and $b = \begin{cases} 0 & \text{if } a = 0 \\ c/a & \text{if } a \neq 0 \end{cases}$
$(\exists x)\ a = 2x + 1$	Diophantine	a is odd
$(\exists x)\ a = x^2$	Diophantine	a is a square
$(\exists x)\ a = b + x$	Diophantine	$a \geq b$
$(\exists x)\ a = b + x + 1$	Diophantine	$a > b$
$(\exists x)\ xa = b$	Diophantine	$a \mid b$
$(\exists x)(\exists y)\ xa = yb + 1$	Diophantine	a and b are coprime
$(\exists x)(\exists y)\ a = (x + 2)(y + 2)$	Diophantine	a is a composite number
$(\exists x)(\exists y)\ a + xc = b + yc$	Diophantine	$a \equiv b \pmod{c}$

Table 4.1: Some polynomial equations and Diophantine expressions

Definition 4.4 (Equivalence)**.** Let Φ and Ψ be logical expressions in which all free variables represent non-negative integers. Then we call Φ and Ψ equivalent, in symbols $\Phi \iff \Psi$, if the bi-implication $\Phi \longleftrightarrow \Psi$ is true for all instantiations of free variables with non-negative integers.

Proposition 4.5. *Let $n \in \mathbb{N}$ be any non-negative integer and let Φ be any logical expression with no free variables other than $X_1, \ldots, X_n$. Then Φ is equivalent to a polynomial equation if and only if there is a polynomial $f \in \mathbb{Z}[X_1, \ldots, X_n]$ in unknowns $X_1, \ldots, X_n$ with integral coefficients such that we have*

$$\Phi \iff f(X_1, \ldots, X_n) = 0.$$

Proof. For the forward implication suppose that Φ is equivalent to a polynomial equation. Take $g, h \in \mathbb{N}[X_1, \ldots, X_n]$ such that we have $\Phi \iff g(X_1, \ldots, X_n) = h(X_1, \ldots, X_n)$ (Definition 4.1). The polynomial $f = g - h \in \mathbb{Z}[X_1, \ldots, X_n]$ satisfies the requirement.

For the implication from right to left suppose that there is a polynomial $f \in \mathbb{Z}[X_1, \ldots, X_n]$ in unknowns $X_1, \ldots, X_n$ with integral coefficients such that we have $\Phi \iff f(X_1, \ldots, X_n) = 0$. Then, by Proposition 3.1, there are polynomials $f^+, f^- \in \mathbb{N}[X_1, \ldots, X_n]$ with non-negative integral coefficients, such that we have

$$\Phi \iff f^+(X_1, \ldots, X_n) = f^-(X_1, \ldots, X_n).$$

The result follows since the right-hand side is a polynomial equation (Definition 4.1). □

Proposition 4.6. *If Φ and Ψ are any two polynomial equations, then the logical expressions $\Phi \wedge \Psi$ and $\Phi \vee \Psi$ are equivalent to polynomial equations.*

Proof. Let X_1 through X_n be all the free variables occurring in Φ and Ψ. By Proposition 4.5 there are polynomials $f, g \in \mathbb{Z}[X_1, \ldots, X_n]$ with integral coefficients such that we have

$$f(X_1, \ldots, X_n) = 0 \iff \Phi \quad \text{and} \quad g(X_1, \ldots, X_n) = 0 \iff \Psi.$$

We see that we have

$$f(X_1,\ldots,X_n)\cdot g(X_1,\ldots,X_n)=0 \iff \Phi\vee\Psi, \text{ and}$$

$$f(X_1,\ldots,X_n)^2+g(X_1,\ldots,X_n)^2=0 \iff \Phi\wedge\Psi.$$

This yields the result by Proposition 4.5. □

Notice that the proof of Proposition 4.6 depends on the existence of polynomials $h,k\in\mathbb{Z}[X,Y]$ with integral coefficients such that for all integers $x,y\in\mathbb{Z}$ we have

$$\begin{aligned} h(x,y)=0 &\longleftrightarrow x=0\vee y=0, \text{ and}\\ k(x,y)=0 &\longleftrightarrow x=0\wedge y=0. \end{aligned}$$

If we look at some analogies of Hilbert's tenth problem, for instance with coefficients in the Gaussian integers $\mathbb{Z}[i]$, the polynomials used in the proof might no longer be appropriate. Then it is sometimes useful to instead look at so-called positive existential expressions, which are obtained from polynomial equations by not only applying the existential quantifier $(\exists\,\cdot\,)$, but also the logical connectives conjunction $\wedge$ and disjunction $\vee$. For more on this see for instance Esther Bod's Master's thesis [1, Sections 1.3, 4.1, and onward].

Lemma 4.7. *Let Φ and Ψ be any two logical expressions in which all free variables represent non-negative integers. Let X be any variable symbol that does not occur in Ψ and let Y be any variable symbol or any non-negative integer. Then we have*

$$[(\exists X)\ \Phi]\wedge\Psi \iff (\exists X)\ [\Phi\wedge\Psi], \tag{4.1}$$

$$[(\exists X)\ \Phi]\vee\Psi \iff (\exists X)\ [\Phi\vee\Psi], \tag{4.2}$$

$$[(\forall X\le Y)\ \Phi]\wedge\Psi \iff (\forall X\le Y)\ [\Phi\wedge\Psi], \textit{ and} \tag{4.3}$$

$$[(\forall X\le Y)\ \Phi]\vee\Psi \iff (\forall X\le Y)\ [\Phi\vee\Psi]. \tag{4.4}$$

Proof. Notice that for every equivalence the set of free variables is the same on either side of the equivalence sign. Let any instantiation of the free variables be given and replace all free variables by their instantiation, so that the equivalences become bi-implications (Definition 4.4). Now the set of free variables in Φ is either $\{X\}$ or $\emptyset$, while the set of free variables in Ψ is $\emptyset$.

If Ψ is false, then the bi-implications (4.2) and (4.4) are true, since after striking out "$\vee\Psi$" the expressions on both sides of the bi-implication are identical. The bi-implications (4.1) and (4.3) are true because both sides are false. This is easily seen everywhere, except possibly on the right-hand side of (4.3). There we note, for the sake of contradiction, that if $(\forall X\le Y)\ [\Phi\wedge\Psi]$ is true, then in particular $\Phi\wedge\Psi$ is true for $X=0$. However $\Phi\wedge\Psi$ is false for any instantiation of X, since Ψ is false. Hence $(\forall X\le Y)\ [\Phi\wedge\Psi]$ is false.

If Ψ is true, then (4.1) and (4.3) are true by striking out "$\wedge\Psi$". The other two bi-implications are true, since both sides are true. This time this is clear everywhere, except maybe at the right-hand side of (4.2). Here we note that taking $X=0\in\mathbb{N}$ (or any other non-negative integer) makes this right-hand side true, since Ψ is true. □

Expression	Equiv. to	Meaning
$(a > b) \vee (a < b)$	Diophantine	$a \neq b$
$(a \leq c) \wedge (c < b)$	Diophantine	$c \in [a, b)$
$c \in [ab, a(b+1))$	Diophantine	$a \neq 0$ and $b = \lfloor c/a \rfloor$
$b = \lfloor (2c+a)/(2a) \rfloor$	Diophantine	$a \neq 0$ and $b = [c/a]$
$(\forall x \leq b)\ xa \neq b$	D	$a \nmid b$
$(\forall x \leq a)(\forall y \leq a)\ a \neq (x+2)(y+2)$	D	a is not composite
$(a \geq 2) \wedge (a$ is not composite$)$	D	a is prime

Table 4.2: Some expressions that are equivalent to Diophantine expressions and D-expressions. The equivalences follow from Proposition 4.8, earlier equivalences in this table, those in Table 4.1, and the fact that expressions remain equivalent to Diophantine expressions if we substitute variables by polynomials with non-negative integral coefficients. The function $[\cdot] : \mathbb{R} \to \mathbb{Z}$ is rounding to the nearest integer and upward in case of ties: $[x] \mapsto \lfloor x + \frac{1}{2} \rfloor$.

Proposition 4.8. *If Φ and Ψ are both Diophantine expressions, then $\Phi \wedge \Psi$ and $\Phi \vee \Psi$ are equivalent to Diophantine expressions. Similarly, if Φ and Ψ are both D-expressions, then $\Phi \wedge \Psi$ and $\Phi \vee \Psi$ are equivalent to D-expressions.*

Proof. Without loss of generality any variable that is bound somewhere in Φ does not occur in Ψ and vice versa. Furthermore, again without loss of generality, Φ contains at least as many quantifiers as Ψ.

We use induction on the total number of quantifiers in Φ and Ψ. If Φ and Ψ both have zero quantifiers, then they are polynomial equations, so that the statement follows by Proposition 4.6. Otherwise Φ begins with a quantifier, and is of one of the forms $(\exists X)\ \Phi'$ and $(\forall X \leq Y)\ \Phi'$ with X and Y variable symbols and Φ' an expression of the same kind as Φ and Ψ. (The case $(\forall X \leq Y)\ \Phi'$ of course only occurs if Φ is a D-expression.)

Take variable X, possibly variable Y, $M = (\exists X)$ or $M = (\forall X \leq Y)$, and expression Φ' such that Φ equals $M\ \Phi'$. Let $\star$ be either of $\wedge$ and $\vee$. Notice that X does not occur in Ψ by our assumptions. By Lemma 4.7 we have

$$\Phi \star \Psi \iff [M\ \Phi'] \star \Psi \iff M\ [\Phi' \star \Psi].$$

By induction $\Phi' \star \Psi$ is equivalent to an expression of the same kind as Φ and Ψ. Hence $M\ [\Phi' \star \Psi]$ is also equivalent to such an expression. (Notice once again that M can only be $(\forall X \leq Y)$ if Φ is a D-expression.) Consequently $\Phi \star \Psi$ is equivalent to an expression of the appropriate kind. □

With Proposition 4.8 in hand we can easily come up with many more statements that are equivalent to Diophantine expressions and D-expressions. Some are listed in Table 4.2. We remark that negations of our expressions are not always equivalent to an expression of the same kind (Subsection 4.5).

4.3 Sets and functions

With the expressions now defined we can describe some sets in $\mathbb{N}^n$. For instance the set $\{x^2 : x \in \mathbb{N}\}$ can be described by $(\exists y)\ x = y^2$ and $\{(x, y) \in \mathbb{N}^2 : x \mid y\}$ by $(\exists z)\ y = zx$. We formalize this in Definition 4.9. Then in Definition

4.10 we carry the qualifications “polynomial”, “Diophantine” and “D-” over to sets, defining polynomial sets, Diophantine sets and D-sets. Finally we define Diophantine functions and D-functions, which are useful for studying these sets. Some examples are presented in Table 4.3.

Definition 4.9 (Representation)**.** Let $n \in \mathbb{N}$ be any non-negative integer and let S be a subset of $\mathbb{N}^n$. Let $X_1, \ldots, X_n$ be different variable symbols. Then an expression Φ with no free variables other than $X_1, \ldots, X_n$ is said to *represent* the set S (or to be a *representation* of S), if we have

$$(X_1, \ldots, X_n) \in S \iff \Phi.$$

We remark that a representation of a set is far from unique (if only because of the choice of variable symbols) and that a single expression represents many sets. For example the sets $\{(x, x^2) : x \in \mathbb{N}\}$, $\{(x^2, x) : x \in \mathbb{N}\}$ and $\{(p, x, q, x^2, r) : p, q, r, x \in \mathbb{N}\}$ are all represented by the expressions $a = b^2$, $d = c^2$ and $f + 3 = k^2 + 3$.

Definition 4.10. Let $n \in \mathbb{N}$ be any non-negative integer. Then a subset S of $\mathbb{N}^n$ is called a *polynomial set*, a *Diophantine set* or a *D-set*, if it has a representation that is a polynomial equation, a Diophantine expression or a D-expression respectively.

Corollary 4.11. *Any polynomial set is a Diophantine set and any Diophantine set is a D-set.*

Proof. This is immediate from Definitions 4.1, 4.2, 4.3, and 4.10. □

Definition 4.12. Let $m, n \in \mathbb{N}$ be any two non-negative integers and let D_f be any subset of $\mathbb{N}^m$. Then a function $f : D_f \to \mathbb{N}^n$ is called a *Diophantine function* or a *D-function* if its graph

$$\{(x_1, \ldots, x_m, y_1, \ldots, y_n) \in D_f \times \mathbb{N}^n \mid f(x_1, \ldots, x_m) = (y_1, \ldots, y_n)\} \subseteq \mathbb{N}^{m+n}$$

is a Diophantine set or a D-set respectively.

Corollary 4.13. *Let $m, n \in \mathbb{N}$ be any two non-negative integers, let D_f be any subset of $\mathbb{N}^m$, and let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be different variable symbols. Then a function $f : D_f \to \mathbb{N}^n$ is a Diophantine function or a D-function if and only if the expression*

$$(X_1, \ldots, X_m) \in D_f \ \wedge \ f(X_1, \ldots, X_m) = (Y_1, \ldots, Y_n)$$

is equivalent to a Diophantine expression or a D-expression respectively.

Definition 4.14 (Juxtaposition)**.** Let $k, m, n \in \mathbb{N}$ be any three non-negative integers, let D_f and D_g be subsets of $\mathbb{N}^k$, and let $f : D_f \to \mathbb{N}^m$ and $g : D_g \to \mathbb{N}^n$ be functions. Then the function

$$D_f \cap D_g \to \mathbb{N}^{m+n}, \quad x \mapsto (f(x), g(x))$$

is said to be obtained from f and g through juxtaposition.

Proposition 4.15. *If f and g in Definition 4.14 are both Diophantine functions or both D-functions, then the function obtained from them through juxtaposition is also a Diophantine function or a D-function respectively.*

Proof. Call the obtained function h. Let $X_1, \ldots, X_k$ and $Y_1, \ldots, Y_{m+n}$ be different variable symbols. Consider the expression

$$\begin{aligned}&[(X_1, \ldots, X_k) \in D_f \ \wedge \ f(X_1, \ldots, X_k) = (Y_1, \ldots, Y_m)] \\ &\wedge [(X_1, \ldots, X_k) \in D_g \ \wedge \ g(X_1, \ldots, X_k) = (Y_{m+1}, \ldots, Y_{m+n})],\end{aligned}$$

which represents the graph of h. By Corollary 4.13 and Proposition 4.8, it is equivalent to a Diophantine expression or a D-expression as appropriate. □

Definition 4.16 (Composition)**.** Let $k, m, n \in \mathbb{N}$ be any three non-negative integers. Let D_f be a subset of $\mathbb{N}^k$ and let D_g be a subset of $\mathbb{N}^m$. Finally let $f : D_f \to \mathbb{N}^m$ and $g : D_g \to \mathbb{N}^n$ be functions. Then the composition of f and g is

$$g \circ f : f^{-1}(D_g) \to \mathbb{N}^n, \quad x \mapsto g(f(x))$$

Proposition 4.17. *If f and g in Definition 4.16 are both Diophantine functions or both D-functions, then their composition $g \circ f$ is also a Diophantine function or a D-function respectively.*

Proof. Let $X_1, \ldots, X_k$, $Y_1, \ldots, Y_m$, and $Z_1, \ldots, Z_n$ be different variable symbols. Consider the expression

$$\begin{aligned}&(\exists Y_1) \cdots (\exists Y_m) \\ &\quad [(X_1, \ldots, X_k) \in D_f \ \wedge \ f(X_1, \ldots, X_k) = (Y_1, \ldots, Y_m)] \\ &\quad \wedge [(Y_1, \ldots, Y_m) \in D_g \ \wedge \ g(Y_1, \ldots, Y_m) = (Z_1, \ldots, Z_n)].\end{aligned}$$

By Corollary 4.13 and Proposition 4.8, it is equivalent to a Diophantine expression or a D-expression as appropriate. Moreover it represents the graph of $g \circ f$, so that by Corollary 4.13, $g \circ f$ is now a Diophantine function or a D-function as required. □

4.4 Computability

In the present subsection we will see that anything that can be computed effectively, can be computed by D-expressions. More precisely we will see that any set in $\mathbb{N}^n$ is recursively enumerable if and only if it is a D-set (Theorem 4.23 and Proposition 4.24). The forward direction of this statement is essential in our proof that Hilbert's tenth problem is unsolvable.

The difficult part of the proof is showing that we can perform some form of recursion, called primitive recursion (Definition 4.18). For this purpose we will use Gödel Coding (Definition 4.19). This allows us to 'store' a sequence of non-negative integers of arbitrary length in just two non-negative integers. Using the function gd, which will be defined below, we can then extract the elements of the sequence again. By means of the bounded universal quantifier, we can express dependencies between the elements of the sequence and make each element represent an iteration (Proposition 4.22).

Some D-functions obtained using primitive recursion are listed in Table 4.4. The methods in this section are those used by Yu. I. Manin [6, Section VI.3].

Domain	Function	Type	Expression
$\mathbb{N}^m$	$(x_1, \dots, x_m) \mapsto x_i$	Dioph.	$y = x_i$
$\{(a, b) : b \mid a\}$	$(a, b) \mapsto a/b$	Dioph.	$zb = a$
$\mathbb{N} \times \mathbb{Z}^+$	$(a, b) \mapsto \lfloor a/b \rfloor$	Dioph.	$a \in [zb, (z + 1)b)$
$\mathbb{N} \times \mathbb{Z}^+$	$(a, b) \mapsto [a/b]$	Dioph.	$(z = \lfloor (2a + b)/(2b) \rfloor)$ $\wedge (b > 0)$
$\{s_1, \dots, s_n\}$	$s_i \mapsto t_i$	Dioph.	$(x = s_1 \wedge z = t_1) \vee \dots$ $\vee (x = s_n \wedge z = t_n)$
$\mathbb{Z}^+ \times \mathbb{Z}^+$	$(a, b) \mapsto \gcd(a, b)$	Dioph.	$(a > 0) \wedge (b > 0)$ $\wedge (z \mid a) \wedge (z \mid b)$ $\wedge (a/z,\ b/z$ coprime$)$
$\mathbb{Z}^+ \times \mathbb{Z}^+$	$(a, b) \mapsto \operatorname{lcm}(a, b)$	Dioph.	$(a > 0) \wedge (b > 0)$ $\wedge z = ab/\gcd(a, b)$
$\mathbb{N}$	$t \mapsto \max_{i=0}^{t} f(i)$	D	$(\exists i)\ [(i \leq t) \wedge z = f(i)]$ $\wedge (\forall i \leq t)\ [f(i) \leq z]$

Table 4.3: Some functions that are easily shown to be either Diophantine functions or D-functions using Propositions 4.8, 4.15, and 4.17 and earlier functions and expressions in Tables 4.1 and 4.2. The function $f : \mathbb{N} \to \mathbb{N}$ is any D-function with domain $\mathbb{N}$.

Definition 4.18 (Primitive recursion)**.** Let $n \in \mathbb{N}$ be any non-negative integer. Let D_f and D_g be any subsets of $\mathbb{N}^n$ and $\mathbb{N}^{n+2}$ respectively. Moreover let functions $f : D_f \to \mathbb{N}$ and $g : D_g \to \mathbb{N}$ be given. Then the function $h : D_h \to \mathbb{N}$ obtained from f and g through *primitive recursion* is recursively defined by the equation

$$h(x, k) = \begin{cases} f(x) & \text{if } k = 0 \\ g(x, h(x, k-1), k) & \text{otherwise,} \end{cases}$$

where $x \in \mathbb{N}^n$ is a vector of length n of non-negative integers and $k \in \mathbb{N}$ is any non-negative integer. The domain D_h of h consist of all vectors of non-negative integers for which this equation makes sense. That is, if we define $D_{h,k}$ recursively for every non-negative integer $k \in \mathbb{N}$ by

$$D_{h,k} = \begin{cases} D_f & \text{if } k = 0 \\ \{x \in D_{h,k-1} : (x, h(x, k-1), k) \in D_g\} & \text{otherwise,} \end{cases}$$

then the domain of h is $D_h = \bigcup_{k \in \mathbb{N}} (D_k \times \{k\})$.

Definition 4.19 (Gödel Coding)**.** Define the Diophantine map $\mathrm{gd} : \mathbb{N}^2 \times \mathbb{Z}^+ \to \mathbb{N}$ by

$$\mathrm{gd}(u, v\,; i) = u \text{ rem } (1 + iv).$$

Lemma 4.20. *Let $n \in \mathbb{N}$ be any non-negative integer and let $v \in \mathbb{Z}$ be any multiple of $n!$. Then we have for any two integers $i, j \in \mathbb{Z}$ with $1 \leq |i - j| \leq n$, that $1 + iv$ and $1 + jv$ are coprime.*

Proof. Suppose, for the sake of contradiction, that $1 + iv$ and $1 + jv$ have a common prime divisor $p \in \mathbb{Z}^+$. We see that we have

$$p \,|(1 + iv) - (1 + jv) = (i - j)v.$$

By $1 \leq |i-j| \leq n$, we have $(i-j) \mid n! \mid v$ and hence $p \mid v^2$. By primality of p we now obtain

$$p \mid v.$$

Consequently p and $1+iv$ are coprime, so that p does not divide $1+iv$. This is a contradiction. We conclude that $1+iv$ and $1+jv$ are coprime. □

Proposition 4.21. *Let $n \in \mathbb{N}$ and $a_1, \ldots, a_n \in \mathbb{N}$ be non-negative integers. Then there are non-negative integers $u, v \in \mathbb{N}$ such that for all i in $\{1, \ldots, n\}$ we have*

$$\operatorname{gd}(u, v\,; i) = a_i.$$

Proof. Take $v \in \mathbb{N}$ to be a multiple of $(n-1)!$ such that we have $v \geq a_i$ for each $i \in \{1, \ldots, n\}$. By Lemma 4.20 the non-negative integers $1+v, 1+2v, \ldots, 1+nv$ are coprime in pairs. Hence, by the Chinese Remainder Theorem, there is a non-negative integer $u \in \mathbb{N}$ such that we have

$$u \equiv a_i \pmod{1+iv}$$

for all i in $\{1, \ldots, n\}$. Take such u. For all i in $\{1, \ldots, n\}$ we have $a_i = u \operatorname{rem} (1+vi) = \operatorname{gd}(u, v\,; i)$ since we have $0 \leq a_i < 1+v \leq 1+iv$. □

Proposition 4.22. *Let $n \in \mathbb{N}$ be any non-negative integer, let D_f and D_g be subsets of $\mathbb{N}^n$ and $\mathbb{N}^{n+2}$ respectively, and let $f : D_f \to \mathbb{N}$ and $g : D_g \to \mathbb{N}$ be D-functions. Then the function $h : D_h \to \mathbb{N}$ obtained from f and g by primitive recursion is also a D-function.*

Proof. We claim to have the following equivalence with free variables $x_1, \ldots, x_n$, and m, in which we write x instead of $x_1, \ldots, x_n$ for the sake of brevity:

$$\begin{aligned}
&((x,m) \in D_h) \wedge (y = h(x,m)) \\
&\iff (\exists u)(\exists v) \\
&\qquad\quad y = \operatorname{gd}(u, v\,; m+1) \\
&\qquad\quad \wedge (\forall k \leq m) \\
&\qquad\qquad [(k=0) \wedge x \in D_f \\
&\qquad\qquad\qquad \wedge \operatorname{gd}(u, v\,; 1) = f(x)] \\
&\qquad\qquad \vee [(k>0) \wedge (x, \operatorname{gd}(u, v\,; k), k) \in D_g \\
&\qquad\qquad\qquad \wedge \operatorname{gd}(u, v\,; k+1) = g(x, \operatorname{gd}(u, v\,; k), k)].
\end{aligned}$$

Since the right-hand side is equivalent to a D-expression by the propositions of the previous subsections, the proving of the equivalence yields that h is a D-function (Corollary 4.13).

For the proof from left to right, suppose that we have non-negative integers $m, x_1, \ldots, x_n, y \in \mathbb{N}$ with $(x,m) \in D_h$ and $y = h(x,m)$. Letting k run from m down to 1, we iteratively find $(x, k-1) \in D_h$, and $(x, h(x, k-1), k) \in D_g$. By Lemma 4.21 there are non-negative integers $u, v \in \mathbb{N}$ such that for each k in $\{0, \ldots, m\}$ we have

$$\operatorname{gd}(u, v\,; k+1) = h(x,k).$$

Now the right-hand side is satisfied.

Function f	Function g	Function h
$a \mapsto 1$	$(a,p,k) \mapsto ap$	$(a,n) \mapsto a^n$
$y(0)$	$(p,k) \mapsto p + y(k)$	$n \mapsto \sum_{k=0}^{n} y(k)$
$y(0)$	$(p,k) \mapsto p \cdot y(k)$	$n \mapsto \prod_{k=0}^{n} y(k)$
0	$(p,k) \mapsto p + 1_{\{k \text{ is prime}\}}$	$n \mapsto \pi(n)$
$J(0,1)$	$(J(a,b),k) \mapsto J(b, a+b)$	$n \mapsto J(\mathrm{fib}(n), \mathrm{fib}(n+1))$

Table 4.4: Some D-functions h obtained from f and g using primitive recursion. The map $y : \mathbb{N} \to \mathbb{N}$ is any D-function, while the map $\pi : \mathbb{N} \to \mathbb{N}$ is the prime counting function. The map $\mathrm{fib} : \mathbb{N} \to \mathbb{N}$ gives the Fibonacci sequence $(0,1,1,2,3,5,8,13,21,\ldots)$. Finally, $J : \mathbb{N}^2 \to \mathbb{N}$ is Cantor's diagonal enumeration [8, page 704] (or any other bijective D-function $\mathbb{N}^2 \to \mathbb{N}$). Notice that since J^{-1} is also a D-function (as its graph is the graph of J with the coordinates permuted) and because we can use existential quantifiers, we obtain from the last row that fib itself is also a D-function.

For the proof from right to left, suppose that we have non-negative integers $m, x_1, \ldots, x_n, y, u, v \in \mathbb{N}$ such that the right-hand side is true. Take $a_k = \mathrm{gd}(u,v\,;k+1)$ for every k in $\{0,\ldots,m\}$. Notice that we have $x \in D_f$ and $a_0 = f(x) = h(x,0)$. Consequently we have $(x,0) \in D_h$. Letting k run over the integers from 1 up to m we find $(x, a_{k-1}, k) \in D_g$, thereby $(x,k) \in D_h$ and finally

$$a_k = g(x, a_{k-1}, k) = g(x, h(x,k-1), k) = h(x,k).$$

In particular we have $(x,m) \in D_h$ and $y = \mathrm{gd}(u,v\,;m+1) = a_m = h(x,m)$, which is what we set out to prove. □

Theorem 4.23. *Any recursively enumerable set is a D-set.*

Proof. For this proof we will use primitive recursive functions. The elementary primitive recursive functions are

$$\begin{array}{ccccc} 1^{(m)} : \mathbb{N}^m \to \mathbb{N}, & \quad \mathrm{pr}_i^m : \mathbb{N}^m \to \mathbb{N}, & \quad \text{and} & \quad \mathrm{suc} : \mathbb{N} \to \mathbb{N}, \\ (x_1,\ldots,x_m) \mapsto 1 & \quad (x_1,\ldots,x_m) \mapsto x_i & & \quad x \mapsto x+1, \end{array}$$

with $i, m \in \mathbb{N}$ any two non-negative integers subject to $i \leq m$. From these functions all primitive recursive functions are obtained by repeated application of juxtaposition, composition, and primitive recursion (Definitions 4.14, 4.16, and 4.18). (This corresponds to Definition V.2.4 in Yu. I. Manin's book [6].)

The elementary primitive recursive functions above are D-functions through the D-expressions $Y = 1$, $Y = X_i$, and $Y = X + 1$ respectively. By Propositions 4.15, 4.17, and 4.22 application of juxtaposition, composition, and primitive recursion to any two D-functions yields a D-function. Hence any primitive recursive function is a D-function.

Now let $n \in \mathbb{N}$ be any non-negative integer and let S be any recursively enumerable subset of $\mathbb{N}^n$. It is known from recursive function theory that there is a primitive recursive function $f : \mathbb{N}^n \to \mathbb{N}$ with $S = \{x \in \mathbb{N}^n : f(x) = 0\}$ [6, Theorem V.4.3]. Take such f. By the previous paragraph, f is a D-function.

Consequently the right-hand side of the equivalence

$$\begin{aligned}&(X_1, \ldots, X_n) \in S\\ &\quad\iff (\exists Y)\ Y = 0\ \wedge\ [(X_1, \ldots, X_n) \in \mathbb{N}^n\ \wedge\ f(X_1, \ldots, X_n) = Y]\end{aligned}$$

is equivalent to a D-expression (Corollary 4.13 and Proposition 4.8). We conclude that S is a D-set. □

The following proposition is not essential to our proof that Hilbert's tenth problem is not solvable.

Proposition 4.24. *Any D-set is a recursively enumerable set.*

Proof. We will shortly define a function

$$b : \{\Phi \in \{\text{D-expressions without free variables}\} \mid \Phi \text{ is true}\} \to \mathbb{N},$$

with the property that if a D-expression Φ is true, then it is also true when every occurrence of $(\exists\,\cdot\,)$ is substituted by $(\exists\,\cdot \leq \ell)$ with $\ell = b(\Phi)$. For now assume that such a function exists.

Let $S \subset \mathbb{N}^n$ be any D-set. Take different variable symbols $X_1, \ldots, X_n$ and a D-expression Φ with no free variables other than $X_1, \ldots, X_n$ such that we have $(X_1, \ldots, X_n) \in S \iff \Phi$ (Definition 4.10). Consider the following algorithm:

```
1  read a vector $(x_1, \ldots, x_n) \in \mathbb{N}^n$
2  take $\Phi'$ to be $\Phi$ with $X_1, \ldots, X_n$ substituted by $x_1, \ldots, x_n$ respectively
3  for $\ell = 0, 1, 2, \ldots$ do
4      take $\Psi$ to be $\Phi'$ with every occurrence of $(\exists\,\cdot\,)$ substituted by $(\exists\,\cdot \leq \ell)$
5      if $\Psi$ is true then
6          return TRUE
```

Notice that Ψ has no free variables, so that the condition in line 5 makes sense semantically. Also notice that all the quantifiers in Ψ are bounded, so that this condition can be verified in finitely many steps (independently of whether the condition holds or not).

Let $x \in \mathbb{N}^n$. If we have $x \in S$, then Φ' as defined in line 2 is true. Hence Ψ is true in iteration $\ell = b(\Phi')$ of the loop, so that the algorithm returns TRUE after finitely many steps. If we have $x \notin S$, then Φ' is false. Hence Ψ is false in every iteration of the loop. Consequently the loop runs infinitely often, so that the algorithm does not terminate. In particular the algorithm does not return TRUE. We conclude that S is recursively enumerable (Definition 2.2).

It remains to show that a function b exists, with the property that if D-expression Φ is true, it remains true if we replace every existential quantifier $(\exists\,\cdot\,)$ in Φ by $(\exists\,\cdot \leq \ell)$ with $\ell = b(\Phi)$.

For any variable symbol X, any D-expression Φ with no free variables other than X, and any non-negative integer $x \in \mathbb{N}$, we denote by $\Phi(x)$ the D-expression Φ with every free occurrence of X substituted by x. Now let b be recursively defined by

$$\begin{aligned} b(E) &= 0,\\ b\big((\forall X \leq m)\ \Phi\big) &= \max\{b(\Phi(x)) \mid x = 0, \ldots, m\},\ \text{and}\\ b\big((\exists X)\ \Phi\big) &= \min\{\max\{x, b(\Phi(x))\} \mid x \in \mathbb{N} \wedge \Phi(x) \text{ is true}\}\end{aligned}$$

where E ranges over the polynomial equations without free variables, X ranges over the variable symbols, $m \in \mathbb{N}$ ranges over the non-negative integers, and Φ ranges over the D-expressions with no free variables other than X.

Let Φ be any D-expression without free variables that is true. We use induction on the number of quantifiers in Φ. If Φ has no quantifiers, then nothing changes if we substitute every occurrence of $(\exists\cdot)$ by $(\exists\cdot \leq \ell)$ with $\ell = b(\Phi)$. Hence the property clearly holds.

Otherwise assume that Φ is of the form $(\forall X \leq m)\ \Phi'$ with X some variable symbol, m some non-negative integer, and Φ' some D-expression with no free variables other than X. Since Φ is true, $\Phi'(x)$ is true for all $x = 0, \ldots, m$. Letting y be in $\{0, \ldots, m\}$, we have by the induction hypothesis that $\Phi'(y)$ is even true when all the existential quantifiers are replaced by ones bounded by $b(\Phi'(y))$. Of course we can also use the looser bound

$$b(\Phi) = \max\{b(\Phi'(x)) \mid x = 0, \ldots, m\} \geq b(\Phi'(y)).$$

We see that the assertion holds for $\Phi = (\forall X \leq m)\ \Phi'$.

Finally assume that Φ is of the form $(\exists X)\ \Phi'$ with X some variable symbol and Φ' some D-expression with no free variables other than X. Since Φ is true, $\Phi'(x)$ is true for some $x \in \mathbb{N}$. Take $x \in \mathbb{N}$ with $\Phi'(x)$ true and $\max\{x, b(\Phi'(x))\}$ minimal. By the induction hypothesis $\Phi'(x)$ is still true when $(\exists\cdot)$ is replaced by $(\exists\cdot \leq \ell)$ with $\ell = b(\Phi'(x))$. As before we can also use the looser bound

$$b(\Phi) = \max\{x, b(\Phi'(x))\} \geq b(\Phi'(x)).$$

We see that $\Phi = (\exists X)\ \Phi'$ is still true when all existential quantifiers are bounded by $b(\Phi)$, namely take $X = x \leq \max\{x, b(\Phi'(x))\} = b(\Phi)$. □

4.5 Negation

We remarked at the end of Subsection 4.2 that negations of Diophantine expressions are D-expressions are not necessarily equivalent to expressions of the same kind. Even though this is not required for the main result, we shall presently prove this.

Corollary 4.25. *There is a D-expression whose negation is not equivalent to a D-expression.*

Proof. The class of D-sets coincides with the class of recursively enumerable sets (Theorem 4.23 and Proposition 4.24). Hence by Corollary 2.6 there is a D-set whose complement is not a D-set. Equivalently there is D-expression of which the negation is not equivalent to a D-expression (Definition 4.10). □

Corollary 4.26. *There is a Diophantine expression of which the negation is not equivalent to a Diophantine expression.*

Proof. Suppose that negations of Diophantine expressions are equivalent to Diophantine expressions. Noticing that the bounded universal quantification $(\forall X \leq Y)\ (\ldots)$ has the same meaning as the quantification $\neg(\exists X)[(X \leq Y) \wedge \neg(\ldots)]$, we see that any D-expression is equivalent to a Diophantine expression. Now, by Corollary 4.25, there is a Diophantine expression of which the negation is not equivalent to a D-expression. Since any Diophantine expression is a D-expression, this negation is a fortiori not equivalent to a Diophantine expression. We have contradicted our assumption. □

The proof of Corollary 4.26 is due to Martin Davis [2, Theorem 2.8] (using arithmetical sets instead of D-sets). To him the result suggested that the classes of Diophantine sets and of D-sets might well coincide. Seventeen years later, in 1970, it was proved that this is indeed the case (Corollary 6.3). Since then constructive proofs of Corollary 4.26 have been developed. See for instance Section 4.6 of Yuri V. Matiyasevich's book [11].

4.6 Relation to Hilbert's tenth problem

Proposition 4.27. *If Hilbert's tenth problem over $\mathbb{N}$ is solvable, then any Diophantine set is recursive.*

Proof. Let $n \in \mathbb{N}$ be any non-negative integer and $S \subseteq \mathbb{N}^n$ any Diophantine set. Take a non-negative integer $m \in \mathbb{N}$ and two polynomials $f, g \in \mathbb{N}[x_1, \ldots, x_n, z_1, \ldots, z_m]$ with non-negative coefficients, such that the Diophantine expression

$$(\exists z_1) \cdots (\exists z_m)\ f(x_1, \ldots, x_n, z_1, \ldots, z_m) = g(x_1, \ldots, x_n, z_1, \ldots, z_m)$$

represents S (Definition 4.10).

We assume that Hilbert's tenth problem over $\mathbb{N}$ is solvable. Hence there is an algorithm that takes a polynomial equation, runs for finitely many steps and then returns whether the equation has a solution in non-negative integers. Take such an algorithm and call it A.

Now consider the following algorithm B, in which in lines 2 and 3 the polynomials f and g are partially evaluated,

```
1  read non-negative integers $x_1, \ldots, x_n \in \mathbb{N}$
2  $f' \leftarrow f(x_1, \ldots, x_n, Z_1, \ldots, Z_m) \in \mathbb{N}[Z_1, \ldots, Z_m]$
3  $g' \leftarrow g(x_1, \ldots, x_n, Z_1, \ldots, Z_m) \in \mathbb{N}[Z_1, \ldots, Z_m]$
4  execute algorithm $A$ on the equation $f'(Z_1, \ldots, Z_m) = g'(Z_1, \ldots, Z_m)$
5  return what algorithm $A$ returned
```

Since algorithm A terminates after a finite number of steps, so does algorithm B. By the properties of algorithm A, algorithm B returns whether input vector $(x_1, \ldots, x_n) \in \mathbb{N}^n$ is in S. Hence the set S is recursive (Definition 2.1). □

5 Exponentiation

In this section we will prove that the functions listed in Table 5.1, which are all related to exponentiation, are Diophantine. These functions will be useful in Section 6 to prove that all D-sets are Diophantine sets.

Our proof is based on proofs by Martin Davis [3, Paragraph 3] and by Yu. I. Manin [6, Proposition VI.5.3]. These proofs are very similar and both use the Pell equation. The proof consists of four phases. First we will use the Pell to construct a Diophantine function (Subsections 5.1 and 5.2). Then we will see that this function is of roughly exponential growth (Subsection 5.3). Using this roughly exponential Diophantine function (which is the topmost function in Table 5.1) we will describe the function $(x, y) \mapsto x^y$ in a Diophantine way (Subsection 5.4). Finally we will construct the other functions in the table using the functions that are listed earlier (also Subsection 5.4).

Domain	Function	Proof
$\mathbb{Z}^+ \times \{a \in \mathbb{Z} : a \geq 2\}$	$(n,a) \mapsto y_n(a)$	Proposition 5.17
$\mathbb{N}^2$	$(x,y) \mapsto x^y$	Proposition 5.21
$\{(n,k) \in \mathbb{N}^2 : n \geq k\}$	$(n,k) \mapsto \binom{n}{k}$	Proposition 5.22
$\mathbb{N}$	$k \mapsto k!$	Proposition 5.23
$\mathbb{N}^3$	$(t,u,v) \mapsto \prod_{i=0}^{t}(u+iv)$	Proposition 5.24
$\mathbb{N}^2$	$(t,a) \mapsto \prod_{j=0}^{t}(a-j)$	Corollary 5.25

Table 5.1: Some Diophantine functions that are related to exponentiation, with references to where their Diophantine nature is proven. The topmost function is defined in Definition 5.15.

The Pell equation is the equation $x^2 - dy^2 = 1$ with $d \in \mathbb{Z}^+$ some positive integer that is not a square (Proposition B.1). For any such d the equation has a solution in positive integers (Remark B.2). The smallest such solution $(x_1, y_1) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ is called the *fundamental solution* (Definition B.4). If d equals $a^2 - 1$ with $a \in \mathbb{Z}$ some integer with $a \geq 2$, then the fundamental solution is $(x_1, y_1) = (a, 1)$ (Proposition B.5). We define the functions $x_., y_. : \mathbb{Z} \to \mathbb{Z}$ by the identity

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n$$

in $\mathbb{Z}[\sqrt{d}] \subset \mathbb{R}$ (Definition B.9 and Corollary B.10). This is well-defined (Proposition B.8). We have $y_n = -y_{-n}$ and $x_n = x_{-n}$ for any integer $n \in \mathbb{Z}$. That is, the map $y_.$ is odd while the map $x_.$ is even (Corollary B.13). The pairs

$$\ldots, (x_{-2}, y_{-2}), (x_{-1}, y_{-1}), (x_0, y_0), (x_1, y_1), (x_2, y_2), \ldots$$

are precisely the solutions of the Pell equation in $\mathbb{N} \times \mathbb{Z}$ in strictly ascending order of y-coordinate with $y_0 = 0$ (Proposition B.15 and Corollaries B.14 and B.20). Consequently we have $y_n \geq n$ for any non-negative integer $n \in \mathbb{N}$ (Corollary B.16).

5.1 Some lemmas on the Pell equation

In this subsection we will prove some additional propositions and lemmas about solutions to the Pell equation, mainly having to do with the solutions modulo integers. We will use these results in our construction of a roughly exponential Diophantine function in Subsection 5.2. Some of the results are formulated in more generality than necessary for the proof, since they might be of interest by themselves.

For the remainder of this subsection, fix a positive non-square integer $d \in \mathbb{Z}^+$, so that[1] the Pell equation has a solution in positive integers. Define the maps $x_.$ and $y_.$ as above.

Lemma 5.1. *For each integer $n \in \mathbb{Z}$ we have that x_n and y_n are coprime.*

Proof. We have $(x_n)x_n + (-dy_n)y_n = 1$. □

[1] Readers who do not wish to depend on Remark B.2, which is not proven in this bachelor thesis, should read ‘such that’ instead of ‘so that’.

Lemma 5.2. *For any two positive integers $n, k \in \mathbb{Z}^+$ we have*

$$y_{kn} \equiv k x_n^{k-1} y_n \pmod{d y_n^3}.$$

Proof. We note that by definition of $x.$ and $y.$ (Definition B.9) we have

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

Hence we find

$$\begin{aligned} y_{kn} = \sum_{\substack{i=0,\ldots,k \\ i \text{ odd}}} \binom{k}{i} x_n^{k-i} y_n^i d^{(i-1)/2} &\equiv \binom{k}{1} x_n^{k-1} y_n^1 d^0 \\ &\equiv k x_n^{k-1} y_n \pmod{d y_n^3}. \end{aligned}$$

$\square$

Corollary 5.3. *For any positive integer $n \in \mathbb{Z}^+$ we have $y_n^2 \mid y_{n y_n}$.*

Proof. Since n is positive, so is y_n (Corollary B.14). Using the lemma (Lemma 5.2) with $k = y_n$, we find

$$y_{y_n n} \equiv y_n x_n^{y_n - 1} y_n \pmod{d y_n^3}.$$

The statement follows by reducing modulo y_n^2. $\square$

Lemma 5.4. *For every positive integer $n \in \mathbb{Z}^+$, there is a positive integer $m \in \mathbb{Z}^+$ such that we have $2y_n^2 \mid y_m$.*

Proof. Notice that $d(2y_n^2)^2$ is not a square, since d is not a square. Since $u^2 - d(2y_n^2)^2 v^2 = 1$ is a Pell equation with non-square positive parameter $d(2y_n^2)^2$, it has a solution $(u, v) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ in positive integers (Remark B.2). Now $(u, 2y_n^2 v) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ is a solution to the Pell equation $x^2 - dy^2 = 1$ under investigation. Take a positive integer $m \in \mathbb{Z}^+$ such that we have $(x_m, y_m) = (u, 2y_n^2 v)$ (Corollary B.20). $\square$

To avoid having to depend on Remark B.2, whose proof is outside the scope of this bachelor thesis, we offer an alternative proof:

Alternative proof for Lemma 5.4. Take $m = 2n y_{2n} \in \mathbb{Z}^+$, so that we have $y_{2n}^2 \mid y_{2n y_{2n}} = y_m$ by Corollary 5.3. From the definition of $x.$ and $y.$ (Definition B.9) we have

$$x_{2n} + y_{2n}\sqrt{d} = (x_n + y_n\sqrt{d})^2 = x_n^2 + d y_n^2 + 2 x_n y_n \sqrt{d}$$

and consequently $y_{2n} = 2 x_n y_n$. We obtain $2y_n^2 \mid 4x_n^2 y_n^2 = y_{2n}^2 \mid y_m$. $\square$

Lemma 5.5. *Let $n, m \in \mathbb{Z}^+$ be positive integers. We have $y_n \mid y_m$ if and only if we have $n \mid m$.*

Proof. For the proof from right to left, suppose that we have $n \mid m$. Take $k = m/n > 0$. We find $y_n \mid y_m = y_{kn}$ by Lemma 5.2.

For the other direction of the proof, suppose that we have $y_n \mid y_m$. Then $(x_m, y_m/y_n) \in (\mathbb{Z}^+)^2$ is a solution to the Pell equation $x^2 - y^2(y_n^2 d) = 1$. This

Pell equation has fundamental solution $(x_n, 1)$ (Proposition B.5). Hence there is a positive integer $k \in \mathbb{Z}^+$ such that we have

$$x_m + \tfrac{y_m}{y_n}\sqrt{y_n^2 d} = (x_n + 1^2\sqrt{y_n^2 d})^k$$

(Corollaries B.20 and B.10). We obtain $x_m + y_m\sqrt{d} = (x_n + y_n\sqrt{d})^k = x_{kn} + y_{kn}\sqrt{d}$, and consequently $m = nk$ as required (Corollary B.10 and Proposition B.15). □

Lemma 5.6. *Let $n, m \in \mathbb{Z}^+$ be positive integers. We have $y_n^2 \mid y_m$ if and only if we have $ny_n \mid m$.*

Proof. For the statement from right to left, suppose that we have $ny_n \mid m$. We obtain $y_n^2 \mid y_{ny_n} \mid y_m$ by Corollary 5.3 and Lemma 5.5.

For the other implication assume that we have $y_n^2 \mid y_m$. We have $n \mid m$ by Lemma 5.5. Take $k = m/n \in \mathbb{Z}^+$. Lemma 5.2 yields $y_m = y_{kn} \equiv kx_n^{k-1}y_n \pmod{dy_n^3}$. By reducing modulo y_n^2 we find $0 \equiv kx_n^{k-1}y_n \pmod{y_n^2}$. That is, we have $y_n^2 \mid kx_n^{k-1}y_n$. Since y_n and x_n are coprime (Lemma 5.1), we find $y_n \mid k$ and hence $ny_n \mid nk = m$. □

Lemma 5.7. *For each positive integer $k \in \mathbb{Z}^+$, we have $y_k \equiv ky_1 \pmod{x_1 - 1}$.*

Proof. Using Lemma 5.2 with $n = 1$, we find $y_k \equiv kx_1^{k-1}y_1 \pmod{dy_1^3}$. Since (x_1, y_1) is a solution to the Pell equation (Definition B.9), we have $dy_1^2 = x_1^2 - 1 = (x_1 - 1)(x_1 + 1)$, and therefore $x_1 - 1 \mid dy_1^3$. By reducing the congruence modulo $x_1 - 1$ we obtain

$$y_k \equiv k1^{k-1}y_1 \equiv ky_1 \pmod{x_1 - 1},$$

which is what we were to prove. □

Corollary 5.8. *Suppose that there is an integer $a \in \mathbb{Z}$ with $a \geq 2$ such that we have $d = a^2 - 1$. Then we have*

$$y_k \equiv k \pmod{a - 1}$$

for each positive integer $k \in \mathbb{Z}^+$.

Proof. For such d the fundamental solution is $(x_1, y_1) = (a, 1)$ (Proposition B.5). The result follows by the lemma (Lemma 5.7). □

Notice now that if we have $d = a^2 - 1$ as in the corollary (Corollary 5.8), we can use a Diophantine equation to extract some information on the index k from the number y_k. This will be essential in our proof that the map $y.$ is Diophantine for such d (Theorem 5.18).

Lemma 5.9. *Let $c, x, y, x', y' \in \mathbb{Z}$ be integers. Then we have*

$$x + y\sqrt{d} \equiv x' + y'\sqrt{d} \pmod{c}$$

in $\mathbb{Z}[\sqrt{d}]$ if and only if we have

$$x \equiv x' \pmod{c} \qquad \textit{and} \qquad y \equiv y' \pmod{c}$$

in $\mathbb{Z}$. (In the first equation $\pmod{c}$ means "modulo the ideal $(c) = c \cdot \mathbb{Z}[\sqrt{d}]$ of $\mathbb{Z}[\sqrt{d}]$", while in the latter ones it means "modulo the ideal $(c) = c\mathbb{Z}$ of $\mathbb{Z}$".)

Proof. Since $\{1, \sqrt{d}\}$ is a $\mathbb{Z}$-basis for $\mathbb{Z}[\sqrt{d}]$ (Corollaries C.2 and C.19), the following are equivalent:

$$\begin{aligned}
&x + y\sqrt{d} \equiv x' + y'\sqrt{d} \pmod{c} \\
&(\exists a, b \in \mathbb{Z})\ (a + b\sqrt{d})c = (x - x') + (y - y')\sqrt{d} \\
&(\exists a, b \in \mathbb{Z})\ ac = x - x' \ \wedge\ bc = y - y' \\
&x \equiv x' \pmod{c} \ \wedge\ y \equiv y' \pmod{c}.
\end{aligned}$$

□

For the remainder of this subsection, fix a positive integer $n \in \mathbb{Z}^+$ and let $[\cdot]$ be the canonical ring homomophism $\mathbb{Z} \to \mathbb{Z}/x_n\mathbb{Z}$. We will investigate the map $[y.] = [\cdot] \circ y.$.

The following propositions (Propositions 5.10 and 5.14) tell us that the map $[y.]$ is in some ways similar to the sine function. Namely, if f is $\sin(\cdot)$ or $[y.]$ respectively, and p is 2π or $4n$ respectively, then $f(i\frac{p}{4} + \cdot\,)$ odd if i is even and even if i is odd. Moreover, if we have $f(x) = f(x')$, then we have $x \equiv \pm x'$ (mod $\frac{p}{2}\mathbb{Z}$).

Proposition 5.10. *The map $[y.]$ is periodic with a period that is a divisor of $4n$. Furthermore, any map $[y_{ni+\cdot}]$ with $i \in \mathbb{Z}$ an integer is odd if i is even and even if i is odd. That is, in symbols we have*

$$[y_{4in+m}] = [y_m] \qquad \textit{and} \qquad [y_{in+m}] = (-1)^{i+1}[y_{in-m}]$$

with $i, m \in \mathbb{Z}$ any two integers.

Proof. Let $i, m \in \mathbb{Z}$ be any two integers. Take $\alpha = x_1 + y_1\sqrt{d}$, so that we have $\alpha^k = x_k + y_k\sqrt{d}$ for each integer $k \in \mathbb{Z}$ (Definition B.9). Noticing that we have $x_n^2 - dy_n^2 = 1$ (Proposition B.11), we obtain the following two congruences in $\mathbb{Z}[\sqrt{d}]$:

$$\begin{aligned}
\alpha^n &= x_n + y_n\sqrt{d} \equiv y_n\sqrt{d} \pmod{x_n}, \\
\alpha^{2n} &= (\alpha^n)^2 \equiv y_n^2 d = x_n^2 - 1 \equiv -1 \pmod{x_n}.
\end{aligned}$$

Letting $j, k \in \mathbb{Z}$ be any two integers, this yields

$$x_{2jn+k} + y_{2jn+k}\sqrt{d} = (\alpha^{2n})^j\alpha^k \equiv (-1)^j(x_k + y_k\sqrt{d}) \pmod{x_n},$$

so that we find (Lemma 5.9)

$$[y_{2jn+k}] = (-1)^j[y_k].$$

We see that we have $[y_{4in+m}] = [y_m]$ as required. Moreover, using that the map $y.$ is odd (Corollary B.13), we obtain

$$\begin{aligned}
[y_{in+m}] &= (-1)^{\frac{i}{2}}[y_m] = -(-1)^{\frac{i}{2}}[y_{-m}] = -[y_{in-m}] && \text{if } i \text{ is even,} \\
[y_{in+m}] &= (-1)^{\frac{i-1}{2}}[y_{n+m}] = (-1)^{\frac{(i-1)+2}{2}}[y_{-n-m}] = [y_{in-m}] && \text{if } i \text{ is odd.}
\end{aligned}$$

That is, we have $[y_{in+m}] = (-1)^{i+1}[y_{in-m}]$ as needed. □

The property $\sin(x) = \sin(x') \longrightarrow x \equiv \pm x' \pmod{\pi\mathbb{Z}}$ of the sine function can be found by using the sine's symmetry properties and the fact that the sine is injective on $(-\frac{\pi}{2}, \frac{\pi}{2}]$. We will prove the analogous property for $[y.]$ (Proposition 5.14) in the same way.

Lemma 5.11. *We have* $-\frac{1}{4}x_n < y_{-n+1}$ *and* $y_n < \frac{3}{4}x_n$*.*

Proof. Since the positive integer d is not a square (Proposition B.1) we have $d \geq 2$. We obtain $x_n^2 > x_n^2 - 1 = dy_n^2 \geq 2y_n^2$ since (x_n, y_n) is a solution to the Pell equation (Proposition B.11). Consequently we have $y_n < x_n/\sqrt{2}$ and thereby

$$y_n < \tfrac{3}{4}x_n.$$

Suppose that $n-1$ is positive, so that not only y_1 but also y_{n-1} is positive (Corollary B.14). Notice from the definition of $x.$ and $y.$ (Definition B.9) that we have

$$x_n = x_1x_{n-1} + dy_1y_{n-1} = \sqrt{1+dy_1^2}\sqrt{1+dy_{n-1}^2} + dy_1y_{n-1} > 2dy_1y_{n-1}.$$

This gives $y_{n-1} < x_n/(2dy_1) \leq \frac{1}{4}x_n$. Since the map $y.$ is odd (Corollary B.13) we obtain

$$-\tfrac{1}{4}x_n < y_{-n+1}.$$

We obtain the same result for $n = 1$ as we then have $-\frac{1}{4}x_n < 0 = y_0 = y_{-n+1}$. □

Lemma 5.12. *Let k and K be any two integers in $\{-n, \dots, n\}$ with $k \neq K$ and $[y_k] = [y_K]$. Then we have $k = -K = \pm n$, $n = 1$ and $d = 3$.*

Proof. By Lemma 5.11, we have that the difference between y_{-n+1} and y_n is less than x_n. Since the map $y.$ is strictly increasing (Proposition B.15), we obtain that $[y.] = y. + x_n\mathbb{Z}$ is injective on $\{-n+1, \dots, n\}$.

Since the map $[y.]$ is odd, it is also injective on $\{-n, \dots, n-1\}$ (Proposition 5.10). Hence if we have $|k - K| < 2n$, then we have $k = K$, which is false. We obtain $|k - K| = 2n$ and therefore $k = -K = \pm n$.

We now have $-y_n = y_{-n} \equiv y_n \pmod{x_n}$. By Lemma 5.11 we have $y_n < \frac{3}{4}x_n$ and therefore $-y_n + x_n = y_n$. This yields $x_n = 2y_n$. Using the Pell equation we obtain $(4-d)y_n^2 = 1$. Taking positivity of n and therefore of y_n into account (Corollary B.14), we find $d = 3$ and $y_n = 1$. Since $(2, 1)$ is the fundamental solution of the Pell equation with this parameter $d = 3$ (Proposition B.5), we find $n = 1$. □

Corollary 5.13. *Let k and K be any two integers in $\{-n, \dots, n\}$ with $[y_k] = [y_K]$. Then we have $k \equiv K \pmod{2n}$.*

Proof. By Lemma 5.12 we have $K = k$ or $k = -K = \pm n$. In the either case the result follows. □

Proposition 5.14. *Let $k, K \in \mathbb{Z}$ be any two integers. Then if we have $[y_k] = [y_K]$, we also have $k \equiv \pm K \pmod{2n}$.*

Proof. By periodicity of $[y.]$ (Proposition 5.10) it is sufficient to prove this for k and K in $\{-n, \dots, 3n\}$. For such k and K define

$$k' = \begin{cases} k & \text{if } k \leq n \\ 2n - k & \text{if } k > n \end{cases}$$

and K' analogously. Since the map $[y_{n+\cdot}]$ is even (Proposition 5.10), we have $[y_{k'}] = [y_k] = [y_K] = [y_{K'}]$. Moreover we have $k', K' \in \{-n, \dots, n\}$, so that Corollary 5.13 yields $k' \equiv K' \pmod{2n}$. The observation that we have $k' \equiv \pm k \pmod{2n}$ and $K' \equiv \pm K \pmod{2n}$ concludes the proof. □

5.2 A Diophantine function

Definition 5.15 ($x_n(a)$ and $y_n(a)$)**.** We define the functions

$$x.(\cdot), y.(\cdot) : \mathbb{Z}^+ \times \{a \in \mathbb{Z} : a \geq 2\} \to \mathbb{Z}^+$$

by the following identity in $\mathbb{R}$ in which $a, n \in \mathbb{Z}^+$ are positive integers with $a \geq 2$:

$$x_n(a) + y_n(a)\sqrt{a^2-1} = (a + \sqrt{a^2-1})^n.$$

This is well-defined by Propositions B.5 and B.8.

Notice that for any $a \in \mathbb{Z}$ with $a \geq 2$, the maps $x.(a)$ and $y.(a)$ are restrictions to $\mathbb{Z}^+$ of the functions $x.$ and $y.$ from the previous subsection, belonging to the Pell equation with parameter $d = a^2 - 1$ and fundamental solution $(a, 1)$. From the definition of $x.(\cdot)$ and $y.(\cdot)$ we might suspect that they display roughly exponential growth. This is indeed the case, as we shall see in the next subsection (Subsection 5.3). We will prove in Theorem 5.18 that $y.(\cdot)$ is Diophantine as a function in two variables.

Lemma 5.16. *Let $a, b, c \in \mathbb{Z}^+$ be positive integers with $a, b \geq 2$ and $a \equiv b \pmod{c}$. For each positive integer $n \in \mathbb{Z}^+$ we have*

$$x_n(a) \equiv x_n(b) \pmod{c} \quad \textit{and} \quad y_n(a) \equiv y_n(b) \pmod{c}.$$

Proof. Let $n \in \mathbb{Z}^+$ be a positive integer. Take $D = a^2 - 1 = b^2 - 1$ in $\mathbb{Z}/c\mathbb{Z}$ and take the following ring homomorphisms (Corollaries C.8 and C.19):

$$\begin{aligned} \phi_a &: \mathbb{Z}[\sqrt{a^2-1}] \to (\mathbb{Z}/c\mathbb{Z})[X]/(X^2 - D), \quad x + y\sqrt{a^2-1} \mapsto x + yX \text{ and} \\ \phi_b &: \mathbb{Z}[\sqrt{b^2-1}] \to (\mathbb{Z}/c\mathbb{Z})[X]/(X^2 - D), \quad x + y\sqrt{b^2-1} \mapsto x + yX. \end{aligned}$$

Recalling that we have $x_n(a) + y_n(a)\sqrt{a^2-1} = (a + \sqrt{a^2-1})^n$ (Definition 5.15) we obtain in $(\mathbb{Z}/c\mathbb{Z})[X]/(X^2 - D)$

$$x_n(a) + y_n(a)X = \phi_a\left[(a + \sqrt{a^2-1})^n\right] = \phi_a\left[a + \sqrt{a^2-1}\right]^n = (a + X)^n,$$

and similarly $x_n(b) + y_n(b)X = (b + X)^n$. Since a and b are equal in $\mathbb{Z}/c\mathbb{Z} \subset (\mathbb{Z}/c\mathbb{Z})[X]/(X^2 - D)$, we have in $(\mathbb{Z}/c\mathbb{Z})[X]/(X^2 - D)$

$$x_n(a) + y_n(a)X = x_n(b) + y_n(b)X.$$

We obtain $x_n(a) \equiv x_n(b) \pmod{c}$ and $y_n(a) \equiv y_n(b) \pmod{c}$ as $\{1, X\}$ is a $\mathbb{Z}/c\mathbb{Z}$-basis for $(\mathbb{Z}/c\mathbb{Z})[X]/(X^2 - D)$ (Corollary C.2). □

Proposition 5.17. *The set*

$$\{(a, n, y) \in \mathbb{N}^3 \mid (a \geq 2) \wedge (n > 0) \wedge (y = y_n(a))\}$$

is Diophantine.

Proof. (See [3, Paragraph 3] and [6, Proposition VI.5.3].) We will prove that the expression $(a \geq 2) \wedge (n > 0) \wedge (y = y_n(a))$ is equivalent to a Diophantine expression, which yields the result by Corollary 4.13. For this purpose we claim to have the following equivalence:

$$\begin{aligned}
&(a \geq 2) \wedge (n > 0) \wedge (y = y_n(a)) \\
&\iff (\exists u)(\exists A)(\exists Y)(\exists x)(\exists v)(\exists X) \\
&\qquad (a \geq 2) \wedge (n > 0) \wedge (n \leq y) && (5.1)\\
&\qquad \wedge (x^2 - (a^2-1)y^2 = 1) \wedge (y > 0) && (5.2)\\
&\qquad \wedge (u^2 - (a^2-1)v^2 = 1) \wedge (v > 0) && (5.3)\\
&\qquad \wedge y^2 \mid v && (5.4)\\
&\qquad \wedge A \geq 2 && (5.5)\\
&\qquad \wedge A \equiv a \pmod{u} && (5.6)\\
&\qquad \wedge 2y \mid A - 1 && (5.7)\\
&\qquad \wedge (X^2 - (A^2-1)Y^2 = 1) \wedge (Y > 0) && (5.8)\\
&\qquad \wedge Y \equiv y \pmod{u} && (5.9)\\
&\qquad \wedge n \equiv Y \pmod{2y}. && (5.10)
\end{aligned}$$

By the propositions of Section 4 the right-hand side of the equivalence is equivalent to a Diophantine expression.

For the proof from right to left suppose that we have non-negative integers $a, n, y, u, A, Y, x, v, X \in \mathbb{N}$ such that equations (5.1) through (5.10) are satisfied. By (5.1), the conditions $a \geq 2$ and $n > 0$ are satisfied, so that it only remains to show that we have $y = y_n(a)$.

By (5.2), (5.3), (5.5), and (5.8) there are positive integers $k, m, K \in \mathbb{Z}^+$ such that we have (Corollary B.20)

$$y = y_k(a), \qquad u = x_m(a), \qquad v = y_m(a), \qquad \text{and} \qquad Y = y_K(A).$$

From (5.6) we see $Y = y_K(A) \equiv y_K(a) \pmod{u}$ (Lemma 5.16), so that (5.9) yields $y_K(a) \equiv Y \equiv y = y_k(a) \pmod{u = x_m(a)}$. By Proposition 5.14 we obtain $K \equiv \pm k \pmod{2m}$. Using (5.4) we obtain $y \mid m$ (Lemma 5.6) and we see

$$K \equiv \pm k \pmod{2y}.$$

By Corollary 5.8 we have $Y \equiv K \pmod{A-1}$, which yields $Y \equiv K \pmod{2y}$ with (5.7). Hence (5.10) gives

$$n \equiv K \equiv \pm k \pmod{2y}.$$

With $0 < n \leq y$ from (5.1) and $0 < k \leq y_k(a) = y$ by Corollary B.16, we obtain $n = k$. Hence we have $y = y_k(a) = y_n(a)$ as required.

For the proof from left to right suppose that we have non-negative integers $a, n, y \in \mathbb{N}$ with $a \geq 2$, $n > 0$, and $y = y_n(a)$. Immediately (5.1) is satisfied, for we have $n \leq y_n(a) = y$ (Corollary B.16). As a consequence, we have $y > 0$ and (5.2) is satisfied by taking $x = x_n(a)$ (Corollary B.20).

Use Lemma 5.4 to take a positive integer $m \in \mathbb{Z}^+$ with $2y^2 \mid y_m(a)$. Now (5.4) and (5.3) hold by taking $u = x_m(a)$ and $v = y_m(a)$ (Corollary B.20). Notice that we have $u^2 = 1 + (a^2-1)v^2 \geq a$ and $u^2 \equiv 1 + (a^2-1)0^2 \equiv 1 \pmod{2y}$. Take $A = a + u^2(u^2 - a)$, so that (5.5) and (5.6) are immediately satisfied. Moreover we have $A \equiv a + 1(1-a) \equiv 1 \pmod{2y}$, satisfying (5.7).

Take $X = x_n(A)$ and $Y = y_n(A)$, so that (5.8) holds. Since we have $A \equiv a \pmod{u}$ by our choice of A, we have $Y = y_n(A) \equiv y_n(a) \equiv y \pmod{u}$ (Lemma 5.16), so that (5.9) holds. Since we have $Y = y_n(A) \equiv n \pmod{A-1}$ (Corollary 5.8), we have (5.10) by (5.7). □

Theorem 5.18. *The function $y_\cdot(\cdot)$ as a function in two variables is Diophantine.*

Proof. Its graph is the Diophantine set in Proposition 5.17. □

5.3 Our Diophantine function is roughly exponential

Proposition 5.19. *Let $d \in \mathbb{Z}^+$ be a positive integer that is not a square and let $(x_1, y_1) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ be the fundamental solution to the Pell equation $x^2 - dy^2 = 1$. Let the maps $x_\cdot : \mathbb{Z} \to \mathbb{Z}$ and $y_\cdot : \mathbb{Z} \to \mathbb{Z}$ be defined as in the beginning of this section (or, equivalently, as in Definition B.9). Then for any non-negative integer $n \in \mathbb{N}$ we have,*

$$\frac{x_{n+1}}{x_1}, \frac{y_{n+1}}{y_1} \in \left[\left(2x_1 - \frac{1}{x_1}\right)^n, (2x_1)^n\right] \subseteq \mathbb{R}.$$

Proof. Call the interval in the statement I_n. That is, for every non-negative integer $n \in \mathbb{N}$, we define the interval $I_n \subset \mathbb{R}$ by $I_n = [(2x_1 - x_1^{-1})^n, (2x_1)^n]$.

We use induction on n. If we have $n = 0$, then the statement is trivial. Otherwise we remark that by definition of $x_\cdot$ and $y_\cdot$ (Definition B.9) we have

$$x_{n+1} + y_{n+1}\sqrt{d} = (x_n + y_n\sqrt{d})(x_1 + y_1\sqrt{d}).$$

This yields $x_{n+1} = x_1 x_n + d y_1 y_n$ and $y_{n+1} = x_1 y_n + y_1 x_n$. Using the induction hypothesis we find

$$\frac{y_{n+1}}{y_1} = x_1 \frac{y_n}{y_1} + x_1 \frac{x_n}{x_1} \in (2x_1) I_{n-1} \subset I_n,$$

which is the statement about y_{n+1}/y_1. For the statement about x_{n+1}/x_1, note that (x_1, y_1) is a solution to the Pell equation (Definition B.9), so that we have $dy_1^2 = x_1^2 - 1$. Again with the induction hypothesis we obtain

$$\frac{x_{n+1}}{x_1} = x_1 \frac{x_n}{x_1} + \frac{dy_1^2}{x_1} \frac{y_n}{y_1} \in \left(x_1 + \frac{x_1^2 - 1}{x_1}\right) I_{n-1} = \left(2x_1 - \frac{1}{x_1}\right) I_{n-1} \subset I_n,$$

which is what remained to be proven. □

Corollary 5.20. *Let $y_\cdot(\cdot) : \mathbb{Z}^+ \times \{a \in \mathbb{Z} : a \geq 2\} \to \mathbb{Z}$ be the function from Definition 5.15. Then for any two non-negative integers $a, n \in \mathbb{N}$ with $a \geq 2$, we have*

$$(2a-1)^n \leq y_{n+1}(a) \leq (2a)^n.$$

Proof. Let $a \in \mathbb{Z}$ with $a \geq 2$. When we take $d = a^2 - 1$ and $(x_1, y_1) = (a, 1)$ (Proposition B.5) in Proposition 5.19, the function $y_\cdot : \mathbb{Z} \to \mathbb{Z}$ from that proposition coincides with the function $y_\cdot(a) : \mathbb{Z} \to \mathbb{Z}$. Hence we obtain

$$(2a - a^{-1})^n \leq y_{n+1}(a) \leq (2a)^n$$

for any non-negative integer $n \in \mathbb{N}$. The result follows by the inequality $2a - 1 \leq 2a - a^{-1}$. □

5.4 Exponentiation and related maps are Diophantine

In this subsection we will show by means of a construction that exponentiation $(x, y) \mapsto x^y$ is Diophantine. For this we will use the roughly exponential Diophantine function $y_\cdot(\cdot)$ from the previous subsections. Then we will proceed to show that several related maps are also Diophantine.

The proofs of Propositions 5.21 and 5.22 are taken from Yu. I. Manin's book [6, Propositions VI.6.1 and VI.7.1]. We have, however, used a different estimate for Proposition 5.21 and provided more details in the proof of Proposition 5.22. The proof of Proposition 5.24 is due to Martin Davis [3, Lemma 4.7].

The construction that we will use depends on the bounds on $y_\cdot(\cdot)$ from Corollary 5.20. It is of historical interest, however, that there also is a construction that works for *any* roughly exponential function (for some precise definition of 'roughly exponential'). This construction is the main subject of Julia Robinson's article [9].

Proposition 5.21. *The function $\mathbb{N}^2 \to \mathbb{N}, (a, n) \mapsto a^n$ is Diophantine.*

Proof. (See [6, Proposition VI.6.1].) We will prove the following equivalence:

$$\begin{aligned} z = a^n \iff & ((n = 0) \wedge (z = 1)) \\ & \vee ((n \geq 1) \wedge (a \leq 1) \wedge (z = a)) \\ & \vee ((n \geq 1) \wedge (a \geq 2) \wedge (\exists N) \\ & \qquad\qquad N \geq y_{n+1}(a + 1) \\ & \qquad\qquad \wedge\, z = [y_{n+1}(aN)/y_{n+1}(N)]), \end{aligned}$$

where $[\cdot]$ is rounding to the nearest integer and up in case of ties. The right-hand side is equivalent to a Diophantine expression by the propositions of Section 4 and by Theorem 5.18. Hence, by Corollary 4.13, proving this equivalence is sufficient to show that the function is Diophantine.

The cases $a = 0$, $a = 1$, $n = 0$ are easily handled. Hence let $a, n \in \mathbb{Z}^+$ be positive integers with $a \geq 2$. By Corollary 5.20 we have

$$(2a - 1)^n \leq y_{n+1}(a) \leq (2a)^n.$$

For each positive integer $N \in \mathbb{Z}^+$ we find

$$\frac{y_{n+1}(aN)}{y_{n+1}(N)} \leq \frac{(2Na)^n}{(2N - 1)^n} = a^n \left(1 + \frac{1}{2N - 1}\right)^n \quad \text{and}$$

$$\frac{y_{n+1}(aN)}{y_{n+1}(N)} \geq \frac{(2Na - 1)^n}{(2N)^n} = a^n \left(1 - \frac{1}{2aN}\right)^n \geq a^n \left(1 - \frac{1}{2N - 1}\right)^n.$$

Hence, for all N larger than some threshold value which may depend on a and n, we can obtain a^n from $y_{n+1}(aN)/y_{n+1}(N)$ by rounding to the nearest integer.

It remains to be proven that we can use $y_{n+1}(a + 1)$ as this threshold value. For this purpose notice that for any real number $x \in \mathbb{R}$ with $|x| \leq 1$ we have

$$|(1 + x)^n - 1| = \left|1 + x \sum_{i=1}^{n} \binom{n}{i} x^{i-1} - 1\right| \leq |x| \sum_{i=0}^{n} \binom{n}{i} |1| = |x| 2^n.$$

Now let any $N \in \mathbb{Z}^+$ with $N \geq y_{n+1}(a+1) > 0$ be given. Since the map $\mathbb{N} \to \mathbb{N}, b \mapsto b^n$ is strictly increasing, we find

$$N \geq y_{n+1}(a+1) \geq (2a+1)^n \geq (2a)^n + 1.$$

This yields $2N - 1 > 2(2a)^n$ and we obtain

$$\left| a^n \left(1 \pm \frac{1}{2N-1}\right)^n - a^n \right| \leq \frac{1}{|2N-1|} \cdot 2^n \cdot a^n < \frac{(2a)^n}{2(2a)^n} = \frac{1}{2}.$$

Consequently rounding $y_{n+1}(aN)/y_{n+1}(N)$ to the nearest integer yields a^n, as required. □

Proposition 5.22. *The function* $\{(n,k) \in \mathbb{N}^2 : n \geq k\} \to \mathbb{N}, (n,k) \mapsto \binom{n}{k}$ *is Diophantine.*

Proof. (See [6, Proposition VI.7.1].) Let $n, k \in \mathbb{N}$ be non-negative integers with $k \leq n$. Let $u \in \mathbb{N}$ be a non-negative integer greater than n^n. For all i from 0 through n, we have $\binom{n}{i} \leq n^n < u$. Hence the number $(u+1)^n = \sum_{i=0}^n \binom{n}{i} u^i$ is written $\binom{n}{n}\binom{n}{n-1}\dots\binom{n}{1}\binom{n}{0}$ in base u. Using the Diophantine function $\mathbb{N}^3 \to \mathbb{N}$, $(q,u,d) \mapsto \lfloor q/u^d \rfloor \text{ rem } u$ (examples in Section 4 and Proposition 5.21), which extracts the d-th base-u-digit of the number q, we can now obtain $\binom{n}{k}$.

More formally and with a looser constraint on u, we claim to have the equivalence

$$\begin{aligned} &(n \geq k) \wedge z = \binom{n}{k} \\ &\iff (n \geq k) \wedge (\exists u) \ \left[(u > n^k) \wedge z = \lfloor (u+1)^n / u^k \rfloor \text{ rem } u\right]. \end{aligned}$$

Since the right-hand side can be rewritten to a Diophantine expression (Section 4 and Proposition 5.21), the statement follows immediately.

To prove this equivalence, let $n, k \in \mathbb{N}$ be any two non-negative integers with $n \geq k$. Take any non-negative integer $u \in \mathbb{N}$ satisfying $u > n^k$. We find

$$(u+1)^n = \sum_{i=0}^{k-1} \binom{n}{i} u^i + \binom{n}{k} u^k + \left(\sum_{i=k+1}^{n} \binom{n}{i} u^{i-k-1} \right) u^{k+1}.$$

The first summation is zero if we have $n = 0$. Otherwise we find

$$\sum_{i=0}^{k-1} \binom{n}{i} u^i \leq \sum_{i=0}^{k-1} n^k u^i \leq \sum_{i=0}^{k-1} (u-1) u^i = u^k - 1 < u^k.$$

We obtain $\lfloor (u+1)^n / u^k \rfloor = \binom{n}{k} + \left(\sum_{i=k+1}^n \binom{n}{i} u^{i-k-1}\right) u$ and therefore

$$\lfloor (u+1)^n / u^k \rfloor \equiv \binom{n}{k} \pmod{u}.$$

Since we have $0 \leq \binom{n}{k} \leq n^k < u$, we find the claimed equality

$$\binom{n}{k} = \lfloor (u+1)^n / u^k \rfloor \text{ rem } u.$$

□

Proposition 5.23. *The function* $\mathbb{N} \to \mathbb{N}, k \mapsto k!$ *is Diophantine.*

Proof. Let $k \in \mathbb{N}$ be any non-negative integer. Notice that we have $1/\binom{n}{k} = k!/\prod_{i=0}^{k-1}(n-i)$ for $n \geq k$. Consider the following functions $\{n \in \mathbb{N} : n \geq k\} \to \mathbb{R}$:

$$n \mapsto \frac{(n-k)^k}{\binom{n}{k}} = \left(\prod_{i=0}^{k-1} \frac{n-k}{n-i}\right) k! \quad \text{and} \quad n \mapsto \frac{n^k}{\binom{n}{k}} = \left(\prod_{i=0}^{k-1} \frac{n}{n-i}\right) k!.$$

For any i in $\{0, \ldots, k-1\}$ we have $\frac{n-k}{n-i} \in [0,1)$ and $\frac{n}{n-i} \in (1, \infty)$, so that the former function is smaller than $k!$ everywhere, while the latter function exceeds $k!$ everywhere. Moreover we have $\lim_{n\to\infty} \frac{n-k}{n-i} = 1$ and $\lim_{n\to\infty} \frac{n}{n-i} = 1$, so that both functions tend to $k!$ as n tends to infinity.

Hence there is $n \geq k$ with $\lceil (n-k)^k/\binom{n}{k} \rceil = \lfloor n^k/\binom{n}{k} \rfloor$. Furthermore, for any such n we have

$$k! = \left\lceil \frac{(n-k)^k}{\binom{n}{k}} \right\rceil = \left\lfloor \frac{n^k}{\binom{n}{k}} \right\rfloor.$$

We obtain the equivalence

$$z = k! \iff (\exists n)\left((n \geq k) \wedge z = \left\lceil \frac{(n-k)^k}{\binom{n}{k}} \right\rceil = \left\lfloor \frac{n^k}{\binom{n}{k}} \right\rfloor\right),$$

which shows that $k \mapsto k!$ is Diophantine (Section 4 and Proposition 5.22). □

Proposition 5.24. *The function* $\mathbb{N}^3 \to \mathbb{N}, (t,u,v) \mapsto \prod_{i=0}^{t}(u+iv)$ *is Diophantine.*

Proof. (See [3, Lemma 4.7].) We claim to have the equivalence

$$\begin{aligned} z = \prod_{i=0}^{t}(u+iv) \iff & [v = 0 \wedge z = u^{t+1}] \\ & \vee \big[v > 0 \wedge (\exists M)(\exists q) \\ & \qquad M = v(u+tv)^{t+1} + 1 \\ & \qquad \wedge\, qv \equiv u \pmod{M} \\ & \qquad \wedge\, z = uv^t(q+t)!/q! \text{ rem } M\big]. \end{aligned}$$

This equivalence implies that the function is Diophantine (Section 4 and Propositions 5.21 and 5.23).

For $v = 0$ the equivalence is clear, so let $u \in \mathbb{N}$ and $v \in \mathbb{Z}^+$ be a non-negative and a positive integer respectively. Take $M = v(u+tv)^{t+1} + 1$. Let $[\cdot]$ denote the canonical ring homomorphism $\mathbb{Z} \to \mathbb{Z}/M\mathbb{Z}$. Since M and v are coprime, the element $[v]$ is a unit in $\mathbb{Z}/M\mathbb{Z}$. Hence there is a non-negative integer $q \in \mathbb{N}$ with $qv \equiv u \pmod{M}$: any non-negative representative of $[q] = [u][v]^{-1} \in \mathbb{Z}/M\mathbb{Z}$ suffices. Take any such $q \in \mathbb{N}$. We obtain

$$\begin{aligned} \prod_{i=0}^{t}(u+iv) &\equiv u\prod_{i=1}^{t}(qv+iv) \pmod{M} \\ &= uv^t \prod_{i=1}^{t}(q+i) = uv^t \frac{(q+t)!}{q!}. \end{aligned}$$

Since $M = v(u+tv)^{t+1}+1$ exceeds $\prod_{i=0}^{t}(u+iv)$, we find that $\prod_{i=0}^{t}(u+iv)$ is the remainder of $uv^t(q+t)!/q!$ after division by M. □

Corollary 5.25. *The function* $\mathbb{N}^2 \to \mathbb{N}, (t,a) \mapsto \prod_{j=0}^{t}(a-j)$ *is Diophantine.*

Proof. We have the equivalence

$$\begin{aligned} &z = \prod_{j=0}^{t}(a-j) \\ &\iff [(a \le t) \wedge (z=0)] \vee \left[(a>t) \wedge \left(z = \prod_{i=0}^{t}((a-t)+i\cdot 1)\right)\right]. \end{aligned}$$

The right-hand side is equivalent to a Diophantine expression by Proposition 5.24 and the propositions of Section 4. Hence the function is Diophantine (Corollary 4.13). □

6 All D-sets are Diophantine

The main theorem of this section (Theorem 6.2) will allow us to eliminate the right-most bounded universal quantifier from any D-expression. By repeatedly applying the theorem, we will be able to transform any D-expression into an equivalent Diophantine expression. This yields that any D-set is a Diophantine set (Corollary 6.3), which will resolve Hilbert's tenth problem in Section 7.

The proof of Theorem 6.2 that we present is due to Martin Davis, Hilary Putnam and Julia Robinson [4, Lemma 3]. It uses functions related to exponentiation that we proved to be Diophantine in Section 5. Since it was not yet known in 1961 that exponentiation can be expressed in a Diophantine manner, the result did not resolve Hilbert's tenth problem at the time; it only showed that Hilbert's tenth problem is unsolvable *if we allow exponentiation.*

The proof uses Lemma 6.1 below. This lemma allows us to trade an expression of the form $(\forall\cdot\le\cdot)(\exists\cdot)\cdots(\exists\cdot)\ g(\ldots)=0$ for an equivalent expression of the form $(\exists T)(\forall\cdot\le T)(\exists\cdot\le T)\cdots(\exists\cdot\le T)\ f(\ldots)=0$, with f and g polynomials with integral coefficients. That is, we can replace the unbounded existential quantifiers by bounded ones at the expense of adding a single unbounded existential quantifier at the front. The usefulness of this, is that we can give a bound on the function values of f that only depends on f itself, the free variables and on variable T. Using this bound we can then select moduli that are sufficiently large to avoid losing information when we use the Chinese Remainder Theorem to evaluate the polynomial f at several points at once. For details see the proof of Theorem 6.2.

Lemma 6.1. *Let $n \in \mathbb{N}$ be any non-negative integer and let $Y, X_1, \ldots, X_n, K$ be different variable symbols. Let Φ be a Diophantine expression with no free variables other than $Y, X_1, \ldots, X_n, K$. Then there are a non-negative integer $m \in \mathbb{N}$, variable symbols $T, Z_1, \ldots, Z_m$ and a polynomial with integral coefficients $f \in \mathbb{Z}[T, Y, X_1, \ldots, X_n, K, Z_1, \ldots, Z_m]$ such that we have*

$$\begin{aligned} (\forall K \le Y)\ \Phi \iff (\exists T)(\forall K \le T)(\exists Z_1 \le T)\cdots(\exists Z_m \le T)& \\ f(T, Y, X_1, \ldots, X_n, K, Z_1, \ldots, Z_m) = 0.& \end{aligned}$$

Proof. For brevity we will write $(\star)$ for $(\forall K \leq Y)\ \Phi$. Moreover we will write $\mathcal{X}_i$ in place of the sequence "$X_1, \ldots, X_i$". After introduction of the appropriate symbols $Z_1, Z_2, \ldots$, we will similarly write $\mathcal{Z}_i$ for the sequence "$Z_1, \ldots, Z_i$".

Since Φ is Diophantine there are a non-negative integer $s \in \mathbb{N}$, distinct variable symbols $Z_1, \ldots, Z_s$ different from $Y, \mathcal{X}_n, K$, and a polynomial with integral coefficients $g \in \mathbb{Z}[Y, \mathcal{X}_n, K, \mathcal{Z}_s]$ such that we have

$$(\star) \iff (\forall K \leq Y)(\exists Z_1) \cdots (\exists Z_s)\ g(Y, \mathcal{X}_n, K, \mathcal{Z}_s) = 0$$

(Definition 4.10 and Proposition 4.5). Let T be a new variable symbol. We introduce it in the right-hand side:

$$\begin{aligned}(\star) \iff (\exists T)\ \big\{&(Y \leq T) \wedge (\forall K \leq Y)(\exists Z_1 \leq T) \cdots (\exists Z_s \leq T)\\ &\qquad g(Y, \mathcal{X}_n, K, \mathcal{Z}_s) = 0\big\}.\end{aligned}$$

The implication from right to left is preserved, as we only introduced extra constraints on the right-hand side. The other direction is also still valid, since we can just pick $T = \max(Y, \mathcal{Z}_s)$. We now change the constraint at $(\forall K \leq \cdot\,)$ to make it uniform with the other constraints. After that we move the quantifiers to the front:

$$\begin{aligned}(\star) &\iff (\exists T)\ \big\{(Y \leq T) \wedge (\forall K \leq T)(\exists Z_1 \leq T) \cdots (\exists Z_s \leq T)\\ &\qquad\qquad [(K > Y) \vee g(Y, \mathcal{X}_n, K, \mathcal{Z}_s) = 0]\big\}\\ &\iff (\exists T)(\forall K \leq T)(\exists Z_1 \leq T) \cdots (\exists Z_s \leq T)\\ &\qquad \big\{(Y \leq T) \wedge [(K > Y) \vee g(Y, \mathcal{X}_n, K, \mathcal{Z}_s) = 0]\big\}.\end{aligned}$$

We write the inequalities explicitly as Diophantine expressions, introducing new variable symbols Z_{s+1} and Z_{s+2}:

$$\begin{aligned}(\star) \iff &(\exists T)(\forall K \leq T)(\exists Z_1 \leq T) \cdots (\exists Z_s \leq T)(\exists Z_{s+1})(\exists Z_{s+2})\\ &\big\{(Y + Z_{s+1} = T) \wedge [(K = Y + Z_{s+2} + 1) \vee g(Y, \mathcal{X}_n, K, \mathcal{Z}_s) = 0]\big\}.\end{aligned}$$

Notice that restricting Z_{s+1} and Z_{s+2} to values in $\{0, \ldots, T\}$ does not affect whether the right-hand side is true. Defining $m = s + 2$ we obtain:

$$\begin{aligned}(\star) \iff &(\exists T)(\forall K \leq T)(\exists Z_1 \leq T) \cdots (\exists Z_m \leq T)\\ &\big\{(Y + Z_{m-1} = T) \wedge [(K = Y + Z_m + 1) \vee g(Y, \mathcal{X}_n, K, \mathcal{Z}_{m-2}) = 0]\big\}.\end{aligned}$$

Since conjunction and disjunctions of polynomial expressions are equivalent to polynomial expressions (Proposition 4.6), there is a polynomial with integral coefficients $f \in \mathbb{Z}[T, Y, \mathcal{X}_n, K, \mathcal{Z}_m]$ such that we have

$$(\star) \iff (\exists T)(\forall K \leq T)(\exists Z_1 \leq T) \cdots (\exists Z_m \leq T)\ f(T, Y, \mathcal{X}_n, K, \mathcal{Z}_m) = 0.$$

This concludes the proof. □

Theorem 6.2. *Let K and X_0 be any two variable symbols and let Φ be any Diophantine expression. Then there is a Diophantine expression Φ' such that we have*

$$(\forall K \leq X_0)\ \Phi \iff \Phi'.$$

Proof. (See [4, Lemma 3].) Pick a non-negative integer $n \in \mathbb{N}$ and distinct variable symbols $X_1, \dots, X_n$ that are not X_0 or K, such that Φ has no free variables other than $X_0, X_1, \dots, X_n, K$. According to Lemma 6.1 with $Y = X_0$, take a non-negative integer $m \in \mathbb{N}$, new distinct variable symbols $T, Z_1, \dots, Z_m$ and a polynomial with integral coefficients $f \in \mathbb{Z}[T, X_0, X_1, \dots, X_n, K, Z_1, \dots, Z_m]$, so that we have the following equivalence

$$\begin{aligned}(\forall K \le X_0)\ \Phi \iff (\exists T)\big\{&(\forall K \le T)(\exists Z_1 \le T) \cdots (\exists Z_m \le T)\\ &f(T, X_0, \dots, X_n, K, Z_1, \dots, Z_m) = 0\big\}.\end{aligned}$$

We will show that the expression on the right-hand side bracketed by curly brackets is equivalent to a Diophantine expression. This is sufficient, as then the whole expression on the right-hand side is also equivalent to a Diophantine expression (Definition 4.2).

For the sake of brevity we will from here on write $\mathcal{X}$ and $\mathcal{Z}$ for $X_0, X_1, \dots, X_n$ and $Z_1, \dots, Z_m$ respectively.

For the purpose of showing that the expression within curly brackets is Diophantine, we will use a polynomial $G \in \mathbb{Z}[T, \mathcal{X}]$ with integral coefficients satisfying

$$G(T, \mathcal{X}) \ge T \qquad \text{and} \qquad G(T, \mathcal{X}) \ge |f(T, \mathcal{X}, K, \mathcal{Z})|$$

for all non-negative integers $T, \mathcal{X}$ and $K, \mathcal{Z}$ in $\{0, \dots, T\}$. To obtain such a polynomial we can for instance write f as a sum of monomials $f(T, K, \mathcal{X}, \mathcal{Z}) = \sum_M c_M M(T, \mathcal{X}, K, \mathcal{Z})$ with coefficients c_M in $\mathbb{Z}$ and take

$$G(T, \mathcal{X}) = T + \sum_M |c_M| M(T, \mathcal{X}, T, T, \dots, T).$$

Take a polynomial $G \in \mathbb{Z}[T, \mathcal{X}]$ satisfying the requirements.

Now take distinct variable symbols $C, V, J, A_1, \dots, A_m$ that are different from $T, \mathcal{X}$. We claim that the following equivalence holds:

$$\begin{aligned}&(\forall K \le T)(\exists Z_1 \le T) \cdots (\exists Z_m \le T)\ f(T, \mathcal{X}, K, \mathcal{Z}) = 0\\ &\iff (\exists C)(\exists V)(\exists A_1) \cdots (\exists A_m)\\ &\qquad V = G(T, \mathcal{X})! && (6.1)\\ &\qquad \wedge\ 1 + CV = \prod_{K=0}^{T} (1 + (K+1)V) && (6.2)\\ &\qquad \wedge\ f(T, \mathcal{X}, C-1, A_1, \dots, A_m) \equiv 0 \pmod{1 + CV} && (6.3)\\ &\qquad \wedge\ 1 + CV \mid \prod_{J=0}^{T} (A_1 - J)\ \wedge\ \cdots\ \wedge\ 1 + CV \mid \prod_{J=0}^{T} (A_m - J). && (6.4)\end{aligned}$$

The right-hand side is equivalent to a Diophantine expression by the propositions is Section 4 and in Subsection 5.4. Hence to prove the statement it is sufficient to prove the equivalence (Definition 4.10). For this purpose let T and $X_0, \dots, X_n$ be any non-negative integers. For the remainder of this proof, every k will be a non-negative integer in $\{0, \dots, T\}$ and every i will be a positive integer in $\{1, \dots, m\}$.

For the implication from right to left, suppose that we have non-negative integers $C, V, A_1, \ldots, A_m \in \mathbb{N}$ that satisfy equations (6.1) through (6.4). For every k take a prime divisor p_k of $1 + (k+1)V$. For every k and i define z_{ki}

$$z_{ki} = A_i \text{ rem } p_k.$$

We will show that the left-hand side of the claimed equivalence is satisfied by taking $Z_1 = z_{k1}, \ldots, Z_m = z_{km}$ when K is k. That is, we will show that we have $z_{ki} \leq T$ for every k and i, and $f(T, \mathcal{X}, k, z_{k1}, \ldots, z_{km}) = 0$ for every k.

Let k and i be given. By the definition of p_k, equation (6.2), and relation (6.4) we have

$$p_k \mid 1 + (k+1)V \mid 1 + CV \mid \prod_{J=0}^{T} (A_i - J).$$

Hence, since p_k is prime, there is some $J \in \{0, \ldots, T\}$ such that we have $p_k \mid A_i - J$, or equivalently $A_i \equiv J \pmod{p_k}$. We find $A_i \in \{0, \ldots, T\} + p_k\mathbb{Z}$. This yields

$$\begin{aligned} z_{ki} = A_i \text{ rem } p_k &\in (\{0, \ldots, T\} + p_k\mathbb{Z}) \cap \{0, \ldots, p_k - 1\} \\ &\subseteq \{0, \ldots, T\}, \end{aligned}$$

so that we find $z_{ki} \leq T$ as required.

We now look at the equations modulo p_k. From equation (6.2) and the definition of p_k we have

$$CV \equiv -1 \equiv (k+1)V \pmod{p_k}.$$

Noticing that V is consequently a unit modulo p_k, we obtain

$$C - 1 \equiv k \pmod{p_k}.$$

From this, the definition of the $z_{k1}, \ldots, z_{km}$ and equation (6.3) we get

$$\begin{aligned} 0 &\equiv f(T, \mathcal{X}, C-1, A_1, \ldots, A_m) \\ &\equiv f(T, \mathcal{X}, k, z_{k1}, \ldots, z_{km}) \pmod{p_k}. \end{aligned} \tag{6.5}$$

Since V is a unit modulo p_k, we have $p_k \nmid V = G(T, \mathcal{X})!$. Therefore we obtain

$$p_k > G(T, \mathcal{X}) \geq |f(T, \mathcal{X}, k, z_{k1}, \ldots, z_{km})|.$$

Together with congruence (6.5) this yields

$$f(T, \mathcal{X}, k, z_{k1}, \ldots, z_{km}) = 0.$$

This concludes the proof of the implication from right to left.

Suppose for the implication from left to right that we have non-negative integers $z_{ki} \in \mathbb{N}$ with $z_{ki} \leq T$ for every i and k. Assume furthermore that they satisfy

$$f(T, \mathcal{X}, k, z_{k1}, \ldots, z_{km}) = 0$$

for every k. Take V and C as equations (6.1) and (6.2) dictate. (Note that these equations are satisfiable.)

Since $T!$ divides $G(T, \mathcal{X})!$ we find by Lemma 4.20, that the numbers $1 + (k+1)V$ and $1 + (\ell+1)V$ are coprime for any k and ℓ in $\{0, \ldots, T\}$ with $k \neq$

ℓ. Consequently, by the Chinese Remainder Theorem, there are non-negative integers $A_1, \ldots, A_m \in \mathbb{N}$ such that we have

$$A_i \equiv z_{ki} \pmod{1 + (k+1)V}$$

for every i and k. Take such non-negative integers $A_1, \ldots, A_m \in \mathbb{N}$.

By the definition of C (equation (6.2)), we have $C - 1 \equiv k \pmod{1+(k+1)V}$ for every k. Hence we find

$$\begin{aligned} 0 &= f(T, \mathcal{X}, k, z_{k1}, \ldots, z_{km}) \\ &\equiv f(T, \mathcal{X}, C-1, A_1, \ldots, A_m) \pmod{1 + (k+1)V}. \end{aligned}$$

for every k. Since $1 + CV$ equals $\prod_{K=0}^{T}(1 + (K+1)V)$ (equation (6.2)), we have by the Chinese Remainder Theorem,

$$f(T, \mathcal{X}, C-1, A_1, \ldots, A_m) \equiv 0 \pmod{1 + CV},$$

which is congruence (6.3).

Let i be given. For every k we have modulo $1 + (k+1)v$,

$$\prod_{J=0}^{T}(A_i - J) \equiv \prod_{J=0}^{T}(z_{ki} - J) \equiv (z_{ki} - z_{ki}) \prod_{\substack{J=0,\ldots,T \\ J \neq z_{ki}}} (z_{ki} - J) \equiv 0.$$

Using the Chinese Remainder Theorem again, we find $\prod_{J=0}^{T}(A_i - J) \equiv 0 \pmod{1 + CV}$, so that the i-th term of conjunction (6.4) is satisfied. This concludes the proof of the implication from left to right. □

Corollary 6.3. *Any D-set is a Diophantine set.*

Proof. Let S be any D-set and Φ a D-expression that represents S. Use Theorem 6.2 repeatedly to eliminate all bounded universal quantifiers in Φ from right to left. The thus obtained expression Φ' contains only existential quantifiers and is therefore Diophantine. Since Φ and Φ' are equivalent, Φ' also represents S. We conclude that S is a Diophantine set. □

7 Hilbert's tenth problem is unsolvable

Theorem 7.1. *Hilbert's tenth problem is unsolvable.*

Proof. By Corollary 6.3 any D-set is a Diophantine set. By Theorem 4.23 any recursively enumerable set is a D-set and hence a Diophantine set. Theorem 2.5 states that there is a recursively enumerable set that is not recursive. Thus there is a Diophantine set that is not recursive. By Proposition 4.27, however, solvability of Hilbert's tenth problem over $\mathbb{N}$ implies that all Diophantine sets are recursive. Hence Hilbert's tenth problem is unsolvable over $\mathbb{N}$. Using Theorem 3.5 we conclude that Hilbert's original tenth problem (over $\mathbb{Z}$) is also unsolvable. □

Appendices

A An alternative definition of some expressions

Instead of the definition of polynomial equations, Diophantine expressions, D-expressions as given in Definitions 4.1 through 4.3, we can use Backus-Naur form (or a variant) to define them. This reveals some ways to generalize Hilbert's tenth problem.

Alternative Definition A.1. We define *polynomial equations*, *Diophantine expressions* and *D-expressions* by means of the following productions:

$$
\begin{aligned}
\langle\text{variable}\rangle &::= \textit{any variable symbol (from some infinite set)}\\
\langle\text{polynomial expr.}\rangle &::= \begin{cases}\textit{any polynomial expression in zero or more un-}\\ \textit{knowns (as produced by } \langle\text{variable}\rangle\textit{) with non-}\\ \textit{negative integral coefficients.}\end{cases}\\
\langle\text{polynomial eq.}\rangle &::= \langle\text{polynomial expr.}\rangle = \langle\text{polynomial expr.}\rangle\\
\langle\text{Diophantine expr.}\rangle &::= \langle\text{polynomial eq.}\rangle\\
&\quad\mid (\exists\langle\text{variable}\rangle)\ \langle\text{Diophantine expr.}\rangle\\
\langle\text{D-expr.}\rangle &::= \langle\text{polynomial eq.}\rangle\\
&\quad\mid (\exists\langle\text{variable}\rangle)\ \langle\text{D-expr.}\rangle\\
&\quad\mid (\forall\langle\text{variable}\rangle \le \langle\text{variable}\rangle)\ \langle\text{D-expr.}\rangle.
\end{aligned}
$$

We remark that we could also use a more formal production than the one given above for $\langle\text{polynomial expr.}\rangle$, for instance

$$
\begin{aligned}
\langle\text{polynomial expr.}\rangle ::= {} & 0 \mid 1 \mid \langle\text{variable}\rangle\\
& \mid (\langle\text{polynomial expr.}\rangle + \langle\text{polynomial expr.}\rangle)\\
& \mid (\langle\text{polynomial expr.}\rangle \cdot \langle\text{polynomial expr.}\rangle).
\end{aligned}
$$

A critical reader will note that if we use this new production in Alternative Definition A.1, different sets of expressions are defined. For instance x^2 would no longer be generated from $\langle\text{polynomial expr.}\rangle$. This difference does not matter for our purposes, as we can always obtain an expression of the same meaning. Instead of x^2 we could for instance write $(x \cdot x)$ or even $((1 \cdot x) \cdot (x + 0))$.

The alternative set of productions for $\langle\text{polynomial expr.}\rangle$

$$
\begin{aligned}
\langle\text{constant}\rangle &::= 0 \mid 1\\
\langle\text{operator}\rangle &::= + \mid \cdot\\
\langle\text{polynomial expr.}\rangle &::= \langle\text{constant}\rangle \mid \langle\text{variable}\rangle\\
&\quad \mid (\langle\text{polynomial expr.}\rangle\langle\text{operator}\rangle\langle\text{polynomial expr.}\rangle)
\end{aligned}
$$

is particularly interesting, as it suggests a way to create analogies of Hilbert's tenth problem, namely by redefining the allowed constants and operators. This is also the way that Hilbert's tenth problem was historically shown to be unsolvable: First it was first proven around 1960 by Martin Davis, Hilary Putnam, and Julia Robinson [4] that Hilbert's tenth problem is unsolvable *if one adds* *an exponentiation operator.* Then in 1970 it was shown by Yuri V. Matiyasevich that exponentiation can be expressed without introducing this operator (Proposition 5.21).

It is also possible to allow different relations than just equality, or to allow the logical symbols $\vee$ (or) and $\wedge$ (and). See for instance Esther Bod's Master's thesis [1] for some explorations of such modifications. (Proposition 4.8 states that in our case we can express $\vee$ and $\wedge$ without introducing these symbols explicitly.)

B The Pell equation

The equation

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

is known as the *Pell equation.* Typically it is to be solved with $x \in \mathbb{N}$ and $y \in \mathbb{N}$ non-negative integers for given non-square positive integer $d \in \mathbb{Z}^+$. We will see that it is particularly easy to find these solutions if we have $d = a^2 - 1$ for given integer $a \in \mathbb{Z}$ with $a \geq 2$. In Section 5 this special case is used to construct a Diophantine function that is roughly exponential.

For the duration of this section fix any positive integer $d \in \mathbb{Z}^+$ such that the Pell equation has at least one solution in $\mathbb{Z}^+ \times \mathbb{Z}^+$.

Proposition B.1. *The positive integer $d \in \mathbb{Z}^+$ is not a square.*

Proof. Suppose for the sake of contradiction that $d \in \mathbb{Z}^+$ is square. Now let $b \in \mathbb{Z}$ be a square root of d. For any solution $(x, y) \in \mathbb{Z}^2$ we find

$$1 = x^2 - dy^2 = x^2 - (by)^2 = (x + by)(x - by).$$

We obtain $x + by = x - by = \pm 1$. This yields $y = 0$ and $x = \pm 1$. Hence there are no solutions in $\mathbb{Z}^+ \times \mathbb{Z}^+$. □

Remark B.2. In fact any non-square positive integer meets the requirements on d. We will, however, not prove this as we will not need this result.

For the remainder of this section we will identify $\mathbb{Z}[X]/(X^2 - d)$ and $\mathbb{Z}[\sqrt{d}]$ through the isomorphism $x + yX \mapsto x + y\sqrt{d}$ of Corollary C.19. This gives us the norm map $N : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}$ of $\mathbb{Z}[\sqrt{d}]$ over $\mathbb{Z}$ given by $x + y\sqrt{d} \mapsto x^2 - dy^2$ (Definition C.14) and the conjugation automorphism $\bar{\cdot} : \mathbb{Z}[\sqrt{d}] \to \mathbb{Z}[\sqrt{d}]$ given by $x + y\sqrt{d} \mapsto x - y\sqrt{d}$ (Definition C.12).

Corollary B.3. *The pair of integers $(x, y) \in \mathbb{Z}^2$ is a solution to the Pell equation if and only if $x + y\sqrt{d} \in \mathbb{Z}[\sqrt{d}]$ is of norm 1.*

Definition B.4 (fundamental solution)**.** We call the solution of the Pell equation in $\mathbb{Z}^+ \times \mathbb{Z}^+$ with minimal first coordinate the *fundamental solution.*

Proposition B.5. *Let $a \in \mathbb{Z}^+$ be any positive integer. Suppose that we have $d = a^2 - 1$, or equivalently that $(a, 1)$ is a solution to the Pell equation. Then $(a, 1)$ is the fundamental solution.*

Proof. Let $(x, y) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ be the fundamental solution. Since we have $y \geq 1$, we have $x^2 = dy^2 + 1 \geq (a^2 - 1) \cdot 1 + 1 = a^2$. Hence, $(a, 1)$ is the fundamental solution. □

Definition B.6 (α)**.** Let $(x, y) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ be the fundamental solution. Define $\alpha \in \mathbb{Z}[\sqrt{d}] \subset \mathbb{R}$ by $\alpha = x + y\sqrt{d}$.

Corollary B.7. *We have $N(\alpha) = 1$ and $\alpha^{-1} = \overline{\alpha} \in \mathbb{Z}[\sqrt{d}]$.*

Proof. This follows by Definition B.6 of α and Corollaries B.3 and C.15. □

Proposition B.8. *Let α be as in Definition B.6. Let $n \in \mathbb{Z}$ be any integer. Then there are unique integers $x_n, y_n \in \mathbb{Z}$ with $x_n + y_n\sqrt{d} = \alpha^n$.*

Proof. Since α is a unit in $\mathbb{Z}[\sqrt{d}]$ (Corollary B.7), α^n is well-defined. By writing α^n on the $\mathbb{Z}$-basis $\{1, \sqrt{d}\}$ of $\mathbb{Z}[\sqrt{d}]$, we find unique integers $x_n, y_n \in \mathbb{Z}$ subject to the equality (Corollaries C.2 and C.19). □

Definition B.9 (x_n, y_n)**.** With α from Definition B.6, we define the two maps

$$\begin{aligned} x_. &: \mathbb{Z} \to \mathbb{Z}, \quad n \mapsto x_n, \\ y_. &: \mathbb{Z} \to \mathbb{Z}, \quad n \mapsto y_n \end{aligned}$$

by the identity

$$x_n + y_n\sqrt{d} = \alpha^n.$$

Corollary B.10. *The fundamental solution is (x_1, y_1). For any integer $n \in \mathbb{Z}$ we have $x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n$.*

Proposition B.11. *For any integer $n \in \mathbb{Z}$, the pair (x_n, y_n) is a solution to the Pell equation.*

Proof. By multiplicativity of the norm N (Definition C.14), we have for each integer $n \in \mathbb{Z}$,

$$N(x_n + y_n\sqrt{d}) = N(\alpha^n) = N(\alpha)^n = 1^n = 1.$$

Now each pair (x_n, y_n) is a solution to the Pell equation (Corollary B.3). □

Proposition B.12. *For each integer $n \in \mathbb{Z}$ we have $2x_n = \alpha^n + \alpha^{-n}$ and $2y_n\sqrt{d} = \alpha^n - \alpha^{-n}$.*

Proof. Let $n \in \mathbb{Z}$ be any integer. We have $\alpha^{-1} = \overline{\alpha}$ (Corollary B.7) and since conjugation is a ring automorphism (Definition C.12), we obtain $\alpha^{-n} = (\alpha^{-1})^n = \overline{\alpha}^n = \overline{\alpha^n}$. We find

$$\alpha^n + \alpha^{-n} = \alpha^n + \overline{\alpha^n} = x_n + y_n\sqrt{d} + (x_n - y_n\sqrt{d}) = 2x_n, \text{ and}$$

$$\alpha^n - \alpha^{-n} = \alpha^n - \overline{\alpha^n} = x_n + y_n\sqrt{d} - (x_n - y_n\sqrt{d}) = 2y_n\sqrt{d}. \qquad □$$

Corollary B.13. *The map $x_.$ is even while the map $y_.$ is odd.*

Proof. For any integer $n \in \mathbb{Z}$ we find

$$x_n = \tfrac{1}{2}(\alpha^n + \alpha^{-n}) = x_{-n}, \quad \text{and} \quad y_n = \tfrac{1}{2\sqrt{d}}(\alpha^n - \alpha^{-n}) = -y_{-n}. \qquad □$$

Corollary B.14. *For any integer $n \in \mathbb{Z}$ we have $x_n > 0$ and $\operatorname{sgn} y_n = \operatorname{sgn} n$.*

Proof. We have $\alpha = x_1 + y_1\sqrt{d} \geq 1 + \sqrt{d} > 1$ (Definition B.6). Hence we obtain the result for $x.$ by $2x_n = \alpha^n + \alpha^{-n} \geq 1 + 0 > 0$ for any integer $n \in \mathbb{Z}$.

For the result about $y.$, notice that for any positive integer $n \in \mathbb{Z}^+$ we have $\alpha^n > 1 > \alpha^{-n}$. This gives $2y_n\sqrt{d} = \alpha^n - \alpha^{-n} > 0$ and we find the result since the map $y.$ is odd. □

Proposition B.15. *The maps*

$$\begin{aligned} &\mathbb{N} \to \mathbb{N}, \quad n \mapsto x_n, \\ &\mathbb{Z} \to \mathbb{Z}, \quad n \mapsto y_n, \ \textit{and} \\ &\mathbb{Z} \to \mathbb{R}, \quad n \mapsto x_n + y_n\sqrt{d} = \alpha^n \end{aligned}$$

are strictly increasing.

Proof. Notice that the respective maps are restrictions to the set of (non-negative) integers of the respective differentiable $\mathbb{R} \to \mathbb{R}$-maps

$$z \mapsto \tfrac{1}{2}(\alpha^z + \alpha^{-z}), \qquad z \mapsto \tfrac{1}{2\sqrt{d}}(\alpha^z - \alpha^{-z}), \qquad \text{and} \qquad z \mapsto \alpha^z.$$

These maps have derivatives

$$z \mapsto \tfrac{\ln\alpha}{2}(\alpha^z - \alpha^{-z}), \qquad z \mapsto \tfrac{\ln\alpha}{2\sqrt{d}}(\alpha^z + \alpha^{-z}), \qquad \text{and} \qquad z \mapsto \alpha^z \ln\alpha.$$

Noticing $\alpha = x_1 + y_1\sqrt{d} \geq 1 + \sqrt{d} > 1$, we have $\ln\alpha > 0$. Consequently the derivatives are positive on $(0,\infty)$, $(-\infty,\infty)$, and $(-\infty,\infty)$ respectively. Hence the result follows. □

Corollary B.16. *For each non-negative integer $n \in \mathbb{N}$ we have $y_n \geq n$.*

Proof. Since $y.$ is strictly increasing, we have $y_n \geq y_0 + n = n$. □

Lemma B.17. *Let $(x,y) \in \mathbb{Z}^2$ be a solution to the Pell equation. For the sets $X, Y \subset \mathbb{Z}$ and $Z \subset \mathbb{R}$ below we have $(x,y) \in X \times Y$ if and only if we have $x + y\sqrt{d} \in Z$.*

Z	$Y = \mathbb{Z}^-$	$Y = \{0\}$	$Y = \mathbb{Z}^+$
$X = \mathbb{Z}^-$	$(-\infty,-1)$	$\{-1\}$	$(-1,0)$
$X = \{0\}$	$\emptyset$	$\emptyset$	$\emptyset$
$X = \mathbb{Z}^+$	$(0,1)$	$\{1\}$	$(1,\infty)$.

Proof. Notice that it is sufficient to prove the forward implications. Since $x = 0$ implies the contradiction $1 = x^2 - dy^2 \leq 0$, the implications for $X = \{0\}$ are true. The case $(X,Y) = (\mathbb{Z}^+,\mathbb{Z}^+)$ follows since we have $x + y\sqrt{d} \geq 1 + \sqrt{d} > 1$ for $x, y \geq 1$. The cases for $Y = \{0\}$ follow since we then have $1 = x^2 - dy^2 = x^2$, which implies $x \in \{\pm 1\}$.

For the case $(X,Y) = (\mathbb{Z}^+,\mathbb{Z}^-)$, let $(x,y) \in \mathbb{Z}^+ \times \mathbb{Z}^-$ be a solution to the Pell equation with x positive and y negative. Notice that $(x,-y)$ is also a solution so that we have $x - y\sqrt{d} \in (1,\infty)$ by the case $(X,Y) = (\mathbb{Z}^+,\mathbb{Z}^+)$. We obtain $x + y\sqrt{d} = 1/(x - y\sqrt{d}) \in 1/(1,\infty) = (0,1)$ by Corollary C.16.

Finally, the cases with $X = \mathbb{Z}^-$ follow from the cases with $X = \mathbb{Z}^+$, as for any solution $(x,y) \in \mathbb{Z}^- \times \mathbb{Z}$ the pair $(-x,-y) \in \mathbb{Z}^+ \times \mathbb{Z}$ is also a solution. □

Lemma B.18. *The only solution* $(x, y) \in \mathbb{Z}^2$ *to the Pell equation with* $1 \leq x + y\sqrt{d} < \alpha$ *is* $(x, y) = (1, 0)$.

Proof. Since $(1, 0)$ is indeed a solution to the Pell equation satisfying the condition, it is sufficient to check that no other solution satisfies the inequality. In the light of Lemma B.17 it is even sufficient to only check this for positive solutions.

Notice that for any positive solution $(x, y) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ to the Pell equation we have

$$x + y\sqrt{d} = x + \sqrt{\tfrac{1}{d}(x^2 - 1)}\sqrt{d} = x + \sqrt{x^2 - 1}.$$

Now suppose for the sake of contradiction that $(x, y) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ is a solution with $1 \leq x + \sqrt{x^2 - 1} < \alpha$. Since

$$f : \mathbb{Z}^+ \to \mathbb{R}, \quad x \mapsto x + \sqrt{x^2 - 1}$$

is strictly increasing and because we have $1 = f(1)$ and $\alpha = f(x_1)$, we obtain $1 \leq x < x_1$. This contradicts the definition of fundamental solution (x_1, y_1) (Definition B.4 and Corollary B.10). □

Proposition B.19. *If* $(x, y) \in \mathbb{N} \times \mathbb{Z}$ *is a solution to the Pell equation with* x *non-negative, then there is an integer* $n \in \mathbb{Z}$ *such that this solution is* (x_n, y_n).

Proof. Notice that we can ignore the case $x = 0$, as $x = 0$ would imply the false assertion $1 = x^2 - dy^2 \leq 0$. Now let $(x, y) \in \mathbb{Z}^+ \times \mathbb{Z}$ be a solution to the Pell equation with x positive. Notice that for every integer $n \in \mathbb{Z}$ we have $(x, y) = (x_n, y_n)$ if and only if we have $x + y\sqrt{d} = x_n + y_n\sqrt{d} = \alpha^n$ (Corollary C.2).

By Lemma B.17 we have $x + y\sqrt{d} \in (0, \infty)$. Now since the function $\mathbb{Z} \to \mathbb{R}, n \mapsto \alpha^n$ is strictly increasing (Proposition B.15) with $\lim_{n\to-\infty} \alpha^n = 0$ and $\lim_{n\to\infty} \alpha^n = \infty$, there is a unique integer $n \in \mathbb{Z}$ such that we have

$$\alpha^n \leq x + y\sqrt{d} < \alpha^{n+1}. \tag{$\star$}$$

Take this $n \in \mathbb{Z}$.

To see that this n suffices, look at $(x', y') \in \mathbb{Z}^2$ given by $x' + y'\sqrt{d} = \alpha^{-n}(x + y\sqrt{d}) \in \mathbb{Z}[\sqrt{d}]$ (Corollaries B.7, C.2, and C.19). By multiplicativity of the norm (Definition C.14) we have

$$N(x' + y'\sqrt{d}) = N(\alpha)^{-n} \cdot N(x + y\sqrt{d}) = 1^{-n} \cdot 1 = 1,$$

so that (x', y') is a solution to the Pell equation (Corollary B.3). Multiplying inequality ($\star$) by $\alpha^{-n} > 0$ we find $1 \leq x' + y'\sqrt{d} < \alpha$. This yields $(x', y') = (1, 0)$ by Lemma B.18. We find $1 = x' + y'\sqrt{d} = \alpha^{-n}(x + y\sqrt{d})$, so that we have

$$x + y\sqrt{d} = \alpha^n,$$

as required. □

Corollary B.20. *The solution set in* $\mathbb{N} \times \mathbb{Z}^+$ *of the Pell equation is precisely*

$$\{(x_n, y_n) : n \in \mathbb{Z}^+\}.$$

In $\mathbb{N} \times \mathbb{Z}$ *it is precisely* $\{(x_n, y_n) : n \in \mathbb{Z}\}$.

Proof. The latter statement follows from Propositions B.11 and B.19. For the former statement notice that the solutions $(x, y) \in \mathbb{N} \times \mathbb{Z}^+$ are precisely those with $x + y\sqrt{d} > 1$ (Lemma B.17). Since $\mathbb{Z} \to \mathbb{R}, n \mapsto \alpha^n = x_n + y_n\sqrt{d}$ maps 0 to 1 and is strictly increasing (Proposition B.15), we obtain from the latter statement that the solutions in $\mathbb{N}^2$ are precisely the pairs (x_n, y_n) with n running through the set of positive integers $\mathbb{Z}^+$. □

C Some algebra

In our analysis of solutions of the Pell equation we use the ring $\mathbb{Z}[X]/(X^2 - d)$ (with $d \in \mathbb{Z}^+$ not a square). In this appendix we develop some theory on this ring and related ones.

Proposition C.1. *Let R be a commutative ring and $f \in R[X]$ a monic polynomial with coefficients in R. Then $R[X]/(f)$ has R-basis*

$$B = \{X^0, \ldots, X^{\deg f - 1}\}.$$

Proof. Let $g \in R[X]$ be any polynomial. Because of division with remainder in the polynomial ring $R[X]$ there are unique polynomials $q, r \in R[X]$ with $\deg r < \deg f$ and $g = qf + r$ [10, Stelling 12.1]. In other words, every coset $g + (f)$ in $R[X]/(f)$ has a unique representative $r \in R[X]$ of strictly smaller degree than f.

Existence of this representative yields that $R[X]/(f)$ is spanned by B. By uniqueness we have that B is R-linearly independent: Suppose that we have λ_0 through $\lambda_{\deg f - 1}$ in R with $\sum_{i=0}^{\deg f - 1} \lambda_i X^i + (f) = 0 + (f)$. Then the polynomials $\sum_{i=0}^{\deg f - 1} \lambda_i X^i$ and 0 represent the same coset and are therefore the same. Hence each λ_i equals zero. We now have that B is an R-basis for $R[X]/(f)$. □

Corollary C.2. *For any commutative ring R and any element d of R, the ring $R[X]/(X^2 - d)$ has R-basis $\{1, X\}$.*

Proposition C.3. *A ring homomorphism $\phi : R \to S$ between rings R and S, induces a ring homomorphism*

$$R[X] \to S[X], \qquad \sum_i c_i X^i \mapsto \sum_i \phi(c_i) X^i$$

between the polynomial rings of R and S.

Proof. It is straightforward to check that addition, multiplication, and the unit element are preserved. Hence the map is indeed a ring homomorphism. □

Definition C.4. We will use the same symbol for the induced map in Proposition C.3 as we use for the original map.

Proposition C.5. *Let $\phi : R \to S$ be a ring homomorphism between commutative rings R and S. Let $I \subseteq R$ be an ideal of R and $J \subseteq S$ an ideal of S containing $\phi(I)$. Then ϕ induces the ring homomorphism*

$$R/I \to S/J, \qquad r + I \mapsto \phi(r) + J.$$

Proof. Consider the ring homomorphism $\phi' : R \to S/J$ given by $r \mapsto \phi(r) + J$ obtained by composing ϕ with the canonical ring homomorphism $S \to S/J$. This homomorphism has kernel $\ker \phi' = \phi'^{-1}(\{0\}) = \phi^{-1}(J) \subseteq R$. Through $\phi(I) \subseteq J$ we obtain $I \subseteq \ker \phi'$. The result follows [10, Homomorfiestelling 11.17]. □

Definition C.6. We will use the same symbol for the induced map in Proposition C.5 as we use for the original map.

Corollary C.7. *Let $\phi : R \to S$ be a ring homomorphism between commutative rings R and S and let $f \in R[X]$ be a monic polynomial. Then the kernel of the ring homomorphism $R[X]/(f) \to S[X]/(\phi(f))$ induced by ϕ is $\ker \phi \cdot (R[X]/(f))$.*

Proof. Notice that $\phi(f)$ is a monic polynomial of the same degree as f. Hence the rings $R[X]/(f)$ and $S[X]/(\phi(f))$ both have $\{X^0, \dots, X^{n-1}\}$ as a basis over their respective base rings. Since the induced map is given by $\sum_{i=0}^{n-1} r_i X^i \mapsto \sum_{i=0}^{n-1} \phi(r_i) X^i$, its kernel is

$$\sum_{i=0}^{n-1} \ker \phi \cdot X^i = \sum_{i=0}^{n-1} \ker \phi \cdot R \cdot X^i = \ker \phi \cdot \sum_{i=0}^{n-1} R X^i = \ker \phi \cdot R[X]/(f). \quad \square$$

Corollary C.8. *If $\phi : R \to S$ is a ring homomorphism between commutative rings R and S and $d \in R$ is an element of R, then the map $R[X]/(X^2 - d) \to S[X]/(X^2 - \phi(d))$ given by $x + yX \mapsto \phi(x) + \phi(y)X$ is a ring homomorphism, with kernel $\ker \phi \cdot R[X]/(X^2 - d)$.*

Corollary C.9. *The map $R/I \to S/J$ from Definition C.6 induced by ϕ has kernel $\phi^{-1}(J)/I$ and image $\phi(R)/(J \cap \phi(R))$.*

Proof. This is immediate from Definition C.6 of the induced map. □

Corollary C.10. *If $\phi : R \to S$ is a ring isomorphism between commutative rings and if $I \subseteq R$ is an ideal of R, then $\phi(I)$ is an ideal of S and the ring homomorphism $R/I \to S/\phi(I)$ induced by ϕ is an isomorphism.*

Proof. The set $\phi(I)$ is the preimage of the ideal I under the ring homomorphism ϕ^{-1}. Hence it is an ideal. The induced ring homomorphism is injective and surjective by Corollary C.9. □

Proposition C.11. *For any commutative ring R and any element d of R, the map $\bar{\cdot} : R[X]/(X^2 - d) \to R[X]/(X^2 - d)$ given by*

$$x + yX \mapsto x - yX$$

is an automorphism of the ring $R[X]/(X^2 - d)$.

Proof. Apply Corollary C.10 to the automorphism $f \mapsto f(-X)$ of $R[X]$ and the ideal $(X^2 - d)$. □

Definition C.12 (Conjugate)**.** We call the map from Proposition C.11 the conjugation automorphism. For an element $f \in R[X]/(X^2 - d)$ we call $\overline{f}$ the conjugate of f.

Proposition C.13. *For any commutative ring R and any element d of R, the map $N : R[X]/(X^2 - d) \to R$ given by*

$$x + yX \mapsto x^2 - dy^2,$$

is multiplicative.

Proof. For any element $f = x + yX$ of $R[X]/(X^2 - d)$, we have

$$f \cdot \overline{f} = (x + yX) \cdot (x - yX) = x^2 - dy^2 = N(f).$$

Now, since $f \mapsto f$ and $f \mapsto \overline{f}$ are ring homomorphisms (Definition C.12), they are multiplicative. Hence so is $f \mapsto N(f)$. □

Definition C.14 (Norm)**.** We call the map $N : R[X]/(X^2 - d) \to R$ from Proposition C.13 the norm map of $R[X]/(X^2 - d)$ over R.

Corollary C.15. *Let R be any commutative ring, d an element of R and f an element of $R[X]/(X^2 - d)$. If $N(f)$ is a unit in R, then f has multiplicative inverse $f^{-1} = \overline{f}/N(f)$.*

Proof. We have $f\overline{f} = N(f)$. The result follows. □

Corollary C.16. *Let $d, x, y \in \mathbb{Z}$ be any three integers. If we have $x^2 - dy^2 = 1$, then the element $x + yX$ of $\mathbb{Z}[X]/(X^2 - d)$ has multiplicative inverse $x - yX$.*

Corollary C.17. *Let R be any commutative ring and let d be any element of R. Let $(R[X]/(X^2 - d))^*$ and R^* be the groups of units of $R[X]/(X^2 - d)$ and R respectively. Then we have*

$$(R[X]/(X^2 - d))^* = N^{-1}(R^*).$$

Proof. The inclusion "$\subseteq$" is immediate from Corollary C.15. Now let $f \in (R[X]/(X^2 - d))^*$ be any unit in $R[X]/(X^2 - d)$. We have $N(f)N(f^{-1}) = N(ff^{-1}) = N(1) = 1$ and therefore $N(f) \in R^*$. We see $(R[X]/(X^2 - d))^* \subseteq N^{-1}(R^*)$ as required. □

Proposition C.18. *For any positive integer $d \in \mathbb{Z}^+$ that is not a square in $\mathbb{Z}$, the map $\mathbb{Z}[X]/(X^2 - d) \to \mathbb{R}$ given by*

$$x + yX \mapsto x + y\sqrt{d}$$

is an injective ring homomorphism.

Proof. Suppose, for the sake of contradiction, that $X^2 - d$ is reducible in $\mathbb{Q}[X]$. Then we have $(X - \alpha)(X + \alpha) = X^2 - \alpha^2 = X^2 - d$ for some $\alpha \in \mathbb{Q}$. By the Lemma of Gauss [10, Lemma 13.5], we have $\alpha \in \mathbb{Z}$. Hence $d = \alpha^2$ is a square in $\mathbb{Z}$, which is false. We conclude that $X^2 - d$ is irreducible, so that $\mathbb{Q}[X]/(X^2 - d)$ and $\mathbb{Q}(\sqrt{d}) \subset \mathbb{R}$ are isomorphic through $x + yX \mapsto x + y\sqrt{d}$ [10, Stelling 21.5]. Composing with the injective ring homomorphism $\mathbb{Z}[X]/(X^2 - d) \to \mathbb{Q}[X]/(X^2 - d)$, $x + yX \mapsto x + yX$ induced by the inclusion $\mathbb{Z} \subset \mathbb{Q}$ (Corollary C.8), we obtain the claimed assertion. □

Corollary C.19. *If $d \in \mathbb{Z}^+$ is a positive integer that is not a square, then the rings $\mathbb{Z}[X]/(X^2 - d)$ and $\mathbb{Z}[\sqrt{d}] \subset \mathbb{R}$ are isomorphic through $x + yX \mapsto x + y\sqrt{d}$.*

Proof. The image of the injective ring homomorphism from Proposition C.18 is $\mathbb{Z}[\sqrt{d}] \subset \mathbb{R}$. □

D List of notation

Notation	Meaning
$\wedge$, $\vee$, $\longrightarrow$, $\longleftrightarrow$	The logical connectives ‘and’, ‘or’, ‘implies’, and ‘if and only if’, respectively
$\neg$	The logical symbol for ‘not’
$\exists$	The existential quantifier
$\forall$	The universal quantifier
$\iff$	Logical equivalence from Definition 4.4
$\circ$	Function composition
$\mid$	Divisibility relation
$\leftarrow$	The assignment operator (in pseudocode)
$\lfloor\cdot\rfloor : \mathbb{R} \to \mathbb{Z}$	The floor function $x \mapsto \max\{y \in \mathbb{Z} : y \leq x\}$
$\lceil\cdot\rceil : \mathbb{R} \to \mathbb{Z}$	The ceiling function $x \mapsto \min\{y \in \mathbb{Z} : y \geq x\}$
$[\cdot] : \mathbb{R} \to \mathbb{Z}$	The function that rounds to the nearest integer and upwards in case of ties: $x \mapsto \lfloor x + \frac{1}{2} \rfloor$
$[\cdot] : R \to R/I$	The canonical ring homomorphism $R \to R/I$ where I is an ideal of R. (Always introduced in the text.)
$\mathbb{C}$	The field of complex numbers
f^+, f^-	Polynomials from Proposition 3.1 with non-negative coefficients so constructed that we have $f = f^+ - f^-$
$\gcd : \mathbb{Z}^2 \to \mathbb{N}$	The function mapping (x, y) to the non-negative generator of $x\mathbb{Z} + y\mathbb{Z}$
$\mathrm{gd} : \mathbb{N}^2 \times \mathbb{Z}^+ \to \mathbb{N}$	The function $\mathrm{gd}(u, v\,; i) = u \text{ rem } (1 + iv)$ from Definition 4.19
im	The image of a function
$\mathrm{lcm} : \mathbb{Z}^2 \to \mathbb{N}$	The function mapping (x, y) to the non-negative generator of $x\mathbb{Z} \cap y\mathbb{Z}$
$\mathbb{N}$	The set of non-negative integers $\{x \in \mathbb{Z} : x \geq 0\}$
$\mathcal{P}(X)$	The power set of X, that is, the set $\{\text{subsets of } X\}$
$\mathbb{Q}$	The field of rational numbers
$\mathbb{R}$	The field of real numbers
rem	The operator given by $x \text{ rem } y = x - \lfloor x/y \rfloor y$ that maps (x, y) to the non-negative remainder of the division x/y
$R[X]$	The polynomial ring in unknown X of the ring R
$\mathrm{sgn} : \mathbb{R} \to \{-1, 0, 1\}$	The sign function $x \mapsto \begin{cases} 0 & \text{if } x = 0 \\ x/\lvert x\rvert & \text{if } x \neq 0 \end{cases}$
$x., y. : \mathbb{Z} \to \mathbb{Z}$	Functions from Definition B.9 enumerating solutions to the Pell equation $x^2 - dy^2 = 1$
$x.(\cdot), y.(\cdot) :$ $\mathbb{Z} \times \mathbb{Z}_{\geq 2} \to \mathbb{Z}$	Functions from Definition 5.15 enumerating solutions to the Pell equation $x^2 + (a^2 - 1)y^2$ where a is the bracketed argument to the function
$\mathbb{Z}$	The set of (rational) integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{Z}^+$	The set of positive integers $\{x \in \mathbb{Z} : x > 0\}$
$\mathbb{Z}^-$	The set of negative integers $\{x \in \mathbb{Z} : x < 0\}$
$\mathbb{Z}[\sqrt{d}]$	The subring of $\mathbb{C}$ generated by 1 and $\sqrt{d} \in \mathbb{C}$

Table D.1: List of notation and meaning

References

[1] Esther Bod. Hilbert's tenth problem and some generalizations. Master's thesis, Utrecht University, January 2009.

[2] Martin Davis. Arithmetical problems and recursively enumerable predicates. *The Journal of Symbolic Logic*, 18(1):33–41, March 1953.

[3] Martin Davis. Hilbert's tenth problem is unsolvable. *The American Mathematical Monthly*, 80(3):233–269, March 1973.

[4] Martin Davis, Hilary Putnam, and Julia Robinson. The decision problem for exponential Diophantine equations. *The Annals of Mathematics*, 74(3):425–436, November 1961.

[5] David Hilbert. Mathematische Probleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 3:253–297, 1900.

[6] Yu. I. Manin. *A Course in Mathematical Logic for Mathematicians.* Graduate Texts in Mathematics. Springer, second edition, 2010.

[7] John C. Martin. *Introduction to Languages and the Theory of Computation.* McGraw-Hill, third edition, 2003.

[8] Julia Robinson. General recursive functions. *Proceedings of the American Mathematical Society*, 1(6):703–718, December 1950.

[9] Julia Robinson. Existential definability in arithmetic. *Transactions of the American Mathematical Society*, 72(3):437–449, May 1952.

[10] Peter Stevenhagen. Algebra 1–3. `http://websites.math.leidenuniv.nl/algebra/`, 2011.

[11] Yuri V. Matiyasevich. *Hilbert's Tenth Problem.* Foundations of Computing. The MIT Press, 1993.