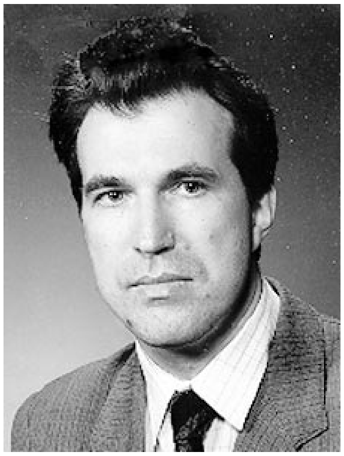
Yuri Matiyasevich



Professor Yuri Matiyasevich was born on March 2, 1947 in Leningrad, the USSR. In 1969 he graduated from *Department of Mathematics and Mechanics* of *Leningrad State University* and continued his study as post-graduate student at the *Steklov Institute of Mathematics, Leningrad Branch (LOMI).* From 1970 till now he works in this Institute, currently as the Head of the Laboratory of Mathematical Logic.

The name of Yuri Matiyasevich became known worldwide in 1970 when he completed the last missing step in the "negative solution" of *Hilbert's tenth problem.* His book on this subject, originally published in Russian, was translated into English (The MIT Press, 1993) and French (Masson, 1995).

Yuri Matiyasevich is *Docteur Honoris Causa* de l'Université d'Auvergne, France (1996) and *Correspondent Member* of the Russian Academy of Sciences (1997).

Hilbert's Tenth Problem: A Two-way Bridge between Number Theory and Computer Science

Historical Background

Number Theory is one of the oldest branches of mathematics, which itself is one of the eldest among the sciences. In contrast, computer science is relatively young. However, it did not emerge on empty ground. Computer Science takes its stems from mathematics, more specifically, from mathematical logic. So if we want to trace the origin of computer science we need to look into the history of mathematics.

Elements of the future computer science can be found at very early stages of the development of mathematics. The word "algorithm" was derived from the name of great arab scientist Al-Khorezmi who lived in the 9th century A. C. However, the idea of an algorithm has a longer history. Every mathematician knows *Euclid's algorithm* for finding the greatest common divisor of two natural numbers. Euclid lived in the 4th century B. C. but the basic algorithms for performing arithmetical operations were invented much before Euclid.

Hilbert's Tenth Problem

Mathematics is a science to a great extent driven by problems. For some of them it took decades of years to be solved.

In 1900 scientists from many countries gathered together in Paris for the *Second International Congress of Mathematicians.* One of the invited lectures was to be delivered by the great German mathematician David Hilbert. It was the last year of the passing century, and he decided to survey the most important, in his opinion, open problems in mathematics which the pending 20th century was to inherit from its predecessor. Hilbert's famous paper *Mathematische Problemen* [18] lists 23 problems (in fact, most of them are collections of related problems).

One, and only one, of these 23 problems can be recognized today as a problem in computer science. This is the tenth problem. The section of Hilbert's paper devoted to this problem is so short that can be reproduced here entirely.

> **10. Entscheidung der Lösbarkeit einer diophantischen Gleichung.**
>
> Eine diophantische Gleichung mit irgendwelchen Unbekannten und mit ganzen rationalen Zahlkoefficienten sei vorgelegt: *man soll ein Verfahren angeben, nach welchen sich mittels einer endlichen Anzahl von Operationen entscheiden läßt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist.*[1]

Of course, Hilbert and his contemporaries viewed this problem as a problem in number theory.

The range of unknowns. Solving equations was a major occupation for mathematicians from ancient times. The equations mentioned in the tenth problem were named after the Greek mathematician Diophantus who lived in the 3rd century A. C. They have the form

$$D(x_1, \ldots, x_m) = 0, \tag{1}$$

[1] **10. Determination of the Solvability of a Diophantine Equation.** Given a diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: *To devise a process according to which it can be determined by a finite number of operations whether the equation is solvable in rational integers.*

where D is a polynomial with integer coefficients. The novelty which Diophantus introduced into the study of polynomial equations was as follows: he restricted the range of admissible values of unknowns to (positive) rational numbers. So for Diophantus the equation $x^2 = 2$ had no solution which contrasted with the previous tradition of solving equations in (positive) real numbers geometrically (represented by segments of the straight line).

In contrast with Diophantus, Hilbert—in the tenth problem—asks about solving Diophantine equations in *"rational integers"*. This terminology may sound strange and misleading to computer scientists. In fact, Hilbert had in mind nothing else but the familiar to everyone integers $0, \pm 1, \pm 2, \ldots$. He used the name *rational integers* because the term *integers* can be understood in a broader sense of *algebraic integers*. While the latter are very interesting and important objects, they are entirely outside of the scope of this paper, and we will use the word *integer* in its "elementary school sense" without risking a terminology confusion.

Very often the range of unknowns is further restricted to *natural numbers*. By the latter one understands either positive integers 1,2,3,... (tradition of number theory) or non-negative integers 0,1,2,... (tradition of mathematical logic). In this paper, we will adhere to the number-theoretical understanding of natural numbers.

Considering Diophantine equations only with unknowns in natural numbers in fact does not restrict the generality of results. It is easy to see that solving the equation

$$D(z_1, \ldots, z_m) = 0 \tag{2}$$

in integers is equivalent to solving equation

$$D(x_1 \Leftrightarrow y_1, \ldots, x_m \Leftrightarrow y_m) = 0 \tag{3}$$

in natural numbers. Reduction in the other direction is a bit less evident: solving an equation

$$D(x_1, \ldots, x_m) = 0 \tag{4}$$

in natural numbers is equivalent to solving the equation

$$D(s_1^2 + t_1^2 + u_1^2 + v_1^2 + 1, \ldots, s_m^2 + t_m^2 + u_m^2 + v_m^2 + 1) = 0 \tag{5}$$

in integers because every positive integer is the sum of the squares of four integers. Technically, it is often more convenient to work with natural numbers only, and from now on we restrict the range of unknowns to positive integers.

Sometimes the range of unknowns is considered as the only distinctive feature of Diophantine equations and then one calls "Diophantine" an equation of arbitrary form as soon as only integer or natural number solutions of it are of interest. However, we will always understand by a "Diophantine equation" an equation of the form (1).

The tenth problem as a decision problem. Why Hilbert considered solving Diophantine equations as an open problem? In fact, since Diophantus time number-theorists have found solutions for plenty of Diophantine equations and also have proved the unsolvability of a large number of other equations. Unfortunately, for different classes of equations, or even for different individual equations, one had to invent different specific methods. In the 10th problem Hilbert asked for a *universal* method for recognizing the solvability of Diophantine equations.

In today's terminology we consider this problem as a *decision problem.* This means that the problem consists of infinitely many subproblems (specified by particular equations) each of which requires an answer "YES" or "NO" ("there is" or "there is no" solution). An expected solution to the problem should be an algorithm applicable to an arbitrary equation and producing the correct answer.

A remark is appropriate here. Hilbert did *not* ask whether such an algorithm exists. On the one hand, he seemed to be an optimist in mathematics and, most likely, he was sure of the existence of such an algorithm (he posed his problems much before the revolutionary works of K. Gödel). On the other hand, Hilbert even did not use the word "algorithm" in the statement of the problem. Instead, he used a rather vague wording "*a process according to which it can be determined by a finite number of operations ...*". He could have used the *word* "algorithm" but it would have not helped much in clarifying the statement, because at that time there was no rigorous *general notion* of an algorithm. What was known were different examples of particular mathematical algorithms like Euclid's algorithm for finding GCD.

The absence of a general definition of an algorithm was not by itself an obstacle for a positive solution of Hilbert's tenth problem. If somebody would have invented the required "*process*" it should have been clear that in fact the process does the job.

The situation is essentially different if there is no required algorithm as it turned out to be the case with Hilbert's 10th problem. To prove this fact, or even to state it rigorously, one need a definition of an algorithm.

So this was a point where computer science could have come for help to number theory in clarifying difficulties arose in solving Diophantine equations. However, there was no computer science at that time, and even the mathematical logic, a connecting link between mathematics and computer science, was yet to mature.

It is interesting to see that Hilbert foresaw the possibility of the future development of the 10th problem. He wrote [18]:

> Occasionally it happens that we seek the solution under insufficient hypotheses or in an incorrect sense, and for this reason do not succeed. The problem then arises: to show the impossibility of the solution under the given hypotheses, or in the sense contemplated. Such proofs of impossibility were effected by the ancients, for instance when they showed that the ratio of the hypotenuse to the side of an isosceles triangle is irrational. In later mathematics, the question as to the impossibility of certain solutions plays a preeminent part, and we perceive in this way that old and difficult problems, such as the proof of the axiom of parallels, the squaring of circle, or the solution of equations of the fifth degree

> by radicals have finally found fully satisfactory and rigorous solutions, although in another sense than that originally intended. It is probably this important fact along with other philosophical reasons that gives rise to conviction (which every mathematician shares, but which no one has as yet supported by a proof) that every definite mathematical problem must necessary be susceptible of an exact settlement, either in the form of an actual answer to the question asked, or by the proof of the impossibility of its solution and therewith the necessary failure of all attempts.

It was possible to prove the unsolvability of the above mentioned classical problems because they contained an explicit description of admissible tools: a fixed axiomatic system, constructions with ruler and compass or expressions in radicals. However, in Hilbert's tenth problem the admissible tools were not fixed. So the failure in finding a universal method for Diophantine equations has been a stimulus for developing a general notion of an algorithm.

Such a definition emerged much later, only in the 30's of this century in the works of Kurt Gödel, Alan Turing, Emil Post, Alonzo Church and other logicians. However, I cannot trace any *direct* influence of Hilbert's tenth problem on this process.

Soon after the appearance of the rigorous notion of algorithm and *Church Thesis*, the first algorithmically undecidable problems were found, first in mathematical logic itself. In 1947 another breakthrough took place. Andrei Markov [33] and Emil Post [49] proved that so called *word problem* for semigroups was undecidable. This problem was stated by Axel Thue [58] in 1914 and is known also as *Thue's problem.* The importance of this result comes from the fact that it was the first decision problem which arose naturally in mathematics and was finally shown undecidable.

Davis' conjecture

After the success in the proof of the undecidability of the Thue problem and all failures to find a decision procedure for Diophantine equations it was quite natural to suspect that Hilbert's tenth problem is undecidable as well. The American mathematician Martin Davis [6, 7] stated at the beginning of 50's a much stronger hypothesis.

Diophantine sets. In order to be able to present Davis's Conjecture we need more terminology. Besides single Diophantine equations, number-theorist consider also *families of Diophantine equations.* Such a family is an equations of the form

$$D(a_1, \dots, a_n, x_1, \dots, x_m) = 0, \tag{6}$$

where D is again a polynomial with integer coefficients, the variables of which are split into two groups: the *parameters* $a_1, \dots, a_n$ and the *unknowns* $x_1, \dots, x_m$. We will suppose that parameters can, similarly to unknowns, assume positive integer values only. For some choice of the values of the parameters $a_1, \dots, a_n$ the equation (6) can have a solution in the unknowns $x_1, \dots, x_m$, for other choices of the values of the parameters it can have no solution. We can consider the set $\mathcal{M}$ of all n-tuples $\langle a_1, \ldots, a_n \rangle$ for which our parametric equation has a solution, that is

$$\langle a_1, \ldots, a_n \rangle \in \mathcal{M} \Longleftrightarrow \exists x_1 \ldots x_m \{D(a_1, \ldots, a_n, x_1, \ldots, x_m) = 0\}. \tag{7}$$

Sets having such representations are called *Diophantine*. An equivalence of the form (7) is called a *Diophantine representation* of the set $\mathcal{M}$. With an abuse of language, one can say that the equation (6) itself is a representation of the set.

Similarly, one says that a relation among natural numbers is Diophantine if so is the set of all tuple of natural numbers satisfying this relation.

Easy examples of Diophantine sets are the following:

- *the set of all squares* represented by equation

$$a \Leftrightarrow x^2 = 0; \tag{8}$$

- *the set of all composite numbers* represented by equation

$$a \Leftrightarrow (x_1 + 1)(x_2 + 1) = 0; \tag{9}$$

- *the set of all positive integers which are not powers of* 2 represented by equation

$$a \Leftrightarrow (2x_1 + 1)x_2 = 0. \tag{10}$$

It is easy to prove that the union and the intersection of two Diophantine sets are also Diophantine. However, the complement (to the set of all n-tuples of corresponding size) of a Diophantine can be non-Diophantine. The latter fact is not trivial at all. It was proved by Martin Davis [7].

In the case of our first example (8) of a Diophantine set, it can be seen that its complement, i.e. *the set of all numbers which are* not *squares*, is Diophantine; it is represented by equation

$$(a \Leftrightarrow (z \Leftrightarrow 1)^2 \Leftrightarrow x)^2 + (z^2 \Leftrightarrow a \Leftrightarrow y)^2 = 0. \tag{11}$$

However, if we ask about the complements of the other two sets, the answers are not evident at all.

- Is *the set of all prime numbers* Diophantine?
- Is *the set of all powers of 2* Diophantine?

It seems that questions like these have never been posed in number theory. There an equation was a primary object of study, and the typical problem was *to describe the set of all solutions of a given Diophantine equation.* The inverse problem of *constructing a Diophantine representation for a given set* was stranger to number theory.

Diophantine sets from a computational point of view. Every Diophantine set is clearly *listable*, or, in another terminology, *recursively enumerable* (r.e.). (So if the complement of every Diophantine set were Diophantine too, then every set described by an arithmetical formula with both existential and universal quantifiers would be listable too, which is not the case. This was Davis's proof of the existence of a Diophantine set with non-Diophantine complement.)

Martin Davis conjectured that *to be listable* is not only necessary but also a sufficient condition for a set of n-tuples of natural numbers *to be Diophantine.*

M. Davis's conjecture. The notions of Diophantine set and recursively enumerable set coincides, i.e. a set is Diophantine if and only if it is recursively enumerable.

The notion of an r.e. set is essentially equivalent to the notion of an algorithm and so it is a fundamental notion of computability theory (see, for example, the book of Martin-Löf [34] where computability theory is developed in terms of listable sets rather than in terms of algorithms). Davis's conjecture states that the general notion of recursively enumerable set is equivalent to the seemingly rather narrow notion of a Diophantine set.

From conjecture to theorem

It took two decades before Davis's conjecture became a theorem, i.e. before it was proved.

Davis' normal form. In the early 50's Martin Davis made the first step towards the proof of his conjecture. Namely he showed that every r.e. set $\mathcal{M}$ of n-tuples of natural numbers has an "almost Diophantine" representation of the form

$$\begin{aligned}&\langle a_1,\dots,a_n\rangle \in \mathcal{M} \Longleftrightarrow \\ &\exists z(\forall y \le z)\exists x_1 \dots x_m\{D(a_1,\dots,a_n,x_1,\dots,x_m,y,z)=0\}.\end{aligned} \tag{12}$$

Note that the universal quantifier in (12) is bounded, so every formula of this kind determines an r.e. set. Representations of the type (12) became known as *Davis normal form.* They were a quantitative improvement over the classical result of Kurt Gödel [16] who demonstrated the existence of similar arithmetical representations with arbitrary number of (bounded) universal quantifiers.

Back to equations. The single remaining universal quantifier was finally eliminated from Davis normal form. At first this was done conditionally and at the cost of extending the set of admissible equations to the so called *exponential Diophantine equations.* They are equations of the form

$$E_{\mathrm{L}}(a_1,\dots,a_n,x_1,x_2,\dots,x_m) = E_{\mathrm{R}}(a_1,\dots,a_n,x_1,x_2,\dots,x_m) \tag{13}$$

where E_{L} and E_{R} are *exponential polynomials*, i.e. expression constructed by traditional rules from the variables and particular natural numbers by addition, multiplication and exponentiations. Namely, Martin Davis and Hilary Putnam [12] proved in 1960 that every r.e. set $\mathcal{M}$ has a purely existential *exponential Diophantine representation*

$$\langle a_1,\dots,a_n\rangle \in \mathcal{M} \Longleftrightarrow \tag{14}$$

$$\exists x_1 \dots x_m \{E_{\mathrm{L}}(a_1,\dots,a_n,x_1,x_2,\dots,x_m) = E_{\mathrm{R}}(a_1,\dots,a_n,x_1,x_2,\dots,x_m)\}$$

under the assumption that *there are arbitrary long arithmetical progressions consisting entirely of prime numbers.*

(The report [12] is not easy to find. When by my request Martin Davis presented a copy of it to me, he said: "Do not read it!")

Thus, the existence of arbitrary many primes in arithmetical progressions would imply an algorithmical result: the undecidability of the weaker form of Hilbert's tenth problem corresponding to a broader class of exponential Diophantine equations. This implication motivated the search of long arithmetical progressions of primes, but why should number-theorist consider this problem? There was an opinion (expressed, as I remember, by physicists not mathematicians) that "prime numbers were born to be multiplied" (for generating all natural numbers) and respectively "additive problems about prime numbers are artificial and need not be considered at all". Nevertheless the problem of how long an arithmetical progressions of primes could be was considered in number theory much before the work of Davis and Putnam. However, "being born for multiplication", prime numbers are very stubborn when one tries to add or subtract them, and we still (1998) do not know whether they form arithmetical progressions of arbitrary large length.

Luckily, Julia Robinson [52] was able to avoid the assumption about primes in progressions, and in 1961 Martin Davis, Julia Robinson and Hilary Putnam published a famous joint paper [11] with an unconditional proof of the existence of an exponential Diophantine representation for every r.e. set.

(The need of prime numbers was caused by the use of Gödel's [16] method of representing n-tuples of natural numbers of arbitrary length based on *Chinese Remainder Theorem.* It is interesting to note that much later I [36] was able to show that primality is not essential at all and prime numbers can be completely avoided.)

Exorcizing exponentiation. With the work of Davis, Putnam and Robinson the hard algorithmical part of the job was done: an analog of Hilbert's tenth problem for exponential Diophantine equation was proved undecidable. All what remained in order to prove the undecidability of the original Hilbert's problem was to learn how to transform an arbitrary exponential Diophantine equation (13) into an equivalent Diophantine equation, i.e. an equation of the form (6) having solutions for the same values of the parameters. Moreover, in order to prove Davis's conjecture in its full form it was sufficient to prove a very particular case of it, namely, to prove that the ternary relation of exponentiation $a = b^c$ was Diophantine. As soon as we have a Diophantine representation for this relation,

$$\langle a,b,c\rangle \in \mathcal{M} \Leftrightarrow a^b = c \Leftrightarrow \exists z_1 \dots z_w \{A(a,b,c,z_1,\dots,z_m) = 0\} \qquad (15)$$

we can easily use the polynomial A for transforming exponential Diophantine equations into Diophantine.

However, as it was mentioned above, the inverse problem of finding an equation for a given set was not popular in number theory, and so in 1961 the existence of (15) was open. Luckily, this problem has already been attacked in mathematical logic. The starting point of this investigation was quite opposite to Davis's conjecture: Alfred Tarski suspected at the end of 40's that even the set of all powers of 2 is *not* Diophantine. Julia Robinson spent some time trying to prove it but then she switched to proving that the relation $a = b^c$ was Diophantine. She failed to do it but found some sufficient conditions. Namely, Julia Robinson proved that the relation of exponentiation is Diophantine provided that there is a Diophantine relation

$$J(u, v) \Leftrightarrow \exists z_1 \dots z_m \{B(u, v, z_1, \dots, z_m) = 0\} \tag{16}$$

such that

1. if $J(u, v)$ holds then $u < v^v$;
2. for every k there are u and v such that $J(u, v)$ holds and $u > v^k$.

Julia Robinson called the relations satisfying the above two conditions *relations of exponential growth*; they became known in the literature as *Julia Robinson relations.*

All it remained to prove Davis conjecture was to find a single relation of exponential growth defined by a Diophantine equation. Surprisingly, among numerous two-parameter equations studied in number theory since Diophantus up to the middle of 20th century there was no Diophantine equation defining a relation of exponential growth.

A few words about words. I began to be involved in investigations on Hilbert's tenth problem at the end of December 1965 when I was a second year student at the department of mathematics and mechanics of Leningrad State University. This subject was suggested to me by my scientific adviser Serguei Maslov (see [13]). However, I studied the works of Martin Davis, Hilary Putnam and Julia Robinson sometime later. Maslov believed that their approach cannot lead to solution because it had not led already. Instead Maslov suggested to try another approach initiated by A. A. Markov.

The idea was as follows. A universal computer science tool for representing information uses words rather than numbers. However, there are many ways to represent words by numbers. One of such methods is naturally related to Diophantine equations. Namely, it is not difficult to show that every 2 × 2 matrix

$$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \tag{17}$$

with the m's being non-negative integers and the determinant equal to 1 can be represented, in an unique way, as a product of matrices

$$M_0 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad M_1 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}. \tag{18}$$

It is evident that any product of such matrices has non-negative integer elements and the determinant equals 1. This implies that we can uniquely represent a word

$a_{i_1} \ldots a_{i_m}$ in a two-letter alphabet by the four-tuple $\langle m_{11}, m_{12}, m_{21}, m_{22} \rangle$ such that

$$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = M_{i_1} \ldots M_{i_m}; \tag{19}$$

the numbers evidently satisfy the Diophantine equation

$$m_{11}m_{22} \Leftrightarrow m_{21}m_{12} = 1. \tag{20}$$

Under this representation of words by matrices, the concatenation of words corresponds to matrix multiplication and thus can be easily expressed as a system of Diophantine equations. This opens a way to transform an arbitrary system of *word equations* into equivalent Diophantine equations. Many decision problems about words had been shown undecidable, so it was quite natural to try to attack Hilbert's tenth problem by proving the undecidability of systems of word equations.

I spent some time trying to show the undecidability of word equations. Much later it became known that this approach was fruitless as G. S. Makanin [29] found an algorithm for word equations. Luckily, I soon abandoned this approach.

My next attempt was to consider a broader class of word equations with additional predicates. Since the ultimate goal was always Hilbert's tenth problem, I could consider only such predicates which (under suitable coding) would be represented by Diophantine equations. In this way I came to what I have called *equations in words and lengths.* Reduction of such equations to Diophantine equations was based on celebrated Fibonacci numbers. It is well-known that every natural number can be represented, in an almost unique way, as the sum of different Fibonacci numbers, none of which are consecutive (so called *Zeckendorf's representation*). Thus we can look at natural numbers as words in two-letter alphabet {0,1} with the additional constraint that there cannot be two consecutive 1's. I [35] managed to show that under this representation of words by numbers both the concatenation of words and the equality of the lengths of two words can be expressed by Diophantine equations.

However, I was unable to show the undecidability of equations in words and length (and this still remains an open problem).

Rabbits strike again. Finally, I switched to Davis-Putnam-Robinson approach and tried to construct a Diophantine relation of exponential growth. Due to my previous work, I realized the importance of Fibonacci numbers for Hilbert's tenth problem. That is why during the summer of 1969 I was reading with great interest the third augmented edition [61] of a popular book on Fibonacci numbers written by N. N. Vorob'ev from Leningrad. It seems incredible that in the 20th century one can still find something new about the numbers introduced by Fibonacci in the 13th century in connection with multiplying rabbits. However, the new edition of the book contained, besides traditional stuff, some original results of the author. In fact, Vorob'ev had obtained them a quarter of a century earlier but he never published anything before. His results attracted my attention at once but I was not able to use them immediately for constructing a Diophantine representation of a relation of exponential growth.

After the summer I had to pass the entrance examinations in order to continue my postgraduate studies at Steklov Institute of Mathematics (Leningrad Branch). Three years are given for the preparation of a Ph. D. thesis; at that time I had already spent 4 years in vain on Hilbert's tenth problem, and there was no real hope to solve it in another 3 years. So I switched to a quite different subject, automatic theorem proving, which was the main area of investigations of Maslov.

My previous work on Hilbert's tenth problem unexpectedly paid in the end. Thanks to it I was considered an expert on the problem, and when Julia Robinson published her new paper [53], it was sent to me for a review for the soviet counterpart of *Mathematical Reviews*. The period when I was not thinking about Diophantine equations, Vorob'ev's theorem and new ideas of Julia Robinson led me to the negative solution of Hilbert's tenth problem. On January 29, 1970 I gave at my institute the first talk on a Diophantine relation of exponential growth. It was the relation *u is the $2v$th Fibonacci number.*

Surprisingly, in order to construct a Diophantine representation for this relation I needed to proof a yet new purely number-theoretical result about Fibonacci numbers, namely, that *kth Fibonacci number is divisible by the square of the lth Fibonacci number if and only if k itself is divisible by the lth Fibonacci number.* This property is not difficult to prove; what is striking is that this beautiful fact has not been discovered, even empirically, since Fibonacci times.

With my example of a Diophantine relation of an exponential growth Davis's conjecture became a theorem which is often referred to as DMPR theorem after Davis-Matiyasevich-Putnam-Robinson. Nowadays detailed and simplified proofs of this theorem can be found in many publications, in particular, in [1, 5, 9, 10, 23, 31, 32, 36, 41, 57].

From Computer Science to Number Theory

DMPR theorem, establishing the equivalence of a number-theoretical notion of Diophantine set and the notion of r.e. set, can serve as a bridge for transporting ideas and methods from computer science to number theory and backward. First we shall survey what number theory gained from DMPR theorem.

The undecidability of Hilbert's Tenth Problem

Of course, the first evident gain is the proof of the undecidability of Hilbert's tenth problem. With it number-theorists have got a "moral right" to abandon further attempts to find a universal method for Diophantine equations and use *ad hoc* tools for particular equations.

DMPR theorem implies an undecidability result stronger than what is required just to "close" Hilbert's tenth problem. Hilbert asked for a universal method suitable for an arbitrary Diophantine equation. Now we can take a particular r.e. but undecidable set $\mathcal{M}$ of natural numbers and construct its Diophantine representation

$$a \in \mathcal{M} \iff \exists x_1 \dots x_m \{M(a, x_1, \dots, x_m) = 0\}. \tag{21}$$

The undecidability of $\mathcal{M}$ implies that there is no universal method to decide, for a given a, whether the equation

$$M(a, x_1, \dots, x_m) = 0 \tag{22}$$

has a solution in $x_1, \dots, x_m$. That is, to get the undecidability we need not consider all Diophantine equations, in any number of unknowns and of arbitrary large degree.

The undecidability of one-parameter Diophantine equation (22) means the following. Suppose that $\mathcal{A}$ is some algorithm hypothetically capable to tell for a given a whether the equation (22) has a solution or not. Now we know that the algorithm $\mathcal{A}$ should fail for some particular number $a_{\mathcal{A}}$, that is, either $\mathcal{A}$ never stops on input $a_{\mathcal{A}}$ or its output, if any, is wrong. The set $\mathcal{M}$ can be chosen in such a way that this counterexample $a_{\mathcal{A}}$ could be effectively found for every algorithm $\mathcal{A}$.

Speeding-up Diophantine equations

The undecidability of Hilbert's tenth problem, being very important for number theory, still cannot be considered as a purely number-theoretical result because its formulation involves notions from computability theory. However, DMPR theorem allows us to obtain many new results about Diophantine equations which are, at least in the form, number-theoretical. We can just take any theorem about r.e. sets and replace words "recursively enumerable" in its statement by "Diophantine". Furthermore, we can use the definition of a Diophantine set and obtain a theorem about Diophantine equations.

Such reformulations of theorems from computer science in terms of Diophantine equations usually give results which are not typical for number theory. As an example, let us consider a Diophantine version of M. Blum's [3] *speed-up theorem* obtained by M. Davis [8]:

For every general recursive function $\alpha(a, w)$ there are Diophantine equations

$$B(a, x_1, \dots, x_n) = 0, \tag{23}$$

$$C(a, y_1, \dots, y_m) = 0 \tag{24}$$

such that:

1. *for every value of a, one and only one of these two equations has a solution;*
2. *if the equations*

$$B'(a, x'_1, \dots, x'_{n'}) = 0, \tag{25}$$

$$C'(a, y'_1, \dots, y'_{m'}) = 0 \tag{26}$$

are solvable exactly for the same values of the parameter a, as equations (23) *and* (24)*, respectively, then there is a third pair of equations*

$$B''(a, x''_1, \dots, x''_{n''}) = 0, \tag{27}$$

$$C''(a, y''_1, \dots, y''_{m''}) = 0 \tag{28}$$

such that:

- ○ *these equations are also solvable exactly for the same values of the parameter* a*, as equations* (23) *and* (24)*, respectively;*
- ○ *for almost all* a*, for every solution of equation* (25) (*equation* (26)) *there is a solution of equation* (27) (*respectively, equation* (28)) *such that*

$$x'_1 + \ldots + x'_{m'} > \alpha(a, x''_1 + \ldots + x''_{m''}) \tag{29}$$

(*or*

$$y'_1 + \ldots + y'_{n'} > \alpha(a, y''_1 + \ldots + y''_{n''}) \tag{30}$$

respectively).

This theorem, in its full generality, is about an arbitrary general recursive function; replacing it by particular (fast growing) functions we obtain theorems which are purely number-theoretical but quite non-standard for number theory.

Universal Diophantine equations

Now we consider a transfer of another idea from computer science to number theory which is of more interest for our considerations.

Jack of all trades. Universal objects like *universal Turing machines*, *universal r.e. sets* and so on are quite ordinary objects in computer science. Now with DMPR theorem we can construct their counterparts in number theory. Namely, for every fixed n, we can construct in an effective way a list

$$D_0, D_1, \ldots \tag{31}$$

consisting of all polynomials with integer coefficients having n parameters and arbitrary number of unknowns. The set $\mathcal{U}_n$ of $(n+1)$-tuples defined by

$$\langle k, a_1, \ldots, a_n \rangle \in \mathcal{U}_n \iff \exists x_1 x_2 \ldots \{D_k(a_1, \ldots, a_n, x_1, x_2, \ldots) = 0\}, \tag{32}$$

being recursively enumerable is Diophantine, so we can construct a single polynomial U_n with $n+1$ parameter such that

$$\langle k, a_1, \ldots, a_n \rangle \in \mathcal{U}_n \iff \exists x_1 \ldots x_m \{U_n(k, a_1, \ldots, a_n, x_1, x_2, \ldots, x_m) = 0\}. \tag{33}$$

The corresponding equation

$$U_n(k, a_1, \ldots, a_n, x_1, x_2, \ldots, x_m) = 0 \tag{34}$$

is *universal* in the following sense: solving an arbitrary Diophantine equation with n parameters

$$D(a_1, \ldots, a_n, x_1, x_2, \ldots) = 0 \tag{35}$$

is equivalent to solving the equation

$$U_n(k_D, a_1, \ldots, a_n, x_1, x_2, \ldots, x_m) = 0 \tag{36}$$

resulting from the single equation (34) by choosing a particular value k_D for the first parameter (this value is, of course, the number of the equation (35) in the list (31)).

The degree and the number of unknowns of the equation (36) is fixed while the equation (35) can have any number of unknowns and can be of arbitrary large degree. As an example let us put $n = 1$ and let s be the number of arithmetical operations required to calculate the value of the *universal polynomial* U_1. Let us take for (35) the equation

$$a = (x_1 + 1) \dots (x_{2^s} + 1). \tag{37}$$

We see that in order to verify that a number a has at least 2^s prime factors it is sufficient to perform only s arithmetical operations!

Is it Number Theory?. To what extent the existence of universal Diophantine equations belongs to computer science and to what extent to number theory? First of all, number-theorists never anticipated such a possibility. It was a surprise even to some logicians familiar with other universal objects. For example, when in 1961 Martin Davis, Hilary Putnam and Julia Robinson proved the existence of exponential Diophantine representations, they immediately got as a corollary the existence of universal exponential Diophantine equations. However, the following paragraph appears in George Kreisel review [26] (for *Mathematical Reviews*) of the celebrated paper by Davis, Putnam and Robinson [11]:

> These results are superficially related to Hilbert's tenth Problem on (ordinary, i.e., non-exponential) Diophantine equations. The proof of the authors' results, though very elegant, does not use recondite facts in the theory of numbers nor in the theory of r.e. sets, and so it is likely that the present result is not closely connected with Hilbert's tenth Problem. Also it is not altogether plausible that all (ordinary) Diophantine problems are uniformly reducible to those in a fixed number of variables of fixed degree, which would be the case if all r.e. sets were Diophantine.

We can look at the existence of universal Diophantine equations as a number-theoretical result *inspired* by computer science. The following question naturally arises: *can the existence of universal Diophantine equations be proved by purely number-theoretical tools*, i.e. without any reference to the notion of r.e. set and constructions of universal r.e. sets? In my book [41] I managed to give what I believe to be such a number-theoretical construction of universal Diophantine equations.

Collapse of Diophantine hierarchy. The possibility to bound both the degree and the number of unknowns in Diophantine representations means that traditional number-theoretical classifications of Diophantine equations as equations in 1, 2, ... unknowns and as equations of degree 1, 2, ... collapse. Of course, it is of considerable interest to find out where exactly this happens.

Let us call a pair $\langle \nu, \mu \rangle$ of natural numbers *a universal Diophantine complexity bound* if every r.e. set can be defined by a Diophantine equation of degree ν with respect to its μ unknowns.

As usual, there is a natural trade-off between ν and μ. The best known today (1998) value of ν is 4. This follows from the existence of universal equations and an old result of Thoralf Skolem [56] who showed that solving an arbitrary Diophantine equation can be easily reduced to solving another Diophantine equation of degree 4 at the cost of introduction of many additional unknowns.

The best known today (1998) value of μ is 9. I obtained this result in [37] but for various reasons (see [41]) I have never published a detailed proof. This was finally done by James P. Jones [21]. He also calculated a value of ν corresponding to $\mu = 9$ and found a number of "intermediate values": the following pairs are all universal Diophantine complexity bounds:

$$\langle 4,58\rangle,\ \langle 8,38\rangle,\ \langle 12,32\rangle,\ \langle 16,29\rangle, \langle 20,28\rangle,\ \langle 24,26\rangle,\ \langle 28,25\rangle,\ \langle 36,24\rangle,$$
$$\langle 96,21\rangle,\ \langle 2668,19\rangle,\ \langle 2\times 10^5,14\rangle, \langle 6.6\times 10^{43},\ 13\rangle,\ \langle 1.3\times 10^{44},12\rangle,$$
$$\langle 4.6\times 10^{44},11\rangle,\ \langle 8.6\times 10^{44},10\rangle,\ \langle 1.6\times 10^{45},9\rangle.$$

It is known that universal Turing machines can be rather small. The above universal bound might produce the impression that writing down a universal Diophantine equation would require hundreds of pages. This is not the case. While the degree of a universal polynomial could be very large, the majority of coefficients can be equal to 0. Here is a universal system of Diophantine equations also constructed in [21]:

$$elg^2 + \alpha = (b \Leftrightarrow xy)q^2, \quad q = b^{5^{60}}, \quad \lambda + q^4 = 1 + \lambda b^5, \quad \theta + 2z = b^5,$$
$$l = u + t\theta, \quad e = y + m\theta, \quad n = q^{16},$$
$$r = \Big[g + eq^3 + lq^5 + \big(2(e \Leftrightarrow z\lambda)(1 + xb^5 + g)^4 + \lambda b^5 + \lambda b^5 q^4\big)q^4\Big][n^2 \Leftrightarrow n]$$
$$+ [q^3 \Leftrightarrow bl + l + \theta\lambda q^3 + (b^5 \Leftrightarrow 2)q^5][n^2 \Leftrightarrow 1],$$
$$p = 2ws^2r^2n^2, \quad p^2k^2 \Leftrightarrow k^2 = 1 = \tau^2, \quad 4(c \Leftrightarrow ksn^2)^2 + \eta = k^2,$$
$$k = r + 1 + hp \Leftrightarrow h, \quad a = (wn^2 + 1)rsn^2, \quad c = 2r + 1 + \phi,$$
$$d = bw + ca \Leftrightarrow 2c + 4a\gamma \Leftrightarrow 5\gamma, \quad d^2 = (a^2 \Leftrightarrow 1)c^2 + 1,$$
$$f^2 = (a^2 \Leftrightarrow 1)i^2c^4 + 1,$$
$$(d + of)^2 = \Big(\big(a + f^2(d^2 \Leftrightarrow a)\big)^2 \Leftrightarrow 1\Big)(2r + 1 + jc)^2 + 1$$

(clearly, a system of Diophantine equations can be easily combined into a single equation by squaring).

Most likely, 9 unknowns is rather far from the least possible values which could be as low as 3 (equations in a single unknown are evidently decidable and great progress for equations in 2 unknowns gives hope to number-theorists to find a decision procedure for such equations). For exponential Diophantine equations I was able to find a much better bound (see [40]): a universal exponential Diophantine equation can have only 3 unknowns.

Prime numbers are of main interest in number theory

With DMPR theorem the inverse problem, i.e. constructing Diophantine equations solvable for prescribed values of the parameters, has got a complete solution. The proof of DMPR theorem was constructive so for every r.e. set we can effectively write down its Diophantine representation.

Of course, the most interesting set in number theory is the set of all prime numbers. According to DMPR theorem, we can find a particular Diophantine equation

$$P(a, x_1, \ldots, x_m) = 0 \tag{38}$$

solvable in $x_1, \ldots, x_m$ if and only if the parameter a is a prime.

Consider now the equation

$$x_0(1 \Leftrightarrow P^2(x_0, x_1, \ldots, x_m)) = a. \tag{39}$$

Clearly, every solution of equation (38) can be extended to a solution of equation (39) just by putting

$$x_0 = a. \tag{40}$$

On the other hand, for every positive a in every solution (in positive integers) of equation (39) the expression in brackets should be positive too. This is possible only when

$$P(x_0, x_1, \ldots, x_m) = 0. \tag{41}$$

But (39) implies (40) and hence (38).

We see that equation (39) is also a representation of the set of all primes. This fact can be stated a bit differently thanks to the special form of (39): *the set of all prime numbers is identical with the set of all positive values assumed by the polynomial* $x_0(1 \Leftrightarrow P^2(x_0, x_1, \ldots, x_m))$, *for positive values of its variables.*

The above trick of passing from (38) to (39) was invented by Hilary Putnam [50] and published in 1960; at that time the possible existence of prime representing polynomials was considered as an informal argument against Davis' conjecture.

The discovery of prime number representing polynomials was quite a surprise for number-theorists. When prominent Soviet mathematician Yu. V. Linnik was told about it, he reacted: "That's wonderful, most likely we shall soon learn a lot of new things about primes. " But then it was explained to him that this was just one corollary of a much more general result about the existence of Diophantine representations for every r.e. set. "It's a pity," Linnik then said. "Most likely, we shall not learn anything new about the primes. "

The generality of DMPR theorem is both its power and the weak point. Hardly we can prove anything new looking at the equation (36) with k_D corresponding to the set of prime numbers. Luckily, for constructing a Diophantine representation of primes we need only part of the machinery developed for the case of arbitrary r.e. set.

Here is a particular prime representing polynomial taken from [24]:

$$\begin{aligned}(k+2)\{\ 1 &\Leftrightarrow [wz + h + j \Leftrightarrow q]^2 \\ &\Leftrightarrow [(gk + 2g + k + 1)(h + j) + h \Leftrightarrow z]^2 \\ &\Leftrightarrow [2n + p + q + z \Leftrightarrow e]^2\end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow\big[16(k+1)^3(k+2)(n+1)^2+1\Leftrightarrow f^2\big]^2\\
&\Leftrightarrow\big[e^3(e+2)(a+1)^2+1\Leftrightarrow o^2\big]^2\\
&\Leftrightarrow\big[(a^2\Leftrightarrow 1)y^2+1\Leftrightarrow x^2\big]^2\\
&\Leftrightarrow\big[16r^2y^4(a^2\Leftrightarrow 1)+1\Leftrightarrow u^2\big]^2\\
&\Leftrightarrow[n+l+v\Leftrightarrow y]^2\\
&\Leftrightarrow\big[\big((a+u^2(u^2\Leftrightarrow a))^2\Leftrightarrow 1\big)\,(n+4dy)^2+1\Leftrightarrow(x+cu)^2\big]^2\\
&\Leftrightarrow\big[(a^2\Leftrightarrow 1)l^2+1\Leftrightarrow m^2\big]^2\\
&\Leftrightarrow\big[q+y(a\Leftrightarrow p\Leftrightarrow 1)+s(2ap+2a\Leftrightarrow p^2\Leftrightarrow 2p\Leftrightarrow 2)\Leftrightarrow x\big]^2\\
&\Leftrightarrow\big[z+pl(a\Leftrightarrow p)+t(2ap\Leftrightarrow p^2\Leftrightarrow 1)\Leftrightarrow pm\big]^2\\
&\Leftrightarrow[ai+k+1\Leftrightarrow l\Leftrightarrow i]^2\\
&\Leftrightarrow\big[p+l(a\Leftrightarrow n\Leftrightarrow 1)+b(2an+2a\Leftrightarrow n^2\Leftrightarrow 2n\Leftrightarrow 2)\Leftrightarrow m\big]^2\ \}\,.
\end{aligned}$$

The above polynomial for primes has 26 variables $a,\ldots,z$ (all letters of the alphabet, including o!) ranging over non-negative integers.

After one of his lectures about this polynomial, J. P. Jones got a telegram saying that this particular polynomial certainly was not representing the primes because itself was a product of two polynomials. J. P. Jones was asked to send as soon as possible a "correct polynomial". But nothing was wrong, it was just Putnam's trick!

For a long time the current (at that moment) known least number of variables in primes representations was smaller than the current (at the same moment) least known number of unknowns required for the representation of an arbitrary r.e. set. Today both records (i.e. for primes and for a general r.e. set) are the same—10 variables. Of course, polynomial constructed especially for primes [38] has much smaller degree than the universal polynomial.

Being unable to construct a Diophantine representation with a smaller number of unknowns for the whole set of prime numbers, we can look for Diophantine representations of its infinite subsets. Such a broader problem is, of course, of less interest for traditional number theory; however, it has direct connections with possible proofs of DMPR theorem. Namely, just before the theorem was proved, Julia Robinson [53] established the following conditional result: *every r.e. set is Diophantine as soon as at least one infinite set of prime numbers is Diophantine.*

J. P. Jones [25] proved that 7 unknowns are sufficient for Diophantine representations of the set of all *Mersenne primes* (i.e. primes of the form $2^n \Leftrightarrow 1$) and for the set of all *Fermat primes* (i.e. odd primes of the form 2^n+1). Up to now, only 5 Fermat primes have been found, and it is considered most likely that there are no others. As for Mersenne primes, (non-rigorous) probabilistic arguments suggest that there are infinitely many such primes but so far nobody was able to give a real proof.

Jones's success in reducing the number of unknowns in the cases of Mersenne and Fermat primes was due to the fact that for numbers of the forms $2^n \pm 1$ there are very efficient tests for primality. I suspected that other classes of primes which can be tested for primality quickly, could also be defined by a Diophantine equation with smaller number of unknowns.

My post-graduate student Maxim Vsemirnov confirmed my supposition. Large prime numbers are vital for modern cryptography, and efficient methods for generating them were developed. In particular, J. Pintz, W. L. Steger, and E. Szemeredi [48] described a (provably) infinite set of prime numbers of special form admitting fast primality test. M. Vsemirnov [62] showed that similar infinite set of primes admits a Diophantine representation with 8 unknowns only.

Non-effective estimates can be non-effectivizable

What are the "duties" of computer science with respect to number theory? One role which could/should be played by computer science is to explain difficulties arising in number theory. Computer science managed to help number theory by proving that Hilbert's tenth problem was undecidable. However, there are other situations in number theory where all attempts fail, and, possibly, because of some deep algorithmical reasons.

Suppose that we have a Diophantine equation

$$D(a, x_1, \dots, x_m) = 0, \tag{42}$$

such that for every value of the parameter a has at most finitely many solutions in $x_1, \dots, x_m$. This fact can be expressed in two form:

1. the equation (42) has at most $\nu(a)$ solutions;
2. in every solution of (42) $x_1 < \sigma(a)$, ..., $x_m < \sigma(a)$

for suitable functions ν and σ.

¿From a mathematical point of view these two statements are equivalent. However, they are rather different from a computational point of view. Having $\sigma(a)$ we can find $\nu(a)$, but not *vice versa*. Number-theorists have found many classes of Diophantine equations with computable $\nu(a)$ for which they fail to compute $\sigma(a)$. In such cases number-theorists say that "the estimate of the size of solutions is *non-effective*".

I was able to show that at least in the theory of exponential Diophantine equations there are estimates which cannot be effective in principle. Namely, we can construct an exponential Diophantine equation

$$E_{\mathrm{L}}(a, x_1, x_2, \dots, x_m) \; = \; E_{\mathrm{R}}(a, x_1, x_2, \dots, x_m) \tag{43}$$

with the following properties:

1. for every value of the parameter a, the equation (43) has at most one solution in $x_1, \dots, x_m$;

2. for every general recursive function σ there is a value of a for which the equation (43) has a solution $x_1, \ldots, x_m$ such that $x_1 > \sigma(a)$.

Improving this result to the case of Diophantine equations remains a challenge for computer science.

From Number Theory to Computer Science

In this section I will show that computer science has also gained something from the collaboration with number theory on Hilbert's tenth problem.

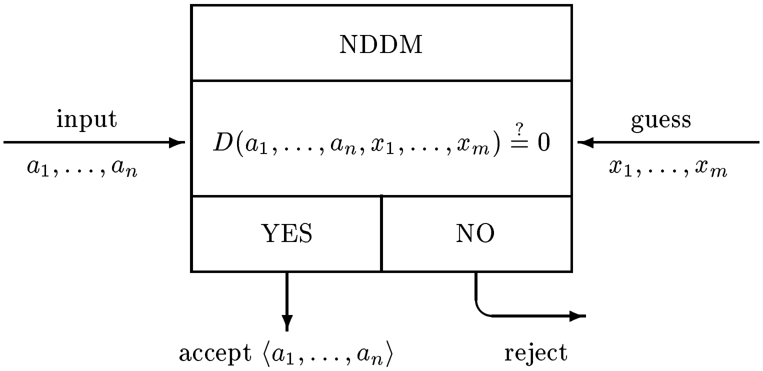
Diophantine equations as computing devices

It was already mentioned above that the notion of r.e. set is as fundamental as the general notion of algorithm. Respectively, we can treat Diophantine equations as computing devices. This was done in a picturesque form by Leonard Adleman and Kenneth Manders in [2, 30]. There they introduced the notion of *Non-Deterministic Diophantine Machine*, NDDM for short.

A NDDM is specified by a Diophantine equation

$$D(a_1, \ldots, a_n, x_1, \ldots, x_m) = 0, \tag{44}$$

and works as follows: on input $a_1, \ldots, a_n$ it guesses the numbers $x_1, \ldots, x_m$ and checks (44); if the equality holds, the n-tuple $\langle a_1, \ldots, a_n \rangle$ is accepted.


NDDM
input
$a_1, \ldots, a_n$
$D(a_1, \ldots, a_n, x_1, \ldots, x_m) \overset{?}{=} 0$
guess
$x_1, \ldots, x_m$
YES
NO
accept $\langle a_1, \ldots, a_n \rangle$
reject

The idea behind the introduction of a new computing device was as follows: in NDDM we have full separation of guessing and deterministic computation, and the latter is very simple—just the calculation of the value of a polynomial.

Now the DMPR theorem states that NDDMs are as powerful as, say, Turing machines: every set recognizable by a Turing machine is recognized by some NDDM, and, of course, *vice versa*.

The original proof of the DMPR theorem was quite constructive, that is, given a Turing machine, we can effectively write down the corresponding Diophantine equation. Unfortunately, such a construction was rather roundabout. Namely, one had:

1. to construct an arithmetical formula with many bounded universal quantifiers describing the work of the Turing machine by the technique introduced by K. Gödel [16];
2. to bring this formula into Davis normal form (12) by gluing bounded universal quantifiers as it is described in Davis' paper [7];
3. to eliminate the last remaining bounded universal quantifier at the cost of passing to exponential Diophantine equations;
4. to eliminate the exponentiation making use of its Diophantine representation (15).

According to a footnote in Davis' paper [7], the idea of obtaining the representation (12) by combining universal quantifiers from a general arithmetic representation was due to the (anonymous) referee of the paper. The original proof of Davis (outlined in [6] and given with details in [10]) was quite different. Technically, it was more involved but in some respects the original construction was very appealing. Namely, Davis managed to arithmetize *Post normal systems* using only one universal quantifier. Now that we know that in fact no universal quantifier is necessary at all, would not it be more natural to try to go further in this direction and obtain, by an arithmetization of a Turing machine, directly a purely existential exponential Diophantine representation (14)?

This idea was very attractive for me and at last I was able to implement it in my paper [37]. There I arithmetized by purely existential formulas the work of a Turing machine. Later it turned out that another kind of computing devices, so called *register machines*, are even more suitable for such existential arithmetization. Register machines were introduced almost simultaneously by several authors: J. Lambek [28], Z. A. Melzak [43], M. L. Minsky [44, 45], and J. C. Shepherdson and H. E. Sturgis [54]. Like Turing machines, register machines have very primitive instructions but, in addition, they deal directly with numbers rather than with words. This leads to a "visual proof" of simulation of register machines by exponential Diophantine equations (see my joint papers with J. P. Jones [22, 23]). Such a proof of DPMR theorem is more suitable for a course in computability theory because its prerequisites in number theory are minimal.

I had a large choice of techniques for proving DMPR theorem during the writing of my book [41]. I decided to simulate Turing machines because they are the classical computing devices. The proof presented in the book differs very much from the first proof via Turing machines given in [37].

Later I found yet another direct way to prove DMPR theorem using induction on the construction of partial recursive functions (see [42]).

I have given talks about Hilbert's tenth problems in many universities, both at departments of mathematics and at departments of computer science. Mathematicians often complained that their salaries are less than salaries of computer scientists. So I often explained to mathematicians that if they are solving Diophantine equations, they are doing some kind of computer science and so they can demand an increase in payment.

The magic of old number theory

The pure number-theoretical proof of the existence of universal equations, the reductions to 9 and 3 unknowns, "visual" simulations of Turing and register machines and many other interesting new results were obtained thanks to a rather old result of always young number theory.

According to the Main Theorem of Arithmetic, the binomial coefficient $\binom{a+b}{b}$, as any natural number, can be represented as the product of powers of prime numbers:

$$\binom{a+b}{a} = 2^{\alpha(2)}3^{\alpha(3)}5^{\alpha(5)}\dots \tag{45}$$

What are these exponents $\alpha(1)$, $\alpha(2)$, $\alpha(3)$, …? A surprising answer was found by Ernst Kummer [27]: *in order to find $\alpha(p)$ one can first write down both numbers a and b in p-base notation and then add them according to school rules; during this addition carries from digit to digit can occur; the number of this carries is exactly $\alpha(p)$.*

The power of Kummer's theorem in constructing Diophantine representations can be informally explained as follows: it connects divisibility properties of numbers with their properties as strings of digits. We can look at numbers as, say, words in the two-letter alphabet $\{0, 1\}$. Binary notation does not forbid two consecutive 1's as it was in above mentioned Zeckendorf representation. We can easily express concatenation by an (exponential) Diophantine equation. Kummer's theorem allows us to express in a Diophantine way many relations among numbers formulated in terms of their binary digits. For example, the property *every binary digit of the number a is less or equal to corresponding binary digit of number b* is equivalent, according to Kummer's theorem, to the congruence

$$\binom{b}{a} \equiv 1 \pmod{2} \tag{46}$$

which can be easily rewritten, with the aid of the binomial theorem, as an exponential Diophantine equation.

Kummer published this beautiful theorem in 1852. At that time this fact did not find many applications in number theory and the theorem was forgotten; several authors rediscovered Kummer's result in this century, and now it is fruitfully used for constructing Diophantine representations of arbitrary r.e. sets.

Diophantine complexity

For Turing machines there are two natural complexity measures: TIME and SPACE. For NDDMs there is only one natural complexity measure which plays the role of both TIME and SPACE. This measure is SIZE, which is the size (in bits) of the smallest solution of the equation. It is not essential whether we define this solution as the one with the smallest possible value of $\max\{x_1,\ldots,x_m\}$, or of $x_1+\ldots+x_m$.

Naturally, SIZE can be used for defining complexity classes. Kosovskii and Vinogradov [60] showed that bounding SIZE by suitable function we can define Grzegorchyk's hierarchy starting from $\mathcal{E}_3$.

We know from the DMPR theorem that NDDMs are as powerful as Turing machines. The main intriguing question is how efficient are the former. Adleman and Manders supposed that NDDN is as efficient as Turing machine. They obtained in [2, 30] the first results in this direction by estimating the SIZE of a NDDM simulating a Turing machine with TIME in special ranges.

They also introduced the class D of all sets $\mathcal{M}$ having representations of the form

$$\begin{aligned}\langle a_1,\ldots,a_n\rangle\in\mathcal{M} \iff {} & \exists x_1\ldots x_m\{D(a_1,\ldots,a_n,x_1,\ldots,x_m)=0\\ & \& x_1+\ldots+x_m\le 2^{|a_1+\ldots+a_n|^k}\},\end{aligned}$$

where $|a|$ denotes, as usual, the (binary) length of a. (Note that in this definition it is not supposed that the polynomial D supplies a Diophantine representation for $\mathcal{M}$, i.e. the equivalence

$$\langle a_1,\ldots,a_n\rangle\in\mathcal{M} \iff \exists x_1\ldots x_m\{D(a_1,\ldots,a_n,x_1,\ldots,x_m)=0\}$$

needs not to hold.) It is easy to see that $\mathbf{D}\subset\mathbf{NP}$. The class $\mathbf{D}$ is known to contain $\mathbf{NP}$-complete problems and Adleman and Manders asked whether in fact $\mathbf{D}=\mathbf{NP}$.

As a partial progress towards proving this equality we can consider the papers [63, 64, 19] where the class $\mathbf{NP}$ and other classes have been defined via an analog of Davis' normal form (with proper bounds on the universal and existential quantifiers). Historically, Davis' normal form was the first step towards the proof of DMPR theorem. Unfortunately, all known methods of eliminating the universal quantifier are too costly in terms of the size of solutions of resulting equations.

I will mention another result of Adleman and Manders which is not technically connected with the works on Hilbert's tenth problem but is formally very close. Namely, they [30] proved that deciding whether a Diophantine equation of the form $ax^2+by=c$ is solvable in positive integers x and y is $\mathbf{NP}$-complete.

Reductions of Hilbert's tenth problem to other decision problems

It was difficult to establish the undecidability of Diophantine equations, and to a great extent this was due to their simple form. This pays off: with the proof of the undecidability of Hilbert's tenth problem computer science got a powerful tool for proving undecidability of other decision problems. Here the simplicity of Diophantine equations is often of great help in reducing Hilbert's tenth problem to other problems.

In this way new, simpler proofs were given for some decision problems already known to be undecidable; also, a number of decision problems were originally shown undecidable by reduction of Hilbert's tenth problem to these problems. Below I list only several such examples just to show the great variety of situations in which we can encounter Diophantine equations, sometimes in disguised form; more than a hundred relevant references are given in my book [41] (the bibliography from the book is available on the Internet [65]).

Straightline programs. It is well-known that the *program equivalence problem* is undecidable: *it is impossible to determine, for arbitrary two computer programs, whether they are equivalent or not.* The undecidability of Hilbert's tenth problem implies that we can restrict the class of programs to very simple ones. For example, in order to get undecidability it is sufficient to compare two programs like

```
if D(x1,x2,x3,x4,x5,x6,x7,x8,x9)=0
   then return 0
   else return 1
```

and

```
return 1
```

(as it was mentioned above, 9 is the best known today (1998) bound for the number of unknowns sufficient for a Diophantine representation of every r.e. set). This trivial corollary can be improved in two directions. First, there is no need for conditional operators or loops, the equivalence problem is undecidable already for straightline programs (see [20]). Second, the required number of variables can be smaller than the number of unknowns required for the undecidability of Diophantine equations. The following is one of several theorems of this type proved by my former student D. V. Shiryaev [55]: *it is impossible to decide, for a given straightline program consisting of operators of the forms* `x ← 1`, `x ← x+y`, `x ← x div y` *whether it always returns 1, even if we restrict ourselves to programs with 3 input and 4 local variables.* For such programs the *problem of simplification* is also undecidable.

Word problem for groups. It was mentioned above that the first decision problem which arose in mathematics and was shown undecidable, was Thue's problem for semigroups. Its counterpart, the *word problem for groups*, was stated a few years earlier by M. Dehn [15], but it turned out much more difficult. The first proofs of its undecidability given by P. S. Novikov [46, 47], W. W. Boone [4] were based on the undecidability of Thue's problem. M. K. Valiev [59] used the undecidability of Hilbert's tenth problem for giving a much simpler proof of the undecidability of Dehn's problem.

Petri nets. One of the important notions in computer science is concurrency. Different tools were proposed for describing and analyzing parallel computations, in particular, *Petri nets.* M. Rabin used the undecidability of (exponential) Diophantine equations in order to prove that the *inclusion problem for Petri nets* is undecidable. He never published this result. A proof can be found in the paper [17] by M. Hack who also proved a stronger result about the undecidability of the *equality problem for Petri nets.*

Calculi. When solving Diophantine equations we seek for solutions in the discrete set of integers. Nevertheless, Hilbert's tenth problem was used for establishing the undecidability of many problems in calculi. Such results put intrinsic limitations on capabilities of systems of computer algebra. For example, it is impossible to determining, for an arbitrary system of polynomial *differential* equations, whether it has a solution on a given interval or whether its solution is unique or not.

Unification problems. The basic idea behind the use of Hilbert's tenth problem for establishing the undecidability of some other decision problems is always essentially the same: we need to find in the latter problem some objects which can represent natural numbers, and express addition and multiplication in the language of the problem which we want to show to be undecidable. If the problem is far from arithmetic, such a reduction might be technically not easy. Sometimes it is difficult to anticipate that Hilbert's tenth problem can be reduced to another problem looking quite differently.

In the winter 1994/1995 I was visiting the Computing (sic!) Science Department of Uppsala University on an invitation of A. Voronkov. At that time he was working on the so called *simultaneous rigid E-unification*, SREU for short. The need for such unification arises in some approaches to automatic deduction of mathematical theorems in first order theories with equality. Several papers were published about SREU proving that this problem is NP-complete, EXPTIME-complete and NEXPTIME-complete; each new paper stated that the previous one was wrong, but was wrong itself too.

Voronkov suggested to search for the undecidability of SREU. It was quite natural for me to try to reduce Hilbert's tenth problem to SREU. However, there was no evident way to simulate multiplication by rigid equations because they deal with arbitrary terms rather than with numbers. I failed, and soon left Uppsala.

A few month later A. Voronkov, in collaboration with A. Degtyarev, proved the undecidability of SREU. Their proof was by reduction of the so called *monadic semi-unification problem* to SREU. The former problem was previously shown to be undecidable by M. Baaz using a reduction of the *second order unification problem.* In its turn, that problem was shown undecidable by W. D. Goldfarb, and that was ultimately done by a reduction of Hilbert's tenth problem! Later, A. Degtyarev and A. Voronkov [14] gave a more direct proof of the undecidability of SREU by reducing Diophantine equations to systems of rigid E-equations.

References

1. Adamowicz Zofia, Zbierski Paveł. *Logic of Mathematics.* John Wiley & Sons, New York a. o., 1997.
2. L. Adleman, K. Manders. Diophantine complexity. In: *17th Annual Symposium on Foundations of Computer Science*, pages 81–88, Houston, Texas, 25-26 October 1976. IEEE.

3. M. Blum. A machine-independent theory of the complexity of recursive functions. *Journal of the ACM*, 14(2):322–336, 1967.
4. W. W. Boone. The word problem. *Ann. Math.* , 70(2):207–265, 1959.
5. E. Börger. *Computability, Complexity, Logic.* North-Holland, Amsterdam, 1989.
6. M. Davis. Arithmetical problems and recursively enumerable predicates (abstract). *J. Symbolic Logic*, 15(1):77–78, 1950.
7. Davis M. Arithmetical problems and recursively enumerable predicates. *Journal of Symbolic Logic*, 18(1):33–41, 1953.
8. M. Davis. Speed-up theorems and Diophantine equations. In Randall Rustin, editor, *Courant Computer Science Symposium 7: Computational Complexity*, pages 87–95. Algorithmics Press, New York.
9. M. Davis. Hilbert's tenth problem is unsolvable. *Amer. Math. Monthly*, 80(3):233–269, 1973 (reprinted in [10]).
10. M. Davis. *Computability and Unsolvability.* Dover Publications, New York, 1982.
11. M. Davis, H. Putnam, and J. Robinson. The decision problem for exponential Diophantine equations. *Ann. Math.* , 74(3):425–436, 1961.
12. M. Davis and H. Putnam. A computational proof procedure; Axioms for number theory; Research on Hilbert's Tenth Problem. O. S. R. Report AFOSR TR59-124, U. S. Air Force, October, 1959.
13. G. V. Davydov, Yu. V. Matijasevich, G. E. Mints, V. P. Orevkov, A. O. Slissenko, A. V. Sochilina, and N. A. Shanin, "Sergei Yur'evich Maslov" (obituary). *Russian Math. Surveys* 39(2) (1984), 133-135. [translated from *Uspekhi Matem. Nauk* 39(236) (1984), 129-131.]
14. A. Degtyarev and A. Voronkov. Simultaneous Rigid E-Unification is Undecidable. CSL'95, Computer Science Logic, 9th International Workshop, Paderborn, Germany, September 1995. (Kleine H. Büuming, editor) *Lecture Notes in Computer Science* 1092, pp. 178-190, 1996.
15. M. Dehn. Über die Topologie des dreidimensionalen Raumes. *Math. Ann.* , 69:137–168, 1910.
16. K. Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. I. *Monatsh. Math. und Phys.* , 38(1):173–198, 1931.
17. M. Hack. The equality problem for vector addition systems is undecidable. *Theoretical Computer Science*, 2(1):77–95, 1976.
18. David Hilbert, Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker Kongress zu Paris 1900, *Nachr. K. Ges. Wiss., Göttingen, Math.-Phys. Kl.* (1900), 253-297. See also *Arch. Math. Phys.* (1901) 44-63, 213-237. See also David Hilbert, *Gesammelte Abhandlungen*, Berlin : Springer, vol. 3 (1935), 310 (Reprinted: New York : Chelsea (1965)). French translation with corrections and additions : *Compte rendu du Deuxième Congrès International des Mathématiciens tenu à Paris du 6 au 12 août 1900*, Gauthier-Villars, 1902, pp. 58-114 (réédition : Editions Gabay, Paris 1992). English translation : *Bull. Amer. Math. Soc.* (1901-1902) 437-479. Reprinted in: *Mathematical Developments arising from Hilbert problems*, Proceedings of symposia in pure mathematics, vol. 28, American Mathematical Society, Browder Ed., 1976, pp. 1-34.
19. Bernard R. Hodgson and Clement F. Kent. A normal form for arithmetical representation of $\mathcal{NP}$-sets. *Journal of Computer and System Sciences*, 27(3):378–388, 1983.
20. O. H. Ibarra and B. S. Leininger. On the simplification and equivalence problems for straight-line programs. *Journal of the ACM*, 30(3):641–656, 1983.
21. James P. Jones. Universal Diophantine equation, *J. Symbolic Logic* 47 (1982), 549-571.

22. J.P. Jones and Y.V. Matijasevič. Register machine proof of the theorem on exponential Diophantine representation of enumerable sets. *J. Symbolic Logic*, 49(3):818–829, 1984.
23. J. P. Jones, Y. V. Matijasevič. Proof of recursive unsolvability of Hilbert's tenth problem. *Amer. Math. Monthly* 98(8):689–709, 1991.
24. James P. Jones, Daihachiro Sato, Hideo Wada, and Douglas Wiens, Diophantine representation of the set of prime numbers. *The American Mathematical Monthly*, 83(6):449–464, 1976.
25. J. P. Jones Diophantine representation of Mersenne and Fermat primes. *Acta Arithmetica*, 35(3):209–221, 1979.
26. G.Kreisel, A3061: Davis, Martin; Putnam, Hilary; Robinson, Julia. The decision problem for exponential Diophantine equations. *Mathematical Reviews*, 24A (6A):573, 1962.
27. E. E. Kummer. Über die Ergänzungssätze zu den allgemeinen Reciprocitätsgesetzen. Journal für die Reine und Angewandte Mathematik, 44:93–146, 1852.
28. J.Lambek. How to program an infinite abacus. *Canad. Math. Bull.* , 4:295–302, 1961.
29. G.S. Makanin. The problem of solvability of equations in a free semigroup. *Math. USSR Sbornik*, 32(2):129–198, 1977.
30. K.L. Manders and L. Adleman. *NP*-complete decision problems for binary quadratics. *J. Comput. System Sci.* , 16(2):168–184, 1978.
31. Yu.I. Manin. *A Course in Mathematical Logic.* Springer, New York; Heidelberg; Berlin, 1977.
32. M. Margenstern. Le théorèm de Matiyassévitch et résultats connexes. In C. Berline, K. McAloon, and J.-P. Ressayre, editors, *Model Theory and Arithmetic*, volume 890 of *Lecture Notes in Mathematics*, pages 198–241. Springer-Verlag, 1981.
33. A. A. Markoff. Impossibility of certain algorithms in the theory of associative systems (in Russian). *Dokl. Akad. Nauk SSSR*, 55(7):587–590, 1947.
34. Per Martin-Löf. *Notes on Constructive Mathematics.* Almqvist and Wiksell, Stockholm, 1970.
35. Yu. Matiyasevich. Svyaz' sistem uravneniĭ v slovakh i dlinakh s 10-ĭ problemoĭ Gil'berta, *Zap. nauch. Seminar. Leningr. otd. Mat. in-ta AN SSSR*, 8:132–144, 1968. English translation: The connection between Hilbert's Tenth Problem and systems of equations between words and lengths. *Seminars in Mathematics, V. A. Steklov Mathematical Institute*, 8:61–67, 1970.
36. Yu. Matiyasevich. Diofantovy mnozhestva. *Uspekhi Mat. Nauk*, 27:5(167),185–222,1972. Translated in: *Russian Mathematical Surveys*, 27(5):124–164, 1972.
37. Yu.V. Matiyasevich. Novoe dokazatel'stvo teoremy ob èksponentsial'no diofantovom predstavlenii perechislimykh predikatov. *Zap. nauchn. seminar. Leningr. otd. Mat. in-ta AN SSSR*, 60:75–92, 1976. Translated in: *J. Soviet Math.* , 14(5):1475-1486, 1980.
38. Yu. Matiyasevich. Prostye chisla perechislyayutsya polinomom ot 10 peremennykh. *Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo Instituta im. V. A. Steklova AN SSSR* , 68:62–82. (Translated as Yu.V. Matijasevič. Primes are nonnegative values of a polynomial in 10 variables. *Journal of Soviet Mathematics*, 15(1):33–44, 1981.)
39. Yu. Matijasevich. Some purely mathematical results inspired by mathematical logic, *Proceedings of Fifth International Congress on Logic, Methodology and Philosophy of Science, London, Ontario, 1975*, Dordrecht : Reidel (1977), 121-127.

40. Yu. V. Matiyasevich. Algorifmicheskaya nerazreshimost' èksponentsial'no diofantovykh uravneniĭ s tremya neizvestnymi. *Issledovaniya po teorii algorifmov i matematicheskoĭ logike*, A. A. Markov and V. I. Homič, Editors, Akademiya Nauk SSSR, Moscow 3:69–78,1979. Translated in: *Selecta Mathematica Sovietica*, 3(3):223–232, 1983/84.

41. Yu. Matiyasevich. *Desyataya Problema Gilberta.* Moscow, Fizmatlit, 1993. English translation: Hilbert's tenth problem. MIT Press, 1993. French translation: Le dixième problème de Hilbert, Masson, 1995

42. Yu. Matiysevich. A direct method for simulating partial recursive functions by Diophantine equations. Annals Pure Appl. Logic, 67, 325–348, 1994.

43. Z. A. Melzak. An informal arithmetical approach to computability and computation. *Canad. Math. Bull.* , 4:279–294, 1961.

44. M. L. Minsky. Recursive unsolvability of Post's problem of "tag" and other topics in the theory of Turing machines. *Ann. of Math. (2)*, 74:437–455, 1961.

45. M. L. Minsky. *Computation: Finite and Infinite Machines.* Prentice Hall, Englewood Cliffs; New York, 1967.

46. P. S. Novikov. On algorithmical undecidability of the word problem in the theory of groups (in Russian). *Dokl. Akad. Nauk SSSR*, 85(4):709–712, 1952.

47. P. S. Novikov. On algorithmical undecidability of the word problem in the theory of groups (in Russian). *Trudy Mat. Inst. Steklov.* , 44, 1955.

48. J. Pintz, W. L. Steiger, E. Szemeredi. Infinite sets of primes with fast pramality test. *Mathematics of Computations*, 53(187):399–406, 1989.

49. E. L. Post. Recursive unsolvability of a problem of Thue. *J. Symbolic Logic*, 12:1–11, 1947.

50. H. Putnam. An unsolvable problem in number theory. *J. Symbolic Logic*, 25(3):220–232, 1960.

51. J. Robinson. Existential definability in arithmetic. *Trans. Amer. Math. Soc.* , 72(3):437–449, 1952.

52. J. B. Robinson. The undecidability of exponential Diophantine equations. *Notices of the American Mathematical Society*, 7(1):75, 1960.

53. J. Robinson. Unsolvable Diophantine problems. *Proceedings of the American Mathematical Society*, 22(2):534–538.

54. J. C. Shepherdson and H. E. Sturgis. Computability of recursive functions, *J. ACM* 10(2):217–255, 1963.

55. D. V. Shiryaev. Nerazreshimost' nekotorykh algoritmicheskikh problem dlya nevetvyashchikhsya program. *Kibernetika*, no. 1:63–66, 1989.

56. Th. Skolem. Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen. *Fundamenta Mathematicae*, 23:150–161 (1934).

57. C. Smoryński. *Logical number theory* I*: An Introduction*, Berlin, Springer-Verlag, 1991.

58. A. Thue. Problem über Veränderungen von Zeichenreihen nach gegebenen Regeln. *Vid. Skr. I. Mat.-natur. Kl.* , 10, 1914. Reprinted in: A. Thue. Selected Mathematical Papers. Oslo, 1977, 493–524.

59. M. K. Valiev. On polynomial reducibility of word problem under embedding of recursively presented groups in finitely presented groups. In J. Bečvář, editor, *Mathematical Foundations of Computer Science 1975*, volume 32 of *Lecture Notes in Computer Science*, pages 432–438. Springer-Verlag, September, 1975.

60. A. K. Vinogradov, N. K. Kosovskiĭ. Ierarkhiya diofantovykh predstavleniĭ primitivno rekursivnykh predikatov. *Vychisl. Tekhn. i Voprosy Kibernet.* , Lenigradskiĭ Gosudarstvennyĭ Universitet, Leningrad 12:99–107, 1975.

61. N. N. Vorob'ev. *Fibonacci Numbers*, 3rd ed., Moscow: Nauka, 1969 (in Russian).
62. M. A. Vsemirnov. Infinite sets of primes with Diophantine representations in eight variables. *Zapiski Nauchnykh Seminarov Peterburgskogo Otdeleniya Matematicheskogo Instituta im. V. A. Steklova RAN* , 220:36–48, 1995.
63. S. Yukna. Arifmeticheskie predstavleniya klassov mashinnoĭ slozhnosti. *Matematicheskaya logika i eë primeneniya*, no. 2:92–107. Institut Matematiki i Kibernetiki Akademii Nauk Litovskoĭ SSR, Vil'nyus, 1982.
64. S. Yukna. Ob arifmetizatsii vychisleniĭ. *Matematicheskaya logika i eë primeneniya*, no. 3:117–125. Institut Matematiki i Kibernetiki Akademii Nauk Litovskoĭ SSR, Vil'nyus, 1983.
65. URL: `http://logic.pdmi.ras.ru/Hilbert10`.