

**Reflections on
Hilbert's Tenth Problem
(with a new conjecture)**

Martin Davis

**Professor Emeritus
Courant Institute, NYU**

**Visiting Scholar
UC Berkeley**

Unless otherwise stated, we'll work with the *natural numbers*:

$$N = \{0, 1, 2, 3, \dots\}$$

Consider a Diophantine equation

$$D(a_1, a_2, \dots, a_n, x_1, x_2, \dots, x_m) = 0$$

Here, a_1, a_2, \dots, a_n are *parameters*, and x_1, x_2, \dots, x_m are *unknowns*. For such a given equation, it is usual to ask:

For which values of the parameters does the equation have a solution in the unknowns? In other words, find:

$$\{ \langle a_1, \dots, a_n \rangle \mid \exists x_1, \dots, x_m [D(a_1, \dots, x_1, \dots) = 0] \}$$

We think of the equation $D = 0$ as furnishing a *definition* of the corresponding set.

Examples

- The Pell equation $x^2 - d(y+1)^2 = 1$ *defines* the set consisting of 0 and the numbers not perfect squares.
- $(x+1)^n + (y+1)^n = (z+1)^n$ defines the set $\{1, 2\}$.
- $a = (x+2)(y+2)$ defines the set of *composite numbers*.
- $a = (2x+3)(y+1)$ defines the set of numbers not powers of 2.

Considering Diophantine equations

$$F(a_1, a_2, \dots, a_n, x_1, x_2, \dots, x_m) = 0$$

as defining the corresponding set

$$\{ \langle a_1, \dots, a_n \rangle \mid \exists x_1, \dots, x_m [F(a_1, \dots, x_1, \dots) = 0] \}$$

we distinguish three classes:

- a set is called *Diophantine* if it has such a definition in which F is a polynomial with integer coefficients. We write \mathcal{D} for the *class of Diophantine sets*.
- a set is called *exponential Diophantine* if it has such a definition in which F is an exponential polynomial with integer coefficients. We write \mathcal{E} for the *class of exponential Diophantine sets*.
- a set is called *recursively enumerable* (or *listable*) if it has such a definition in which F is a computable function. We write \mathcal{R} for the *class of recursively enumerable sets*. (“Recursively enumerable” is abbreviated: r.e.)

Evidently:

$$\mathcal{D} \subseteq \mathcal{E} \subseteq \mathcal{R}$$

Converse inclusions?

Remark

The system of equations:

$$E_1 = 0$$

$$E_2 = 0$$

... ..

$$E_n = 0$$

is equivalent to the single equation

$$E_1^2 + E_2^2 + \dots + E_n^2 = 0$$

So, a system of equations is as good as a single equation for giving a Diophantine definition.

Hilbert's 10th problem: Find an algorithm to determine whether a given polynomial Diophantine equation with integer coefficients has an integer solution.

It's equivalent to the analogous problem for solutions in natural numbers:

- $p(x_1, \dots, x_n) = 0$ has a solution in integers if and only if at least one of the 2^n equations

$$p(\pm x_1, \dots, \pm x_n) = 0$$

has a solution in natural numbers.

- $p(x_1, \dots, x_n) = 0$ has a solution in natural numbers if and only if

$$p(q_1^2 + r_1^2 + s_1^2 + t_1^2, \dots, q_1^2 + r_1^2 + s_1^2 + t_1^2) = 0$$

has an integer solution.

Theorem (Church, Post, Turing) *There is a set $K \subseteq N$ such that $K \in \mathcal{R}$, but K is not computable, i.e., there is no algorithm for testing membership in K .*

MRDP (=DPRM) Theorem: $\mathcal{D} = \mathcal{R}$.

$$K = \{a \in N \mid \exists x_1, \dots, x_n [\pi(a, x_1, \dots, x_n) = 0]\}$$

with π a polynomial. So, there is no algorithm to determine, for given a , whether the corresponding equation has a solution .

Hence, Hilbert's 10th problem is unsolvable.

History of MRDP Theorem

Davis 1950: For every $S \in \mathcal{R}$, there is a polynomial p such that

$$S = \{a \mid \exists y \forall k \leq y \exists x_1, \dots, x_n [p(a, k, y, x_1, \dots, x_n) = 0]\}$$

Julia Robinson's Hypothesis (JR) 1950:

There is a function $f \in \mathcal{D}$ such that $f(x) = O(x^x)$ but $f(x) \neq O(x^k)$ for any positive integer k .

Definition: $\text{exp} = \{ \langle a, b, c \rangle \mid c = a^b \}$.

Robinson 1950: $\text{JR} \Rightarrow \text{exp} \in \mathcal{D} \Rightarrow \mathcal{D} = \mathcal{E}$.

Davis, Putnam, Robinson 1961: $\mathcal{E} = \mathcal{R}$.
Hence, $\text{JR} \iff \mathcal{D} = \mathcal{R}$

Matiyasevich (1970): $F_{2n} \in \mathcal{D}$ (where F_n is the n th Fibonacci number). **Hence JR.**

The positive numbers a for which

$$p(a, x_1, \dots, x_n) = 0$$

has a solution is the positive part of the range of the polynomial $a(1 - p^2(a, x_1, \dots, x_n))$. So as Hilary Putnam remarked: *The set of positive integers in an r.e. set is always representable as the positive part of the range of a polynomial.*

Theorem:(Jones,Sato,Wada,Wiens 1976) The positive prime numbers are the positive part of the range of:

$$\begin{aligned}
 (k + 2)\{ & 1 - [wz + h + j - q]^2 \\
 & - [(gk + 2g + k + 1)(h + j) + h - z]^2 \\
 & - [2n + p + q + z - e]^2 \\
 & - [16(k + 1)^3(k + 2)(n + 1)^2 + 1 - f^2]^2 \\
 & - [e^3(e + 2)(a + 1)^2 + 1 - o^2]^2 \\
 & - [(a^2 - 1)y^2 + 1 - x^2]^2 \\
 & - [16r^2y^4(a^2 - 1) + 1 - u^2]^2 \\
 & - [n + \ell + v - y]^2 \\
 & - [((a + u^2(u^2 - a))^2 - 1)(n + 4dy)^2 + 1 - (x + cu)^2]^2 \\
 & - [(a^2 - 1)\ell^2 + 1 - m^2]^2 \\
 & - [q + y(a - p - 1) + s(2ap + 2a - p^2 - 2p - 2) - x]^2 \\
 & - [z + p\ell(a - p) + t(2ap - p^2 - 1) - pm]^2 \\
 & - [ai + k + 1 + \ell - i]^2 \\
 & - [p + \ell(a - n - 1) + b(2an + 2a - n^2 - 2n - 2) - m]^2\}
 \end{aligned}$$

Theorem (Davis 1972) Let $\mathcal{C} = \{0, 1, 2, \dots, \aleph_0\}$. Let $\mathcal{A} \subseteq \mathcal{C}$ where $\mathcal{A} \neq \emptyset$ and $\mathcal{A} \neq \mathcal{C}$. Then, there is no algorithm to determine of a given polynomial Diophantine equation whether the number of solutions of that equation belongs to \mathcal{A} .

Corollary There is no algorithm to determine of a given polynomial Diophantine equation whether the number of solutions of that equation is a prime number, is infinite, is the sum of two squares, etc.

Universal Diophantine Equation

Theorem There is a polynomial Diophantine equation

$$p(a, n, x_1, \dots, x_m) = 0 \quad (1)$$

such for every r.e. set S of natural numbers, there is an n such that $a \in S$ if and only if (1) has a solution x_1, \dots, x_m .

How small can the degree of p be? By a device of Skolem, 4.

What about m , the number of unknowns? Matiyasevich-Robinson: m can be 13. Even (Matiyasevich) 9.

James Jones has investigated the tradeoff between the degree (δ) and the number of unknowns (ν) in a universal equation.

ν	δ	ν	δ	ν	δ
58	4	28	20	21	96
38	8	26	24	19	2668
32	12	25	28	13	6.6×10^{43}
29	16	24	36	9	1.6×10^{45}

There are interesting analogues of Hilbert's tenth problem for many rings, in particular the ring of rational numbers, and the ring of integers of an algebraic number field. The technique typically used to prove the unsolvability of Hilbert's tenth problem for various rings is to provide a Diophantine definition of the rational integers over such rings. Some examples:

1. The ring of integers of any totally real algebraic number field (*Denef, Lipschitz*).
2. The ring of integers of any algebraic number field admitting exactly one pair of conjugate complex embeddings (*Shlapentokh, Pheidas*).
3. The ring of integers of any algebraic number field whose Galois group over the rationals is abelian (*Shapiro, Shlapentokh*).
4. The field of rational functions in one indeterminate over a field of finite characteristic $\neq 2$ (*Pheidas*).
5. Certain rings of elements of algebraic function fields (*Shlapentokh*).

HOLD THE PRESSES!

Barry Mazur and Karl Rubin recently posted a preprint to the ArXiv that may be of interest to FOM readers. They show that if the Shafarevich-Tate group of an elliptic curve over a number field is always finite (actually they assume something weaker than this), then Hilbert's Tenth Problem has a negative answer over the ring of integers of any number field.

<http://arxiv.org/abs/0904.3709>

(They acknowledge Poonen and Shlapentokh)

Poonen's Theorem

We write N for the set of natural numbers, and \mathcal{Q} for the set of rationals. As usual let $\pi(x)$ stand for the number of prime numbers $\leq x$. If A is a set of primes, let $\pi_A(x)$ be the number of elements of A that are $\leq x$.

Making use of elliptic curves, Bjorn Poonen has proved:

Theorem: (Poonen 2003) *There is a computable set A of prime numbers such that*

$$\lim_{x \rightarrow \infty} \frac{\pi_A(x)}{\pi(x)} = 1$$

and Hilbert's 10th problem is unsolvable over the ring \mathcal{U} of all rational numbers whose denominators are products of primes in A .

Definition: *Let \mathcal{R} be a subring of \mathcal{Q} . A set $W \subseteq \mathcal{R}^m$ is Diophantine over \mathcal{R} if there is a polynomial p with coefficients in \mathcal{R} such that*

$$W = \{a \in \mathcal{R}^m \mid (\exists x \in \mathcal{R}^k)[p(a, x) = 0]\}$$

Poonen's Lemma. *There is a computable map $n \rightarrow y_n$ of N into \mathcal{U} such that the sets of triples $\{ \langle y_a, y_b, y_{a+b} \rangle \mid a, b \in N \}$ and $\{ \langle y_a, y_b, y_{ab} \rangle \mid a, b \in N \}$ are both Diophantine over \mathcal{U} .*

Let

$$S = \{a \in N \mid (\exists x \in N^k)[q(a, x) = 0]\},$$

where q is a polynomial with integer coefficients, be a Diophantine definition of some recursively enumerable (r.e.) set. Using Poonen's Lemma we see that there exists a polynomial p with coefficients in \mathcal{U} such that

$$S = \{a \in N \mid (\exists x \in \mathcal{U}^\ell)[p(y_a, x) = 0]\}.$$

Writing

$$\hat{S}_p = \{a \in N \mid (\exists x \in \mathcal{Q}^\ell)[p(y_a, x) = 0]\},$$

we have $S \subseteq \hat{S}_p \subseteq N$.

An r.e set $S \subseteq N$ is *simple* if $N - S$ is infinite but contains no infinite r.e. subset.

Applying the above to a simple set S we have: **either \hat{S}_p is itself simple or $N - S_p$ is finite.**

Conjecture: *There is a Diophantine definition of a simple set S for which $N - \hat{S}_p$ is infinite.*

It is easy to see that this conjecture implies the unsolvability of H10 over \mathcal{Q} .

Why the Conjecture is Plausible

It is easy to construct simple sets, and there are a number of ways to do so.

But: if the conjecture is false, no matter how S is constructed, and no matter what Diophantine definition of S is provided, \hat{S}_p will differ from N by only finitely many elements. Because the additional primes permitted in denominators in the transition from \mathcal{U} to \mathcal{Q} form a sparse set, this seems implausible.