Almost All Diophantine Sets Are Undecidable

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
vladik@utep.edu

Abstract

The known 1970 solution to the 10th Hilbert problem says that no algorithm is possible that would decide whether a given Diophantine equation has a solution. In set terms, this means that not all Diophantine sets are decidable. In a posting to the Foundations of Mathematica mailing list, Timothy Y. Chow asked for possible formal justification for his impression that most Diophantine equations are not decidable. One such possible justification is presented in this paper.

1 Formulation of the Problem

Decidability of Diophantine equations: a brief reminder. A *Diophantine equation* is an equation of the type $P(x_1, \ldots, x_n, c_1, \ldots, c_m) = 0$, where P is a polynomial with integer coefficients, $x_1, \ldots, x_n$ are natural-valued unknowns, and $c_1, \ldots, c_m$ are natural-valued parameters.

For many equations:

- for some values of the parameters $c = (c_1, \ldots, c_m)$, there is a solution, while
- for other values $c = (c_1, \ldots, c_m)$, the equation has no solution.

For some Diophantine equations, an algorithm is known that, given the tuples of parameters $c = (c_1, \ldots, c_m)$, tell us whether the corresponding equation has a solution. One of the problems that David Hilbert formulated in 1900 – as challenges for the 20 century mathematicians – was to find a general algorithm for telling whether a given instance of a Diophantine equation has a solution; this was Problem No. 10; see, e.g., [1, 2].

The answer to this problem turned out to be negative: in 1970, Yuri Matiyasevich proved that no such general algorithm is possible; see, e.g., [3]. Moreover, he proved the following stronger result.

Namely, one can easily check that for each Diophantine equation, the corresponding *Diophantine set* D_P – i.e., the set

$$D_P \stackrel{\text{def}}{=} \{(c_1, \ldots, c_m) : \exists x_1 \ \ldots \ \exists x_n \, (P(x_1, \ldots, x_n, c_1, \ldots, c_m) = 0)\}$$

of all the tuples $c = (c_1, \ldots, c_m)$ for which there the corresponding instance of the equation has a solution is *computationally enumerable* in the sense that there exists an algorithm that eventually enumerates all the elements of this set.

Indeed, we can order all possible tuples $(x_1, \ldots, x_n, c_1, \ldots, c_m)$ into a single sequence and for each of these tuples, check whether $P(x_1, \ldots, x_n, c_1, \ldots, c_m) = 0$. If this equality is satisfied, we produce the tuple $(c_1, \ldots, c_m)$.

One can easily check that this procedure will produce all the tuples from the Diophantine set D_P – and only these tuples.

What Matiyasevich proved is that, vice versa, every computationally enumerable set is Diophantine, i.e., for every computationally enumerable set S, there exists a polynomial P for which $S = D_P$.

Question. In a 2017 posting to the Foundations of Mathematics mailing list, Timothy Y. Chow asked the following question: *The impression I've gotten—although I don't think I've seen it explicitly asserted anywhere—is that in some sense "most" Diophantine sets are* not *computable. Are there any results, or even heuristic arguments, in this direction?*

Why this equation is difficult. This question would have been more precise – and potentially having a direct answer – if there was a natural probability measure on the set of all Diophantine equations. If such a measure existed, we would check what is the probability of the set of all decidable Diophantine equations.

However, no such natural measure is known, so we have to come up with a less direct answer to this question.

What we do in this paper. In this paper, we prove a simple argument formally justifying this impression.

2 Towards Our Explanation

The first auxiliary result that we use in our justification. For every two sets A and B, we can form their *disjoint union* $A \sqcup B \stackrel{\text{def}}{=} (\{0\} \times A) \cup (\{1\} \times B)$. Let us show that the disjoint union of two Diophantine sets is also a Diophantine set.

Indeed, let the first Diophantine set is corresponds to the polynomial $P(x, c)$ and the second one to the polynomial $P'(x, c)$. Let us that their disjoint union can be represented by the equation

$$P'' \stackrel{\text{def}}{=} c'' \cdot (P(x, c))^2 + (1 - c'') \cdot (P'(x, c))^2 + (c'' \cdot (1 - c''))^2 = 0,$$

for some auxiliary natural-valued parameter c'' (which will be equal to 0 or 1).

Indeed:

- If for some c, the equation $P(x,c) = 0$ has a solution, then by setting $c'' = 1$ we get a solution to the above equation.
- Similarly, if for some c, the equation $P'(x,c) = 0$ has a solution, then we get a solution to the above equation by setting $c'' = 0$ and thus, $1 - c'' = 1$.

Vice versa, if the above equation has a solution, then, since the sum of three non-negative terms forming the polynomial P'' is equal to 0, all three terms must be equal to 0. From the fact that the third term $(c'' \cdot (1 - c''))^2$ is equal to 0, we conclude that either $c'' = 1$ or $c'' = 0$.

- In the first case, we get $P(x,c) = 0$, hence $c \in A$.
- In the second case, we get $P'(x,c) = 0$ and thus, $c' \in B$.

Thus, indeed, the Diophantine set $D_{P''}$ corresponding to the polynomial P'' has the form $(\{0\} \times A) \cup (\{1\} \times B)$, i.e., is the disjoint union of the sets A and B.

A simple consequence of this auxiliary result. By induction, we can now easily prove that for every tuples $A, \ldots, B$ or Diophantine sets, their disjoint union $A \sqcup \ldots \sqcup B$ is also a Diophantine set

Preliminary construction. Let us effectively enumerate all Diophantine sets into a sequence $S_1, \ldots, S_n, \ldots$

Let us now pick some natural number n, and consider sets $S_1, \ldots, S_n$.

Let us apply disjoint union to the preliminary construction. For any natural number d, if we consider disjoint unions $S_{i_1} \sqcup \ldots \sqcup S_{i_d}$ of d sets S_{i_k} $(1 \le i_k \le n)$, then we get n^d possible product sets.

The whole class of Diophantine sets can be obtained if we tend both n and d to infinity.

The second auxiliary result that we use in our justification. It is easy to see that the disjoint union $A \sqcup \ldots \sqcup B$ of several sets $A, \ldots, B$ is decidable if and only if all the sets A, $\ldots$, B are decidable.

Explanation. Now, we are ready for the desired explanation.

Let $m(n)$ denote the number of decidable sets among the first n Diophantine sets S_1, $\ldots$, S_n. Then out of n^d possible disjoint unions, $(m(n))^d$ are decidable, so the proportion of decidable products is $\left(\dfrac{m(n)}{n}\right)^d$.

Since some Diophantine sets are undecidable, for large n, we get $m(n) < n$ and thus, $\dfrac{m(n)}{n} < 1$. Thus, for all sufficiently large n, when d tends to infinity, we have $\left(\dfrac{m(n)}{n}\right)^d \to 0$, i.e., we conclude that the proportion of decidable sets tends to 0.

This is true for every n, so the limit of this limit when n tends to infinity is also 0.

In this sense, the proportion of decidable sets is 0, and thus, *almost all Diophantine sets are undecidable.*

Acknowledgments

This work was supported in part by the National Science Foundation grant HRD-1242122 (Cyber-ShARE Center of Excellence).

The author is thankful to Martin Davis and to all the participants of the Foundations of mathematics mailing list for valuable discussions.

References

[1] F. E. Browder (ed.), *Mathematical Developments Arising from Hilbert's Problems*, American Mathematical Society, Providence, Rhode Island, 1976.

[2] D. Hilbert, "Mathematical problems, lecture delivered before the International Congress of Mathematics in Paris in 1900", *Bulletin of the American Mathematical Society*, 1902, Vol. 8, pp. 437–479.

[3] Yu. V. Matiyasevich, *Hilbert's Tenth Problem*, MIT Press, Cambridge, Massachusetts, 1993.