Undecidability in number theory

Bjorn Poonen

University of California at Berkeley
(on sabbatical at Harvard and MIT in Fall 2007)

September 20, 2007

Examples of polynomial equations

Do there exist integers x, y, z such that

$$x^3 + y^3 + z^3 = 29?$$

Examples of polynomial equations

Do there exist integers x, y, z such that

$$x^3 + y^3 + z^3 = 29?$$

Yes: $(x, y, z) = (3, 1, 1)$.

Examples of polynomial equations

Do there exist integers x, y, z such that

$$x^3 + y^3 + z^3 = 30?$$

Examples of polynomial equations

Do there exist integers x, y, z such that

$$x^3 + y^3 + z^3 = 30?$$

Yes: $(x, y, z) = (-283059965, -2218888517, 2220422932)$.

(discovered in 1999 by E. Pine, K. Yarbrough, W. Tarrant, and M. Beck, following an approach suggested by N. Elkies.)

Examples of polynomial equations

Do there exist integers x, y, z such that

$$x^3 + y^3 + z^3 = 33?$$

Examples of polynomial equations

Do there exist integers x, y, z such that

$$x^3 + y^3 + z^3 = 33?$$

Unknown.

Examples of polynomial equations

Do there exist integers x, y, z such that

$$536x^{287196896} - 210y^{287196896} + 777x^3y^{16}z^{4732987}$$
$$-1111x^{54987896} - 2823y^{927396} + 27x^{94572}y^{9927}z^{999}$$
$$-936718x^{726896} + 887236y^{726896} - 9x^{24572}y^{7827}z^{13}$$
$$+89790876x^{26896} + 30y^{26896} + 987x^{245}y^6z^{6876}$$
$$+9823709709790790x^{28} - 1987y^{28} + 1467890461986x^2y^6z^4$$
$$+80398600x^2z^{12} - 27980186xy + 3789720156y^2 + 9328769x$$
$$-1956820y - 2758932498572709879076864584689 8z = 389?$$

Hilbert's tenth problem

D. Hilbert, in the 10th of the list of 23 problems he published after a famous lecture in 1900, asked his audience to find a method that would answer all such questions.

Hilbert's tenth problem (H10)

Find an algorithm that solves the following problem:

input: *a multivariable polynomial* $f(x_1, \ldots, x_n)$ *with integer coefficients*

output: *YES or NO, according to whether there exist integers* $a_1, a_2, \ldots, a_n$ *such that* $f(a_1, \ldots, a_n) = 0$.

More generally, one could ask for an algorithm for solving a system of polynomial equations, but this would be equivalent, since

$$f_1 = \cdots = f_m = 0 \iff f_1^2 + \cdots + f_m^2 = 0.$$

Hilbert's tenth problem

Hilbert's tenth problem (H10)

Find a Turing machine that solves the following problem:

input: *a multivariable polynomial $f(x_1, \ldots, x_n)$ with integer coefficients*

output: *YES or NO, according to whether there exist integers $a_1, a_2, \ldots, a_n$ such that $f(a_1, \ldots, a_n) = 0$.*

Theorem (Davis-Putnam-Robinson 1961 + Matiyasevich 1970)

No such algorithm exists.

In fact they proved something stronger...

Diophantine sets

Definition

$A \subseteq \mathbb{Z}$ is diophantine if there exists

$$p(t, \vec{x}) \in \mathbb{Z}[t, x_1, \ldots, x_m]$$

such that

$$A = \{\, a \in \mathbb{Z} : (\exists \vec{x} \in \mathbb{Z}^m)\ p(a, \vec{x}) = 0 \,\}.$$

Example

The subset $\mathbb{N} := \{0, 1, 2, \ldots\}$ of $\mathbb{Z}$ is diophantine, since for $a \in \mathbb{Z}$,

$$a \in \mathbb{N} \iff (\exists x_1, x_2, x_3, x_4 \in \mathbb{Z})\ x_1^2 + x_2^2 + x_3^2 + x_4^2 - a = 0.$$

Listable sets

Definition

$A \subseteq \mathbb{Z}$ is listable (recursively enumerable) if there is a Turing machine such that A is the set of integers that it prints out when left running forever.

Example

The set of integers expressible as a sum of three cubes is listable.

> *(Print out $x^3 + y^3 + z^3$ for all $|x|, |y|, |z| \leq 10$, then print out $x^3 + y^3 + z^3$ for $|x|, |y|, |z| \leq 100$, and so on.)*

Negative answer to H10

What Davis-Putnam-Robinson-Matiyasevich really proved is:

DPRM theorem: Diophantine $\iff$ listable

(They showed that the theory of diophantine equations is rich enough to simulate any computer!)

The DPRM theorem implies a negative answer to H10:

- Recursion theory (a diagonal argument) provides a listable set for which no algorithm can decide membership.
- So there exists a *diophantine* set for which no algorithm can decide membership.
- Thus H10 has a negative answer.

More fun consequences of the DPRM theorem

"Diophantine $\iff$ listable" has applications beyond the negative answer to H10:

- Prime-producing polynomials
- Diophantine statement of the Riemann hypothesis

The set of primes equals the set of positive values assumed by the 26-variable polynomial

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

as the variables range over nonnegative integers (J. Jones, D. Sato, H. Wada, D. Wiens).

Riemann hypothesis

The DPRM theorem gives an explicit polynomial equation that has a solution in integers if and only if the Riemann hypothesis is false.

Sketch of proof.

- One can write a computer program that, when left running forever, will detect a counterexample to the Riemann hypothesis if one exists.
- Simulate this program with a diophantine equation. □

H10 over $\mathbb{Q}$

- It is not known whether there exists an algorithm that decides whether a multivariable polynomial equation has a solution in rational numbers.
- If $\mathbb{Z}$ is diophantine over $\mathbb{Q}$, then the negative answer for $\mathbb{Z}$ implies a negative answer for $\mathbb{Q}$.
- But there is a conjecture that implies that $\mathbb{Z}$ is *not* diophantine over $\mathbb{Q}$:

Conjecture (Mazur 1992)

For any polynomial equation $f(x_1, \ldots, x_n) = 0$ with rational coefficients, if S is the set of rational solutions, then the closure of S in $\mathbb{R}^n$ has at most finitely many connected components.

First-order sentences over $\mathbb{Z}$

- In terms of logic, H10 asks for an algorithm to decide the truth of positive existential sentences

$$(\exists x_1 \exists x_2 \cdots \exists x_n)\ p(x_1, \ldots, x_n) = 0.$$

  in the language of rings, where the variables run over integers.
- More generally, one can ask for an algorithm to decide the truth of arbitrary first-order sentences, in which any number of bound quantifiers $\exists$ and $\forall$ are permitted: a typical such sentence is

$$(\exists x)(\forall y)(\exists z)(\exists w) \quad (x \cdot z + 3 = y^2) \ \vee \ \neg(z = x + w)$$

- Long before DPRM, the work of Church, Gödel, and Turing in the 1930s made it clear that there was no algorithm to solve the harder problem of deciding the truth of first-order sentences over $\mathbb{Z}$.

First-order sentences over $\mathbb{Q}$

Though it is not known whether $\mathbb{Z}$ is diophantine (i.e., definable by a positive existential formula) over $\mathbb{Q}$, we have

Theorem (J. Robinson 1949)

One can characterize $\mathbb{Z}$ as the set of $t \in \mathbb{Q}$ such that a particular first-order formula of the form

$$(\forall \vec{x})(\exists \vec{y})(\forall \vec{z})(\exists \vec{w})\ p(t, \vec{x}, \vec{y}, \vec{z}, \vec{w}) = 0$$

is true, when the variables range over rational numbers.

Corollary

There is no algorithm to decide the truth of a first-order sentence over $\mathbb{Q}$.

Using quaternion algebras, one can improve J. Robinson's result to

Theorem (P. 2007)

It is possible to define $\mathbb{Z}$ in $\mathbb{Q}$ with a formula with 2 universal quantifiers followed by 7 existential quantifiers.

Corollary

There is no algorithm for deciding, given an algebraic family of morphisms of varieties, whether there exists one that is surjective on rational points.

Theorem (P. 2007)

The set $\mathbb{Z}$ equals the set of $t \in \mathbb{Q}$ such that

$$\begin{gathered}(\forall a, b)(\exists x_1, x_2, x_3, x_4, y_2, y_3, y_4) \\ (a + x_1^2 + x_2^2 + x_3^2 + x_4^2)(b + x_1^2 + x_2^2 + x_3^2 + x_4^2) \\ \cdot \Big[\left(x_1^2 - a x_2^2 - b x_3^2 + a b x_4^2 - 1\right)^2 \\ + \prod_{n=0}^{2309} \left((n - t - 2x_1)^2 - 4a y_2^2 - 4b y_3^2 + 4ab y_4^2 - 4\right)^2 \Big] \\ = 0\end{gathered}$$

is true, when the variables range over rational numbers.

H10 over subrings of $\mathbb{Q}$

Let $\mathcal{P} = \{2, 3, 5, \ldots\}$. There is a bijection

$$\{\text{subsets of } \mathcal{P}\} \leftrightarrow \{\text{subrings of } \mathbb{Q}\}$$
$$S \mapsto \mathbb{Z}[S^{-1}].$$

Examples:

- $S = \emptyset$, $\mathbb{Z}[S^{-1}] = \mathbb{Z}$, *answer is negative*
- $S = \mathcal{P}$, $\mathbb{Z}[S^{-1}] = \mathbb{Q}$, *answer is unknown*

- What happens for S in between?
- How large can we make S (in the sense of density) and still prove a negative answer for H10 over $\mathbb{Z}[S^{-1}]$?
- For finite S, a negative answer follows from work of Robinson, who used the Hasse-Minkowski theorem (local-global principle) for quadratic forms.

H10 over subrings of $\mathbb{Q}$, continued

Theorem (P., 2003)

There exists a recursive set of primes $S \subset \mathcal{P}$ of density 1 such that

1. *There exists a curve E such that $E(\mathbb{Z}[S^{-1}])$ is an infinite discrete subset of $E(\mathbb{R})$. (So the analogue of Mazur's conjecture for $\mathbb{Z}[S^{-1}]$ is false.)*
2. *There is a diophantine model of $\mathbb{Z}$ over $\mathbb{Z}[S^{-1}]$.*
3. *H10 over $\mathbb{Z}[S^{-1}]$ has a negative answer.*

The proof takes E to be an elliptic curve (minus ∞), and uses properties of integral points on elliptic curves.

Ring	H10	1st order theory
$\mathbb{C}$	YES	YES
$\mathbb{R}$	YES	YES
$\mathbb{F}_q$	YES	YES
p-adic fields	YES	YES
$\mathbb{F}_q((t))$	?	?
number field	?	NO
$\mathbb{Q}$	?	NO
global function field	NO	NO
$\mathbb{F}_q(t)$	NO	NO
$\mathbb{C}(t)$	?	?
$\mathbb{C}(t_1, \ldots, t_n)$, $n \geq 2$	NO	NO
$\mathbb{R}(t)$	NO	NO
$\mathcal{O}_k$	?	NO
$\mathbb{Z}$	NO	NO

increasing arithmetic complexity