A talk given at Dalian Comb. Conf. (December 12, 2020)

# Combinatorial Ideas related to Hilbert's Tenth Problem

Zhi-Wei Sun

Nanjing University
Nanjing 210093, P. R. China
zwsun@nju.edu.cn
http://math.nju.edu.cn/∼zwsun

December 12, 2019

# Abstract

Hilbert's Tenth Problem (HTP) asks for an effective algorithm to test whether an arbitrary polynomial equation

$$P(x_1, \ldots, x_n) = 0$$

(with integer coefficients) has solutions over the ring of integers. This was finally solved by Matiyasevich in 1970 negatively.

Hilbert's Tenth Problem (HTP) asked for an algorithm to test wether an arbitrary polynomial equation with integer coefficients has solutions over the ring of integers. This was finally solved negatively by Yu. Matiyasevich in 1970. In this talk we introduce combinatorial ideas in the solution of HTP and its later developments. In particular, we focus on Matiyasevich-Robinson's Relation-Combining Theorem and another Relation-Combining Theorem given by the speaker.

# Part I. Hilbert's Tenth Problem

# Hilbert's Tenth Problem

In 1900, at the Paris conference of ICM, D. Hilbert presented 23 famous mathematical problems. He formulated his tenth problem as follows:

*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 in a finite number of operations whether the equation is solvable in rational integers.`

In modern language, Hilbert's Tenth Problem (HTP) asked for an effective algorithm to test whether an arbitrary polynomial equation

$$P(z_1, \ldots, z_n) = 0$$

(with integer coefficients) has solutions over the ring $\mathbb{Z}$ of the integers.

However, at that time the exact meaning of algorithm was not known.

# Diophantine equations over $\mathbb{N}$ and $\mathbb{Z}$

Throughout this talk, variables always range over $\mathbb{Z}$.

Let $P(z_1, \ldots, z_n) \in \mathbb{Z}[z_1, \ldots, z_n]$. Then

$$\begin{aligned} & \exists z_1 \ldots \exists z_n [P(z_1, \ldots, z_n) = 0] \\ \iff & \exists x_1 \geqslant 0 \ldots \exists x_n \geqslant 0 \left[ \prod_{\varepsilon_1, \ldots, \varepsilon_n \in \{\pm 1\}} P(\varepsilon_1 x_1, \ldots, \varepsilon_n x_n) = 0 \right]. \end{aligned}$$

On the other hand, by Lagrange's four-square theorem (each $m \in \mathbb{N}$ can be written as the sum of four squares), we have

$$\begin{aligned} & \exists x_1 \geqslant 0 \ldots \exists x_n \geqslant 0 [P(x_1, \ldots, x_n) = 0] \\ \iff & \exists u_1 \exists v_1 \exists y_1 \exists z_1 \ldots \exists u_n \exists v_n \exists y_n \exists z_n \\ & [P(u_1^2 + v_1^2 + y_1^2 + z_1^2, \ldots, u_n^2 + v_n^2 + y_n^2 + z_n^2) = 0] \end{aligned}$$

So HTP has the following equivalent form (HTP over $\mathbb{N}$): *Is there an algorithm to decide for any polynomial $P(x_1, \ldots, x_n)$ with integer coefficients whether the Diophantine equation $P(x_1, \ldots, x_n) = 0$ has solutions with $x_1, \ldots, x_n \in \mathbb{N}$ ?*

# Diophantine relations and Diophantine sets

A relation $R(a_1, \ldots, a_m)$ with $a_1, \ldots, a_m \in \mathbb{N}$ is said to be *Diophantine* if there is a polynomial $P(t_1, \ldots, t_m, x_1, \ldots, x_n)$ with integer coefficients such that

$$R(a_1, \ldots, a_m) \iff \exists x_1 \geqslant 0 \ldots \exists x_n \geqslant 0[P(a_1, \ldots, a_m, x_1, \ldots, x_n) = 0].$$

For example, the relation $<$ is Diophantine since
$a < b \iff \exists x \geqslant 0[a + x + 1 = b]$.

A set $A \subseteq \mathbb{N}$ is Diophantine if and only if the predicate $a \in A$ is Diophantine.

A relation $R(a_1, \ldots, a_m)$ with $a_1, \ldots, a_m \in \mathbb{N}$ is said to be *exponential Diophantine* if there is a polynomial $P(t_1, \ldots, t_m, x_1, \ldots, x_n, y_1, \ldots, y_k, z_1, \ldots, z_k)$ with integer coefficients such that $R(a_1, \ldots, a_m)$ holds if and only if there are $x_1, \ldots, x_n, y_1, \ldots, y_k, z_1, \ldots, z_k \in \mathbb{N}$ such that

$$P(a_1, \ldots, a_m, x_1, \ldots, x_n, y_1^{z_1}, \ldots, y_k^{z_k}) = 0.$$

# Systems of Diophantine equations

A system of finitely many Diophantine equations is equivalent to a single Diophantine equation.

In fact, if $P_i(z_1, \ldots, z_n) \in \mathbb{Z}[z_1, \ldots, z_n]$ for all $i = 1, \ldots, k$, then

$$\begin{aligned} & P_1(z_1, \ldots, z_n) = 0 \wedge \ldots \wedge P_k(z_1, \ldots, z_n) = 0 \\ \iff & P_1^2(z_1, \ldots, z_n) + \ldots + P_k^2(z_1, \ldots, z_n) = 0. \end{aligned}$$

# $z = \binom{n}{k}$ is exponential Diophantine

If $0 < k \leqslant n$ and $u > 2^n$, then

$$\frac{(u+1)^n}{u^k} = \binom{n}{k} + u \sum_{k<m\leqslant n} \binom{n}{m} u^{m-k-1} + \sum_{0\leqslant i<k} \binom{n}{i} \frac{u^i}{u^k}$$

by the binomial theorem, hence

$$\left\lfloor \frac{(u+1)^n}{u^k} \right\rfloor \equiv \binom{n}{k} \pmod{u}$$

and thus $\binom{n}{k}$ is the least nonnegative residue of $\lfloor (u+1)^n / u^k \rfloor$ modulo $u$.

For $z \geqslant 0$ and $n \geqslant k > 0$, the relation $z = \binom{n}{k}$ holds if and only if there are $u, v, w, x, y \in \mathbb{N}$ such that

$$u > v,\ v = 2^n,\ z \equiv w \pmod{u},\ z < u,$$
$$x = (u+1)^n,\ y = u^k,\ yw \leqslant x < (w+1)y.$$

# $z = n!$ is exponential Diophantine

If $n > 0$ and $m > (2n)^{n+1}$, then

$$n! < \frac{m^n}{\binom{m}{n}} = \frac{n!}{\prod_{r=1}^{n-1}(1 - r/m)} < \frac{n!}{(1 - n/m)^n} < n! \left(1 + \frac{2n}{m}\right)^n < n!+1$$

and thus $n! = \lfloor m^n / \binom{m}{n} \rfloor$.

For $z \geqslant 0$ and $n > 0$, the relation $z = n!$ holds if and only if there are $u, v, w, x, y \in \mathbb{N}$ such that

$$u > v,\ v = w^{n+1},\ w = 2n,\ x = u^n,\ y = \binom{u}{n},\ yz \leqslant x < (z+1)y.$$

Therefore, $z = n!$ is exponential Diophantine!

# Eliminate bounded universal quantifier

**Theorem** (M. Davis, H. Putnam and J. Robinson [Annals of Math. 1961]). There is no algorithm to test whether the equation

$$P(x_1, \ldots, x_n, 2^{x_1}, \ldots, 2^{x_n}) = 0$$

has natural number solutions, where $P$ is an arbitrary polynomial with integer coefficients.

To prove this, one needs the following crucial result.

**Theorem** (Davis-Putnam-Robinson). Let $b \in \mathbb{Z}^+$, $P(y, x_1, \ldots, x_m) \in \mathbb{Z}[y, x_1, \ldots, x_m]$, and $B(b, w) = P^*(b, w, \ldots, w)$ with $P^*(y, x_1, \ldots, x_m)$ obtained by replacing each coefficient in $P(y, x_1, \ldots, x_m)$ by its absolute value. Then

$$\forall 0 \leqslant y < b \exists x_1 \geqslant 0 \ldots \exists x_m \geqslant 0 [P(y, x_1, \ldots, x_m) = 0]$$

$\iff$ there exist $q, w, z_1, \ldots, z_m \in \mathbb{N}$ such that

$$q \equiv -1 \pmod{b!(b + w + B(b, w))!}, \text{ and}$$

$$\binom{q}{b} \text{ divides } \binom{z_1}{w}, \ldots, \binom{z_m}{w} \text{ and } P(q, z_1, \ldots, z_m).$$

# A Lemma

**Lemma**. Let $b, w \in \mathbb{Z}^+$. Suppose that $q \in \mathbb{N}$, $q \equiv -1 \pmod{b!(b+w+B(b,w))!}$ and

$$q_y = \frac{q+1}{y+1} - 1 = \frac{q-y}{y+1} \quad \text{for } y = 0, 1, \ldots, b-1.$$

Then $q_0, \ldots, q_{b-1}, w!$ are pairwise coprime, and $\prod_{y=0}^{b-1} q_y = \binom{q}{b}$.

*Proof.* It is easy to verify the last equality. Let $y \in \{0, \ldots, b-1\}$. As $(b!)^2 \mid q+1$, for any prime $p \leqslant b$ we have

$$q_y = b! \frac{b!}{y+1} \cdot \frac{q+1}{(b!)^2} - 1 \equiv -1 \not\equiv 0 \pmod{p}.$$

If $p$ is a common prime divisor of $q_y$ and $w!$, then $p$ divides $q+1-(q-y) = y+1$, hence $p \leqslant b$ and $p \nmid q_y$ which leads a contradiction. So $p_y$ is coprime to $w!$. If $y' \in \{0, \ldots, b-1\}$ with $y' \neq y$, and $p$ is a prime dividing $q_y$ and $q_{y'}$, then $p$ divides $q-y-(q-y') = y'-y$, hence $p \leqslant b$ and $p \nmid q_y$, which leads a contradiction. So, $q_0, \ldots, q_{b-1}$ are pairwise coprime.

# Proof of the Theorem

$\Rightarrow$: For each $y = 0, \ldots, b-1$ there are $x_{1,y}, \ldots, x_{m,y} \in \mathbb{N}$ with $P(y, x_{1,y}, \ldots, x_{m,y}) = 0$. Take a positive integer $w > \max\{x_{i,y} : \ 1 \leqslant i \leqslant m \ \& \ 0 \leqslant y < b\}$. Then

$$|P(y, x_{1,y}, \ldots, x_{m,y})| \leqslant B(b, w) \quad \text{for all } y = 0, \ldots, b-1.$$

Choose $q \in \mathbb{N}$ and define $q_y$ for $0 \leqslant y < b$ as in the Lemma. Then $q_0, \ldots, q_{b-1}, w!$ are pairwise coprime by the Lemma.

By the Chinese Remainder Theorem, for each $i = 1, \ldots, m$ there is an integer $z_i \geqslant 0$ such that

$$z_i \equiv x_{i,y} \pmod{q_y} \quad \text{for all } y = 0, \ldots, b-1.$$

Thus

$$P(q, z_1, \ldots, z_m) \equiv P(y, x_{1,y}, \ldots, x_{m,y}) = 0 \pmod{q_y}$$

for all $y = 0, \ldots, b-1$, and hence $\binom{q}{b} = \prod_{y=0}^{b-1} q_y$ divides $P(q, z_1, \ldots, z_m)$. As $0 \leqslant x_{i,y} < w$ for each $0 \leqslant y < b$, $q_y$ divides $\prod_{k=0}^{w-1}(z_i - k) = w!\binom{z_i}{w}$ and hence $q_y \mid \binom{z_i}{w}$. Thus $\binom{q}{b} = \prod_{y=0}^{b-1} q_y$ divides $\binom{z_i}{w}$.

# Proof of the Theorem

$\Leftarrow$: For each $y = 0, \ldots, b-1$ define $q_y$ as in the Lemma. By the Lemma, $q_0, \ldots, q_{b-1}, w!$ are pairwise coprime. Let $p_y$ be a prime divisor of $q_y$. Then

$$p_y \mid q_y \mid \binom{q}{b} \mid \binom{z_i}{w},$$

and hence $p_y \mid z_i - x_{i,y}$ for some $0 \leqslant x_{i,y} < w$.

Let $0 \leqslant y < b$. If $p_y \leqslant B(b, w)$, then $p_y$ divides $q + 1 - (q - y) = y + 1$, hence $p_y \leqslant b$ and $p_y \nmid q_y$ which leads a contradiction. Thus $p_y > B(b, w)$. As $p_y \mid q_y \mid q - y$ and $p_y \mid q_y \mid \binom{q}{b}$, we have

$$P(y, x_{1,y}, \ldots, x_{m,y}) \equiv P(q, z_1, \ldots, z_m) \equiv 0 \pmod{p_y}.$$

Since

$$|P(y, x_{1,y}, \ldots, x_{m,y})| \leqslant B(b, w) < p_y,$$

we obtain $P(y, x_{1,y}, \ldots, x_{m,y}) = 0$.

# Julia Robinson's Hypothesis

A. Tarski conjectured in 1948 that $\{2^n : \ n \in \mathbb{N}\}$ is not a Diophantine set. His PhD student J. Robinson did not succeed in proving this with serious efforts.

**JR Hypothesis** (J. Robinson, 1950). There is a Diophantine relation $R(a, b)$ with $a, b \in \mathbb{N}$ such that

$$R(a, b) \Rightarrow b < a^a$$

and

$$\forall k > 0 \exists a \geqslant 0 \exists b \geqslant 0 [R(a, b) \ \& \ a^k < b].$$

Under this hypothesis, J. Robinson showed that the exponential relation $a^b = c$ is Diophantine and hence all r.e. sets are Diophantine. So, the JR Hypothesis implies the negative solution of HTP.

J. Robinson tried to prove her JR Hypothesis but got no success. This made her depressed and doubt her Hypothesis.

# Davis' approach

In 1968 M. Davis showed that if the equation

$$9(u^2+7v^2)^2-7(x^2+7y^2)^2=2\ (u,v,x,y\in\mathbb{N})$$

only has finitely many solutions then the relation $a^b=c$ is Diophantine.

In 1972, Shanks found the first nontrivial solution of the equation with

$$u=525692038369576,\ v=155632703919101 3,$$
$$x=2484616164142152,\ y=1381783865776981.$$

Up to now, nobody can show that the Diophantine equation

$$9(u^2+7v^2)^2-7(x^2+7y^2)^2=2\ (u,v,x,y\in\mathbb{N})$$

only has finitely many solutions.

# Matiyasevich's Theorem

Recall that the Fibonacci sequence $(F_n)_{n\geqslant 0}$ defined by

$$F_0 = 0,\ F_1 = 1,\ \text{ and } F_{n+1} = F_n + F_{n-1}\ (n = 1, 2, 3, \ldots)$$

increases exponentially.
In 1970 Yu. Matiyasevich, a 23-year-old Russian, confirmed the JR Hypothesis by showing that the relation $y = F_{2x}$ (with $x, y \in \mathbb{N}$) is Diophantine! It follows the exponential relation $a^b = c$ (with $a, b, c \in \mathbb{N}$, $a > 1$ and $c > 0$) is Diophantine, i.e. there exists a polynomial $P(a, b, c, x_1, \ldots, x_n)$ with integer coefficients such that

$$a^b = c \iff \exists x_1 \geqslant 0 \ldots \exists x_n \geqslant 0[P(a, b, c, x_1, \ldots, x_n) = 0].$$

This, together with the Davis-Putnam-Robinson work in 1961, led the following result.

**Matiyasevich's Theorem** (or MDPR Theorem) (1970). HTP has a negative solution!

# Part II. Combinatorial Ideas to Reduce Unknowns

# Small $\nu$ with $\exists^\nu$ over $\mathbb{N}$ undecidable

For a set $S \subseteq \mathbb{Z}$ we let $\exists^n$ over $S$ denote the set of formulas

$$\exists x_1 \in S \ldots \exists x_n \in S[P(x_1, \ldots, x_n) = 0]$$

with $P(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$.

It is interesting to find the least $\nu \in \mathbb{Z}^+ = \{1, 2, 3, \ldots\}$ such that $\exists^\nu$ over $\mathbb{N}$ is undecidable.

$\nu < 200$ (Matiyasevich, Summer of 1970)
$\nu \leqslant 35$ (J. Robinson, 1970)
$\nu \leqslant 24$ (Matiyasevich and Robinson, 1970)
$\nu \leqslant 14$ (Matiyasevich and Robinson, 1970)
$\nu \leqslant 13$ (Matiyasevich and Robinson, 1973 [Acta Arith. 27(1975)])
$\nu \leqslant 9$ (Matiyasevich, 1975; details in Jones [J. Symbolic Logic, 1982])

# A lemma

Let $\square = \{x^2 : \ x \in \mathbb{N}\}$ and $A_1, \ldots, A_k \in \mathbb{Z}$. Then

$$\begin{aligned} & A_1, \ldots, A_k \in \square \\ \iff & \exists x_1 \geqslant 0 \ldots \exists x_k \geqslant 0 (x_1^2 = A_1 \wedge \ldots \wedge x_k^2 = A_k) \\ \iff & \exists x_1 \geqslant 0 \ldots \exists x_k \geqslant 0 [(x_1^2 - A_1)^2 + \ldots + (x_k^2 - A_k)^2 = 0]. \end{aligned}$$

To reduce the number of unknowns needed, Matiyasevich and J. Robinson found the following lemma.

# A lemma

Let $\square = \{x^2 : \ x \in \mathbb{N}\}$ and $A_1, \ldots, A_k \in \mathbb{Z}$. Then

$$\begin{aligned} & A_1, \ldots, A_k \in \square \\ \iff & \exists x_1 \geqslant 0 \ldots \exists x_k \geqslant 0 (x_1^2 = A_1 \wedge \ldots \wedge x_k^2 = A_k) \\ \iff & \exists x_1 \geqslant 0 \ldots \exists x_k \geqslant 0 [(x_1^2 - A_1)^2 + \ldots + (x_k^2 - A_k)^2 = 0]. \end{aligned}$$

To reduce the number of unknowns needed, Matiyasevich and J. Robinson found the following lemma.

**Lemma** (Matiyasevich-Robinson, 1975). The polynomial $J_k(x_1, \ldots, x_k, x)$ given by

$$\prod_{\varepsilon_1, \ldots, \varepsilon_k \in \{\pm 1\}} \left( x + \varepsilon_1 \sqrt{x_1} + \varepsilon_2 \sqrt{x_2} X + \ldots + \varepsilon_k \sqrt{x_k} X^{k-1} \right).$$

with $X = 1 + \sum_{j=1}^{k} x_j^2$ has integer coefficients. Moreover, $A_1, \ldots, A_k \in \mathbb{Z}$ are all squares if and only if $J_k(A_1, \ldots, A_k, x) = 0$ for some $x \in \mathbb{N}$.

This can be proved by induction on $k$ and using Galois theory.

# Matiyasevich-Robinson's Relation-Combining Theorem

**Matiyasevich-Robinson's Relation-Combining Theorem** [Acta Arith. 27(1975)] Let $A_1, \ldots, A_k$ and $R, S, T$ be integers with $S \neq 0$. Then

$$\begin{aligned} & A_1 \in \square \wedge \ldots \wedge A_k \in \square \wedge S \mid T \wedge R > 0 \\ \iff & \exists n \geqslant 0 [M_k(A_1, \ldots, A_k, S, T, R, n) = 0], \end{aligned}$$

where $\square = \{x^2 : \ x \in \mathbb{N}\}$, and

$$\begin{aligned} & M_k(x_1, \ldots, x_k, w, x, y, z) \\ = & \prod_{\varepsilon_1, \ldots, \varepsilon_k \in \{\pm 1\}} \left( x^2 + w^2 z - w^2(2y-1)\left( x^2 + X^k + \sum_{j=1}^{k} \varepsilon_j \sqrt{x_j} X^{j-1} \right) \right) \\ = & (w^2(1-2y))^{2^k} J_k\left( x_1, \ldots, x_k, x^2 + X^k + \frac{x^2 + w^2 z}{w^2(1-2y)} \right) \end{aligned}$$

with $X = 1 + \sum_{j=1}^{k} x_j^2$, and $J_k(x_1, \ldots, x_k, x)$ being

$$\prod_{\varepsilon_1, \ldots, \varepsilon_k \in \{\pm 1\}} \left( x + \varepsilon_1 \sqrt{x_1} + \varepsilon_2 \sqrt{x_2} X + \ldots + \varepsilon_k \sqrt{x_k} X^{k-1} \right).$$

# Proof of the Relation-Combining Theorem

Let $W = 1 + \sum_{j=1}^{k} A_j^2$. Recall that

$$\begin{aligned} &M_k(A_1, \ldots, A_k, S, T, R, n) \\ =&(S^2(1-2R))^{2^k} J_k\left(A_1, \ldots, A_k, T^2 + W^k + \frac{S^2 n + T^2}{S^2(1-2R)}\right). \end{aligned}$$

If $A_1, \ldots, A_k \in \square$, $S \mid T$ and $R > 0$, then

$$\begin{aligned} n =&(2R-1)(T^2 + W^k - \sqrt{A_1} - \sqrt{A_2}W - \ldots - \sqrt{A_k}W^{k-1}) - \frac{T^2}{S^2} \\ \geqslant& T^2 + W^k - \sum_{i=0}^{k-1}(W-1)W^i - \frac{T^2}{S^2} \geqslant W^k - (W^k - 1) \geqslant 0 \end{aligned}$$

and $M_k(A_1, \ldots, A_k, S, T, R, n) = 0$.

Now suppose that $M_k(A_1, \ldots, A_k, S, T, R, n) = 0$ for some $n \in \mathbb{N}$. Then $\alpha = T^2 + W^k + (S^2 n + T^2)/(S^2(1-2R))$ is a rational zero of the monic polynomial $J_k(A_1, \ldots, A_k, x)$. As rational algebraic integers lie in $\mathbb{Z}$, we have $\alpha \in \mathbb{Z}$, hence $A_1, \ldots, A_k \in \square$ and $S \mid T$. As $\alpha \leqslant \sum_{i=0}^{k-1} \sqrt{A_i} W^{i-1} \leqslant W^k - 1$ and $n \geqslant 0$, we have $R > 0$.

# Coding idea of Matiyasevich and Robinson (1975)

Let $b \in \mathbb{N}$, $\delta \in \mathbb{Z}^+$, and

$$P(z_0, \ldots, z_\nu) = \sum_{\substack{i_0,\ldots,i_\nu \in \mathbb{N} \\ i_0+\ldots+i_\nu \leqslant \delta}} a_{i_0,\ldots,i_\nu} z_0^{i_0} \cdots z_\nu^{i_\nu}.$$

$$B > 2\delta!(1+b^\delta)\left(1 + \sum_{i_0+\ldots+i_\nu \leqslant \delta} a_{i_0,\ldots,i_\nu}^2\right),$$

$$D(x) = x^{(\delta+1)^{\nu+2}} + \sum_{i_0+\ldots+i_\nu \leqslant \delta} c_{i_0,\ldots,i_\nu} a_{i_0,\ldots,i_\nu} x^{(\delta+1)^{\nu+1} - \sum_{s=0}^{\nu} i_s (\delta+1)^s}$$

with $c_{i_0,\ldots,i_\nu} = i_0! \ldots i_\nu! (\delta - i_0 - \ldots - i_\nu)!$. Then

$P(z_0, \ldots, z_\nu) = 0$ for some $z_0, \ldots, z_\nu \in [0, b]$

$\Longleftrightarrow$ there is a number $c$ of the form $1 + \sum_{i=0}^{\nu} c_i B^{(\delta+1)^i}$ with $c_i \in [0, b]$

such that the coefficient of $x^{(\delta+1)^{\nu+1}}$ in $\left(1 + \sum_{i=0}^{\nu} c_i x^{(\delta+1)^i}\right)^\delta D(x)$

is zero.

# Matiyasevich's idea to use binary representations

For $a, b \in \mathbb{N}$ written in base $p$ with $p$ prime, let $\tau_p(a, b)$ denote the number of carries occurring in the addition of $a$ and $b$. Kummer noted that $\tau_p(a, b) = \operatorname{ord}_p\binom{a+b}{a}$.

Let $b, B \in 2\uparrow = \{2^n : \ n \in \mathbb{N}\}$ with $b \leqslant B$. Let $\delta, \nu \in \mathbb{Z}^+$. For $c = \sum_{j=0}^{(\delta+1)^\nu} c_j B^j$ with $c_j \in [0, B)$, and $M = \sum_{j=0}^{(\delta+1)^\nu} m_j B^j$ with

$$m_j = \begin{cases} B - b & \text{if } j = (\delta+1)^s \text{ for some } s = 1, \ldots, k, \\ B - 1 & \text{otherwise,} \end{cases}$$

$$\tau_2(c, M) = 0 \iff \tau_2(c_j, m_j) = 0 \text{ for all } j = 0, \ldots, (\delta+1)^\nu$$

$$\iff c = \sum_{i=1}^{\nu} z_i B^{(\delta+1)^i} \text{ for some } z_1, \ldots, z_k \in [0, b)$$

If $N \in 2\uparrow$ and $S, T \in [0, N)$, then

$$\tau_2(S, T) = 0 \iff N^2 \mid \binom{2R}{R}$$

where $R = (N-1)((S+T+1)N + T + 1)$.

# The 9 Unknowns Theorem

The above ideas, together with some other works in the 1975 paper of Matiyasevich and Robinson, led Mtijasevich obtain the following celebrated theorem.

**Matiyasevich's 9 Unknowns Theorem**: $\exists^9$ over $\mathbb{N}$ is undecidable!

The detailed proof of this theorem appeared in Jones [J. Symbolic Logic, 1982].

Up to now, no one has shown that $\exists^\nu$ over $\mathbb{N}$ is undecidable for some $\nu < 9$, although A.Baker, Matiyasevich and Robinson all believed that $\exists^3$ over $\mathbb{N}$ might be undecidable.

# Putnam's trick

**H. Putnam** [J. Symbolic Logic 25(1960)]: Let $A \subseteq \mathbb{N}$. Suppose that for any $a \in \mathbb{N}$ we have

$$a \in A \iff \exists x_1 \geqslant 0 \ldots \exists x_n \geqslant 0[P(a, x_1, \ldots, x_n) = 0].$$

Then

$$A = \{\tilde{P}(x_0, x_1, \ldots, x_n) : \ x_0, \ldots, x_n \in \mathbb{N}\} \cap \mathbb{N},$$

where $\tilde{P}$ is a suitable polynomial with integer coefficients.

*Proof.* Define

$$\tilde{P}(x_0, x_1, \ldots, x_n) = (x_0 + 1)(1 - P(x_0, x_1, \ldots, x_n)^2) - 1.$$

If $x_0, \ldots, x_n \in \mathbb{N}$, then

$$\tilde{P}(x_0, x_1, \ldots, x_n) \geqslant 0 \iff P(x_0, \ldots, x_n) = 0 \Rightarrow \tilde{P}(x_0, \ldots, x_n) = x_0 \in A.$$

Thus

$$\begin{aligned} A =& \{x_0 \in \mathbb{N} : \ \exists x_1 \geqslant 0 \ldots, \exists x_n \geqslant 0[P(x_0, \ldots, x_n) = 0]\} \\ =& \{\tilde{P}(x_0, x_1, \ldots, x_n) : \ x_0, \ldots, x_n \in \mathbb{N}\} \cap \mathbb{N}. \end{aligned}$$

# The set of all primes

By Wilson's theorem, an integer $p > 1$ is prime if and only if $(p-1)! \equiv -1 \pmod{p}$. In view of this, the set of all primes is Diophantine, and Matiyasevich obtained the following surprising result.

**Matiyasevich (1975)**: There is a polynomial $P(x_0, \ldots, x_9)$ with integer coefficients such that

$$\{P(x_0, x_1, \ldots, x_9) : \ x_0, \ldots, x_9 \in \mathbb{N}\} \cap \mathbb{N}$$

coincides the set of all primes.

**Remark**. This looks incredible to number theorists!

There is no non-constant polynomial $P(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$ such that $P(x_1, \ldots, x_n)$ with $x_1, \ldots, x_n \in \mathbb{N}$ are all primes. For, if $P(x_1, \ldots, x_n)$ is a prime $p$, then

$$P(x_1 + py_1, \ldots, x_n + py_n) \equiv 0 \pmod{p}$$

for all $y_1, \ldots, y_n \in \mathbb{N}$.

# $\exists$ over $\mathbb{Z}$ is decidable

**Matiyasevich and Robinson** [Acta Arith. 27(1975)]: If $a_0, a_1, \ldots, a_n$ and $z$ are integers with $a_0 z \neq 0$ and $\sum_{i=0}^{n} a_i z^{n-i} = 0$, then

$$|z|^n \leqslant |a_0 z^n| \leqslant \sum_{i=1}^{n} |a_i||z|^{n-i} \leqslant \sum_{i=1}^{n} |a_i||z|^{n-1}$$

and hence

$$|z| \leqslant \sum_{i=1}^{n} |a_i|.$$

Thus $\exists$ over $\mathbb{N}$ and $\exists$ over $\mathbb{Z}$ are decidable (in polynomial time).

It is not known whether $\exists^2$ over $\mathbb{Z}$ is decidable. But A. Baker proved in 1968 that if $P(x, y) \in \mathbb{Z}[x, y]$ is homogenous, irreducible and of degree at least three then for any $m \in \mathbb{Z}$ there is an effective algorithm to determine whether $P(x, y) = m$ for some $x, y \in \mathbb{Z}$.

# Relative results

For any $m \in \mathbb{Z}$, by Lagrange's four-square theorem

$$m \geqslant 0 \iff \exists z_1 \exists z_2 \exists z_3 \exists z_4 [m = z_1^2 + z_2^2 + z_3^2 + z_4^2].$$

Thus

$$\exists^n \text{ over } \mathbb{N} \text{ is undecidable} \Rightarrow \exists^{4n} \text{ over } \mathbb{Z} \text{ is undecidable.}$$

By the Gauss-Legendre theorem on sums of three squares,

$$\mathbb{N} \setminus \{x^2 + y^2 + z^2 : \ x, y, z \in \mathbb{Z}\} = \{4^k(8l + 7) : \ k, l \in \mathbb{N}\}.$$

If $n \in \mathbb{N}$, then $4n + 1 = (2x)^2 + (2y)^2 + (2z + 1)^2$ for some $x, y, z \in \mathbb{Z}$, and hence $n = x^2 + y^2 + z^2 + z$. Thus, for any $m \in \mathbb{Z}$,

$$m \geqslant 0 \iff \exists x \exists y \exists z [m = x^2 + y^2 + z^2 + z].$$

It follows that

$$\exists^n \text{ over } \mathbb{N} \text{ is undecidable} \Rightarrow \exists^{3n} \text{ over } \mathbb{Z} \text{ is undecidable.}$$

Thus $\exists^{27}$ over $\mathbb{Z}$ is undecidable by the 9 unknowns theorem, as pointed out by S.P. Tung in [Japan J. Math., 11(1985)].

# A new relation-combining theorem

Tung (1985) asked whether $\exists^\nu$ over $\mathbb{Z}$ is undecidable for some $\nu < 27$.

**New Relation-Combining Theorem** (Z.-W. Sun [Z. Math. Logik Grundlag. Math. 38(1992)]): Let $A_1, \ldots, A_k, B, C_1, \ldots, C_n, D, E$ be integers with $D \neq 0$. Then

$$\begin{aligned} & A_1, \ldots, A_k \in \square \wedge B \neq 0 \wedge C_1, \ldots, C_n \geqslant 0 \wedge D \mid E \\ \Longleftrightarrow\ & \exists z_1 \ldots \exists z_{n+2} [P(A_1, \ldots, A_k, B, C_1, \ldots, C_n, D, E, z_1, \ldots, z_{n+2}) = 0], \end{aligned}$$

where $P$ is a suitable polynomial with integer coefficients.

This implies that

$$\exists^n \text{ over } \mathbb{N} \text{ is undecidable} \Rightarrow \exists^{2n+2} \text{ over } \mathbb{Z} \text{ is undecidable.}$$

So $\exists^{20}$ over $\mathbb{Z}$ is undecidable by the 9 unknowns theorem.

# A lemma

**An Observation of R. M. Robinson**: For any $m \in \mathbb{Z}$, we have

$$m \neq 0 \iff \exists x \exists y [m = (2x+1)(2y+1)].$$

Note that if $m \in \mathbb{Z} \setminus \{0\}$ then we can write
$m = \pm 3^a(3y+1) = (2x+1)(3y+1)$ with $x, y \in \mathbb{Z} \setminus \{0\}$.

If $d \in \mathbb{Z}^+$ is **not** a perfect square, then the Pell equation

$$y^2 - dx^2 = 1$$

has infinitely many integer solutions; in particular $dx^2 + 1 \in \square$ for some $x \in \mathbb{Z} \setminus \{0\}$.

In 1992, I made use of this fact from number theory to give the following lemma.

**Lemma**. Let $m \in \mathbb{Z}$. Then

$$m \geqslant 0 \iff \exists x \neq 0((4m+2)x^2 + 1 \in \square).$$

# Proof of the New Relation-Combning Theorem

Observe that

$$
\begin{aligned}
& B \neq 0 \wedge C_1, \ldots, C_n \geqslant 0 \\
\iff & C_1, \ldots, C_{n-1} \geqslant 0 \wedge B^2(C_n + 1) > 0 \\
\iff & \exists x_1 \ldots \exists x_{n-1}[(4C_1 + 2)x_1^2 + 1, \ldots, 4C_{n-1} + 2)x_{n-1}^2 + 1 \in \square \\
& \wedge B^2(C_n + 1)x_1^2 \ldots x_{n-1}^2 > 0] \\
\iff & \exists x_1 \ldots \exists x_{n-1}[(4C_1 + 2)x_1^2 + 1, \ldots, 4C_{n-1} + 2)x_{n-1}^2 + 1 \in \square \\
& \wedge B^2(C_n + 1)x_1^2 \ldots x_{n-1}^2 > 0].
\end{aligned}
$$

Thus, by applying Matiyasevich-Robinson's Relation-Combining Theorem, we get

$$
\begin{aligned}
& A_1, \ldots, A_k \in \square \wedge B \neq 0 \wedge C_1, \ldots, C_n \geqslant 0 \wedge D \mid E \\
\iff & \exists x_1 \ldots \exists x_{n-1}[A_1, \ldots, A_k, (4C_1 + 2)x_1^2 + 1, \ldots, 4C_{n-1} + 2)x_{n-1}^2 + 1 \in \square \\
& \wedge B^2(C_n + 1)x_1^2 \ldots x_{n-1}^2 > 0 \wedge D \mid E] \\
\iff & \exists x_1 \ldots \exists x_{n-1} \exists x \exists y \exists z[M_{k+n-1}(A_1, \ldots, A_k, (4C_1 + 2)x_1^2 + 1, \ldots, \\
& 4C_{n-1} + 2)x_{n-1}^2 + 1, D, E, B^2(C_n + 1)x_1^2 \ldots x_{n-1}^2, x^2 + y^2 + z^2 + z) = 0].
\end{aligned}
$$

# My Main Result on HTP

I actually obtained the following result in 1990 (thirty years ago).

**Main Theorem** (Z.-W. Sun, arXiv:1704.03504). (i) There is no algorithm to determine for any $P(z_1,\ldots,z_9)\in\mathbb{Z}[z_1,\ldots,z_9]$ whether the equation

$$P(z_0,\ldots,z_9)=0$$

has integral solutions with $z_9\geqslant 0$ (or $z_1+\ldots+z_9\geqslant 0$).

(ii) There is no algorithm to determine for any $Q(z_1,\ldots,z_{10})\in\mathbb{Z}[z_1,\ldots,z_9]$ whether the equation

$$Q(z_0,\ldots,z_{10})=0$$

has integral solutions with $z_{10}\neq 0$ (or $z_1+\ldots+z_{10}\neq 0$).

*Remark*. Let $z_9'=z_9-z_1-\ldots-z_8$. Then

$$\begin{aligned}&P(z_1,\ldots,z_8,z_9')=0 \text{ with } z_1+\ldots+z_8+z_9'\geqslant 0\\ \iff &P(z_1,\ldots,z_8,z_9)=0 \text{ with } z_9\geqslant 0.\end{aligned}$$

# $\exists^{11}$ over $\mathbb{Z}$ is undecidable

Recall that

$$m \geqslant 0 \iff \exists x \exists y \exists z [m = x^2 + y^2 + z^2 + z].$$

So,

$$\begin{aligned} &\exists z_1 \ldots \exists z_8 \exists z_9 \geqslant 0 [P(z_1, \ldots, z_8, z_9) = 0] \\ \iff &\exists z_1 \ldots \exists z_{11} [P(z_1, \ldots, z_8, z_9^2 + z_{10}^2 + z_{11}^2 + z_{11}) = 0]. \end{aligned}$$

Similarly, in view of R.M. Robinson's observation

$$m \neq 0 \iff \exists x \exists y [m = (2x + 1)(2y + 1)],$$

we have

$$\begin{aligned} &\exists z_1 \ldots \exists z_9 \exists z_{10} \neq 0 [Q(z_1, \ldots, z_9, z_{10}) = 0] \\ \iff &\exists z_1 \ldots \exists z_{11} [Q(z_1, \ldots, z_9, (2z_{10} + 1)(3z_{11} + 1)) = 0]. \end{aligned}$$

Therefore, both parts of the Main Theorem implies the undecidability of $\exists^{11}$ over $\mathbb{Z}$.

# $\exists^{17}$ over $\square$ is undecidable

**Theorem** (Z.-W. Sun, arXiv:1704.03504). There is no algorithm to decide for any polynomial $P(x_1,\ldots,x_{17})\in\mathbb{Z}[x_1,\ldots,x_{17}]$ the equation

$$P(x_1^2,\ldots,x_{17}^2)=0$$

has integer solutions.

**Lemma**. Any integer can be written as $2^\delta(x^2-y^2)$ with $\delta\in\{0,1\}, x, y\in\mathbb{Z}$.

Note that $x=(\frac{x+1}{2})^2-(\frac{x-1}{2})^2$.

**Conjecture** (Z.-W. Sun, arXiv:1704.03504). There is no algorithm to decide for any polynomial $P(x,y,z)\in\mathbb{Z}[x,y,z]$ the equation

$$P(x^2,y^2,z^2)=0$$

has integer solutions.

# References

For main sources of my work mentioned here, you may look at:

1. Z.-W. Sun, *Reduction of unknowns in Diophantine representations*, Sci. China Math. 35(1992), 257–269.

2. Z.-W. Sun, *A new relation-combining theorem and its application*, Z. Math. Logik Grundlag. Math. 38(1992), 209-212.

3. Z.-W. Sun, *Further results on Hilbert's Tenth Problem*, Sci. China Math., in press. See also `arXiv:1704.03504`.

# Thank you!