EPOCH AI

FrontierMath Open Problems

Epoch AI

Solve the finiteness problem for all equations of size $H \leq 24$

This story began in 2018, when Zidane asked on MathOverflow[1] what the smallest unsolved Diophantine equation is.

To state the problem precisely, we recall several definitions. A *monomial* in variables $x_1, \dots, x_n$ with an integer coefficient is an expression of the form $M = ax_1^{k_1} \cdots x_n^{k_n}$, where a is a nonzero integer and the exponents $k_1, \dots, k_n$ are nonnegative integers. The *degree* of this monomial is the sum $d = k_1 + \cdots + k_n$. Two monomials $M = ax_1^{k_1} \cdots x_n^{k_n}$ and $M' = a'x_1^{k_1'} \cdots x_n^{k_n'}$ are said to be *similar* if they have identical exponent patterns, that is, if $k_i = k_i'$ for every i.

A *polynomial* with integer coefficients in the variables $x_1, \dots, x_n$ is a finite sum of such monomials. A *polynomial Diophantine equation* is an equation of the form

$$P(x_1, \dots, x_n) = 0, \tag{1}$$

where P is a non-constant polynomial with integer coefficients. Without loss of generality, we may assume that P is written in *reduced form*, meaning that no two of its monomials are similar.

Suppose that P contains k distinct (pairwise non-similar) monomials with integer coefficients $a_1, \dots, a_k$ and corresponding degrees $d_1, \dots, d_k$. Zidane defined the *size* of (1) by

$$H(P) = \sum_{i=1}^{k} |a_i| \, 2^{d_i}. \tag{2}$$

For example, for the equation

$$z^2 + y^2 z + x^3 - 2 = 0$$

the size is

$$H = 2^2 + 2^3 + 2^3 + 2 = 22.$$

It follows directly from the definition that, up to relabelling of variables, there are only finitely many polynomial Diophantine equations of any fixed size. Hence we can order all polynomial Diophantine equations by size. Zidane asked: what is the first unsolved Diophantine equation in this order?

We decided to investigate this question by solving the equations in increasing order of size, proceeding as far as possible. The project turned out to be surprisingly rich: it begins with completely trivial equations, then quickly moves to olympiad-level problems, then to research-level mathematics, and finally to genuinely open questions.

It is straightforward to write a computer program that outputs all equations of small size. The program identifies equations related by renaming or permuting variables, multiplication by a nonzero constant, or substitutions of the form $x_i \mapsto -x_i$. The smallest equations it outputs are $x = 0$ and $x + 1 = 0$, of sizes $H = 2$ and $H = 3$, respectively. These are trivial one-variable equations. Since integer solutions to *all* one-variable equations are easy to determine, we modify the program to exclude them. The next smallest equations are then $x + y = 0$ and $xy = 0$, both of size $H = 4$. Again, both are trivial: the first is linear, and the second has P in (1) reducible over $\mathbb{Q}$. All linear equations can be excluded, and reducible equations can be reduced to irreducible ones, so they can be excluded as well. After these exclusions, the smallest remaining equation is $xy + 1 = 0$, of size $H = 5$.

We continue in this manner: solve the smallest equation not eliminated so far, identify a general solvable family to which it belongs, exclude that family, then move to the next smallest remaining equation, and so on.

[1] https://mathoverflow.net/questions/316708

All equations of size $H \leq 8$ turn out to be very easy: either we can list all integer solutions directly, or we can provide a polynomial parametrization of all solutions. For instance, all integer solutions to

$$x^2 - yz = 0$$

can be parametrized as

$$(x, y, z) = (uvw,\, uv^2,\, uw^2)$$

for integers u, v, w. If we insist on polynomial parametrizations, however, the problem suddenly jumps to research level at $H = 9$. Indeed, the question of whether the set of integer solutions to

$$xy - zt = 1 \tag{3}$$

admits a polynomial parametrization was posed by Skolem in the 1930s, remained open for over 70 years, and was resolved by Vaserstein [4] in 2010. He proved that every integer solution to (3) can be written as

$$(x, y, z, t) = \big(X(u),\, Y(u),\, Z(u),\, T(u)\big), \qquad u = (u_1, \ldots, u_{46}),$$

where X, Y, Z, T are polynomials in 46 variables with integer coefficients such that $XY - ZT$ is identically 1. Using this deep theorem, we proved in [2] that the solution sets to all equations of size $H \leq 12$ can be described via (one or many) polynomial parametrizations.

The bound $H \leq 12$ is best possible in this sense. At size $H = 13$, there are equations such as $yz = x^3 + 1$ for which the existence of a polynomial parametrization is open, as well as equations such as

$$x^2 - 2y^2 = \pm 1, \tag{4}$$

for which polynomial parametrizations of all solutions are known not to exist. The equations (4) are the classical Pell equations: all solutions can be described by recurrence relations, so we must accept such descriptions of the integer solutions sets. As we move to larger H, we encounter equations whose solution sets are accessible only through methods such as Vieta jumping, addition of rational points on elliptic curves, or more advanced techniques that we do not discuss here. Starting around $H \geq 17$, for equations such as $y^2 + z^2 = x^3 \pm 1$, it becomes unclear what it should even mean to "solve" an equation, and what kinds of descriptions of infinite solution sets we are willing to accept.

This motivates a shift to the following relaxed, but completely rigorous, problem, which we call the *finiteness problem*.

Problem 1. *Given a Diophantine equation, determine whether its set of integer solutions is finite or infinite. If it is finite, find all integer solutions.*

Of course, Problem 1 is much easier than describing *all* integer solutions: if we can produce just one infinite family of solutions, we are done. For example, if there exist polynomials $Q_1(t), \ldots, Q_n(t)$ with integer coefficients, not all constant, such that

$$P\big(Q_1(t), \ldots, Q_n(t)\big) \equiv 0, \tag{5}$$

then (1) has infinitely many integer solutions, and Problem 1 is resolved for this equation. Deciding whether such polynomials exist is undecidable in general. Nevertheless, we can search for identities of the form (5) with small degrees and coefficients, and exclude equations for which we find suitable Q_i.

After applying this exclusion (together with the earlier exclusions of easily solvable equations), the smallest surviving equation is

$$y(x^2 - y) = z^2 + 1,$$

of size $H = 17$. This equation has no integer solutions: quadratic reciprocity implies that all positive divisors of $z^2 + 1$ are congruent to 1 or 2 modulo 4, while a simple analysis modulo 4 shows that this is impossible for the left-hand side. After formulating this obstruction in a general way and excluding equations solvable by it, the next smallest equation is

$$y^2 + z^2 = x^3 - 2, \tag{6}$$

of size $H = 18$. Here the substitution $x = t^2 + 2$ and $y = t^3 + 3t$ reduces (6) to $z^2 = 3t^2 + 6$, a two-variable quadratic equation for which it is easy to produce infinitely many integer solutions. More generally, if there exist polynomials $Q_1(u, v), \dots, Q_n(u, v)$ with integer coefficients, not all constant, such that the two-variable equation

$$P\big(Q_1(u, v), \dots, Q_n(u, v)\big) = 0 \tag{7}$$

has infinitely many integer solutions (u, v), then (1) has infinitely many integer solutions as well. As before, finding such substitutions $x_i = Q_i(u, v)$ is difficult in general, but can sometimes be done in specific cases.

We then continue this strategy: resolve Problem 1 for the currently smallest remaining equations, understand which further equations fall to the same method and exclude them, and proceed. When we could not solve an equation, we asked for help on MathOverflow. The methods evolve from completely elementary to quite sophisticated. For example, resolving Problem 1 for

$$xyz = x^3 + y^2 + 2 \tag{8}$$

(of size $H = 22$) required a nontrivial idea introduced by Mordell and further developed by Schinzel [3]; see `https://mathoverflow.net/questions/411958/`.

The joint efforts of MathOverflow users have resolved Problem 1 for all equations of size $H \le 21$. In fact, there is only one remaining equation of size $H = 22$, and only nine remaining equations of size $H \le 24$. These are listed in Table 1.

H	Equation	H	Equation	H	Equation
22	$z^2 + y^2z + x^3 - 2 = 0$	23	$z^2 + y^2z + x^3 + x + 1 = 0$	24	$z^2 + y^2z + x^3 - x - 2 = 0$
23	$z^2 + y^2z + x^3 - x - 1 = 0$	23	$z^2 + y^2z + x^3 - 3 = 0$	24	$z^2 + y^2z + x^3 - x + 2 = 0$
23	$z^2 + y^2z + x^3 + x - 1 = 0$	23	$z^2 + y^2z + x^3 + 3 = 0$	24	$z^2 + y^2z - z + x^3 + 2 = 0$

Table 1: Equations of size $H \le 24$ for which Problem 1 remains open.

All these equations have the shape

$$z^2 + (y^2 + a)z + P(x) = 0 \qquad \text{for some integer } a \text{ and some cubic polynomial } P(x), \tag{9}$$

and the problem reduces to determining whether the discriminant

$$(y^2 + a)^2 - 4P(x)$$

is a perfect square for infinitely many integers (x, y). Numerical experiments with many small solutions, together with heuristic considerations, give us $> 99.9\%$ confidence that the answer is "yes" for all nine equations, but a proof remains elusive.

This is the concrete problem we propose:

Prove that every equation in Table 1 has infinitely many integer solutions.

This would resolve Problem 1 for all equations of size $H \le 24$. More importantly, we are about 99% confident that solving this problem requires the introduction of a new and powerful method for proving that a Diophantine equation has infinitely many integer solutions. It is certainly possible that some of the equations in Table 1 can be handled by existing techniques. For example, one could attempt a deeper search for substitutions $x = Q_1(t)$, $y = Q_2(t)$, $z = Q_3(t)$ satisfying (5), or substitutions $x = Q_1(u, v)$, $y = Q_2(u, v)$, $z = Q_3(u, v)$ such that (7) has infinitely many integer solutions in (u, v). However, our fairly extensive searches have found no such parametrizations, and it seems highly unlikely that *all nine* equations in Table 1 can be resolved in this manner. We therefore expect that a genuinely new approach is needed—one that would likely apply to all equations of the form (9), and possibly to many others beyond this family. In particular, we hope that such a new method would also resolve at least some of the hundreds of other open equations arising in this project, as listed in [1].

References

[1] Bogdan Grechuk. "A systematic approach to Diophantine equations: open problems". In: *arXiv preprint arXiv:2404.08518* (2024).

[2] Bogdan Grechuk. *Polynomial Diophantine equations—a systematic approach*. With contributions by Ashleigh Wilcox. Springer, Cham, [2024] ©2024, pp. xxii+807. ISBN: 978-3-031-62948-8; 978-3-031-62949-5. DOI: 10.1007/978-3-031-62949-5. URL: https://doi.org/10.1007/978-3-031-62949-5.

[3] A. Schinzel. "On the congruence $f(x)+g(y)+c \equiv 0 \pmod{xy}$ (completion of Mordell's proof)". In: *Acta Arith.* 167.4 (2015), pp. 347–374. ISSN: 0065-1036,1730-6264. DOI: 10.4064/aa167-4-3. URL: https://doi.org/10.4064/aa167-4-3.

[4] Leonid Vaserstein. "Polynomial parametrization for the solutions of Diophantine equations and arithmetic groups". In: *Ann. of Math. (2)* 171.2 (2010), pp. 979–1009. ISSN: 0003-486X,1939-8980. DOI: 10.4007/annals.2010.171.979. URL: https://doi.org/10.4007/annals.2010.171.979.