Open Problems Column
Edited by William Gasarch[1]

This Issues Column! This issue's Open Problem Column is by Bogdan Grechuk is titled *On the smallest open Diophantine equations*. It searches for the smallest Diophantine equations that are non-trivial to solve.

Request for Columns! I invite any reader who has knowledge of some area to contact me and arrange to write a column about open problems in that area. That area can be (1) broad or narrow or anywhere in between, and (2) really important or really unimportant or anywhere inbetween.

On the smallest open Diophantine equations
by Bogdan Grechuk[2]

Abstract

This paper reports on the current status of the project in which we order all polynomial Diophantine equations by an appropriate version of "size", and then solve the equations in that order. We list the "smallest" equations that are currently open, both unrestricted and in various families, like the smallest open symmetric, 2-variable or 3-monomial equations. All the equations we discuss are amazingly simple to write down but some of them seem to be very difficult to solve.

Key words: Diophantine equations, Hilbert's tenth problem, Hasse principle, quadratic reciprocity, sum of squares.

1 Introduction.

In many areas of mathematics researchers order instances of difficult problems by some natural parameters and try to solve the problems at least for "small" instances, while in other areas such systematic approach is currently missing. To illustrate this point, let us look at some recent breakthroughs in knot theory and in number theory.

One of the most amazing recent results in knot theory is the 2020 theorem of Lisa Piccirillo, stating that the Conway knot is not slice [23]. Here, a knot is called (smoothly) *slice* if it bounds a smoothly embedded 2-dimensional disk in 4-dimensional space, and the Conway knot is a certain knot labeled 11n34 in the standard knot tables. Here, "11" refers to the fact the the diagram of this knot has 11 crossings. The question whether the Conway knot is slice or not has been opened for about 50 years, and has attracted a lot of attention because *all other knots of under* 13 *crossings were classified according to whether they are slice knots or not, and the Conway knot has been the only exception.*

One of the most amazing results in number theory is the Andrew Wiles' proof [32] of Fermat Last Theorem. The Theorem states that exponential Diophantine equation $x^n + y^n = z^n$ has no non-trivial integer solutions for $n \geq 3$, has been an open question for many centuries, and has attracted a lot of attention because... Fermat claimed that he has a beautiful proof of this fact which he never wrote down, and many generations of mathematicians were trying to find the missing proof.

A more recent example of an amazing result in the area of Diophantine equations is the 2021 paper of Booker and Sutherland [5], which studies equations in the form

$$x^3 + y^3 + z^3 = k. \tag{1}$$

Note that (1) has no solution if $k \equiv \pm 4 \pmod 9$. It is known that there are infinitely many integer solutions for $k = 1$ and 2. For $k = 3$, (1) has the solutions $(1,1,1)$ and permutations of $(4,4,-5)$. Miller and Woolett [20] asked if there is always a solution when $k \not\equiv \pm 4 \pmod 9$. Mordell [21] asked if there are other solutions when $k = 3$. Brooker and Sutherland showed that

[1] Universit of Maryland, Dept of CS, gasarch@umd.edu

[2] School of Computing and Mathematical Sciences, University of Leicester, LE1 7RH, UK; bg83@leicester.ac.uk

(i) $x^3+y^3+z^3=42$ has a solution (combined with prior work this shows that for $|k|\leqslant 100$, $k\not\equiv\pm 4 \pmod 9$, there is always a solution), and

(ii) there is another solution to $x^3+y^3+z^3=3$, namely: $x=569,936,821,221,962,380,720$, $y=-569,936,821,113,563,493,509$, $z=-472,715,493,453,327,032$.

As you can see, in knot theory researchers ordered all knots by a natural parameter, number of crossings, and then try to answer questions of interest *systematically* for all knots with the given number of crossings. A similar approach is taken in many other areas of mathematics: if a problem is difficult or undecidable in general, researchers order the instances of the problem in some natural way, and then try to "solve" at least "small" instances. For example, the Halting problem has been solved for all 2-symbol Turing machines with 2, 3 and 4 states, so the next open case is 5-state machines. In contrast, Diophantine equations are not studied in order, and often attract interest just because some time ago a famous mathematician wrote down an equation and asked to solve it.

One reason for this is that it is not obvious how to order all Diophantine equations is a natural way. It is easy to arrange in order some specific infinite families of equations. For example, Fermat Last Theorem can be treated as one exponential equation or as an infinite family of polynomial Diophantine equations ordered by the value of parameter n. Similarly, equations (1) can be naturally ordered by parameter k. However, it is less easy to order all polynomial Diophantine equations, that is, all equations of the form

$$P(x_1,\dots,x_n)=0, \tag{2}$$

where P is a polynomial with integer coefficients. For this, we need to assign to every Diophantine equation a "size" parameter, such that for any bound B there is only a finite number of equations of size at most B. Many natural notions of "size" or "height" do not satisfy this condition. For example, if we define the height of the equation as the maximum (or sum) of absolute values of its coefficients, then there are infinitely many equations of height 1, e.g. $x^n=0$, $n=1,2,\dots$.

It is natural to define the "size" of equation (2) as a sum of sizes of monomials of P, so it is left to define the size of a monomial $ax_1^{k_1}\dots x_n^{k_n}$. If we consider an equation as an input to a computer program which then tries to solve it, then standard way to measure the size of the input is the number of bit needed to describe it. For simplicity, let us assume that we do not use the power symbol, and write $x_1^{k_1}$ as $x_1x_1\dots x_1$ (k_1 times) and so on. Then we need $d=k_1+\dots+k_n$ symbols to write $x_1^{k_1}\dots x_n^{k_n}$, where d is the degree of the monomial. We also need about $\log_2|a|$ symbols to write the coefficient a in binary, so we can define the length of the monomial as $l=\log_2|a|+d$. This is not an integer, but ordering the monomials by l is equivalent to ordering them by $H=2^l=|a|2^d$, which is an integer. Hence, let us define the size of equation (2) as

$$H(P)=\sum_{i=1}^{k}|a_i|2^{d_i}, \tag{3}$$

provided that polynomial P consists of k monomials with integer coefficients $a_1,\dots,a_k$ and degrees $d_1,\dots,d_k$, respectively. For example, for the equation (1) with $k=3$ we have $H=2^3+2^3+2^3+3=27$. This notion of size has been suggested by anonymous mathoverflow user Zidane who asked in 2018 what is the smallest open Diophantine equation, see [33]. Note that H is always a non-negative integer, and, for any B, there is a finite number of equations with $H\leq B$. Hence, we may list all equations with sizes $H=0,1,2,\dots$, try to solve them in this order, and report the smallest equations we cannot solve. This is exactly the purpose of this paper.

In [15], this project is implemented for the Hilbert 10th problem, that is, the problem of determining whether an equation has any integer solution. In this paper, we consider more general problem of determining all integer solutions, but also re-iterate the open questions posted in [15]. Section 2 considers the most general problem of solving the equations completely, including the description of the solution set even if it is infinite. Section 3 considers this problem for the 2-variable equations only. In Section 4, we consider an easier problem of determining whether the solution set is finite, and if so, list all the solutions. In Section 5, we study an even easier problem of determining whether a given equation has any integer solution or not. For each of these

problems, we have identified the smallest equations for which the problem is non-trivial, and challenge the readers to try to solve these equations.

2 Describing all solutions: polynomial families

We have exactly one equation $0 = 0$ of size $H = 0$, and exactly two equations $\pm 1 = 0$ of size $H = 1$. More generally, for every $H > 0$, we have two equations $\pm H = 0$ with no variables and no solutions. If we do not count these as "equations" and insist that every equation should have at least one variable, then the smallest equations are $\pm x = 0$ of size $H = 2$ and the nest smallest are $\pm x \pm 1 = 0$ of size $H = 3$. All these equations are trivial to solve. More generally, for any equation in one variable

$$a_m x^m + \cdots + a_1 x + a_0 = 0 \tag{4}$$

with integer coefficients $a_m, \ldots, a_0$, any non-zero integer solution must be a divisor of a_k, where k is the smallest integer such that $a_k \neq 0$. This gives an algorithm to list all integer solutions of (4). Moreover, this can be done in polynomial time, see [8].

Further, we will consider only equations in at least two variables. The smallest such equations are $\pm x \pm y = 0$ and $\pm xy = 0$ of size $H = 4$. To avoid solving essentially the same equations multiple times, we call two equations equivalent if one can be transformed into another after multiplication by -1 and/or substitutions in the form $x_i \to -x_i$. Obviously, it suffices to consider only one equation from each equivalence class. With this convention, the only equations of size $H = 4$ we need to consider are

$$x + y = 0 \tag{5}$$

and

$$xy = 0. \tag{6}$$

Both these equations have infinitely many integer solutions, which can be presented in a parametric form. For equation (5), the solutions are $(x, y) = (u, -u)$, where u is an integer parameter. This is a special case of polynomial family, as defined below.

Definition 2.1. *We say that a subset $S \subset \mathbb{Z}^n$ is a polynomial family if there exist polynomials $P_1, \ldots, P_n$ is k variables $u_1, \ldots, u_k$ such that $(x_1, \ldots, x_n) \in S$ if and only if there exists integers $u_1, \ldots, u_k$ such that $x_i = P_i(u_1, \ldots, u_k)$, $i = 1, \ldots, n$.*

In this terminology, the solution set of the equation (5) is a polynomial family with $k = 1$ parameter u, $P_1(u) = u$ and $P_2(u) = -u$. More generally, the solution set of any equation in the form

$$x_n + Q(x_1, \ldots, x_{n-1}) = 0 \tag{7}$$

can be represented as polynomial family $x_i = u_i$, $i = 1, \ldots, n-1$, $x_n = -Q(u_1, \ldots, u_{n-1})$.

The solution set to the equation (6) is $(x, y) = (0, u)$ or $(x, y) = (u, 0)$ for any integer u. Note that it is not a polynomial family but a union of two polynomial families. Obviously, if we can represent the solution set as a union of any finite number of polynomial families, then we can classify the equation as "solved". Note that if the solutions sets of equations $P_1 = 0$ and $P_2 = 0$ are finite unions of polynomial families, then the same is true for the equation

$$P_1 \cdot P_2 = 0. \tag{8}$$

From now on, we will exclude the equations of the forms (7) and (8) from further analysis.

For $H = 5$, we would like to mention equation

$$2x + 1 = 0. \tag{9}$$

It is a one-variable equation (4), and we have already discussed and excluded all such equations. However, it is interesting as the smallest equation for which the left-hand side is always odd and therefore cannot be equal

to 0. More generally, if there exists an integer $m \geq 2$ such that $P(x_1, \ldots, x_n)$ is never divisible by m, then equation (2) has no integer solutions. Hence, we may exclude such equations.

The only non-excluded equation of size $H = 5$ is

$$xy + 1 = 0 \tag{10}$$

whose solutions are $(x, y) = (1, -1)$ and $(x, y) = (-1, 1)$. There are no other integer solutions because every potential solution must be a divisor of 1. More generally, any equation of the form

$$P_1 \cdot P_2 = c \tag{11}$$

for some integer c can be solved by enumerating divisors d of c, and, for every divisor, solve the system of equations $P_1 = d, P_2 = c/d$. From now on, we will exclude the equations of the form (11) as well. Families (4), (7) and (11) cover all the equations of size $H \leq 7$.

For $H = 8$, there are some equations not considered so far, for example, equation

$$x^2 + y^2 = 0, \tag{12}$$

whose only integer solution is $x = y = 0$ because this is the only real solution. More generally, we can exclude all equations (2) in n variables whose set of real solutions is a bounded region in $\mathbb{R}^n$, because any bounded region has at most a finite number of integer points, and we may use the direct substitution to check which of these points are the solutions to (2).

Another equation not excluded so far is

$$xy + 2z = 0. \tag{13}$$

Because xy is even, then either x or y must be even. In the first case, let us write $x = 2u$ for some integer u, while in the second case $y = 2v$ for some integer v. Hence, we get two families of integer solutions $(x, y, z) = (2u, v, -uv)$ and $(x, y, z) = (u, 2v, -uv)$ for some integers u, v. More generally, any equation of the form

$$ax_n + Q(x_1, \ldots, x_{n-1}) = 0, \tag{14}$$

where $a \neq 0$ is an integer, can be solved by enumerating all possible $|a|^{n-1}$ remainders $x_1, \ldots, x_{n-1}$ can give after division by a, and in each case representing $x_i = au_i + r_i$, $i = 1, \ldots, n-1$, with $0 \leq r_i < |a|$. Then each case when $Q(au_1 + r_1, \ldots, au_{n-1} + r_{n-1})$ is divisible by a lead to a polynomial family of the solutions of (14).

All equations we have considered so far are completely trivial. The first equation that deserves to be given to students as an exercise is the equation

$$x^2 - yz = 0. \tag{15}$$

The question is, of course, how to represent all solutions as a polynomial family. To solve this exercise, the students should note that any integers y and z can be represented as $y = uv^2$ and $z = u'w^2$ with u, u' square-free. Now, for yz to be a perfect square we must have $u = u'$, hence $z = uw^2$, from which we find $x = uvw$. Conversely, for any u, v, w (not necessarily square-free), the triple $(x, y, z) = (uvw, uv^2, uw^2)$ is a solution to (15).

A bit more difficult exercise is to write as a polynomial family the set of all solutions to the equation

$$xy - zt = 0. \tag{16}$$

The answer is $(x, y, z, t) = (u_5u_1u_2, u_5u_3u_4, u_5u_1u_3, u_5u_2u_4)$ for integers u_1, u_2, u_3, u_4, u_5. It is trivial to check that the given family satisfies (16). We leave it to the reader to show that, conversely, all solutions to (16) are covered by this parameterization. This finishes the analysis of all equations of size $H \leq 8$.

With $H = 9$, the problem suddenly jumps from student-exercise level to research level. The question whether the set of integer solutions of the equation

$$xy - zt = 1 \tag{17}$$

is a polynomial family has been first asked by Skolem [28] in the 1930's, remained open for over 70 years, and has been answered by Vaserstein [30] in 2010, who proved that it is indeed a polynomial family with 46 parameters. As a corollary of this result, Vaserstein also showed that, for any integer c, the solution set of the equation

$$x^2 - yz = c \tag{18}$$

is the union of a finite number of polynomial families. In particular, this covers the equations $x^2 - yz = \pm 1$, and finishes the analysis of all equations of size $H \leq 9$.

Vaserstein also proved that, for any c, the solution set of the equation

$$xy - zt = c \tag{19}$$

is the union of a finite number of polynomial families. His theorem is applicable to many other equations. For example, let x, y, z be a solution to the equation

$$x^2 + x - yz = 0. \tag{20}$$

Because $x^2 + x = x(x+1)$, the prime factors of y are distributed between x and $x+1$, hence y can be written as $y = ab$, where a and b are divisors of x and $x+1$, respectively. With $c = \frac{x}{a}$ and $d = \frac{x+1}{b}$, we get $db - ac = (x+1) - x = 1$, which is exactly (17), hence the set of such (a, b, c, d) is a polynomial family. Then $(x, y, z) = (ac, ab, cd)$ is also a polynomial family.

This finishes the analysis of all equations of size $H \leq 10$. With $H = 11$, we meet some easy equations with 2 variables that we will discuss in the next section, as well as the equations

$$x^2 + x \pm 1 - yz = 0. \tag{21}$$

We leave them to the reader as the first open question of this paper.

Open Question 2.2. *Are the solution sets to the equations (21) polynomial families? If not, are they unions of a finite number of polynomial families?*

3 Describing all solutions: equations in 2 variables

As we have seen in the previous section, describing the solution set of a 3-variable equation may be a quite non-trivial problem even with $H \leq 11$. In this section, we will restrict out attention to 2-variable equations, for which many powerful results and techniques are available. The smallest 2-variable equations not of the form (11) or (14) are the equations

$$y^2 = x^2 + x \pm 1$$

of size $H = 11$. If $|x| > 2$, then

$$(|x| - 1)^2 < x^2 + x \pm 1 < (|x| + 1)^2,$$

and also $x^2 + x \pm 1 \neq |x|^2$, hence $x^2 + x \pm 1$ cannot be a perfect square. By checking cases $|x| \leq 2$, we may easily list all the solutions. For $H = 12$, we meet similar equations $y^2 = x^2 + x \pm 2$ and $y^2 = x^2 + 2x$, that can be solved by exactly the same method. A little bit different is the equation

$$y^2 + y = x^2 + x,$$

which, after multiplication by 4 and adding 1, can be reduced to $(2y+1)^2 = (2y+1)^2$, from which we derive two families of solutions $(x, y) = (u, u)$ and $(x, y) = (u, -u-1)$. A notable equation is

$$y^2 = 2x^2 \tag{22}$$

whose only integer solution is $(x, y) = (0, 0)$. If it would have a non-zero integer solution (x, y), then rational number $t = \frac{y}{x}$ would satisfy $t^2 = 2$, hence the non-existence of non-zero integer solutions to (22) is equivalent to the irrationality of $\sqrt{2}$, a famous old question with rich history. Similarly, equations

$$x^2 + xy \pm y^2 = 0$$

reduce to non-existence of rational solutions to $t^2 + t \pm 1 = 0$. Finally, equation

$$2xy + x + y = 0$$

reduces to the question for which x the ratio $\frac{x}{2x+1}$ can be an integer. It is easy to see that this is possible only for $x = -1$ and $x = 0$, leading to the solutions $(x, y) = (-1, -1)$ and $(x, y) = (0, 0)$. The more general equation

$$Q(x)y + R(x) = 0 \tag{23}$$

reduces to the question when $\frac{R(x)}{Q(x)}$ is an integer, which can be easily answered in full generality.

For $H = 13$, we meet more interesting equations, such as

$$y^2 - 2x^2 = 1 \tag{24}$$

and

$$y^2 - 2x^2 = -1, \tag{25}$$

which are known as Pell equation and negative Pell equation, respectively. These equations are interesting because if a pair (x, y) of positive integers solves any of them, then ratio $\frac{y}{x}$ is an approximation to $\sqrt{2}$ with a good trade-off between the quality of the approximation and the size of the denominator. For this reason, these equations has been studied since ancient times. It is easy to check that both equations have infinitely many integer solutions. For example, equation (24) has a solution $(x_0, y_0) = (1, 0)$, and direct substitution shows that if (x_n, y_n) is a solution, then so is

$$x_{n+1} = 3x_n + 2y_n, \quad y_{n+1} = 4x_n + 3y_n.$$

This gives an infinite sequence of solutions defined by recurrence relations. A bit more difficult to show that this sequence gives *all* solutions with $x \geq 0$ and $y \geq 0$ (and then all integer solutions can be obtained by changing signs), but this is also well-known, see, for example, Theorem 3.2.1 in [2]. Similarly, all the non-negative solutions to (25) can be obtained starting from $(x_0, y_0) = (1, 1)$ and applying the same recurrence relations. As a side note, we remark that Pell equations (24) and (25) are the simplest examples of the equations whose solution sets are known to be *not* finite unions of polynomial families, see [30].

All the equations we have considered so far in this section are quadratic equations. In fact, there is a general algorithm that, given integers a, b, c, d, e, f as an input, solves the general 2-variable quadratic equation

$$ax^2 + bxy + cy^2 + dx + ey + f = 0. \tag{26}$$

The algorithm is implemented online at [1]. For an equation (26), it lists all the solutions if there are finitely many of them, and otherwise describes all solutions as a union of polynomial families or in the form of linear recurrence relations.

Based on this, we can exclude all quadratic equations (26) from further analysis. Together with the previously excluded families, this eliminates all the equations of size $H \leq 13$ with two exceptions:

$$y^2 = x^3 + 1 \tag{27}$$

and

$$y^2 = x^3 - 1. \tag{28}$$

Equations of the form $y^2 = x^3 + k$ are known as Mordell's equations and are well studied. It is known that there is a finite number of integer solutions for each k, and there is an algorithm that, given k, outputs all the

solutions. See [14] for the description of the algorithm and for the explicit list of all solutions in the range $|k| \le 10,000$.

More generally, there are known practical algorithms for finding integer solutions to the equations in the form

$$y^2 + axy + cy = x^3 + bx^2 + dx + e \tag{29}$$

under some minor conditions on the integer coefficients a, b, c, d, e that guarantee that the solution set is finite. One such algorithm is implemented in an open-source and free to use computer algebra system SageMath [34], that can be run online at `https://sagecell.sagemath.org/`. To solve (29), we run the command

$$sage: EllipticCurve([a,b,c,d,e]).integral\_points() \tag{30}$$

For example, command

$$sage: EllipticCurve([0,0,0,0,1]).integral\_points()$$

returns $[(-1:0:1),(0:1:1),(2:3:1)]$, which means that the only integer solutions to (27) are $(x,y) = (-1,0)$, $(0,1)$, and $(2,3)$. In a similar way, we find that the only integer solution to (28) is $(x,y) = (1,0)$.

This finishes the analysis of the equations of size $H \le 13$. Starting with $H \ge 14$, we will exclude all the equations in the form (29). After this, the only remaining equation of size $H = 14$ is

$$y^2 + yx^2 + x = 0. \tag{31}$$

This equation is not directly in the form (29), but can be easily reduced to it. Indeed, after multiplication by $4y$ and adding 1 to both sides, we can rewire the equation as $4y^3 + (4x^2y^2 + 4xy + 1) = 1$, or $(2xy+1)^2 = -4y^3 + 1$. With new variable $z = 2xy + 1$, this simplifies to $z^2 = -4y^3 + 1$. Now multiply both sides by 16 to get $(4z)^2 = (-8y)^3 + 16$, or

$$Y^2 = X^3 + 16 \tag{32}$$

with new variables $Y = 4z = 4(2xy+1)$ and $X = -8y$. Note that if x, y are integers then so is X, Y. Now, command

$$sage: EllipticCurve([0,0,0,0,16]).integral\_points()$$

shows that the only integer solutions to (32) are $(X,Y) = (0,\pm 4)$. For these solutions, $y = -X/8 = 0$ happen to be an integer, and substitution $y = 0$ in (31) returns $x = 0$. Hence, $(x,y) = (0,0)$ is the only integer solution to (31).

Computer algebra system Maple has a command Weierstrassform that helps to transform a broad range of equations to the form (29). In this example, command

$$Weierstrassform(x + x^2y + y^2, x, y, X, Y)$$

returns $[X^3 - 1/4 + Y^2, y, x*y + 1/2, (-1+2*Y)/(2*X), X]$, that shows that the equation can be transformed to $X^3 - 1/4 + Y^2$ after substitutions $X = y$, $Y = xy + 1/2$. The remaining steps can be easily done by hand. A combination of Weierstrassform and EllipticCurve commands allows to solve all 2-variable equations of size $H \le 15$, and many equations afterwords.

The algorithms we have discussed so far are the special cases of much more general algorithms applicable to much broader classes of 2-variable equations

$$P(x,y) = 0. \tag{33}$$

To introduce them, we need a few definitions. A polynomial P with integer coefficients is called absolutely irreducible if it cannot be written as a product $P = P_1 \cdot P_2$ of non-constant polynomials, even if we allow complex coefficients. It is known that if P is irreducible over $\mathbb{Q}$ but not absolutely irreducible, then all integer solutions to (33) can be determined easily, see e.g. [15]. On the other hand, if P is not irreducible over $\mathbb{Q}$, then equation (33) reduces to equations of the same form for each of the factors. Hence, we may assume that P in

(33) is absolutely irreducible. In this case, the set of all complex solutions to (33) forms a connected surface. The *genus* g of such surface is the maximum number of cuttings that can be made along non-intersecting closed simple curves on the surface without making it disconnected. The genus–degree formula

$$g \leq \frac{1}{2}(d-1)(d-2), \tag{34}$$

where d is the degree of P, implies that all quadratic polynomials have genus 0, while all cubic polynomials have genus at most 1. Poulakis [24, 25] developed practical algorithm to solve all 2-variable equations of genus $g = 0$. The algorithm can decide whether a given equation has finite or infinite number of solutions, list all solutions in the former case, and describes them in the parametric form and/or using recurrence relations in the latter case.

Hence, it suffices to consider equations with $g \geq 1$. In this case, there is always a finite number of integer solutions [27]. In 1970, Baker [4] developed an effective upper bound for the absolute value of all possible solutions as an explicit function of the coefficients of P, provided that $g = 1$. This gives an algorithm to list all the solutions of an arbitrary genus 1 equation. In particular, by the genus–degree formula (34), this result covers all 2-variable cubic equations. While Baker's bounds are enormous and the corresponding algorithm is impractical, a practical method for finding all integer solutions to genus 1 equations was later developed by Stroeker and Tzanakis [29].

Further, Baker [3] developed in 1969 a general method for solving equations in the form

$$y^2 = P(x), \tag{35}$$

where $P(x)$ is a polynomial of arbitrary degree that has at least three simple (possibly complex) zeros. As proved in [15], this implies the method to determine all integer solutions to the equation

$$a(x)y^2 + b(x)y + c(x) = 0. \tag{36}$$

where $a(x)$, $b(x)$ and $c(x)$ are arbitrary polynomials with integer coefficients. Indeed, if (36) has an integer solution, then $b^2(x) - 4a(x)c(x)$ must be a perfect square, and we can apply Baker's algorithm to determine all such x, see [15] for details.

This allows us to focus on the equations of degree at least 4 that are at least cubic in each of the variables. The simplest examples of such equations are, say, $y^3 = x^4 + 1$ or $y^3 = x^4 + x + 1$. However, such equations are covered by another theorem of Baker [3], who developed an algorithm for listing all the solutions of the equation

$$y^m = P(x), \tag{37}$$

provided that $m \geq 3$, and $P(x)$ is a polynomial with integer coefficients of degree at least 3 with at least two simple zeros. In 1984, Brindza [6] showed the the conditions on P can be significantly relaxed. It is easy to see [15] that this result also implies the algorithm for solving equation

$$ay^m = P(x). \tag{38}$$

Baker's and Brindza's methods for solving equations (35) and (37) are impractical even for the equations with small coefficients. However, there are practical methods for which we do not have proof that they work in general, but which seem to work for any individual equation in this form. For example, Bruin and Stoll [7] decided the solvability in rationals of all the equations (35) where P is a square-free, has degree at most 6, and has integral coefficients of absolute value at most 3. More recently, Hashimoto and Morrison [18] determined the set of all rational solutions for a large family of the equations in the form (37).

Based on this, we eliminate equations of the form (38) from further analysis. After this, the smallest non-eliminated ones are

$$x^3y + y^3 \pm x = 0 \tag{39}$$

of size $H = 26$ and the next-smallest are

$$x^3y + y^3 \pm x + 1 = 0 \tag{40}$$

of size $H = 27$. These equations can be easily solved directly (exercise!), but we instead note that all such equations are covered by another deep result. Let $\mathcal{F}$ be a family of polynomials

$$P(x,y) = \sum_{i=0}^{m} \sum_{j=0}^{n} a_{ij} x^i y^j$$

with integer coefficients a_{ij} of degree $m > 0$ in x and $n > 0$ in y which are irreducible over $\mathbb{Q}[x,y]$, and such that either

(C1) there exists a coefficient $a_{ij} \neq 0$ of P such that $ni + mj > mn$,

or

(C2) the sum of all monomials $a_{ij}x^iy^j$ of P for which $ni + mj = nm$ can be decomposed into a product of two non-constant relatively prime polynomials in $\mathbb{Z}[x,y]$.

In 1887, Runge [26] proved if $P \in \mathcal{F}$ then equation $P = 0$ has at most finitely many integer solutions. In 1992, Walsh [31] developed an effective upper bound for the size of possible solutions, which implies the existence of an algorithm for listing all the solutions. Note that equations (39) and (40) satisfy (C1), because in this case $n = m = 3$, there is a non-zero coefficient $a_{31} = 1$, and, for $i = 3$ and $j = 1$, we have $ni + mj = 3 \cdot 3 + 3 \cdot 1 > 3 \cdot 3 = mn$. Walsh's theorem allows us to exclude all equations that satisfy either (C1) or (C2) from further analysis.

This allows to exclude all equations of size $H \leq 27$, and all equations of size $H = 28$ with three exceptions:

$$x^4 + xy + y^3 = 0, \tag{41}$$

$$y^3 + y = x^4 + x \tag{42}$$

and

$$y^3 - y = x^4 + x. \tag{43}$$

The listed equations does not satisfy (C1) because for them we have $m = 4$, $n = 3$, and there is no non-zero coefficient a_{ij} with $3i + 4j > 12$. Equality $3i + 4j = 12$ holds for coefficients a_{40} and a_{03}, and polynomials $x^4 \pm y^3$ are irreducible, hence (C2) also fails.

Equation (41) has genus 2. Computer algebra system Magma has built-in method (called Chabauty) for finding all rational solutions for some genus 2 equations, and the method happens to work for this particular equation, returning that the only rational solution is $x = y = 0$. Equations (42) and (43) have genus 3, and this Magma function is not applicable to them. By Siegel1's Theorem [27], the set of their integer solutions is finite. The direct search returns solutions $(x,y) = (1,0)$, $(0,0)$, and $(1,1)$ for (42) and $(x,y) = (-1,-1)$, $(-1,0)$, $(-1,1)$, $(0,-1)$, $(0,0)$ and $(0,1)$ for (43), but the problem is to prove that no other solutions exists. We leave this to the reader as open questions.

Open Question 3.1. *Find all integer solutions to* (42).

Open Question 3.2. *Find all integer solutions to* (43).

4 Finding the solution set if it is finite

Now let us return to Diophantine equations in 3 or more variables. As we have seen in Section 2, in this case the problem of describing the solution set can be quite non-trivial even for simple equations. In general, this problem is not even well posed: if the solution set is infinite but not a finite union of polynomial families and cannot be described by recurrence relations, then was counts as an "acceptable description" of this solution set? For the sets with no obvious "structure" this problem is more philosophical than mathematical, and we will not discuss it further. Instead, we will focus on the following problem, which is completely well-defined.

Problem 4.1. *Given a polynomial Diophantine equation, decide whether its solution set is finite, and if so, list all the solutions.*

Note that proving that the solution set is infinite completely solves Problem 4.1, and no further analysis is required. If we focus on Problem 4.1 only, equations (21) in the Open Question 2.2 are trivial, because they obviously has infinitely many integer solutions. Indeed, we may choose $z = 1$, take arbitrary x, and set $y = x^2 + x \pm 1$. This finishes the analysis of all equations of size $H \leq 11$.

With $H = 12$, we cannot resist mentioning famous equation

$$x^2 + y^2 = z^2 \tag{44}$$

whose integer solutions are known as Pythagorean triples. A standard approach of solving this equation is noting that if $z = 0$ then $x = y = 0$ and otherwise the equation can be written as $(x/z)^2 + (y/z)^2 = 1$ and reduces to finding rational points in the circle $x^2 + y^2 = 1$. To find them, we can choose any one rational point, say $(x, y) = (1, 0)$, and draw all possible lines through this point with rational slope k. Any such line intersects the circle at $(1, 0)$ and in another point which can be easily seen to be a rational point. This way we get parameterization of all rational points with rational parameter k. From this, we can easily write down all integer solutions to (44). However, in the context of Problem 4.1, equation (44) is trivial, because it has infinitely many integer solutions for $y = 0$. As another example, equation

$$x^2 + y^2 = z^2 + 1 \tag{45}$$

has infinitely many integer solutions (take $x = z$) for $y = 1$. From now on, we will exclude from consideration all equations that have infinitely many integer solutions for some fixed value of one of the variables.

After this, the only non-excluded equation of size $H \leq 13$ is

$$x^2 + y^2 = z^2 - 1 \tag{46}$$

This equation has at most a finite number of integer solutions for any fixed x, y, or z, but still has infinitely many integer solutions. Indeed, for any integer t, we have a solution $x = 2t^2$, $y = 2t$, $z = 2t^2 + 1$. Integer solutions to (45) and (46) are known as "almost polynomial triples" and "nearly polynomial triples", respectively. See [10] for the complete description of the solution sets to these equations.

More generally, for any integer a, equation

$$x^2 + y^2 = z^2 + a \tag{47}$$

has infinitely many integer solutions. Indeed, we can rewrite the equation as $x^2 - a = y^2 - z^2 = (y - z)(y + z)$. For simplicity, assume that $y - z = 1$, so that $x^2 - a = y + z = 2z + 1$, from which we can find $z = (x^2 - a - 1)/2$. Now, if $a = 2k - 1$ is odd, take $x = 2t$, $z = 2t^2 - k$, and $y = z + 1 = 2t^2 - k + 1$, while if $a = 2k$ is even, take $x = 2t + 1$, $z = 2t^2 + 2t - k$, and $y = z + 1 = 2t^2 + 2t - k + 1$. As a side note, we remark that a complete description of the solution set to (47) with $a = 3$ has been an open question until Vaserstein [30, Example 15] proved that it is the union of two polynomial families.

More generally, if there exist polynomials $Q_1(t), \ldots, Q_n(t)$ with integer coefficients, not all constant, such that

$$P(Q_1(t), \ldots, Q_n(t)) \equiv 0 \tag{48}$$

then equation $P(x_1, \ldots, x_n) = 0$ has infinitely many integer solutions. In general, deciding the existence of such polynomials is a quite non-trivial problem. However, we can at least verify (48) for polynomials with small degree and coefficients, and exclude the equations for which we managed to find the corresponding Q_i. This method allows us to solve all the remaining equations of size $H \leq 16$.

The first equation of size $H = 17$ we discuss is

$$y^2 + z^2 = x^3 - 1 \tag{49}$$

We will prove that it has infinitely many integer solutions. Equivalently, there exists infinitely many integers x such that x^3-1 is the sum of two squares. Identity

$$(a^2+b^2)(c^2+d^2)=(ac+bd)^2+(ad-bc)^2$$

shows that if two integers can be represented as sum of squares, then so is their product. Because $x^3-1=(x-1)(x^2+x+1)$, it suffices to find x such that $x+1=u^2$ and $x-1=v^2+w^2$, or $v^2+w^2=u^2-2$. But this is equation (47) with $a=-2$ which has infinitely many integer solutions. In fact, this leads to explicit parametric family $x=3+8t+12t^2+8t^3+4t^4$, $y=5+20t+38t^2+40t^3+24t^4+8t^5$ and $z=-1-8t-28t^2-44t^3-44t^4-24t^5-8t^6$ of the solutions to (49), but the coefficients in these polynomials are too large for the direct search. This is the reason why (49) has not been excluded automatically and required explicit argument.

Another equation of size $H=17$ that requires attention is

$$y(x^2-y)=z^2+1 \tag{50}$$

We will prove that it has no integer solutions. For this, we will need a well-known fact [15] that all odd prime factors of a sum of squares z^2+1 must be congruent to 1 modulo 4. Hence the same is true for the odd prime factors of positive integers y and x^2-y. Because the product of any number of such primes is again 1 modulo 4, this implies that if z^2+1 is odd, then both y and x^2-y are congruent to 1 modulo 4, but then x^2 is congruent to 2 modulo 4, a contradiction. If z^2+1 is even, its prime factorization contains exactly one factor of 2, which goes to either y or x^2-y, resulting in x^2 being 3 modulo 4, again a contradiction.

This finishes the analysis of equations of size $H=17$. For $H=18$, we start with the equation similar to (49),

$$y^2+z^2=x^3-2 \tag{51}$$

Unlike x^3-1, x^3-2 does not factorise, so different technique is required. We will present a solution given by Max Alekseyev in the comment to mathoverflow question[3]. Let $x=t^2+2$ for some integer t. Then $x^3-2=(t^2+2)^3-2=(t^3+3t)^2+(3t^2+6)$. It is left to select $y=t^3+3t$ and note that equation $z^2=3t^2+6$ has infinitely many integer solutions. This can be checked directly or using the Gauss theorem [22, p. 57], that states that the general quadratic equation with integer coefficients

$$az^2+bzt+ct^2+dz+et+f=0$$

such that $D=b^2-4ac>0$, D is not a perfect square, and $\Delta=4acf+bde-ae^2-cd^2-fb^2\neq 0$ has either no integer solutions or infinitely many of them. In our case, $D=-4\cdot 3\cdot(-1)=12>0$, $\Delta\neq 0$, and there is an integer solution, say $z=3$, $t=1$. This finishes the proof.

Another equation of size $H=18$ is

$$x^2+y^2+xyz-2=0 \tag{52}$$

This equation requires completely different techniques called Vieta jumping. The idea is that if (x,y,z) is any solution to (52), then $t=x$ is a solution to the quadratic equation

$$t^2+yzt+y^2-2=0,$$

and this equation has another solution $t'=-yz-x=\frac{y^2-2}{x}$, which is also an integer. Hence, any solution (x,y,z) to (52) produces another solution $(-yz-x,y,z)$, and, by a similar argument, one more solution $(x,-xz-y,z)$. The technique suggests to consider a solution with $|x|+|y|+|z|$ minimal and either prove that there is no such solution, or find such minimal solution and then produce infinitely many other solutions by the transformations above.

Returning to equation (52), it has solutions $(x,y,z)=(\pm 1,\pm 1,0)$ with $z=0$. Let us prove that there are no other solutions. Assume the contrary and let (x,y,z) be a solution with $z\neq 0$ such that $|x|+|y|+|z|$ is

[3] https://mathoverflow.net/questions/409857/representing-x3-2-as-a-sum-of-two-squares

minimal. By symmetry, we may assume that $|x| \geq |y|$. By Vieta jumping, there is another solution (x', y, z) with $|x'| \geq |x|$ and $xx' = y^2 - 2$. Then $|y^2 - 2| = |x||x'| \geq x^2 \geq y^2$, hence either $y^2 = 0$ or $y^2 = 1$. But $y^2 = 0$ implies $x^2 = 2$ which is impossible, while $y^2 = 1$ implies that $z = 0$, a contradiction. Hence, the only solutions are $(x, y, z) = (\pm 1, \pm 1, 0)$.

To apply this technique to general equation $P(x_1, \ldots, x_n) = 0$, let us denote S the set of all variables x_i for which the equation can be written as

$$a_i x_i^2 + Q_i x_i + R_i = 0$$

where $|a_i| = 1$ and Q_i and R_i are polynomials in other variables. We can then use any computer algebra system to solve optimization problem of maximizing t over $(x_1, \ldots, x_n, t) \in \mathbb{R}^{n+1}$ subject to constraints $P = 0$, $|x_i| \geq t$ for each i, and $|-(Q_i/a_i) - x_i| \geq |x_i|$ for each variable $x_i \in S$. If the optimal value t^* to this optimization problem is infinite, then the method does not work for this equation. But if $t^* < \infty$, then we have $\min\{|x|, |y|, |z|\} \leq t^* < \infty$ for any solution (x, y, z) with $|x| + |y| + |z|$ minimal. So, we next check, for each integer t such that $|t| \leq t^*$ and each $i = 1, \ldots, n$, whether the equation has any integer solutions with $x_i = t$. If there are no such solutions, then the equation has no integer solutions at all. If there are such solutions, we next check whether any of them produces an infinite chain of solutions via Vieta jumping. In the rest of the paper, we will exclude the equations solvable by this method.

This finishes the analysis of all equations of size $H \leq 18$. Starting with $H = 19$, we get several equations in the form

$$ax^2 + bx + c + dy + exyz = 0,$$

such as, for example, $1 + x - x^2 + 2y + xyz = 0$, $1 + x^2 + 3y + xyz = 0$, etc. These equations are not covered by techniques discussed so far, but are in fact easy. Indeed, solutions with $x = 0$ or $y = 0$ can be found by direct substitution. Otherwise fraction $\frac{dy+c}{x}$ is an integer, hence we have $|dy + e| \geq |x|$, or $|x/y| \leq |d + e/y| \leq |d| + |e|$. Then $|z| = |(ax^2 + bx + c + dy)/(exy)| \leq |(ax^2 + bx + c + dy)/(xy)| = |ax/y + b/y + c/(xy) + d/x| \leq |a|(|d| + |e|) + |b| + |c| + |d|$. Now direct search in this region returns the full set of solutions.

A more interesting equation that require a new idea is

$$y^2 + z^2 = x^3 + 3 \tag{53}$$

To solve it, recall that a positive integer is the sum of two squares if and only if all its prime factors congruent to 3 modulo 4 enters its prime factorization an even number of times, see e.g. [15]. In particular, this implies that if a and b do not share prime factors congruent to 3 modulo 4, and ab is the sum of two squares, then so are both a and b. Now note that $x(x^3 + 3) = (x^2 - 1)^2 + (2x^2 + 3x - 1)$. Let x be any integer such that $2x^2 + 3x - 1$ is a perfect square (there are infinitely many such integers). Then x is not divisible by 3 (otherwise $2x^2 + 3x - 1$ would be 2 mod 3 and could not be a perfect square), hence x and $x^3 + 3$ are co-prime. But their product $x(x^3 + 3)$ is the sum of two squares, hence $x^3 + 3$ is a sum of two squares as well.

The same method allows to solve many other similar equations, such as $y^2 + z^2 = x^3 + x + 1$, $y^2 + z^2 = x^3 - x - 1$, $y + y^2 + z^2 = x^3 - 1$, etc. (The last equation after multiplication by 4 can be rewritten as $(2y + 1)^2 + (2z)^2 = 4x^3 - 3$, so it suffices to prove that $4x^3 - 3$ is the sum of squares infinitely often, and then the same method applies). We will exclude any further equations solvable by this method. This finishes the analysis for $H \leq 19$.

The only new equations of size $H = 20$ are homogeneous quadratic equations like

$$x^2 + y^2 = 3z^2.$$

The only integer solution to this equation is $(x, y, z) = (0, 0, 0)$. Indeed, if there is any other solution then we can divide it by any common factor and obtain a new solution for which (x, y, z) are co-prime. However, the sum of squares $x^2 + y^2$ is divisible by 3 only if both x and y are divisible by 3. But in this case $x^2 + y^2$ is divisible by 9, hence z is divisible by 3, a contradiction with the co-primality assumption. Famous Hasse–Minkowski theorem (Hasse principle) states that if a homogeneous quadratic equation has non-zero real solutions but no non-zero integer solutions, then this can always be proved by divisibility analysis modulo some p as above. This allows us to exclude such equations as well.

For $H = 21$, the only equation of different type is

$$y(x^2+2) = 2z^2 - 1. \tag{54}$$

So far we have used only the information about prime factors of sum of two squares, while this equation requires the analysis of prime factors of other quadratic polynomials, in this case x^2+2 and $2z^2-1$. As shown in [15], all odd prime factors of x^2+2 must be 1 or 3 modulo 8, while all prime factors of $2z^2-1$ must be 1 or 7 modulo 8. A combination of these facts imply that if (x,y,z) solves (54), then all prime factors of x^2+2 are congruent to 1 modulo 8. But then x^2+2 must be itself congruent to 1 modulo 8, which is a contradiction. We refer to [15] how to apply this method in general, but here will not list further equation solvable in this way.

For $H = 22$, we start to meet equations like

$$y^2 + yz + z^2 = x^3 - x \tag{55}$$

that require the analysis of which integers can be represented in the form y^2+yz+z^2. Let S be the set of all such integers. It is known that S is also the set of integers representable as $3y^2+z^2$, and also the set of integers n such that every prime p of the form $p = 3k+2$ enters the prime factorization of n in the even power. We need to prove that x^3-x belongs to S for infinitely many x. Choose any odd x such that $2x^2-2x-4 = 3t^2$ for some integer t (there are infinitely many such x). Then $(x^3-x)(x-2) = (x^2-x-2)^2 + (2x^2-2x-4)$ belongs to S. Because (x^3-x) and $(x-2)$ do not share any prime factors in the form $p = 3k+2$, this implies that $x^3-x \in S$. The same method allows to solve other equations of this type, such as $y^2+yz+z^2 = x^3+x$ and $y^2+yz+z^2 = x^3-2$.

The next equation we discuss is

$$y(z^2-y) = x^3 + 2.$$

We present a solution given by Mathoverflow user Tomita[4]. By considering the equation as quadratic in y, we conclude that it has infinitely many integer solutions if and only if the determinant $D = (-z^2)^2 - 4(x^3+2) = z^4 - 4x^3 - 8$ is a perfect square infinitely often. Now assume that $x = -3t^2-2t-2$ and $z = 3t+1$ for some integer t. Then $D = (3t+1)^4 - 4(-3t^2-2t-2)^3 - 8 = (12t^2+8t+25)(3t^2+2t+1)^2$. It is left to remark that $12t^2+8t+25$ is a perfect square for infinitely many integers t.

The same method solves another equation,

$$xyz = x^3 + y^2 - 2.$$

We need $D = x^2z^2 - 4x^3 + 8$ to be a perfect square. Select $x = 6t^2+1$ and $z = 6t$, then $D = 4(6t^2-1)^2(3t^2+1)$. It is left to note that there are infinitely many integers t such that $3t^2+1$ is a perfect square.

However, we currently do not see how to use these (or other) methods to solve similar equations

$$y(z^2-y) = x^3 - 2 \tag{56}$$

and

$$xyz = x^3 + y^2 + 2. \tag{57}$$

These are the only remaining open equations of size $H \le 22$. A computer search for polynomials $x = Q(t)$ and $z = R(t)$ with small degree and coefficients returns no polynomials for which the same method works. Hence, we need either a deeper search for polynomials with large coefficients, or a new idea. We will leave these equations to the readers as open questions.

Open Question 4.2. *Are there infinitely many integer solutions to (56)?*

Open Question 4.3. *Are there infinitely many integer solutions to (57)?*

[4] https://mathoverflow.net/questions/411958

One may also study Problem 4.1 for some restricted families of polynomials. For example, if we restrict the number of variables and consider 2-variable equations only, then the smallest equations for which Problem 4.1 is open are equations (42) and (43) of size $H = 28$, see open questions 3.1 and 3.2.

Another nice class of equations we may consider are symmetric equations, ones that are invariant after cyclic shirt of variables. The smallest symmetric equation not directly solvable by the methods described above turns out to be the equation

$$x^2y + y^2z + z^2x = 1 \tag{58}$$

of size $H = 25$.

Open Question 4.4. *Solve Problem 4.1 for the equation* (58).

Finally, we may also restrict the number of monomials. It is easy to solve all 2-monomial equations [15], hence the first interesting case is 3-monomial ones. The smallest 3-monomial equation which seems to be not solvable by the described methods is the equation

$$x^3y^2 = z^3 + 2 \tag{59}$$

of size $H = 42$. This equation has obvious solutions $(x, y, z) = (1, \pm 1, -1)$. Note that any integer n can be represented in the form x^3y^2 if and only if for every prime number p dividing n, p^2 also divides n. Such integers are called powerful numbers. So, the question is to find all integers z such that $z^3 + 2$ is a powerful number.

Open Question 4.5. *Find all integer solutions to the equation* (59).

5 Existence of solutions: Hilbert 10th problem

In addition to Problem 4.1, one may consider the following problem with Yes/No answer.

Problem 5.1. *Given a polynomial Diophantine equation, check whether it has any integer solution.*

Hilbert's 10th problem asks for a general method for solving Problem 5.1 for all Diophantine equations. Building on the work of Davis, Putnam and Robinson [9], Matiyasevich [19] proved in 1970 that no such general algorithm exists. See excellent recent surveys of Gasarch [13, 11, 12] for a detailed discussion for which Diophantine equations Problem 5.1 can be solved, and in which cases it is known to be undecidable. For all the equations we left open in the previous sections, Problem 5.1 is trivial because these equations have some obvious small solutions.

In [15], we found the smallest Diophantine equation for which Problem 5.1 is currently open. This is the equation

$$y(x^3 - y) = z^3 + 3 \tag{60}$$

of size $H = 31$.

Open Question 5.2. *Do there exist integers x, y, z satisfying* (60)*?*

The same question can also be asked for restricted families of equations. Among the 2-variable equations, the smallest open are

$$y^3 + xy + x^4 + 4 = 0, \tag{61}$$

$$y^3 + xy + x^4 + x + 2 = 0, \tag{62}$$

$$y^3 + y = x^4 + x + 4 \tag{63}$$

and

$$y^3 - y = x^4 - 2x - 2 \tag{64}$$

of size $H = 32$.

Open Question 5.3. *Determine whether each of the equations* (61)-(64) *have any integer solution.*

The smallest open symmetric equation is

$$x^3 + y^3 + z^3 + xyz = 5 \tag{65}$$

of size $H = 37$.

Open Question 5.4. *Do there exist integers* x, y, z *satisfying* (65)*?*

Finally, the smallest open 3-monomial equation is

$$x^3y^2 = z^3 + 6 \tag{66}$$

of size $H = 46$.

Open Question 5.5. *Do there exist integers* x, y, z *satisfying* (66)*?*

In addition, we may consider the smallest open equations with respect to alternative measures of size. As noted in the introduction, a natural measure of "length" of a monomial M of degree d with coefficient a is $l(M) = \log_2 |a| + d$. Then we can define the length $l(P)$ of a polynomial P consisting of k monomials with coefficients $a_1, \ldots, a_k$ and degrees $d_1, \ldots, d_k$, respectively, as

$$l(P) = \sum_{i=1}^{k} (\log_2 |a_i| + d_i). \tag{67}$$

Then, instead of ordering the equations by H, we may order them by length l, or, equivalently, by an integer

$$L(P) := 2^{l(P)} = \prod_{i=1}^{k} |a_i| \cdot 2^{\sum_{i=1}^{k} d_i}.$$

Note that the formula for $L(P)$ is the same as the formula (3) for $H(P)$, except that the summation is replaced by a product.

As established in [15], the shortest equations for which Problem 5.1 is open are the equations

$$y(x^3 - y) = z^4 + 1, \tag{68}$$

$$2y^3 + xy + x^4 + 1 = 0 \tag{69}$$

and

$$x^3y^2 = z^4 + 2 \tag{70}$$

that have length $l = 10$.

Open Question 5.6. *Do there exist integers* x, y, z *satisfying* (68)*?*

Open Question 5.7. *Do there exist integers* x, y *satisfying* (69)*?*

Open Question 5.8. *Do there exist integers* x, y, z *satisfying* (70)*?*

6 Conclusions

We have ordered all Polynomial Diophantine equations by a parameter H defined in (3) and tried to solve the equations in that order. We have considered the following problems in the decreasing level of difficulty.

- Completely solve the equation: list all solutions if there are finitely many and describe all solutions (for example, as a union of polynomial families) if the solution set is infinite.

- Determine whether the solution set is finite, and if yes, list all the solutions.
- Check whether an equation has any integer solution.

For each of the problems, we have identified the smallest equations for which the problem is open. In some cases, we also identified the smallest open equations is certain families, such as the smallest open 2-variable, symmetric, or 3-monomial equations. The list of current smallest open equations can also be found on Mathoverflow [16, 17], where the plan is to always keep the list up-to-dated.

We suspect that some of the open equations listed in this paper are relatively easy, and are suitable for the first research project of a graduate or even undergraduate student. On the other hand, we are confident that some of our equations are quite difficult and may stimulate the development of new methods and techniques in number theory.

Acknowledgments

I thank anonymous mathoverflow user Zidane whose question [33] inspired this work. I also thank other mathoverflow users, including but not limited to Will Sawin, Fedor Petrov, Andrew R. Booker, Victor Ostrik, Jeremy Rouse, Max Alekseyev and Tomita for very helpful discussions on this topic and for solving some equations in this project that was previously listed as open. I thank William Gasarch for the interest to publish this paper in SIGAST NEWS Open Problem Column and for proofreading an earlier draft of this paper and providing me with helpful list of comments and suggestions. I also thank Aubrey de Grey for writing his own version of the computer program for enumerating equations, which provides an independent validation that none of the equations of small size has been accidentally missing.

References

[1] Dario Alpern. Generic two integer variable equation solver. https://www.alpertron.com.ar/QUAD.HTM, 2020 (accessed June 12, 2020).

[2] Titu Andreescu and Dorin Andrica. *Quadratic Diophantine Equations*. Springer, 2015.

[3] Alan Baker. Bounds for the solutions of the hyperelliptic equation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 65(2):439–444, 1969.

[4] Alan Baker and John Coates. Integer points on curves of genus 1. *Mathematical Proceedings of the Cambridge Philosophical Society*, 67(3):595–602, 1970.

[5] Andrew R Booker and Andrew V Sutherland. On a question of Mordell. *Proceedings of the National Academy of Sciences*, 118(11), 2021.

[6] B Brindza. On s-integral solutions of the equation $y^m = f(x)$. *Acta Mathematica Hungarica*, 44(1):133–139, 1984.

[7] Nils Bruin and Michael Stoll. Deciding existence of rational points on curves: an experiment. *Experimental Mathematics*, 17(2):181–189, 2008.

[8] Felipe Cucker, Pascal Koiran, and Steve Smale. A polynomial time algorithm for Diophantine equations in one variable. *Journal of Symbolic Computation*, 27(1):21–29, 1999.

[9] Martin Davis, Hilary Putnam, and Julia Robinson. The decision problem for exponential Diophantine equations. *Annals of Mathematics*, 74:425–436, 1961.

[10] Orrin Frink. Almost Pythagorean triples. *Mathematics Magazine*, 60(4):234–236, 1987.

[11] William Gasarch. Hilbert's tenth problem: Refinements and variations. *arXiv preprint arXiv:2104.07220*, 2021.

[12] William Gasarch. Hilbert's tenth problem: Refinements and variations. *SIGACT News*, 52(2), 2021.

[13] William Gasarch. Hilbert's tenth problem for fixed d and n. *Bulletin of EATCS*, 1(133), 2021.

[14] Josef Gebel, A Pethö, and Horst G Zimmer. On Mordell's equation. *Compositio Mathematica*, 110(3):335–367, 1998.

[15] Bogdan Grechuk. Diophantine equations: a systematic approach. *arXiv preprint arXiv:2108.08705*, 2021.

[16] Bogdan Grechuk. Can you solve the listed smallest open Diophantine equations? https://mathoverflow.net/questions/400714/, 2021 (accessed December 31, 2021).

[17] Bogdan Grechuk. On the smallest open Diophantine equations: beyond Hilbert's 10 problem. https://mathoverflow.net/questions/411958/, 2021 (accessed December 31, 2021).

[18] Sachi Hashimoto and Travis Morrison. Chabauty-coleman computations on rank 1 picard curves. *arXiv preprint arXiv:2002.03291*, 2020.

[19] Ju V Matijasevic. Enumerable sets are Diophantine. In *Soviet Math. Dokl.*, volume 11, pages 354–358, 1970.

[20] Jeffrey CP Miller and Michael FC Woollett. Solutions of the Diophantine equation: $x^3 + y^3 + z^3 = k$. *Journal of the London Mathematical Society*, 1(1):101–110, 1955.

[21] LJ Mordell. On the integer solutions of the equation $x^2 + y^2 + z^2 + 2xyz = n$. *Journal of the London Mathematical Society*, 1(4):500–510, 1953.

[22] Louis Joel Mordell. *Diophantine equations*. Academic Press, 1969.

[23] Lisa Piccirillo. The Conway knot is not slice. Ann. of Math., 191(2):581–591, 2020.

[24] Dimitrios Poulakis. Points entiers sur les courbes de genre 0. *Colloquium Mathematicae*, 66(1):1–7, 1993.

[25] Dimitrios Poulakis and Evaggelos Voskos. Solving genus zero Diophantine equations with at most two infinite valuations. *Journal of Symbolic Computation*, 33(4):479–491, 2002.

[26] Carl Runge. Ueber ganzzahlige lösungen von gleichungen zwischen zwei veränderlichen. *Journal fur die reine angewandte Mathematik*, 1887.

[27] Carl L Siegel. Über einige anwendungen diophantischer approximationen. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, 1:14–72, 1929.

[28] Thoralf Skolem. *Diophantische gleichungen*. Berlin, 1938.

[29] Roel Stroeker and Nikolaus Tzanakis. Computing all integer solutions of a genus 1 equation. *Mathematics of computation*, 72(244):1917–1933, 2003.

[30] Leonid Vaserstein. Polynomial parametrization for the solutions of Diophantine equations and arithmetic groups. *Annals of mathematics*, pages 979–1009, 2010.

[31] P Walsh. A quantitative version of Runge's theorem on Diophantine equations. *Acta Arithmetica*, 62:157–172, 1992.

[32] Andrew Wiles. Modular elliptic curves and Fermat's last theorem. *Annals of mathematics*, 141(3):443–551, 1995.

[33] Zidane. What is the smallest unsolved Diophantine equation? https://mathoverflow.net/questions/316708/, 2018 (accessed June 12, 2020).

[34] Paul Zimmermann, Alexandre Casamayou, Nathann Cohen, Guillaume Connan, Thierry Dumont, Laurent Fousse, François Maltey, Matthias Meulien, Marc Mezzarobba, Clément Pernet, et al. *Computational mathematics with SageMath*. SIAM, 2018.