Creating decidable diophantine equations

Robert DOUGHERTY-BLISS, Charles KENNEY, and Doron ZEILBERGER

February 6, 2024

Abstract

We extract ideas implicit in Matiyasevich's (negative) resolution of Hilbert's tenth problem, to construct many non-trivial infinite families for which we can not just decide that they have solutions, but in fact explicitly construct all the solutions. We can also come up with many families for which we can prove that no solutions exist.

1 Preface

In 1970, 23-year-old Yuri Matiyasevich, standing on the shoulders of Julia Robinson, Martin Davis, and Hilary Putnam, shocked the world of mathematics by showing that David Hilbert's dream of finding an algorithm that inputs *any* polynomial

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

with integer coefficients, and outputs **true** or **false** if $P = 0$ has, or does not have, solutions in integers, can never come to be.

Of course, for *specfic* equations, and even, many specific infinite families, one can often decide, but there is no *magic bullet* that, can decide all of them.

For example, Pythagoras got very upset when Hippasus of Metapontum discovered that the diophantine equation

$$x^2 - 2y^2 = 0 \quad ,$$

has **no solution**, and Hippasus (probably) gave a fully rigorous proof (**not** the usual one but a geometrical version of the **reduction formula** if (x, y) $x > y > 0$ is a solution so is $(y, 2x - y$ and since $(1, 0)$ is not a solution **qed**. Going backwards, using the fact that if (x, y) is a solution of *Pell's equation*

$$x^2 - 2y^2 \; = \; \pm 1$$

then so is $(x+2y, x+y)$, we can prove that there are **infinitely many** solutions. A little more effort will show that these are the only ones. (see [Z]).

Much harder is the fact proved by Sir Andrew Wiles [W], that for every $n > 2$ the diophantine equation

$$x^n + y^n = z^n \quad ,$$

has no solution.

On the positive side, Noam Elkies [E] famously proved that

$$A^4 + B^4 + C^4 = D^4 \quad ,$$

has *infinitely* many solutions.

Wouldn't it be nice to be able to manufacture, *at will*, many examples of diophantine equations for which we can explicitly construct *all solutions*?

There is a cheap way to do this. As most of us know, the triple

$$a = A^2 - B^2 \quad , \quad , b = 2AB \quad , \quad c = A^2 + B^2 \quad ,$$

satisfies

$$a^2 + b^2 = c^2$$

A little more challenging is to prove that all solutions of $a^2 + b^2 = c^2$ with $\gcd(a, b) = 1$ are of this form. But this is true, and the same idea applies more generally.

Take any $m+1$ polynomials in m variables with integer coefficients

$$P_i(a_1, \ldots, a_m) \quad , \quad 1 \leq i \leq m+1$$

and define

$$X_i = P_i(a_1, \ldots, a_m) \quad , \quad 1 \leq i \leq m+1 \quad ,$$

using, e.g **Gröbner bases** (the Buchberger algorithm) we can *eliminate* $a_1, \ldots, a_m$ and get a polynomial equation (with integer coefficients)

$$Q(X_1, \ldots, X_m) \; = \; 0 \quad ,$$

that has a parametric solution as above. Alas, in general this leads to monster equations and unlike the case with Pythagorean triples, it is not clear that there aren't other solutions.

So it would be nice to be able to generate, in a systematic way, many examples of simple diophantine equations for which we know infinitely many solutions, and to be able to prove that these are all of them. It would also be nice to manufacture not too complicated, but non-trivial, diophantine equations for which we can conclusively prove that there aren't any solutions.

2 Diophantine equations from recurrences

In Matiyasevich's proof (we use the versions in [JM] and [M]) a central role is played by *Pell's equation* and the fact ([M], pp. 19-20) that two consecutive terms $x = a_b(n), y = a_b(n+1)$ of the sequence of integers defined by the **second-order** linear recurrence

$$a_b(0) = 0 \quad , \quad a_b(1) = 1 \quad , \quad a_b(n+2) \, = \, ba_b(n+1) - a_b(n) \quad ,$$

satisfy the diophantine equation

$$x^2 - bxy + y^2 = 1 \quad ,$$

and conversely if $x > y$ satisfies it, then there must be an n such that $x = a_b(n+1)$, $y = a_b(n)$. This gave us the idea to consider higher-order recurrences.

If F_n is the nth Fibonacci number, then taking the determinant of the well-known matrix identity

$$\begin{bmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}^n ,$$

yields Cassini's identity $F_{n+1}F_{n-1} - F_n^2 = (-1)^n$. If we square this and apply the Fibonacci recurrence, then we obtain $P(F_{n-1}, F_n) = 1$ for some polynomial P. From another perspective, starting with the Fibonacci numbers we created a polynomial diophantine equation $P(x, y) = 1$ with infinitely many solutions. Our goal is to repeat this for a wider class of recurrences.

Consider the linear recurrence

$$a(n) = c_1 a(n-1) + \cdots + c_d a(n-d). \tag{1}$$

Our recipe has two parts. First, the matrix

$$B = \begin{bmatrix} c_1 & c_2 & \cdots & c_d \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

satisfies the "forward identity"

$$\begin{bmatrix} a(n+1) \\ a(n) \\ \vdots \\ a(n-d+2) \end{bmatrix} = B \begin{bmatrix} a(n) \\ a(n-1) \\ \vdots \\ a(n-d+1) \end{bmatrix}$$

for any sequence which satisfies (1). Second, (1) has d "fundamental solutions" which form a basis for all solutions. They are the solutions whose initial conditions are all zero except for a single entry, which is instead one. By luck, the columns of the identity matrix are exactly the initial conditions of these fundamental solutions. If we call the fundamental solutions $e_0(n), e_1(n), \ldots, e_{d-1}(n)$, then

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & \cdots & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} e_0(d-1) & e_1(d-1) & \cdots & e_{d-1}(d-1) \\ e_0(d-2) & e_1(d-2) & \cdots & e_{d-1}(d-2) \\ e_0(d-3) & e_1(d-3) & \cdots & e_{d-1}(d-3) \\ & & \cdots & \\ e_0(0) & e_1(0) & \cdots & e_{d-1}(0) \end{bmatrix}.$$

This implies a relation between B^n and the fundamental solutions:

$$B^n = B^n I = \begin{bmatrix} e_0(n+d-1) & e_1(n+d-1) & \cdots & e_{d-1}(n+d-1) \\ e_0(n+d-2) & e_1(n+d-2) & \cdots & e_{d-1}(n+d-2) \\ e_0(n+d-3) & e_1(n+d-3) & \cdots & e_{d-1}(n+d-3) \\ & & \cdots & \\ e_0(n) & e_1(n) & \cdots & e_{d-1}(n) \end{bmatrix}. \tag{2}$$

From the elementary theory of difference equations, every solution to (1)—including the fundamental ones—can be expressed as a linear combination of the sequences $e_0(n)$, $e_0(n+1)$, …, $e_0(n+d-1)$. Therefore every entry in the right-hand side of (2) is actually a linear combination of shifts of $e_0(n)$. By taking determinants in (2) it follows that

$$P(e_0(n), e_0(n+1), \ldots, e_0(n+d-1)) = (\det B)^n$$

for some polynomial P. Laplace expansion implies $\det B = (-1)^d c_d$, so setting $c_d = (-1)^d$ makes the right-hand side 1.

The previous considerations lead to the following proposition.

Proposition 1 *For any integers $c_1, c_2, \ldots, c_{d-1}$, there is a nonzero polynomial $P(x_1, x_2, \ldots, x_d)$ such that the diophantine equation*

$$P(x_1, x_2, \ldots, x_d) = 1$$

has infinitely many solutions. In particular, the points $(a(n), a(n+1), \ldots, a(n+d-1))$ are solutions, where $a(n)$ satisfies

$$a(n) = \sum_{k=1}^{d-1} c_k a(n-k) + (-1)^d a(n-d)$$

and has initial conditions $0, 0, \ldots, 0, 1$.

Our goal is to show that the diophantine equations in Proposition 1 are sometimes solved by *only* the recurrence solutions. This goal is too lofty in general, but we have arguments which apply to an infinite family of recurrences, and one detailed case study concerning the Tribonacci numbers.

3 Tribonacci numbers

We begin with the Tribonacci numbers as a detailed example. The main idea is to show that all solutions to the associated diophantine equation are generated, in some sense, by increasing solutions, and then to construct all increasing solutions.

Definition 1 Define the numbers T_n by

$$\begin{aligned} T_0 &= T_1 = 0 \\ T_2 &= 1 \\ T_n &= T_{n-1} + T_{n-2}, \end{aligned}$$

the polynomial P_T by

$$P_T(x,y,z) = x^3 + 2x^2y + x^2z + 2xy^2 - 2xyz - xz^2 + 2y^3 - 2yz^2 + z^3,$$

and the map R_T by

$$R_T(x,y,z) = (y, z, x+y+z).$$

Note that P_T is invariant under R_T, i.e., $P_T \circ R_T = P_T$.

Proposition 2 *If $P_T(x,y,z) = 1$ for integers (x,y,z), then (x,y,z) is the result of repeatedly applying R_T or its inverse to a nonnegative increasing solution. That is, there exist integers $0 \le a \le b \le c$ and a positive integer n such that $P(a,b,c) = 1$ and (x,y,z) is $R_T^n(a,b,c)$ or $R_T^{-n}(a,b,c)$.*

Proof Repeatedly applying R_T to our initial point (x,y,z) produces a sequence $a(n)$ which satisfies

$$a(n) = a(n-1) + a(n-2) + a(n-3)$$

with initial conditions $(a(0), a(1), a(2)) = (x,y,z)$. Because P_T is invariant under R_T we have $P_T(a(n), a(n+1), a(n+2)) = 1$ for all n. The elementary theory of difference equations implies $a(n) \sim c \cdot \alpha^n$ where $\alpha = 1.8393$ is the unique real root of $X^3 - X^2 - X - 1$ and

$$c = \alpha \frac{(\alpha^2 - \alpha - 1)a(0) + (\alpha - 1)a(1) + a(2)}{\alpha^2 + 2\alpha + 3}.$$

Note that c is real. If $c < 0$, then we eventually obtain a strictly negative solution, which is impossible because $P_T(x, y, z) \leq 0$ if $x, y, z \leq 0$. If $c = 0$ then $(\alpha^2 - \alpha - 1)a(0) + (\alpha - 1)a(1) + a(2) = 0$, and this is impossible because $\{1, \alpha, \alpha^2\}$ is linearly independent over the rationals. The remaining possibility is $c > 0$, which implies that we eventually have $0 < a(n) < a(n + 1) < a(n + 2)$, and we get back to (x, y, z) by applying the inverse map R_T^{-1}. ■

Proposition 3 *If $P_T(x, y, z) = 1$ for integers $0 \leq x \leq y \leq z$, then $(x, y, z) = (T_n, T_{n+1}, T_{n+2})$ for some integer $n \geq 0$.*

Proof The map $R_T^{-1}(x, y, z) = (z - x - y, x, y)$ takes solutions to other solutions. Note that if $0 \leq z - x - y \leq x$, then the new solution is also nonegative and increasing, and in fact strictly smaller unless $x = y = 0$. (If $x = y = 0$ then $z = 1$ is the unique solution.) We will show that $0 \leq z - x - y \leq x$ for all increasing solutions with sufficiently large z.

If we divide both sides of the equation $P_T(x, y, z) = 1$ by z^3, and make the change of variables $(t, s) = (x/z, y/z)$, then we obtain

$$2s^3 + 2s^2t + 2st^2 + t^3 - 2st + t^2 - 2s - t + 1 = \frac{1}{z^3}. \tag{3}$$

Call the left-hand side of this equation $f(s, t)$ and note that it is a cubic defined on the unit square. It is a routine calculus exercise to show that the minimum of $f(s, t)$ on the region $1 - t - s < 0$ is

$$\frac{398 - 68\sqrt{34}}{27}.$$

Therefore we cannot have both (3) and $1 - t - s < 0$ for

$$z > \left(\frac{398 - 68\sqrt{34}}{27}\right)^{-1/3} = 2.6235.$$

It follows that $0 \leq z - x - y$ for all increasing solutions to $P_T(x, y, z) = 1$ with $z \geq 3$. By an analogous argument on the region $1 - t - s > t$, all increasing solutions to $P_T(x, y, z) = 1$ with $z \geq 5$ satisfy $z - x - y \leq x$.

Repeatedly applying the "backwards" map R_T^{-1} produces smaller, nonnegative, increasing solutions as long as $z \geq 5$, and so this process terminates at a solution with $0 \leq x \leq y \leq z < 5$. It is simple to check that all such solutions return to the point $(0, 0, 1)$ under the map R_T^{-1}, and so *all* increasing nonnegative solutions come from applying the "forward" map R_T to $(0, 0, 1)$. This produces exactly the Tribonacci numbers. ■

See Figure **??** for a visual representation of the maps and regions in Proposition 3.

Theorem 1 *If $P_T(x, y, z) = 1$ for integers x, y, z, then $(x, y, z) = (T_n, T_{n+1}, T_{n+2})$ for some integer n.*

Proof By the previous two propositions, every solution comes from applying the maps $(x, y, z) \mapsto (y, z, x + y + z)$ and $(x, y, z) \mapsto (z - x - y, x, y)$ to the solution $(0, 0, 1)$, which produces exactly the Tribonacci numbers with positive and negative indices. ■

4 Uniqueness in general

The arguments from the previous section carry over almost verbatim to the general third-order recurrence. The main difficulty is in establishing the minimum of the analogous cubic (3). For any specific recurrence it is completely routine to check whether the proof of Proposition 3 works, but Proposition 5 gives a weaker statement about an infinite family.

Definition 2 For any positive integers a and b, define the polynomial $P_{ab}(x, y, z)$ as

$$a^2y^2z+abxyz+aby^3+b^2xy^2+ax^2z+axy^2-2ayz^2+2bx^2y-bxz^2-by^2z+x^3-3xyz+y^3+z^3$$

Proposition 4 *Let a and b be positive integers such that X^3-aX^2-bX-1 is irreducible over $\mathbf{Q}$ and has a single largest root which is real and greater than 1. Then all integer solutions to $P_{ab}(x, y, z) = 1$ are generated by applying the map $(x, y, z) \mapsto (z - ay - bx, x, y)$ or its inverse to a nonnegative, increasing solution.*

Proof The argument is the same as in Proposition 2. The irreducibility of $X^3 - aX^2-bX-1$, with largest root $\alpha > 1$, implies the linear independence of $\{1, \alpha, \alpha^2\}$ over the rationals and gives the correct asymptotics. ■

Proposition 5 *Fix positive integers a and b and consider the recurrence*

$$u(n) = au(n - 1) + bu(n - 2) + u(n - 3). \tag{4}$$

If a is sufficiently large relative to b, then all solutions $0 \le x \le y \le z$ to the diophantine equation $P_{ab}(x, y, z) = 1$ are generated by applying (4) to finitely many solutions.

It should be noted that while the following proof is non-constructive, the *method* is not. Carrying out the proof for any *specific* integers a and b will determine an exact bound under which the finitely many initial conditions can be found.

Proof The polynomial $P_{ab}(x, y, z)$ is invariant under the map

$$(x, y, z) \mapsto (z - ay - bx, x, y), \tag{5}$$

so it takes solutions to solutions. In particular, the new solution is strictly closer to the origin unless $x = y = 0$ (which yields the unique solution $z = 1$). We will show that the new solution is also nonnegative and increasing for sufficiently large z.

If we divide both sides of $P_{ab}(x, y, z) = 1$ by z^3 and make the change of variables $(t, s) = (x/z, y/z)$, then we obtain $f_{ab}(t, s) = z^{-3}$ where f_{ab} is a cubic in t and s on the unit square. Because f_{ab} is a cubic, it is possible to exactly compute its critical points on the unit square, as well as the critical points of boundary functions such as $f_{ab}(0, s)$ and $f_{ab}(1, s)$. If we treat b as a constant and perform asymptotic expansions as $a \to \infty$ of these critical points, it turns out that the minimum of f_{ab} on the region $\{1 - as - bt < 0\} \cup \{1 - as - bt > t\}$ occurs on the line $1 - as - bt = 0$, and it equals

$$\frac{1}{a^6} - \frac{b^2}{4a^7} - \frac{9b}{2a^8} + O(a^{-9}).$$

So $f_{ab}(t, s) = z^{-3}$ fails if

$$z > a^2 + \frac{b^2}{12}a + \frac{3b}{2} + \frac{b^4}{72} + O(a^{-1}). \tag{6}$$

It follows that $0 < 1 - as - bt < t$, also known as $0 < z - ay - bx < x$ for any solution $0 \le x \le y \le z$ with sufficiently large z. We may therefore iterate (5) on such a solution until we reach one where z is below the bound implied by (6), and there are only finitely many of these. ■

Theorem 2 *Let a and b be positive integers such that*

1. *$X^3 - aX^2 - bX - 1$ is irreducible over the rationals and has a single largest root which is real and greater than 1; and*

2. *a is sufficiently large relative to b (in the non-constructive sense of proposition 4).*

Then all integer solutions to $P_{ab}(x, y, z) = 1$ are generated by applying (4) *forwards or backwards to finitely many initial solutions.*

Note that the first condition is not very restrictive. The cubic $X^3 - aX^2 - bX - 1$ has a rational root only if $b = a + 2$.

5 Examples

A single family The characteristic equation

$$X^3 - 10X^2 - 3X - 1$$

leads to the diophantine equation

$$x^3 + 6x^2y + 10x^2z + 19xy^2 + 27xyz - 3xz^2 + 31y^3 + 97y^2z - 20yz^2 + z^3 = 1.$$

Theorem 2 (along with explicit arguments from Proposition 5) shows that all solutions to this equation are generated by applying the maps $(x, y, z) \mapsto (y, z, 10z + 3y + x)$ and $(x, y, z) \mapsto (z - 10y - 3x, x, y)$ to the initial solution $(0, 0, 1)$.

Multiple families The characteristic equation

$$X^3 - 2X^2 - 3X - 1$$

leads to the diophantine equation

$$x^3 + 6x^2y + 2x^2z + 11xy^2 + 3xyz - 3xz^2 + 7y^3 + y^2z - 4yz^2 + z^3 = 1.$$

Theorem 2 (along with explicit arguments from Proposition 5) shows that all solutions to this equation are generated by applying the maps $(x, y, z) \mapsto (y, z, 3z + 2y + x)$ and $(x, y, z) \mapsto (z - 3y - 2x, x, y)$ to the initial solutions

$$(0, 0, 1), (0, 1, 3), (0, 2, 7), (1, 1, 4).$$

A failure The characteristic equation

$$X^3 - X^2 - 3X - 1 = (X - 1)(X^2 - 2X - 1)$$

corresponds to setting $a = 1$ and $b = 3$, which leads to the diophantine equation

$$x^3 + 6x^2y + x^2z + 10xy^2 - 3xz^2 + 4y^3 - 2y^2z - 2yz^2 + z^3 = 1.$$

Our method fails here on two counts. First, the proof of Proposition 5 does not go through ($a = 1$ is not big enough relative to $b = 3$). Second, this recurrence has degenerate integer solutions like $(-1)^n$ which do not have the correct asymptotics.