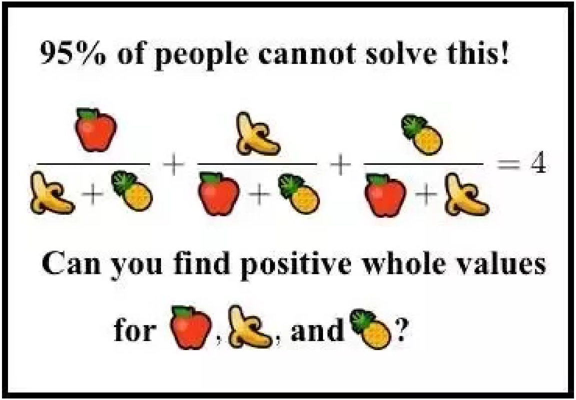
Can we solve Diophantine equations?

David Ang

University College London

Wednesday, 28 May 2025

Can you solve this?



Smallest positive whole values:

154476802108746166441951315019919837485664325669565431700026634898253202035277999

36875131794129999827197811565225474825492979968971970996283137471637224634055579

4373612677928697257861252602371390152816537558161613618621437993378423467772036

Diophantine equations

A **Diophantine equation**, named after Diophantus of Alexandria, is a *polynomial* equation with *integer* coefficients in *two or more* unknown variables.

For instance, the equation:

$$\frac{X}{Y+Z} + \frac{Y}{X+Z} + \frac{Z}{X+Y} = 4$$

is essentially equivalent to the polynomial equation:

$$X^3 + Y^3 + Z^3 = 3X^2Y + 3XY^2 + 3X^2Z + 3XZ^2 + 3Y^2Z + 3YZ^2 + 5XYZ$$

To **solve** a Diophantine equation means to find all its *integer* solutions. Are there any? Can we write one down? Are there infinitely many? Can we generate them systematically? How are they distributed?

Some examples

Here are some famous Diophantine equations.

- Pythagoras's equation $X^2 + Y^2 = Z^2$. This has solutions:

$$X = (m^2 - n^2)k \qquad Y = 2mnk \qquad Z = (m^2 + n^2)k$$

- Pell's equation $X^2 - nY^2 = 1$ for fixed $n \in \mathbb{Z}$.
  - For $n = 60$, the smallest solution is $X = 31$ and $Y = 4$.
  - For $n = 61$, the smallest solution is $X = 1766319049$ and $Y = 226153980$.
  - For $n = 62$, the smallest solution is $X = 63$ and $Y = 8$.
- Mordell's equation $Y^2Z = X^3 - nZ^3$ for fixed $n \in \mathbb{Z}$.
  - For $n = -1$, the only solutions are $(-1, 0), (0, \pm 1), (2, \pm 3)$.
  - For $n = 1$, the only solutions are $(1, 0)$.
  - For $n = 2, 4, 11$, there are infinitely many solutions.
  - For $n = \pm 6, \pm 7$, there are no solutions.
- The Erdös–Straus conjecture says there are positive integer solutions to $\frac{4}{n} = \frac{1}{x} + \frac{1}{y} + \frac{1}{z}$ for fixed $n \in \mathbb{Z}$. This is still an open problem!

Sum of three cubes

What $n \in \mathbb{Z}$ can be represented as $X^3 + Y^3 + Z^3 = n$?

- Does $n = 1$ work? Yes:

$$1^3 + 1^3 + (-1)^3 = 1 \qquad 9^3 + 10^3 + (-12)^3 = 1 \qquad \ldots$$

- Does $n = 16$ work? Yes:

$$2^3 + 2^3 + 0^3 = 16 \qquad (-511)^3 + (-1609)^3 + 1626^3 = 16 \qquad \ldots$$

- Do all $n \in \mathbb{Z}$ work? No:

$$4, 5, 13, 14, 22, 23, 31, 32, 40, 41, 49, 50, 58, 59, 67, 68, \ldots \text{ all fail}$$

- Does $n = 42$ work? Yes:

$$12602123297335631^3 + 80435758145817515^3 + (-80538738812075974)^3 = 42$$

  This was only discovered in September 2019!
- Does $n = 114$ work? Nobody knows as of May 2025.

Fermat's last theorem

In 1637, Pierre de Fermat claimed the following theorem.

Conjecture (Fermat's last theorem)

The only integer solutions to $X^n + Y^n = Z^n$ for some $n > 2$ satisfy $XYZ = 0$.

"I have discovered a truly marvelous proof of this, which this margin is too narrow to contain."

In 1995, Andrew Wiles published the first complete proof, which involved *very advanced* 20th century mathematics.

I think Fermat was mistaken.

Why are Diophantine equations so difficult?

Hilbert's tenth problem

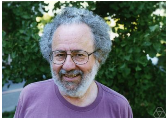
In 1900, David Hilbert published a list of 23 unsolved problems ranging over all areas of mathematics.

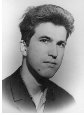
Question (Hilbert)

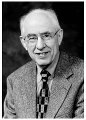
Is there an algorithm to solve *any* Diophantine equation?

Answer (Davis, Matiyasevich, Putnam, Robinson)

No.







We have to get creative!

Linear equations

Observe that an integer solution gives a solution modulo n for any $n \in \mathbb{N}$.

Question

Is there an integer solution to $15X + 21Y = 35$?

Answer

No, because $15X + 21Y \equiv 0 \mod 3$, but $35 \equiv 2 \mod 3$.

Theorem (Bézout's identity)

There is an integer solution to $aX + bY = c$ iff $\gcd(a, b) \mid c$.
Furthermore, there is an algorithm to determine all of its solutions.

Proof.

Refer to MATH0006 Algebra 2. $\square$

Bézout's identity

Question

Can we write down an integer solution to $15X + 21Y = 36$?

Answer

Yes, because 36 is divisible by $\gcd(15, 21) = 3$. By the division algorithm:

$$
\begin{aligned}
21 &= 1 \cdot 15 + 6 && \text{divide 21 by 15} \\
15 &= 2 \cdot 6 + 3 && \text{divide 15 by 6}
\end{aligned}
$$

By reversing the division algorithm:

$$
\begin{aligned}
3 &= 15 - 2 \cdot 6 && \text{substitute 3} \\
&= 15 - 2 \cdot (21 - 1 \cdot 15) && \text{substitute 6} \\
&= 3 \cdot 15 - 2 \cdot 21 && \text{rearrange}
\end{aligned}
$$

Thus $X = \frac{36}{3} \cdot 3 = 36$ and $Y = \frac{36}{3} \cdot -2 = -24$ works!

Quadratic equations

Can we do something similar for quadratic equations $X^2 + Y^2 = b$?

Question

Is there an integer solution to $X^2 + Y^2 = 7^5$?

Answer

No, because $X^2, Y^2 \equiv 0, 1 \mod 4$, but $7^5 \equiv 3 \mod 4$.

Theorem (Sum of two squares theorem)

There is an integer solution to $X^2 + Y^2 = b$ iff b is not divisible by a prime congruent to 3 modulo 4 with odd exponent.

Proof.

Refer to MATH0034 Number Theory. $\square$

Sum of two squares theorem

Question

Can we write down an integer solution to $X^2 + Y^2 = 5^3$?

Answer

Yes, because 5 is a prime congruent to 1 modulo 4. In particular, 5^3 is not divisible by any prime congruent to 3 modulo 4 with odd exponent.
In the ring of Gaussian integers $\mathbb{Z}[i]$:

$$5^3 = X^2 + Y^2 = (X + iY)(X - iY)$$

By *unique factorisation in* $\mathbb{Z}[i]$, write $X \pm iY = (W \pm iZ)^3$. Then:

$$5^3 = ((W + iZ)(W - iZ))^3 = (W^2 + Z^2)^3$$

Now $W = 2$ and $Z = 1$ is an integer solution to $W^2 + Z^2 = 5$. Moreover:

$$X + iY = (W + iZ)^3 = (W^3 - 3WZ^2) + i(3W^2Z - Z^3)$$

Thus $X = W^3 - 3WZ^2 = 2$ and $Y = 3W^2Z - Z^3 = 11$ works!

Number rings

Can we do something similar for quadratic equations $X^2 + aY^2 = b$?

Question

Is there an integer solution to $X^2 + 2Y^2 = 7^2$?

Answer

Consider the number ring $R := \mathbb{Z}[\sqrt{-2}]$. Factorise:

$$7^2 = X^2 + 2Y^2 = (X + \sqrt{-2}Y)(X - \sqrt{-2}Y)$$

By unique factorisation in R, write $X \pm \sqrt{-2}Y = (W \pm \sqrt{-2}Z)^2$. Then:

$$7^2 = ((W + \sqrt{-2}Z)(W - \sqrt{-2}Z))^2 = (W^2 + 2Z^2)^2$$

There are no integer solutions to $W^2 + 2Z^2 = 7$!

Note that $W^2 + 2Z^2 > 0$, so it is easy to rule out solutions.

Failure of unique factorisation

Solving the quadratic equation $X^2 + aY^2 = b$ seems to rely on unique factorisation in the ring $R := \mathbb{Z}[\sqrt{-a}]$, but this might fail.

Examples

- In $R = \mathbb{Z}[\sqrt{-5}]$, we have $6 = 2 \cdot 3 = (1 + \sqrt{-5}) \cdot (1 - \sqrt{-5})$.
- In $R = \mathbb{Z}[\sqrt{10}]$, we have $10 = 2 \cdot 5 = \sqrt{10} \cdot \sqrt{10}$.

The solution is to replace $X + \sqrt{-a}$ with the **ideal**:

$$\langle X + \sqrt{-a} \rangle := \{(X + \sqrt{-a})r : r \in \mathbb{Z}[\sqrt{-a}]\},$$

This has unique factorisation into **prime ideals** if $a \not\equiv 3 \mod 4$.

The failure of unique factorisation into *primes* is measured by the **ideal class group** $\mathrm{Cl}(R)$. For some $\mathrm{Cl}(R)$, a similar argument still works!

For more details, refer to MATH0035 Algebraic Number Theory.

Cyclotomic rings

In the 19th century, Ernst Kummer proved Fermat's last theorem for many exponents using this approach.

Theorem (Kummer)

If p is a regular odd prime, then the only integer solutions to $X^p + Y^p = Z^p$ satisfy $XYZ = 0$.

Call a prime p **regular** if it does not divide the size of $\mathrm{Cl}(R)$.

His idea was to consider the **cyclotomic ring** $R := \mathbb{Z}[\zeta_p]$ for $\zeta_p := e^{\frac{2\pi i}{p}}$, where a similar argument works for the factorisation:

$$Z^p = X^p + Y^p = (X + Y) \cdot (X + \zeta_p Y) \cdot (X + \zeta_p^2 Y) \cdot \dots \cdot (X + \zeta_p^{p-1} Y)$$

Conjecturally, about 61% of all primes are regular.

Rational projective plane

Observe that $X^n + Y^n = Z^n$ is **homogeneous** of degree n.

In particular, this *almost* gives a correspondence:

$$\begin{array}{rcl}
\{(X,Y,Z) \in \mathbb{Z}^3 : X^n + Y^n = Z^n\} & \leftrightsquigarrow & \{(x,y) \in \mathbb{Q}^2 : x^n + y^n = 1\} \\
(X,Y,Z) & \mapsto & (\frac{X}{Z}, \frac{Y}{Z}) \\
(xz, yw, wz) & \leftarrow\!\shortmid & (\frac{x}{w}, \frac{y}{z})
\end{array}$$

This correspondence is not quite bijective:

- Both (X, Y, Z) and $(\lambda X, \lambda Y, \lambda Z)$ map to $(\frac{X}{Z}, \frac{Y}{Z})$.
- Where does $(X, Y, 0)$ map to?

Both of these issues can be fixed by working in the **projective plane**.

- Replace the left hand side with equivalence classes up to scaling.
- Supplement the right hand side with "points at infinity".

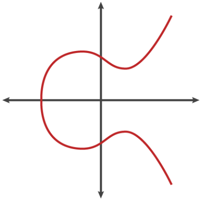
For more details, refer to MATH0076 Algebraic Geometry.

Fermat curves

By working in the projective plane:

$$\left\{ \begin{array}{c} \textit{integer} \text{ solutions of} \\ X^n + Y^n = Z^n \end{array} \right\} \leftrightsquigarrow \left\{ \begin{array}{c} \textit{rational} \text{ solutions of} \\ x^n + y^n = 1 \end{array} \right\}$$

When $n = 3$, the cubic equation $x^3 + y^3 = 1$ defines an object in algebraic geometry called an **elliptic curve**, which lives in the projective plane.



In particular, rational solutions of the equation correspond to **rational points** on the curve.

In fact, the fruit equation

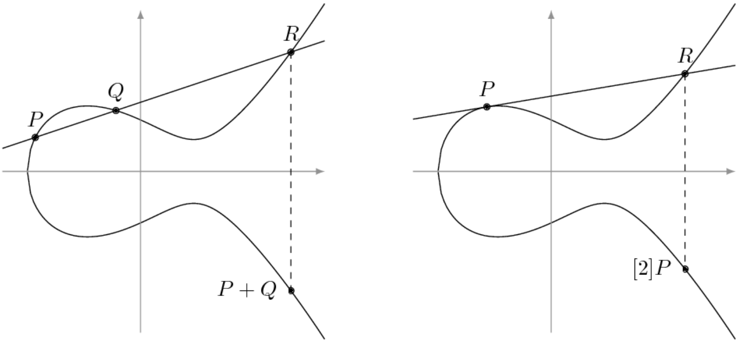
$$X^3 + Y^3 + Z^3 = 3X^2Y + 3XY^2 + 3X^2Z + 3XZ^2 + 3Y^2Z + 3YZ^2 + 5XYZ$$

also defines an elliptic curve!

Elliptic curves

The set of rational points on an elliptic curve forms an abelian group:

$$P + Q + R = 0 \qquad \Longleftrightarrow \qquad P, Q, R \text{ are collinear}$$



This gives a way to generate new rational solutions from old ones!

Cubic equations

Question

Can we write down two rational solutions to $x^3 - y^2 = 4$?

Answer

This defines an elliptic curve, with a rational solution $x = 2$ and $y = 2$. The tangent of $e(x, y) = x^3 - y^2 - 4$ at the rational point $(2, 2)$ is:

$$\frac{\partial e}{\partial x}(2) \cdot (x - 2) + \frac{\partial e}{\partial y}(2) \cdot (y - 2) = 0$$

This simplifies as $y = 3x - 4$, which substitutes into $e(x, y) = 0$ to yield:

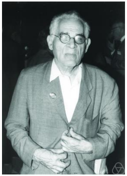
$$0 = x^3 - (3x - 4)^2 - 4 = (x - 2)^2(x - 5)$$

Thus $y = 3(5) - 4 = 11$, so $(5, 11)$ works!

In fact, *adding* the rational point $(2, 2)$ to itself repeatedly generates the *only* infinite family of rational solutions to $x^3 - y^2 = 4$.

Mordell's theorem

In 1922, Louis Mordell classified the abstract group structure of rational points on elliptic curves.



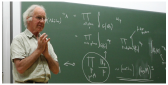
Theorem (Mordell)

The rational points on an elliptic curve can be generated from a finite set of initial rational points.

Associated to an elliptic curve E is a complex-analytic **L-function** $L_E(s)$.

Conjecture (Birch, Swinnerton-Dyer)

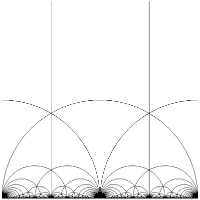
An elliptic curve E has infinitely many rational points iff $L_E(1) = 0$.



For more details, refer to MATH0036 Elliptic Curves.

Modular forms

Andrew Wiles proved Fermat's last theorem by studying properties of general L-functions.



Another object with an associated L-function is a **modular form**, which is a highly symmetric function on the upper half $\mathcal{H}$ of the complex plane.

Conjecture (Shimura, Taniyama, Weil)

Elliptic curves are related to modular forms.

Newforms

The modular forms of interest are the so-called **level-N newforms**.

These are functions $f : \mathcal{H} \to \mathbb{C}$ satisfying the **modular condition**:

$$f\left(\frac{az+b}{cz+d}\right) = (cz+d)^2 \cdot f(z)$$

for any $a, b, c, d \in \mathbb{Z}$ such that $ad - bc = 1$ and $N \mid c$.

Theorem (Valence formula)

For fixed N, there are finitely many level-N newforms.

In fact, there are *no* level-N newforms for:

$$N \in \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 16, 18, 22, 25, 28, 60\}$$

For more details, refer to MATH0104 Modular Forms.

Modularity theorem

Also associated to a modular form f is its **Hecke L-function** $L_f(s)$.

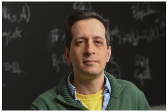
Call an elliptic curve E **modular** if there is a level-N newform f such that $L_E(s) = L_f(s)$ for some N.

Theorem (Wiles)

For squarefree N, all elliptic curves are modular.

Theorem (Breuil, Conrad, Diamond, Taylor)

All elliptic curves are modular.



Fermat's last theorem

Fermat's last theorem can now be deduced from the modularity theorem.

Assume for a contradiction that $X^n + Y^n = Z^n$ has an integer solution not satisfying $XYZ = 0$. Consider the elliptic curve E given by:

$$y^2 = x(x - X^n)(x + Y^n)$$

This is called the **Frey curve** associated to the triple (X, Y, Z).

The modularity theorem says that E corresponds to a level-N newform f.

Theorem (Ribet)

f can be "level-lowered" to a level-2 newform.

There are no level-2 newforms, hence a contradiction!

Formalising Fermat

My PhD supervisor Kevin Buzzard started a massive project to teach the modularity theorem to a computer.

This means *formally* defining all the relevant objects (elliptic curves, modular forms) and *rigorously* verifying all the details of the proof.

`https://imperialcollegelondon.github.io/FLT/`

This is a *huge* amount of work, and we need *all* the help we can get!

To get started, check out MATH0109 Theorem Proving in Lean!

LEAN
THEOREM PROVER