An unusual cubic representation problem

David Kurniadi Angdinata

Wednesday, 16 January 2019

Abstract

Can you find positive whole values for a, b, c satisfying

$$\frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b} = 4?$$

A Bremner and A Macleod recently showed in [2] that such an equation has a solution, but the smallest solution involves three eighty-digit natural numbers,

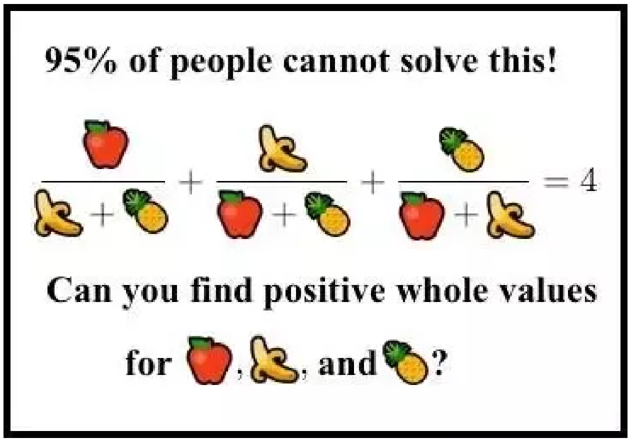
$$\begin{aligned} a &= 154476802108746166441951315019919837485664325669565431700026634898253202035277999, \\ b &= 36875131794129999827197811565225474825492979968971970996283137471637224634055579, \\ c &= 4373612677928697257861252602371390152816537558161613618621437993378423467772036. \end{aligned}$$

This report aims to describe a procedure to verify these values, with an attempt to develop the rich theory of elliptic curves.

1 Introduction

1.1 Background and motivation

The bulk of the theory was based on [3] with some facts taken from [4], while the presentation delivery was based on [1] with some facts taken from [2]. The paper was a slight generalisation of the following problem.


95% of people cannot solve this!
+
+
+
+
+
= 4
Can you find positive whole values
for , , and ?

Now the original meme does not have the additional requirement of positive whole values, which do not really mean anything. If values refer to those of the complex field, then the problem is trivial, by the fundamental theorem of algebra. Alternatively, substituting $b = 1$ and $c = 0$ also gives a real quadratic solution $a = 2 + \sqrt{3}$. The requirement of integers makes the problem slightly less trivial, but an observant eye would detect a trivial solution pretty easily, say $a = 11$, $b = 4$, and $c = -1$. This also generates a family of twelve related solutions, namely by flipping all the signs and permuting the roles of a, b, c.

1.2 Problem restatement

The problem can be written as

$$\frac{a}{b+c}+\frac{b}{a+c}+\frac{c}{a+b}=4, \qquad a,b,c\in\mathbb{Z}, \tag{1}$$

where integers are considered for the sake of an easier argument, and checking $a,b,c>0$ for a valid solution to (1) will be done later. This would have a family of *trivial solutions*

$$\begin{array}{rcl}(a,b,c) & = & (11,4,-1)\\ & = & (-11,-4,1)\\ & = & (1,-4,-11)\\ & = & \ldots\end{array}.$$

Clearing denominators, (1) can be reduced to

$$a^3+b^3+c^3-5abc-3\left(a^2b+ab^2+a^2c+ac^2+b^2c+bc^2\right)=0, \qquad a,b,c\in\mathbb{Z}. \tag{2}$$

There is an issue here needing to be addressed, that is the multiplication of denominators, which would be illegal if they are zero. This would now have several *invalid solutions*

$$\begin{array}{rcl}(a,b,c) & = & (1,-1,0)\\ & = & (-1,1,0)\\ & = & (-1,1,-1)\\ & = & \ldots\end{array}.$$

These extraneous solutions, in a sense complete the space of solutions of (1), and will make the procedure described later much easier. Again, there is now a need to check that $a+b, a+c, b+c>0$ for a valid solution to (1) later. This clearing of denominators is known as a *birational transformation.*

1.3 Dimensionality of solution space

An immediate observation points out that (1) and (2) are *homogeneous*, which is to say that the polynomial defining it has terms of the same degree. For any given solution

$$(a,b,c)=(a_0,b_0,c_0),$$

scaling it by a non-zero rational number λ also produces another solution

$$(a,b,c)=(\lambda a_0,\lambda b_0,\lambda c_0),$$

which can be verified explicitly by collecting terms. These new solutions clearly do not help in understanding the space of solutions to (2), as they do not contribute in the generation of more solutions. It would thus be conducive to consider the equivalence class of solutions

$$(a,b,c)=[a_0,b_0,c_0],$$

defined by the equivalence relation

$$(a_0,b_0,c_0)\sim(a_0',b_0',c_0') \qquad \Longleftrightarrow \qquad (a_0,b_0,c_0)=(\lambda a_0',\lambda b_0',\lambda c_0') \text{ for some } \lambda\in\mathbb{Q}^*.$$

Considering only the equivalence classes where $c\neq 0$, a natural representative can be designated to each of them. For instance, letting $\lambda=\frac{1}{c}$ for any c will force $c=1$ in each equivalence class, cutting down (2) to one of only two variable dimensions a and b, both of which can be arbitrary rational numbers. In other words, (2) only pretends to have three-dimensional when it actually behaves more like two-dimensional, plus the case where $c=0$. To respect this property, (2) can and will be rewritten in the form

$$a^3+b^3+1-5ab-3\left(a^2b+ab^2+a^2+a+b^2+b\right)=0, \qquad a,b\in\mathbb{Q}, \tag{3}$$

which is cubic of two variables. This is known as the *dehomogenisation* of a homogeneous equation.

2 Elliptic curves

2.1 Informal definition

Elliptic curves are a family of homogeneous cubic equations with the aforementioned property, where they have three variables, yet in a sense behave more like two-dimensional. These are usually represented by an equation of the form

$$y^2 = x^3 + Ax + B,$$

where A and B are in some field. There is an additional requirement that $4A^3 + 27B^2 \neq 0$, which is to ensure that the graph of the equation is drawn without any kinks. Elliptic curves are typically one of the simplest non-trivial structures in algebraic geometry, and are used widely in and out of applications.

- For instance, one of the *Millennium Prize Problems* set by the *Clay Mathematics Institute* is the *Birch and Swinnerton-Dyer conjecture*, which questions properties of the *rank* of a rational elliptic curve.
- In Wiles' proof of *Fermat's last theorem*, he deduced a contradiction to *Fermat's equation* having any non-trivial solutions, by asserting the existence of an impossible *modular elliptic curve.*
- There are also extremely efficient algorithms of performing primality testing and integer factorisation involving elliptic curves, both of which considered state-of-the-art techniques.
- Finally, the advent of *elliptic curve cryptography* also allowed for relatively small keys for encryption and decryption schemes, revolutionalising many cryptosystems.

2.2 Formal definition

An *elliptic curve* over a *perfect field* K is a *smooth projective plane algebraic curve* E of *genus one* with a *flex* K*-rational base point* $\mathcal{O}_E$.

- An *algebraic curve* is an *algebraic variety of dimension one*, which is to say that it is the space of solutions to one specified equation.
- Being embedded in the *projective plane* refers to the notion of having three variables but behaving more like two-dimensional, through considering equivalence classes of solutions.
- *Smoothness* is a degeneracy condition that avoids any nodes, cusps, self-intersections, or isolated points.
- *Genus one* basically means a degree three cubic curve via the *genus-degree formula*, but the term is used here as the usual invariant for algebraic curves. It can be defined algebraically via a fundamental theorem in algebraic geometry known as the *Riemann-Roch theorem*, but it was motivated from topology, where the *topological genus* refers to the number of holes in some surface. In this case, a complex elliptic curve can be thought of as a *Riemann surface*, which can be deformed into a torus with one hole via its *fundamental parallelogram.*
- An additional requirement of a base point distinguishes elliptic curves from general cubic curves. Elliptic curves are often considered over the *algebraic closure* of its base field, but K*-rationality* ensures that the base point has coordinates in K. Being a *flex* means that the *intersection multiplicity* of its tangent with the ambient curve is exactly three, which can be informally thought of as slight perturbations of its tangent locally giving exactly three intersections with the ambient curve. This last condition is not usually considered by virtue of the *chord-tangent method*, but it simplifies definitions.
- Finally, *perfectness* of the base field is also a degeneracy condition that requires every *algebraic extension* to be *separable.*

Fortunately, these definitions can be translated, again via the *Riemann-Roch theorem*, to say that any elliptic curve over $\mathbb{Q}$ is of the form

$$\left\{(x, y) \in \mathbb{Q}^2 \mid x^3 + Ax + B\right\} \cup \{\mathcal{O}\},$$

for some $A, B \in \mathbb{Q}$ such that its *discriminant* $4A^3 + 27B^2$ is non-zero, where $\mathcal{O} = [0, 1, 0]$ is a formal symbol representing an invalid solution. This definition also works for any perfect field K of characteristic *char* $(K) \notin \{2, 3\}$.

2.3 Weierstrass equations

With this in mind, (3) can also be transformed in several steps into a standard form of an elliptic curve, via algorithms detailed in [5] and [6]. Coincidentally, it happens that the curve given by (3) is smooth, so that the following transformations are actually *isomorphisms of algebraic varieties*. This does not matter in the grand scheme of things, as otherwise they are birational transformations that would still preserve most of the relevant structure in consideration.

- The initial transformation sends any invalid solution, say $[1,-1,0]$, to the *point at infinity* $\mathcal{O}$, and its tangent, $c = 6a + 6b$, to the *line at infinity* $z = 0$. This could be done by choosing a different point in the tangent, say $[1, 0, 6]$, and sending it to different point in the line at infinity, say $[1, 0, 0]$. A third linearly independent finite point, say $[1, 0, 0]$, can then be sent to a third linearly independent finite point, say $[0, 0, 1]$, to allow the *projective transformation* matrix to be invertible. In other words, the change of variables is induced by a matrix in $PGL_3(\mathbb{Q})$. This produces the cubic curve
$$\left\{(x,y) \in \mathbb{Q}^2 \mid 6y^2 + 6xy + 6y = -91x^3 + 141x^2 + 15x - 1\right\} \cup \{\mathcal{O}\}.$$
This is followed by an appropriate transformation $z \mapsto -\frac{6}{91}z$ to give a *long Weierstrass equation*
$$\left\{(x,y) \in \mathbb{Q}^2 \mid y^2 + xy - \tfrac{91}{6}y = x^3 + \tfrac{47}{2}x^2 - \tfrac{455}{12}x - \tfrac{8281}{216}\right\} \cup \{\mathcal{O}\}.$$
This elliptic curve is in fact a *minimal model* with smallest possible discriminant.

- The next transformation traces a simplified version of *Laska's algorithm* for rational elliptic curves, which imposes several conditions on the coefficients of the equation after applying isomorphisms. This produces a unique *restricted* and *integral* minimal model
$$\left\{(x,y) \in \mathbb{Q}^2 \mid y^2 + xy + y = x^3 - 234x + 1352\right\} \cup \{\mathcal{O}\},$$
which also has small integral coefficients. This particular form can also be looked up in the *L-functions and modular forms database* under the label $910.a4$.

- Another transformation, which works for fields of characteristic not two, completes the square with the transformation $y \mapsto -\frac{1}{2}x + y - \frac{1}{2}z$ to give a *medium Weierstrass equation*
$$\left\{(x,y) \in \mathbb{Q}^2 \mid y^2 = x^3 + \tfrac{1}{4}x^2 - \tfrac{467}{2}x + \tfrac{5409}{4}\right\} \cup \{\mathcal{O}\},$$
which is a minimal model but does not have integral coefficients. Clearing denominators,
$$\left\{(x,y) \in \mathbb{Q}^2 \mid y^2 = x^3 + 109x^2 + 224x\right\} \cup \{\mathcal{O}\}$$
is an integral model but will not have minimal discriminant anymore.

- Finally, another transformation, which works for fields of characteristic not two or three, completes the cube with the transformation $x \mapsto x - \frac{1}{12}z$ to give a *short Weierstrass equation*
$$\left\{(x,y) \in \mathbb{Q}^2 \mid y^2 = x^3 - \tfrac{11209}{48}x + \tfrac{1185157}{864}\right\} \cup \{\mathcal{O}\},$$
which is a minimal model but does not have integral coefficients. Clearing denominators again,
$$\left\{(x,y) \in \mathbb{Q}^2 \mid y^2 = x^3 - 302643x + 63998478\right\} \cup \{\mathcal{O}\},$$
is an integral model but will not have minimal discriminant anymore.

All of these transformations preserve the *j-invariant* of the rational elliptic curve, a quantity that characterises the isomorphism classes of elliptic curves over algebraically closed fields. The focus now shifts to the simplest model of the same elliptic curve, namely the equation representing the last curve,

$$y^2 = x^3 + Ax + B, \qquad A = -302643, \qquad B = 63998478. \tag{4}$$

The overall transformation can be recorded in one change of variables

$$\begin{cases} a = \frac{1}{72}x + \frac{1}{216}y - \frac{277}{24} \\ b = \frac{1}{72}x - \frac{1}{216}y - \frac{277}{24} \\ c = \frac{1}{6}x - \frac{95}{2} \end{cases} \qquad \begin{cases} x = \dfrac{1710a + 1710b - 831c}{6a + 6b - c} \\ y = \dfrac{-9828a + 9828b}{6a + 6b - c} \end{cases},$$

which are clearly inverses.

2.4 Group law

An elliptic curve is also an *abelian variety of dimension one*, which makes it an *algebraic group*. In other words, an abelian group could be defined over points on the elliptic curve. The following definition for (4) will suffice for most purposes, but more complicated definitions exist for fields of characteristic two and three.

- The identity point is $\mathcal{O} \in E$.
- The inverse of a point is obtained by reflecting the point about the x-axis, where $-(x, y) = (x, -y)$.
- The sum of two points is obtained by inverting the third point of intersection between the curve and the line joining the two points, where

$$P + Q = \begin{cases} S & P = (x, y)\,,\ Q = (x', y')\,,\ x \neq x' \\ R & P = Q = (x, y)\,,\ y \neq 0 \\ P & Q = \mathcal{O} \\ \mathcal{O} & P = Q = (x, 0) \end{cases} \qquad ,$$

$$S = \left(\frac{\left(A+xx'\right)\left(x+x'\right)+2\left(B-yy'\right)}{\left(x-x'\right)^2}, \frac{\left(Ay'-x'^2y\right)\left(3x+x'\right)+\left(x^2y'-Ay\right)\left(x+3x'\right)-4B\left(y-y'\right)}{\left(x-x'\right)^3} \right),$$

$$R = \left(\frac{x^4-2Ax^2-8Bx+A^2}{4y^2}, \frac{x^6+5Ax^4+20Bx^3-5A^2x^2-4ABx-A^3-8B^2}{8y^3} \right).$$

The explicit equations here describe the four cases involved and show that it is possible to program the computations despite its geometric definition. For it to be well-defined, there are several things to check.

- Firstly, point addition needs to be well-defined over the base field. *Bézout's theorem* states that two projective algebraic curves C and D over an algebraically closed field intersect at $\deg C \deg D$ points counted with intersection multiplicity. It follows that the third point of intersection in the above definition is well-defined over the algebraic closure of the base field. Since the explicit equations are all rational functions, this point would also be well-defined over the base field.
- Next, commutativity of point addition holds by symmetry of its definition.
- Finally, associativity of point addition involves more complications and is practically the only non-trivial axiom to check. There are at least three ways to verify this fact.
  - An extremely tedious approach considers all possible cases in the explicit equations to be verified sequentially. This could be done with a computer algebra system that might take several hours.
  - A more geometric approach considers ten points in *general position* without loss of generality. The *Cayley-Bacharach* theorem states that three projective algebraic cubic curves over an algebraically closed field that coincide at eight points counted with intersection multiplicity would coincide at the ninth point given by Bézout's theorem. Now consider the following diagram, where $P, Q, R, S \in E$ are points in general position.

$$\begin{tikzcd} Q \arrow[r, dash] \arrow[d, dash] & R \arrow[r, dash] \arrow[d, dash] & -(Q+R) \arrow[d, dash] \\ P \arrow[r, dash] \arrow[d, dash] & S \arrow[r, dash] \arrow[d, dash] & Q+R \arrow[d, dash] \\ -(P+Q) \arrow[r, dash] & P+Q \arrow[r, dash] & \mathcal{O} \end{tikzcd}$$

    The union of the three horizontal lines and the union of the three vertical lines give two cubic curves, both of which passes through eight points in the ambient elliptic curve. It follows that $-(P+(Q+R)) = S = -((P+Q)+R)$, proving associativity.
  - A more algebraic approach considers the *Picard group* $Pic^0(E)$ of an elliptic curve E, which is in set-theoretic bijection with E via the *summation map*. This also induces an isomorphism of abelian groups that coincides with the geometric definition, which has associativity as a consequence.

This completes the proof of the group law of an elliptic curve.

3 Procedure

3.1 Generation algorithm

The following algorithm then details the procedure to generate new solutions from old solutions.

- Choose an initial solution for

$$a^3 + b^3 + c^3 - 5abc - 3\left(a^2b + ab^2 + a^2c + ac^2 + b^2c + bc^2\right) = 0, \qquad a, b, c \in \mathbb{Z}.$$

- Apply the change of variables

$$\begin{cases} x = \dfrac{1710a + 1710b - 831c}{6a + 6b - c} \\ y = \dfrac{-9828a + 9828b}{6a + 6b - c} \end{cases}.$$

- Compute multiples of point in

$$y^2 = x^3 + Ax + B, \qquad (x, y) \in \mathbb{Q}^2.$$

- Apply the change of variables

$$\begin{cases} a = \frac{1}{72}x + \frac{1}{216}y - \frac{277}{24} \\ b = \frac{1}{72}x - \frac{1}{216}y - \frac{277}{24} \\ c = \frac{1}{6}x - \frac{95}{2} \end{cases}.$$

- Terminate or repeat.

3.2 Computation failure

It is entirely possible that the algorithm fails for a few reasons. The solutions obtained may not have positive whole values, and running the procedure over and over again may never result in any such solutions. Such a failure occurs when considering the multiples of any invalid solution, such as

$$(a, b, c) = (-1, 1, -1).$$

Applying the change of variables gives

$$(x, y) = (831, 19656),$$

so consider the cyclic subgroup generated by this point.

$$\begin{aligned} 1\,(x, y) &= (831, 19656), \\ 2\,(x, y) &= (363, 1404), \\ 3\,(x, y) &= (327, 0), \\ 4\,(x, y) &= (363, -1404), \\ 5\,(x, y) &= (831, -19656), \\ 6\,(x, y) &= \mathcal{O}. \end{aligned}$$

The algorithm hits a *torsion subgroup* of order six, which will never result in any valid solutions, no matter the number of repetitions.

3.3 Computation success

Fortunately, this is not always the case, provided no invalid solutions are chosen in the beginning. Considering any trivial solution will eventually result in success, such as for

$$(a, b, c) = (11, 4, -1)\,.$$

Applying the change of variables gives

$$(x, y) = (291, -756)\,,$$

so again consider the cyclic subgroup generated by this point. Duplicating the point once gives

$$2\,(x, y) = \left(\tfrac{22107}{49}, -\tfrac{1506492}{343}\right),$$

where reverting the change of variables would give

$$(a, b, c) = (-8784, 5165, 9499)\,.$$

This is in fact a point of infinite order by the *Nagell-Lutz theorem*, so continuing would be conducive. Triplicating the point once gives

$$3\,(x, y) = \left(-\tfrac{2694138}{11881}, -\tfrac{14243306490}{1295029}\right),$$

where reverting the change of variables would give

$$(a, b, c) = (679733219, -375326521, 883659076)\,.$$

Triplicating this point once more gives

$$\begin{aligned} 9\,(x, y) = (&\tfrac{3823387580080160076063605209061052603963389916327719142}{13514400292716288512070907945002943352692578000406921} \\ &-\tfrac{1587622549247318249299172296638373895912313166958011719500537215259315694916502670}{1571068668597978434556364707291896268838086945430031322196754390420280407346469}), \end{aligned}$$

where reverting the change of variables would finally give the required solutions. This last point is said to have a large *height*, which measures the complexity of the coordinates from the numerators and denominators.

3.4 Miscellaneous considerations

The general equation being considered in the original paper was

$$\frac{a}{b+c} + \frac{b}{a+c} + \frac{c}{a+b} = N, \qquad a, b, c \in \mathbb{N}^*, \qquad N \in \mathbb{Z},$$

which considers a similar procedure as aforementioned. The torsion subgroup and rank of the elliptic curves *birationally equivalent* to these cubic curves can be computed systematically, such that, in conjunction with *Mordell's theorem*, they are all isomorphic as abelian groups to the finitely generated abelian group

$$\mathbb{Z}^r \oplus \begin{cases} \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/6\mathbb{Z} & N = 2 \\ \mathbb{Z}/6\mathbb{Z} & \text{otherwise} \end{cases}, \qquad r \in \mathbb{N}^*.$$

The curve for $N = 4$ has $r = 1$, which ensures the existence of a point of infinite order, hence giving a termination point for the algorithm. The termination point detailed above does indeed give the smallest solution, which can be proven with the theory of heights. It is entirely possible that $r = 0$, or that $r > 0$ but yet no positive whole solutions exist no matter the number of repetitions. Such is the case for all odd N, which can be proven using ad-hoc congruences. For even N, it is also possible that no solutions exist, such is the case for $N = 8$. When there are solutions, they may take arbitrarily large point multiples to obtain, and the resulting solution may have arbitrarily large positive values. In the case for $N = 178$, the corresponding initial point has to be multiplied over a hundred thousand times, and the resulting solution has over four hundred million digits.

4 References

[1] A Amit's Quora answer on *How do you find the positive integer solutions to*

$$\frac{x}{y+z}+\frac{y}{z+x}+\frac{z}{x+y}=4?$$

[2] A Bremner and A Macleod's 2014 paper on *An unusual cubic representation problem*

[3] J Silverman's 1986 book *The arithmetic of elliptic curves*

[4] R Hartshorne's 1977 book on *Algebraic geometry*

[5] N Duif's 2011 implementation on *Transforming a general cubic elliptic curve equation to Weierstrass form*

[6] M Laska's 1982 paper on *An algorithm for finding a minimal weierstrass equation for an elliptic curve*