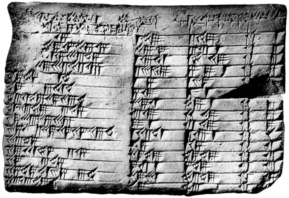
Sums of three cubes

Andrew Sutherland
Massachusetts Institute of Technology

(joint work with Andrew Booker, University of Bristol)



Number Theory Web Seminar
May 7, 2020

A Diophantine problem

Many of the oldest problems in number theory involve equations of the form

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

where P is a polynomial with integer coefficients and k is a fixed integer.
We seek integer solutions in $x_1, \ldots, x_n$. Some notable examples:

- $x^2 + y^2 = z^2$ [Babylonians?]
  $(119, 120, 169), (4601, 4800, 6649), \ldots$ [Babylonians ∼1800 BCE]
- $x^2 - 4729494y^2 = 1$ [Archimedes 251 BCE]
  $776 \ldots 800$ cattle [Amthor 1880, German-Williams-Zarnke, 1965]
- $x^3 + y^3 = z^3$ [Fermat 1637]
  No solutions with $xyz \neq 0$. [Euler 1753]
- $v^5 + w^5 + x^5 + y^5 = z^5$ [Euler 1769]
  $(27, 84, 110, 133, 144)$ [Lander-Parkin 1966]
- $w^4 + x^4 + y^4 = z^4$ [Euler 1769]
  $(2682440, 15365639, 18796760, 20615673)$ [Elkies 1986]

Algorithm to find (or determine existence of) solutions?

Q: Is there an algorithm that can answer all such questions? [Hilbert 1900]
A: No! [Davis, Robinson, Davis-Putnam, Robinson, Matiyasevich 1970]

What if we restrict the degree of the polynomial P? Maybe that will help!

Q: How about degree one? [Euclid $\sim$250 BCE, Diophantus $\sim$250]
A: Yes! [Euclid $\sim$250 BCE, Brahmagupta 628]

Q: How about degree two? [Babylonians, Diophantus, Hilbert 1900]
A: Yes! [Babylonians, Diophantus, Fermat, Euler, Legendre, Lagrange] [Siegel 1972]

Q: How about degree three? [Waring 1770]

A: We have no idea.

Sums of two cubes

Q: Which primes are sums of two cubes?
A: The prime 2 and primes of the form $3x^2 - 3x + 1$ for some integer x.

This list of primes begins $2, 7, 19, 37, 61, 127, 271, 331, 397, 547, 631, 919, \ldots$
We believe this list to be infinite, but this is not known.

Proof:

- $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$, so either $x + y = 1$ or $x^2 - xy + y^2 = 1$.
- If $x^2 - xy + y^2 = 1$ then $x = y = 1$, in which case $x^3 + y^3 = 2$.
- If $x + y = 1$ then $x^2 - x(1 - x) + (1 - x)^2 = 3x^2 - 3x + 1$ must be prime.

There are infinitely many primes of the form $x^3 + 2y^3$ [Heath-Brown 2001].
This implies that infinitely many primes are the sum of three cubes.

Digression

What happens if we allow rational cubes? For example

$$13 = \left(\frac{2}{3}\right)^3 + \left(\frac{7}{3}\right)^3$$

is a sum of rational cubes, but 13 is not a sum of integer cubes.

This amounts to finding rational points on the elliptic curve $x^3 + y^3 = n$, which can also be written as $E_n \colon Y^2 = X^3 - 432n^2$.

We know that $E(\mathbb{Q}) \simeq T \oplus \mathbb{Z}^r$, where $\#T \leq 16$ and $r := r(E)$ is the *rank* of E. Under the Birch and Swinnerton-Dyer conjecture, $r(E) > 0$ if and only if

$$L_E(s) := \prod_p (1 - a_p p^{-s} + \chi(p) p^{1-2s})^{-1}$$

has a zero at $s = 1$ (here $a_p := p + 1 - \#E(\mathbb{F}_p)$ and $\chi(p) = 1$ for $p \nmid \Delta(E)$).
If $p \equiv 4, 7, 8 \bmod 9$ then $r(E_p) > 0$ and if $p \equiv 2, 5 \bmod 9$ then $r(E_p) = 0$.[1]
The case $p \equiv 1 \bmod 9$ is more complicated, but fairly well understood.

[1] Here we assume BSD, but see [Kriz20] for recent progress on removing this assumption.

Sums of two cubes

Let us now consider an arbitrary integer k. If we have

$$k = x^3 + y^3 = (x+y)(x^2 - xy + y^2),$$

then we can write $k = rs$ with $r = x + y$ and $s = x^2 - xy + y^2$.
If we now put $y = r - x$, we obtain the quadratic equation

$$s = 3x^2 - 3rx + r^2,$$

whose integer solutions we can find using the quadratic formula.

This yields an algorithm to determine all integer solutions to $x^3 + y^3 = k$:

- Factor the integer k.
- Use this factorization to enumerate all $r, s \in \mathbb{Z}$ for which $k = rs$.
- If $t := \sqrt{12s - 3r^2} \in \mathbb{Z}$ then output $x = (3r + t)/6$ and $y = (3r - t)/6$.

Example:

For $k = 1729 = 19 \cdot 91$ we find $t = 3$, yielding $x = 10$ and $y = 9$.
For $k = 1729 = 13 \cdot 133$ we find $t = 33$, yielding $x = 12$ and $y = 1$.

Sums of four or more cubes

Every integer has infinitely many representations as the sum of five cubes. This follows from the identity

$$6m = (m+1)^3 + (m-1)^3 - m^3 - m^3.$$

If we write $k = 6a + r$, then $r^3 \equiv r \bmod 6$ and, we can apply this identity to $m = f(n) := (k - (6n+r)^3)/6$ for any integer n, yielding the parameterization

$$k = (6n+r)^3 + (f(n)+1)^3 + (f(n)-1)^3 - f(n)^3 - f(n)^3.$$

A more complicated collection of similar identities (and extra work in one particularly annoying case) shows that all $k \not\equiv \pm 4 \bmod 9$ can be represented as a sum of four cubes in infinitely many ways [Demjanenko 1966].

It is conjectured that in fact every integer k has infinitely many representations as a sum of four cubes [Sierpinski], but the case $k \equiv \pm 4 \bmod 9$ remains open.

Sums of three cubes

Not every integer is the sum of three cubes. Indeed, if $x^3 + y^3 + z^3 = k$ then

$$x^3 + y^3 + z^3 \equiv k \bmod 9$$

The cubes modulo 9 are $0, \pm 1$; there is no way to write ± 4 as a sum of three. This rules out all $k \equiv \pm 4 \bmod 9$, including $4, 5, 13, 14, 22, 23, 31, 32, \ldots$

There are infinitely many ways to write $k = 0, 1, 2$ as sums of three cubes. For all $n \in \mathbb{Z}$ we have

$$\begin{aligned} n^3 + (-n)^3 + 0^3 &= 0, \\ (9n^4)^3 + (3n - 9n^4)^3 + (1 - 9n^3)^3 &= 1, \\ (1 + 6n^3)^3 + (1 - 6n^3)^3 + (-6n^2)^3 &= 2. \end{aligned}$$

Multiplying by m^3 yields similar parameterizations for k of the form m^3 or $2m^3$. For $k \not\equiv \pm 4 \bmod 9$ not of the form m^3 or $2m^3$ the question is completely open.

Remark 1: The parameterizations above are not exhaustive [Payne,Vaserstein 1992].
Remark 2: Every $k \in \mathbb{Z}$ is the sum of three rational cubes [Ryley 1825].

Mordell's challenge

There are two easy ways to write 3 as a sum of three cubes:

$$1^3 + 1^3 + 1^3 = 3,$$
$$(-5)^3 + 4^3 + 4^3 = 3.$$

In a 1953 paper Mordell famously wrote:

> *I do not know anything about the integer solutions of* $x^3 + y^3 + z^3 = 3$ *beyond the existence of. . . it must be very difficult indeed to find out anything about any other solutions.*

This remark sparked a 65 year search for additional solutions.

None were found, but researchers did find solutions for many other values of k in the process of trying to answer Mordell's challenge.

20th century timeline for sums of three cubes

Progress on $x^3 + y^3 + z^3 = k$ with $k > 0$ and $|x|, |y|, |z| \leq N$:

- 1908 Werebrusov finds a parametric solution for $k = 2$.
- 1936 Mahler finds a parametric solution for $k = 1$.
- 1942 Mordell proves any other parameterization has degree at least five (likely none exist).
- 1953 Mordell asks about $k = 3$.
- 1955 Miller, Woollett check $k \leq 100$, $N = 3200$, solve all but nine $k \leq 100$.
- 1963 Gardiner, Lazarus, Stein: $k \leq 1000$, $N = 2^{16}$, crack $k = 87$, all but seventy $k \leq 1000$.
- 1992 Heath-Brown, Lioen, te Riele crack $k = 39$.
- 1992 Heath-Brown conjectures infinity of solutions for all $k \not\equiv \pm 4 \bmod 9$.
- 1994 Koyama checks $k \leq 1000$, $N = 2^{21} - 1$, finds 16 new solutions.
- 1994 Koyama checks $k \leq 1000$, $N = 3414387$, finds 2 new solutions.
- 1994 Conn, Vaserstein crack $k = 84$.
- 1995 Jagy cracks $k = 478$.
- 1995 Bremner cracks $k = 75$ and $k = 768$.
- 1995 Lukes cracks $k = 110$, $k = 435$, and $k = 478$.
- 1996 Elkies checks $k \leq 1000$, $N = 10^7$ finding several new solutions (follow up by Bernstein).
- 1997 Koyama, Tsuruoka, Sekigawa check $k \leq 1000$, $N = 2 \cdot 10^7$ finding five new solutions.
- 1999-2000 Bernstein checks $k \leq 1000$, $N \geq 2 \cdot 10^9$, cracks $k = 30$ and ten other $k \leq 1000$.
- 1999-2000 Beck, Pine, Tarrant, Yarbrough Jensen also crack $k = 30$, and $k = 52$.

At the end of the millennium, only 33, 42, 74 and twenty-four other $k \leq 1000$ were open.

Poonen's challenge

To add further fuel to the fire, Bjorn Poonen opened his AMS Notices article "Undecidability in number theory" with the following paragraph:

> *Does the equation* $x^3 + y^3 + z^3 = 29$ *have a solution in integers? Yes:* $(3, 1, 1)$*, for instance. How about* $x^3 + y^3 + z^3 = 30$*? Again yes, although this was not known until 1999: the smallest solution is* $(283059965, -2218888517, 2220422932)$*. And how about 33? This is an unsolved problem.*

This spurred another 10 years of searches, with 33 nearly as desirable as 3.

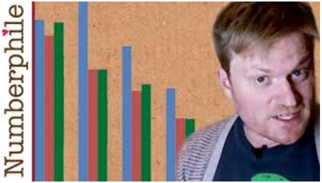
Elsenhans and Jahnel searched to $N = 10^{14}$ cracking nine more $k \leq 1000$. Huisman pushed on to $N = 10^{15}$ and cracked $k = 74$ in 2016.

In spring 2019 Andrew Booker finally answered Poonen's challenge with

$$8866128975287528^3 - 8778405442862239^3 - 2736111468807040^3 = 33,$$

leaving 42 as the only unresolved case below 100 (and ten other $k \leq 1000$). But still no progress on Mordell's challenge, even with $N = 10^{16}$ [Booker].

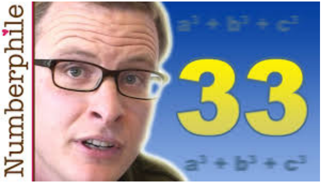
Popularization



Brady Haran

74 is Cracked!

(Sander Huisman)



The uncracked problem with 33

(Tim Browning)

42 is the new 33

(Andrew Booker)

Mathematician solves 64-year-old ‘Diophantine puzzle’ (Newsweek)

“. . . the mathematician [Andrew Booker] is now working with [S] of MIT in an attempt to find the solution for the final unsolved number below a hundred: 42.”

The significance of 42 [Douglas Adams]

“O Deep Thought computer. . . We want you to tell us....The Answer.”
“The Answer to what?” asked Deep Thought.
“Life!” urged Fook. “The Universe!” said Lunkwill.
“Everything!” they said in chorus.

Deep Thought paused for a moment’s reflection. . .
“There is an answer. But, I’ll have to think about it.”

seven and a half million years pass

“Good Morning,” said Deep Thought at last. “Er...good morning, O Deep Thought” said Loonquawl nervously, “do you have...”
“An Answer for you?” interrupted Deep Thought. “Yes, I have.”

“Forty-two,” said Deep Thought, with infinite majesty and calm.

Deep Thought designs Earth to compute the Ultimate Question whose answer is 42.
Mice (the most intelligent beings on earth) take charge of this ten million year project.
Unfortunately, Earth is destroyed by the Vogons before the project is completed.

Search algorithms

We seek solutions to $x^3 + y^3 + z^3 = k$ for some fixed k (such as $k = 3$ or $k = 42$). How long does it take to check all $x, y, z \in \mathbb{Z}$ with $\max(|x|, |y|, |z|) \leq N$?

1. Naive brute force: $O(N^3)$ arithmetic operations.
2. Less naive brute force (is $x^3 + y^3 - k$ a cube?): $O(N^{2+o(1)})$.
3. Apply sum of two cubes algorithm to $k - z^3$: $O(N^{1+o(1)})$ (expected).

None of these is fast enough to go past $N = 10^{16}$ in a reasonable time frame.

We instead follow an approach suggested by Heath-Brown, Lioen, and te Riele, that seeks solutions for a fixed value of k (in contrast to Elkies' approach, which seeks solutions to $x^3 + y^3 + z^3 \leq b$ with b small).

With suitable optimizations this gives a heuristic complexity of $O(N(\log\log N)^{1+o(1)})$ arithmetic operations (in our range of interest these are 64-bit or 128-bit word operations using 1-3 clock cycles).

The setup and the strategy

Assume $x^3 + y^3 + z^3 = k > 0$, $|x| > |y| > |z| \geq \sqrt{k}$, $k \equiv \pm 3 \bmod 9$ cube free.

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

Define $d := |x + y|$ so that z is a cube root of k modulo d.

$$\{x, y\} = \left\{ \frac{\operatorname{sgn}(k - z^3)}{2} \left(d \pm \sqrt{\frac{4|k - z^3| - d^3}{3d}} \right) \right\},$$

Thus d, z determine x, y, and one finds that $d < \alpha|z|$, where $\alpha := \sqrt[3]{2} - 1 \approx 0.26$. One also finds that $3 \nmid d$ and $\operatorname{sgn}(z)$ is determined by $d \bmod 3$ and $k \bmod 9$.

Given N, our strategy is to enumerate all $d \in \mathbb{Z} \cap (0, \alpha N)$ coprime to 3, and for each d enumerate all $z \in \mathbb{Z}$ satisfying $z^3 \equiv k \bmod d$ with $|z| \leq N$ such that

$$3d(4\operatorname{sgn}(z)(z^3 - k) - d^3) = \square \qquad (1)$$

is a square. Every such (d, z) yields a solution (x, y, z), and we will find all solutions satisfying our assumptions with $|z| \leq N$ (even if $|x|, |y| > N$).

Elliptic curves again

With k fixed and d as above, if we put $B_d := -2(6d)^3(d^3 + 4\operatorname{sgn}(z)k)$, then the solutions to (1) are precisely the affine integral points on the elliptic curve

$$E_d\colon Y^2 = X^3 + B_d.$$

For small values of d it may be feasible to determine the integral points on E_d.

Doing so addresses infinitely many possibilities for z in one fell swoop.
But this is typically feasible only when d is quite small, say $d \leq 40$.
($d \leq 100$ using Bremner's 3-isogeny trick, $d \leq 20,000$ under GRH).

The problem of finding integral representations for k as a sum of three cubes can thus be reduced to the problem of finding integral points on a one-parameter family of CM elliptic curves over $\mathbb{Q}$.

This does not make the problem any easier, it only highlights the challenge.
Finding the integral points on even a single E_d with $d \approx 10^{16}$ is almost never feasible, and there are 10^{16} such E_d to consider.

Complexity obstacles

problem: To compute cube roots of k mod d we need the factorization of d.

solution: Enumerate d combinatorially, as a product of prime powers along with cube roots of k mod d (also lets us efficiently skip useless d).

problem: There are $\Omega(N \log N)$ pairs (d, z) we potentially need to consider.

solution: For $d \leq N^{3/4}$ (say) we sieve arithmetic progressions of z mod d using small auxiliary primes $p \nmid d$. Each p reduces the number of pairs (d, z) by a factor of about 2, and $O(\log \log N)$ such p suffice.

We don't literally sieve, we use the CRT to lift progressions modulo d to progressions modulo pd, but only use the lifts that yield solutions modulo p (about half, on average, and we can select p that give less than half).

With this approach the total number of pairs (d, z) with $d \leq N^{3/4}$ we need to consider becomes $o(N)$, and for $d > N^{3/4}$ we heuristically expect $O(N)$.

Cube roots mod primes, prime powers, composites

For $p \equiv 2 \bmod 3$ cubing is 1-to-1, since 3 is invertible modulo $p-1$, and

$$z \equiv k^{(2p-1)/3} \bmod p \quad \Longleftrightarrow \quad z^3 \equiv k^{2p-1} \equiv k \bmod p$$

Compute $k^{(2p-1)/3} \bmod p$ using $O(\log p)$ multiplications (square-and-multiply).

For $p \equiv 1 \bmod 3$ cubing is 3-to-1. Let $p = 3^w m + 1$ with $3 \nmid m$, and $b \equiv k^m \bmod p$.
Compute $b^3, b^{3^2}, \ldots b^{3^v} \bmod p$ until $b^{3^v} \equiv 1 \bmod p$. Cube roots exist iff $v < w$.

Pick random x until $a := x^m \bmod p$ has order 3^w, then compute $n := \log_a b$,
so that $a^n \equiv b \bmod p$. Then $3|n$, and $a^{n/3}$ is a cube root of b modulo p.

Now chose $e \in \{1, 2\}$ so that $3|(p - em)$. Then $a^{3ne} \equiv b^e \equiv k^{em} \bmod p$ and

$$z \equiv a^{ne} k^{(p-em)/3} \quad \Longrightarrow \quad z^3 \equiv a^{3ne} k^{p-em} \equiv k^p \equiv k \bmod p.$$

The other two cube roots are $\zeta_3 z$ and $\zeta_3^2 z$, where $\zeta_3 := a^{3^{w-1}}$.

Use Hensel lifting for prime powers, CRT for products of prime powers.

CRT sieving

For $k = 33$ and $d = 5$ we must have $z \equiv 2 \bmod d$ and $\operatorname{sgn}(z) = +1$.
But we also know $z \equiv k + d \equiv 0 \bmod 2$, and only $z \equiv 0 \bmod 7$ satisfies

$$3d(4\operatorname{sgn}(z)(z^3 - k) - d^3) = \square \bmod 7.$$

p	modulus	residue classes	$\lvert z\rvert \le 10^{16}$ to check
	5	1	2.0×10^{15}
2	10	1	1.0×10^{15}
7	70	1	1.4×10^{14}
13	910	3	3.3×10^{13}
17	15470	27	1.7×10^{13}
23	355810	324	9.1×10^{12}
29	10318490	4860	4.7×10^{12}
43	443695070	92340	2.1×10^{12}
67	29727569690	2493180	8.4×10^{11}
103	3061939678070	107206740	3.5×10^{11}

Cubic reciprocity constraints allow only 14 residue classes modulo $27k = 891$, and this further reduces the number of z to check by another factor of 63.6.

This leaves only 5.5×10^9 values of z to check, which takes about a minute.

Implementation

- Heavily optimized C code using GCC intrinsics to access particular features of the Intel instruction set (including 80-bit long doubles).
- Batch modular inversions (a la Montgomery), Montgomery and Barrett modular reduction (Montgomery for exponentiation, Barrett for CRT).
- `smalljac/ffpoly` finite field implementation to compute cube roots modulo primes and lift them modulo prime powers.
- `primesieve` library to enumerate primes [Walisch].
- `gmp` multiprecision library for testing solutions over $\mathbb{Z}$, but only after passing precomputed filters (bitmap checks) modulo auxiliary primes.
- `cygwin` to create a Microsoft Windows compatible executable so we can take full advantage of Charity Engine's crowd-sourced compute grid.
- Parallelization is achieved by partitioning d by largest prime factor. We split the work into jobs that only take a few hours (millions of jobs).
- We used two cores on each compute node and try to keep the memory footprint under 1GB per core (share all precomputed tables).

The conjecture of Heath-Brown

Heath-Brown's conjecture uses products of local densities to estimate

$$R_k(N_1, N_2) := \#\{(x,y,z) \in \mathbb{Z}^3 : x^3 + y^3 + z^3 = k,\ N_1 \le \max(|x|,|y|,|z|) \le N_2\}$$

as $N \to \infty$. Assume k is cube free, and p prime and $n \ge 1$ define

$$N(p^n) := \#\{(x,y,z) \bmod p^n : x^3 + y^3 + z^3 \equiv k \bmod p^n\},$$

$$\sigma_p := \frac{N(p)}{p^2}\ (p \ne 3), \quad \sigma_3 = \frac{N(9)}{81}, \quad \sigma_\infty := 6\int_{N_1}^{N_2}\int_0^z \frac{dy}{3(z^3 - y^3)^{2/3}} dz = c\log\frac{N_2}{N_1},$$

where $c = \frac{2\Gamma(1/3)^2}{3\Gamma(2/3)} \approx 3.5332$. For $N_2 \gg N_1 \gg 0$ we should then expect

$$R_k(N_1, N_2) \sim \prod_{p \le \infty} \sigma_p = \delta_k \log\frac{N_2}{N_1},$$

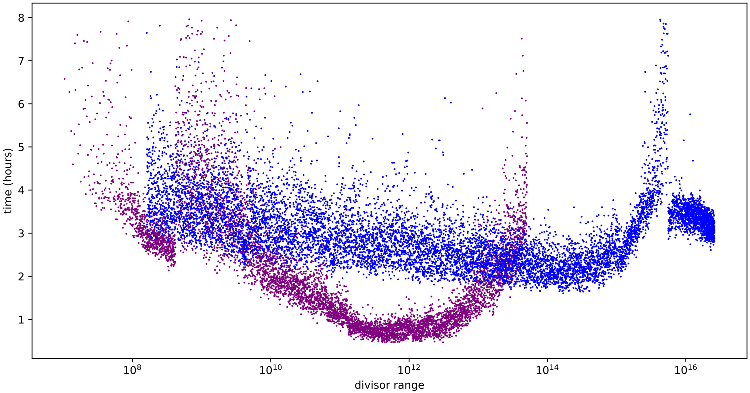
where δ_k is an explicit constant that depends only on k.

For $k = 3$ we should expect one solution $|x| > |y| > |z|$ in $[N, \alpha N]$ for $\alpha \approx 10^7$

Heath-Brown vs Huisman for $3 \le k < 100$

			$N = 10^5$		$N = 10^{10}$		$N = 10^{15}$	
k	$\delta_k/6$	N_0	expect	actual	expect	actual	expect	actual
3	0.061	12969857	0.7	2	1.4	2	2.1	2
93	0.072	1185438	0.8	2	1.6	3	2.5	3
74	0.086	106692	1.0	0	2.0	0	3.0	1
33	0.089	77368	1.0	0	2.0	0	3.1	0
30	0.090	68020	1.0	0	2.1	1	3.1	3
39	0.090	68358	1.0	0	2.1	1	3.1	1
12	0.100	22518	1.1	1	2.3	2	3.4	2
87	0.104	14593	1.2	1	2.4	2	3.6	3
75	0.112	7287	1.3	0	2.6	1	3.9	4
42	0.113	6728	1.3	0	2.6	0	3.9	0
60	0.119	4531	1.4	3	2.7	5	4.1	8
...								
37	0.335	20	3.9	3	7.7	6	11.6	8
82	0.406	12	4.7	3	9.3	8	14.0	13
9	0.427	11	4.9	3	9.8	8	14.8	15
44	0.434	11	5.0	1	10.0	7	15.0	16
7	0.437	10	5.0	3	10.1	11	15.1	18
57	0.458	9	5.3	10	10.6	17	15.8	23
...								
62	1.000	3	11.5	10	23.0	21	34.6	33
97	1.074	3	12.4	10	24.7	24	37.1	37
63	1.200	3	13.8	8	27.6	18	41.4	26
83	1.210	3	13.9	16	27.9	32	41.8	49
90	1.854	2	21.3	20	42.7	36	64.0	48
99	1.989	2	22.9	21	45.8	35	68.7	56

The search for 42



Each dot represents 50 cores, approximately 90 core-years.

The result for 42

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

$$d = |x + y| = 11 \cdot 43 \cdot 215921 \cdot 1008323 = 102980666258459 \approx 1.030 \times 10^{14}$$

$$x \approx -8.053873 \times 10^{16}, \quad y \approx 8.043575 \times 10^{16}, \quad z \approx 1.260212 \times 10^{16}$$

$$\begin{array}{lr} & -522413599036979150280966144853653247149764362110424 \\ & +520412211582497361738652718463552780369306583065875 \\ + & \underline{2001387454481788542313426390100466780457779044591} \\ & 42 \end{array}$$

The result for 3

$$569936821221962380720^3 - 569936821113563493509^3 - 472715493453327032^3 = 3$$

$$d = |x+y| = 167 \cdot 649095133 = 108398887211 \approx 1.084 \times 10^{11}$$

$$x \approx 5.699368 \times 10^{20}, \quad y \approx -5.699368 \times 10^{20}, \quad z \approx -4.727155 \times 10^{17}$$

$$\begin{array}{r} 185131426470358721030003064550489120286063150089838997749248000 \\ -185131426364725746289073278168542399539619802127338908944671229 \\ -\underline{\quad\quad 105632974740929786381946720746443347962500088804576768} \\ 3 \end{array}$$

Heath-Brown vs Huisman $100 \leq k < 1000$ (selected)

k	$\delta_k/6$	N_0	$N = 10^5$ expect	$N = 10^5$ actual	$N = 10^{10}$ expect	$N = 10^{10}$ actual	$N = 10^{15}$ expect	$N = 10^{15}$ actual
858	0.029	1720798182665417	0.3	1	0.7	2	1.0	2
276	0.032	42958715811596	0.4	1	0.7	1	1.1	2
390	0.033	15332443619105	0.4	0	0.8	0	1.1	0
516	0.033	13632255817671	0.4	0	0.8	1	1.1	1
663	0.033	12076668982001	0.4	0	0.8	1	1.1	1
975	0.039	163996624946	0.5	0	0.9	0	1.3	0
165	0.040	90472906051	0.5	0	0.9	0	1.4	0
555	0.043	14746456526	0.5	1	1.0	2	1.5	2
921	0.044	6885076231	0.5	0	1.0	0	1.5	0
348	0.045	5369191063	0.5	2	1.0	2	1.5	3
906	0.050	536676769	0.6	0	1.1	0	1.7	0
366	0.051	324767552	0.6	0	1.2	0	1.8	1
579	0.051	348505529	0.6	0	1.2	0	1.8	0
654	0.057	46795226	0.7	2	1.3	2	2.0	3
114	0.058	26824751	0.7	0	1.3	0	2.0	0
705	0.062	8959243	0.7	1	1.4	2	2.2	2
732	0.063	7553865	0.7	0	1.5	0	2.2	0
402	0.079	321328	0.9	1	1.8	2	2.7	3
633	0.080	282820	0.9	0	1.8	0	2.8	0
537	0.089	80345	1.0	2	2.0	3	3.1	3
795	0.089	71223	1.0	0	2.1	0	3.1	0
641	0.128	2519	1.5	1	2.9	1	4.4	2
627	0.130	2248	1.5	0	3.0	0	4.5	0
956	0.217	102	2.5	3	5.0	6	7.5	8
782	0.453	10	5.2	3	10.4	5	15.7	11
855	2.641	2	30.4	27	60.8	51	91.2	77

A better search strategy

To check $|z| \le N$ we need to check $d \le B := (\sqrt[3]{2} - 1)N \approx N/4$.

The value of B determines the number of arithmetic progressions (about $B/2$). The value of N/B determines the length of these arithmetic progressions.

It is much cheaper to increase N than it is to increase B. Indeed, for $N \gg B$ the cost is $O(N^\epsilon)$ (due to sieving) versus $O(B)$.

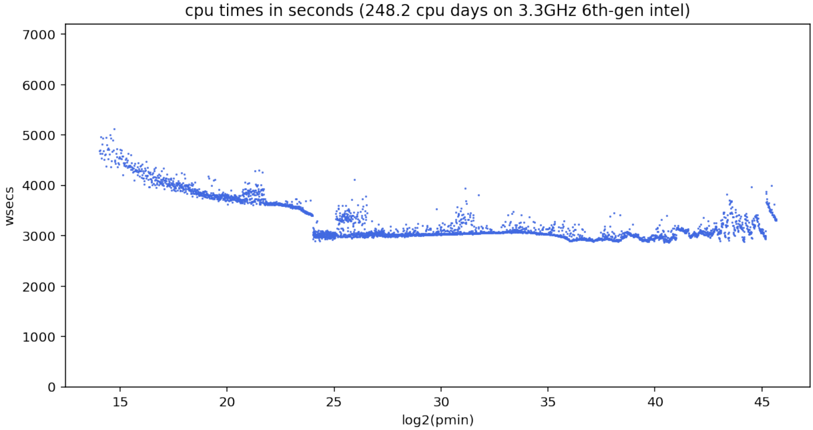
On the other hand, one heuristically expects the density of solutions to decay exponentially with N/B. This leads to an optimization problem. We want to choose $R := N/B$ to minimize $T(B,N) = T(B,RB)$. The optimal R should satisfy

$$T_B(B,RB)\frac{\partial B}{\partial R} + T_N(B,RB)(B + R\frac{\partial B}{\partial R}) = 0,$$

where T_B and T_N denote partial derivatives of $T(B,N)$. We typically want $R \in [50, 250]$ (this depends heavily on the implementation, and also on k).

We should also skip prime values of d close to B, which produce few progressions (an average of one for $p > B/2$). Better to wait for larger B.

The search for 42 redux



Each dot represents 2 cores, approximately 0.7 core years.

The search for 42 redux