Taxicab numbers

Avi Zeff

1. Introduction

(This is based on a project in college joint with Kevin Beuchot and Sanath Devalapurkar.)

There is a famous story of Hardy visiting Ramanujan in the hospital and remarking that he had arrived in taxicab number 1729, which he found to be a very boring number. Ramanujan immediately replied that on the contrary it was a very interesting number: it is the smallest number that can be written in two different ways as the sum of two cubes, i.e. $1729 = 12^3 + 1^3 = 10^3 + 9^3$. Hardy asked him about the same question for fourth powers, but he did not know off the top of his head.

This is not surprising: the smallest example for fourth powers is $635318657 = 158^4 + 59^4 = 134^4 + 133^4$. A property like this which first appears for fairly large numbers is interesting, and you might ask some related questions: are there any other such numbers that can be written in two ways as the sum of two cubes? If so, how many are there (perhaps infinitely many)? What about for fourth powers, or fifth powers and so forth? (Lower powers are less interesting: for 0th or 1st powers the question is trivial, and for squares it is pretty easy to see using Pythagorean triple-type arguments that there are infinitely many solutions.)

A computer search can quickly find many more examples of the first kind: for example, $4104 = 15^3 + 9^3 = 16^3 + 2^3$, $13832 = 20^3 + 18^3 = 24^3 + 2^3$, $20683 = 27^3 + 10^3 = 24^3 + 19^3$, ... This might lead you to guess that there are infinitely many.

For fourth powers, examples are rarer, but we can still find a few more without too much work: after 635318657, the next example is $3262811042 = 227^4 + 157^4 = 239^4 + 7^4$, then $8657437697 = 257^4 + 256^4 = 193^4 + 292^4$. Again, while these seem to be rarer than for cubes, we still might guess that there are infinitely many.

What about for fifth powers or higher? In this case, by contrast, there seem to be no examples at all! At the least, by a computer search I did yesterday I can tell you there are no examples less than 10^{15}. Since the first example in the case of fourth powers is so much larger than in the case of cubes, it's not unreasonable to think that there might be some very large examples that we can't find, but there's heuristic reason to believe that there really are no examples. I won't say too much about this heuristic, but basically the idea is we imagine that nth powers are distributed randomly such that there's about $\sqrt[n]{N}$ of them less than N, and so we can estimate the probability of pairs colliding like this; we could then use that probability to estimate the number of examples less than N, which comes out to be an increasing function of N as $N \to \infty$ for $n = 3$ or 4 but bounded for $n = 5$, so we'd expect only finitely many examples and all of them to be quite small; the fact that there are no small examples suggests that there are none at all.

This sounds like a really nice heuristic, and you might wonder why I'm not saying more about it. The answer is that it actually gives an estimate for the number of solutions to $a^n + b^n = c^n + d^n$ for (a, b) and (c, d) distinct (even after switching) when $n = 3$ or $n = 4$, but unfortunately at least in the case $n = 4$ this estimate is provably wrong! We can find more examples (in an infinite family) than this estimate predicts should exist. Thus it's worth taking with a grain of salt, but since we can't find any examples and all our methods will fail with $n > 4$ it seems pretty safe. Thus we have the following conjecture.

Conjecture. *Let $n \geq 3$ be an integer, and consider solutions to $N = a^n + b^n = c^n + d^n$ with (a, b) and (c, d) distinct.*

(a) If $n \leq 4$, there are infinitely many such solutions.

(b) If $n \geq 5$, there are no solutions.

It's actually pretty easy to "prove" part (a) of this conjecture: we can write down explicit families of solutions. Let r be any integer: then

$$\begin{aligned} a &= -27r \\ b &= r^6 - 9r^3 + 27 \\ c &= 3r^5 - 27r^2 \\ d &= r^6 - 18r^3 + 27 \end{aligned}$$

(among other families) satisfies $a^3 + b^3 = c^3 + d^3$, and

$$\begin{aligned} a &= r^6 + 3r^5 - 2r^4 + r^2 + 1 \\ b &= r^7 + r^5 - 2r^3 - 3r^2 + r \\ c &= r^6 - 3r^5 - 2r^4 + r^2 + 1 \\ d &= r^7 + r^5 - 2r^3 + 3r^2 + r \end{aligned}$$

satisfies $a^4 + b^4 = c^4 + d^4$. These can be proved by plugging these formulas into the equation and doing some algebra.

These are incredibly unsatisfying, though: where did any of this come from? Why should we care about these particular families? Why would you think to write this down?

To answer these questions, we'll bring in some tools from algebraic geometry. These will prove unable to handle the case $n \geq 5$, though, and in fact we won't be able to prove part (b) of our conjecture (this is an open problem!). What we will do, though, is try to bound the number of such solutions in the case $n \geq 5$: first, we'll figure out what a "trivial bound" is, and then we'll use some sieve theory to improve on that bound.

2. Algebraic geometry

Instead of thinking of solutions to $a^n + b^n = c^n + d^n$, we can more straightforwardly think of the problem as finding at least two (distinct) solutions to $N = a^n + b^n$ for some fixed N: in particular, are there infinitely many N for which this is possible?

Let's focus on the case $n = 3$. In this case, $N = a^3 + b^3$ is a smooth cubic curve in a and b. This sort of thing has another name: it is an elliptic curve. Elliptic curves have a very useful property: given two points on an elliptic curve, we can add them together to get a third point, and if the two points we start with have rational coordinates, so will the third point. (Draw out and explain) These don't have to be different points: we can start with a single point P and add it to itself by taking a tangent line, to get a point $2P$.

In particular, suppose that we had a point $P = (a_0, b_0)$ on our curve $N = a^3 + b^3$, i.e. $a_0^3 + b_0^3 = N$ for some rational numbers a_0, b_0. Then we could construct another point $2P = P + P$, and a third point $3P = 2P + P$, and so on.

Now, this doesn't show that $N = a^3 + b^3$ necessarily has infinitely many *integer* solutions; it actually doesn't necessarily show that it has infinitely many rational solutions either, because sometimes there may (at least in principle) be *no* solutions, so we couldn't start this process. However, let's just take the first two points $P = (a_0, b_0)$ and $2P = (a_1, b_1)$, and write $a_i = \frac{x_i}{y_i}$, $b_i = \frac{z_i}{w_i}$. We have

$$N = \frac{x_0^3}{y_0^3} + \frac{z_0^3}{w_0^3} = \frac{x_1^3}{y_1^3} + \frac{z_1^3}{w_1^3},$$

so clearing denominators we find

$$(y_0 y_1 w_0 w_1)^3 N = (y_1 w_0 w_1 x_0)^3 + (y_0 y_1 w_1 z_0)^3 = (y_0 w_0 w_1 x_1)^3 + (y_0 y_1 w_0 z_1)^3.$$

Thus although it isn't the number we started with, we've found a new number $(y_0 y_1 w_0 w_1)^3 N$ which can be written as the sum of two cubes in two different ways.

This was only using the two points P and $2P$. We could add in a third point $3P$ and do something similar to find a new number which can be written as the sum of two cubes in *three* different ways, and in particular take a pair other than this first pair here to find a new solution; iterating, we find infinitely many solutions. Thus to start with we only need to find some N for which there is at least one solution; an easy choice is $N = 9$, with rational point given by $9 = 1^3 + 2^3$. (For the experts, technically we need to check that this point has infinite order, but this is easy to do directly: this elliptic curve actually has rank 1 and trivial torsion.)

Okay, so we're done with the $n = 3$ case! Well, sort of. We did technically find infinitely many examples, but for example this will miss 1729 (the first solution generated in this way is $3087 = 7^3 + 14^3 = 20^3 + (-17)^3$), and it kind of seems like we're cheating: another apparently just as good way to get infinitely many solutions would be to start with $1729 = 12^3 + 1^3 = 10^3 + 9^3$ and multiply through by any cube. Our method really is better, though: by choosing different pairs from the increasingly many we get by taking higher and higher multiples of P, we get solutions which are not multiples of a lower solution, and the method works for infinitely many N which need not themselves have two (or indeed any) integer solutions.

Another method uses a little more geometry, so I'll describe it in less detail. Instead of fixing an N, let's look at the equation $a^n + b^n = c^n + d^n$. Up to scalar multiplication, we can think of all of the coordinates as rational numbers instead of integers and divide by one of them, so there are only three free coordinates and one equation: this means there are two degrees of freedom, so this describes an algebraic surface, which we'll call H_n. If we do a coordinate change to (x, y, z, w) by

$$a = x + y, \qquad b = z - w, \qquad c = x - y, \qquad d = z + w$$

something interesting happens: our equation becomes

$$(x + y)^n + (z - w)^n = (x - y)^n + (z + w)^n,$$

which in the case $n = 3$ simplifies to

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

and for $n = 4$ gives

$$x^3y + xy^3 = z^3w + wz^3.$$

Since we only care about solutions up to scalar multiplication, we can divide through by any scalar to assume that one of the coordinates, say w, is equal to 1. We have a map given by sending a tuple (x, y, z, w) to one of the coordinates, say x; this has image in the coordinate line, which in this case is a projective line $\mathbb{P}^1$. If we look at the fibers of this map, i.e. the equation above where we fix $w = 1$ and $x = x_0$ as a constant, then in both cases the resulting equation is degree three in the remaining variables! In other words, the map $H_n \to \mathbb{P}^1$ gives H_n as (roughly) a *pencil* of cubic curves over $\mathbb{P}^1$ for either $n = 3$ or $n = 4$, which turn out to be elliptic curves. If we went up to degree 5, though, the curves would be degree 5, which is too large for us to deal with.

We know that elliptic curves have magical properties that can help us, so this is a good sign for the case $n \leq 4$. We can then do a similar process with this pencil of elliptic curves, where starting from two rational points we can find a third, and this actually lets us build whole families of rational points on H_n which we can find explicitly to give solutions like the expressions given in the introduction.

3. Sieving

That pretty well takes care of the case $n \leq 4$; we've proven part (a) of the conjecture. As mentioned, part (b) in the case $n \geq 5$ is an open problem which we are not going to solve today, but hopefully we can say something about it. Let $R_n(N)$ be the number of positive integers $x \leq N$ which can be written in at least two different ways as the sum of two nth powers. We know that $R_n(N) \to \infty$ as N grows for $n \leq 4$, and we believe that in fact $R_n(N) = 0$ for all N for $n \geq 5$, but we don't know how to prove that; instead we'd like to bound $R_n(N)$ somehow.

One way to do this would be to ask an easier question: what is the number $R_n^*(N)$ of positive integers $x \leq N$ which can be written in at least *one* way as the sum of two nth powers? This definitely grows without bound, but certainly we should have $R_n(N) \leq R_n^*(N)$; this is the "trivial" bound. Thus we'd at least like to do better than that:

Conjecture. *For any $n \geq 5$, as $N \to \infty$ we have $\frac{R_n(N)}{R_n^*(N)} \to 0$.*

This is the least we could ask for, but at least it's nontrivial. Even this is surprisingly hard to prove: a lot of reasonably sophisticated techniques actually give worse bounds for $R_n(N)$ than the trivial bound ($R_n^*(N)$ is about $N^{2/n}$, because it depends on the choice of two integers a and b each of order at most $N^{1/n}$).

We actually study the number $r_n(M)$ of tuples (a, b, c, d) of positive integers bounded by M such that $a^n + b^n = c^n + d^n$ and (a, b) and (c, d) distinct; since $N = a^n + b^n$ is then roughly of order at most M^n, we can bound $R_n(N)$ using bounds on $r_n(M)$.

The idea is this. Counting integer solutions to polynomial equations, at least of high degree, is hard, as we've seen. However, if instead we work modulo p, things get much easier: there's a whole field of math around giving strong bounds for solutions to algebraic equations over finite fields. Therefore we can give some sort of restriction on the solutions

modulo each prime, and then put all this information back together to get a global bound using a sieve.

The idea of a sieve is exactly this: we have some condition from each prime number, and we want to say something about how these combine together. The classic example is the sieve of Erasthotenes, where we use the fact that each prime p rules out $1/p$ of the numbers above that from being prime. (Describe) In practice the sieve of Erasthotenes doesn't give much in the way of bounds, but there are many ways of modifying it to be very useful. A classic result along these lines is that there aren't "too many" twin primes, which I will not get into but is a good application.

The problem is that in those cases, we rule out only one class for each prime, and in most applications it's not more than one or two. Here we actually want to rule out many classes: the number of allowable classes turns out to be a positive proportion of all of them for many primes.

To solve this, we use a device called the large sieve, which I could write out as a formula but I'll just say is the result of some very clever manipulations by Gallagher in 1973. We actually want to use a higher-dimensional version of the sieve, which isn't too bad. The way this works is this: first, we use some coordinate transformations and algebraic manipulations to change the form of our equation $a^n + b^n = c^n + d^n$, just like for the surface over $\mathbb{P}^1$ (we actually have to assume n is even for this, but there are ways around this). The point of these manipulations is that if we then hold two of our new variables constant, we get an equation in the other two variables which describes a certain lattice, so we are sieving on the lattice for certain conditions modulo many prime numbers. This is the sort of thing the two-dimensional large sieve can handle, and if we choose our parameters carefully we get a bound; we then sum back over the variables we fixed, and the best choice of parameters turns out to give a bound of

$$r_n(M) \ll M^{2-\frac{1}{20}+\epsilon}.$$

This translates to

$$R_n(N) \ll N^{\frac{2-\frac{1}{20}}{n}+\epsilon}.$$

This is better than the trivial bound $R_n^*(N) \sim N^{2/n}$! In particular this means that

$$\frac{R_n(N)}{R_n^*(N)} \ll N^{-\frac{1}{20n}+\epsilon},$$

which goes to 0 as $N \to \infty$.

Finally, let me say a little bit about the state of the art. The initial result of this type is due (as far as I know) to Hooley, who used a similar method slightly more carefully to get a slightly better constant. Heath-Brown has a paper from 2002 where he shows that

$$R_n(N) \ll N^{1+\epsilon}$$

for $n \geq 13$, and Browning shortly afterwards improved it to

$$R_n(N) \ll N^{2/3+\epsilon}$$

for $n \geq 27$; and in 2006 Heath-Brown eliminated the main term to get

$$R_n(N) \ll N^{\frac{3}{\sqrt{n}}+\frac{2}{n-1}+\epsilon}$$

for all $n \geq 2$. There may be more recent improvements, but I am not aware of them.

This problem generalizes Fermat's last theorem, for which the approach is entirely abstract and algebro-geometric rather than sieve-theoretic—although sophisticated methods combining geometry and sieve theory can give nontrivial bounds, as we've seen and as Heath-Brown and Browning have shown, sieving is still very hard and not well-suited to proving the nonexistence of any nontrivial solutions. This would probably require a better understanding of algebraic surfaces over the integers than we currently have, but it is a fast-moving field and so far as I know no one has tried.

This is also an instance of a more general problem: is there a polynomial $f(x, y)$ of degree at least 5 which never takes the same value twice? It turns out that the answer to this is yes if we assume a more general conjecture, the Bombieri-Lang conjecture, which is a statement about there not being "too many" points on algebraic surfaces, which is an analogue of Faltings's theorem and is related to a whole web of conjectures and theorems about distribution and density of rational points on various kinds of varieties.