18.821 Project 1 First Draft
Sums of Cubes

Julianne Flusche, Luis Modes, Daniel Santiago

March 22, 2024

Abstract

We explore the positive integer solutions to the equation $N = x^3 + y^3 = z^3 + w^3$, as well as their density. We prove that there are infinitely many solutions to this equation with $\gcd(x, y, z, w) = 1$, and conjecture that the number of solutions to $x^3 + y^3 = z^3 + w^3$ with $x, y, z, w \leq n$ grows superlinearly in n. We also give bounds for the number of integer solutions to $N = x^3 + y^3$ based on the prime decomposition of N, and briefly discuss finding solutions to the equation $N = x^n + y^n = z^n + w^n$, where $n = 2, 4$, and 5.

Contents

1 Introduction 1

2 Infinitely many primitive solutions 2

3 Density of taxicab numbers 6

4 Bounding the number of solutions to $N = x^3 + y^3$ 7

5 Other powers 10

6 Individual Contributions 11

7 Appendix 12

1 Introduction

One day in 1918, British mathematician G. H. Hardy visited his friend and collaborator, Indian mathematician Srinivasa Ramanujan, in hospital. Upon arriving, Hardy remarked that the license number of the taxi he came in, 1729, was very boring. Ramanujan then replied, "not at all — that is the first integer that is the sum of two (positive) cubes in two different ways." As a result, 1729 became known as the Hardy-Ramanujan number, the smallest of a set of numbers called *taxicab numbers*.

Definition 1.1. A *taxicab number* is a number N that can be expressed as a sum of two cubes of positive integers in at least two different ways. That is,

$$N = a^3 + b^3 = c^3 + d^3, \tag{1}$$

where $a, b, c, d \in \mathbb{Z}_{>0}$ and $\{a, b\} \neq \{c, d\}$. We call the tuple (a, b, c, d, N) a *solution*.

See appendix I for the lowest and highest fifty taxicab numbers N with a bound of 10,000 on a, b, c, d. Some sources [1] define the n-th taxicab number as the smallest integer that can be expressed as a sum of two cubes in n different ways. For the sake of this paper, we will define taxicab numbers as above.

The main question of interest is whether there are an infinite number of taxicab numbers. We know there must be an infinite number of taxicab numbers given that we can scale any taxicab number by a cube to get another taxicab number. That is, $1729k^3$ is a taxicab number for any integer k because $1729k^3 = (9k)^3 + (10k)^3 = k^3 + (12k)^3$ for any k. To form a more interesting research question, we consider only taxicab numbers that are not scalar multiples of smaller ones. We refer to these taxicab numbers as *primitive*. Let us make this notion more precise.

Definition 1.2. We say that a solution (a, b, c, d, N) and the number N in this tuple are *primitive* if $\gcd(a, b, c, d) = 1$. That is, there does not exist an integer $k > 1$ such that k divides a, b, c, and d.

The question we now ask is if there are infinitely many primitive taxicab numbers. The following theorem, which we prove in section 2, shows that this is the case.

Theorem 1.3 (See Theorem 2.5). There are infinitely many primitive solutions to the equation $N = x^3 + y^3 = z^3 + w^3$.

We prove this theorem in section 2. To prove the theorem, we begin with the curve $C : x^3+y^3-N = 0$, where $N = a^3 + b^3$. We then draw the tangent at (a, b) and extend it until it cuts C again in (c, d). By Bézout's theorem [2], we know this intersection will exist, and we can also show c, d are rational. Next, we draw the tangent at (c, d) until it cuts C again at (e, f). The key idea is that if we choose suitable values for a, b, we can get a primitive solution N from some combination of the values c, d, e, and f.

Once we have shown that there are an infinite number of primitive taxicab numbers, a logical next question to ask is what is the density of taxicab numbers in the integers. In section 3 of this paper we explore the rate of growth of the number of taxicab numbers in the integers as we increase the bound on a, b, c, and d. Through experimentation, we arrive at the following conjecture.

Conjecture 1.4 (See Conjecture 3.2). The number of positive integer solutions to $x^3 + y^3 = z^3 + w^3$ with $x, y, z, w \leq n$ grows quicker than any linear function on n.

From this conjecture, we make a conclusion about the distance between successive taxicab numbers in the integers.

In section 4, we switch our focus from proving there are infinitely many taxicab numbers to finding restrictions on the set of possible taxicab solutions. To do so, we bound the number of integer solutions (x, y) to the equation $N = x^3 + y^3$ based on the prime factorization of N. In the process, we find some infinite families of numbers that cannot be taxicab numbers, namely primes, numbers of the form $3p$ with p a prime, and numbers of the form $2p$ with p a prime congruent to 1 mod 3.

After studying the equation $x^3 + y^3 = z^3 + w^3$, it is natural to wonder what happens when we consider any exponent n, as opposed to just $n = 3$. In section 5, we consider the equation $N = a^n + b^n = c^n + d^n$ for powers n other than 3. We investigate whether there are infinitely many numbers N that satisfy this property for $n = 2, 4$, and 5.

2 Infinitely many primitive solutions

Let us prove that there are infinitely many primitive taxicab numbers. In our proof, we will start with a point (a, b) in the curve $C : x^3 + y^3 - N = 0$. Then, we draw the tangent to C through (a, b) until it cuts C again at the point (c, d). In order to be able to compute c and d explicitly in terms of (a, b), we work out the formula of the tangent line in Lemma 2.1 and the formula for the point of intersection in Lemma 2.2. Then, we draw the tangent to C through (c, d) until it cuts C again at the point (e, f). As (c, d) and (e, f) both lie in C, we have that $c^3 + d^3 = e^3 + f^3$. If we choose suitable values a and b, we are guaranteed to get a positive integer primitive solution from $c^3 + d^3 = e^3 + f^3$ by clearing denominators. The precise conditions for "suitable" will show up in Lemma 2.3. Lemma

2.4 will help us to check that we did not add any extra factor when we cleared denominators. With this process, we will be able to generate infinitely many solutions, as desired.

Lemma 2.1. Consider the curve $C : x^3 + y^3 - N = 0$. If $(a, b) \in C$, then the equation of the line through (a, b) tangent to C is given by $T : y = -\frac{a^2}{b^2}x + \frac{N}{b^2}$.

Proof. Recall that the slope of the tangent line to the curve $C : f(x, y) = 0$ at the point $(a, b) \in C$ is given by $\frac{\partial y}{\partial x}\Big|_{(a,b)}$. Taking the derivative of the curve $C : x^3 + y^3 - N = 0$ with respect to x, we get that $\frac{\partial y}{\partial x} = -\frac{x^2}{y^2}$. Thus, as T goes through (a, b) and has slope $\frac{\partial y}{\partial x}\Big|_{(a,b)} = -\frac{a^2}{b^2}$, it is given by the equation $\frac{y-b}{x-a} = -\frac{a^2}{b^2}$, which we can rewrite as

$$y = -\frac{a^2}{b^2}x + \frac{a^3}{b^2} + b,$$

that is,

$$y = -\frac{a^2}{b^2}x + \frac{N}{b^2},$$

as desired. □

Lemma 2.2. Consider the curve $C : x^3+y^3-N = 0$. If $N = a^3+b^3$, with $a \neq b$, and $(a, b) \in C$, then the tangent line T to C at (a, b) cuts C at the point with rational coordinates $(c, d) = \left(-\frac{a^4+2ab^3}{b^3-a^3}, \frac{b^4+2a^3b}{b^3-a^3}\right)$.

Proof. By Lemma 2.1, the equation for the tangent line is $y = -\frac{a^2}{b^2}x + \frac{N}{b^2}$. Now, consider the cubic polynomial

$$P(x) = x^3 + \left(-\frac{a^2}{b^2}x + \frac{N}{b^2}\right)^3 - N = \left(\frac{b^6 - a^6}{b^6}\right)x^3 + \frac{3a^4N}{b^6}x^2 - \frac{3a^2N^2}{b^6}x + \frac{N^3}{b^6} - N.$$

Note that a is a root with multiplicity 2 of $P(x)$, so it must have a third real root c. If c is this real root, by Vieta's formulas [3],

$$\begin{aligned} 2a + c &= -\frac{\frac{3a^4N}{b^6}}{\left(\frac{b^6-a^6}{b^6}\right)} \\ c &= -\frac{3a^4(b^3 + a^3)}{(b^3 + a^3)(b^3 - a^3)} - 2a \\ c &= -\frac{a^4 + 2ab^3}{b^3 - a^3}, \end{aligned}$$

so plugging this value of c into the equation for our tangent line, we get

$$d = -\frac{a^2}{b^2} \cdot \left(-\frac{a^4 + 2ab^3}{b^3 - a^3}\right) + \frac{N}{b^2} = \frac{a^6 + 2a^3b^3}{b^2(b^3 - a^3)} + \frac{(b^3 + a^3)(b^3 - a^3)}{b^2(b^3 - a^3)} = \frac{b^4 + 2a^3b}{b^3 - a^3}.$$

□

Lemma 2.3. There are infinitely many pairs $(a, b) \in \mathbb{Z}^2_{>0}$ satisfying the following properties:

1. $b > a$,
2. $\gcd(a^4 + 2ab^3, a^3 + b^3) = 1$,
3. $\gcd(a^4 + 2ab^3, b^3 - a^3) = 1$, and

4. $2(a^4+2ab^3)^3 > (b^4+2ba^3)^3$.

Proof. We claim that any pair $(a,b)=(n,n+1)$ with n sufficiently large works, in which case it would be clear that there are infinitely many of these. In particular, as $\lim_{n\to\infty}\left(\frac{n+1}{n}\right)^{12}=1$, there exists some N such that if $n \geq N$, then $\left(\frac{n+1}{n}\right)^4 \leq \left(\frac{n+1}{n}\right)^8 \leq \left(\frac{n+1}{n}\right)^{12} < 2$. Let $n \geq N$ for this fixed N be our definition of sufficiently large. Now, let us show that all these pairs satisfy all the required properties.

1. Note that $b=n+1>n=a$.

2. As a,b are consecutive integers, $\gcd(a,b)=1$, which implies $\gcd(a,a^3+b^3)=\gcd(b,a^3+b^3)=1$, so by the properties of the greatest common divisor,

$$\begin{aligned}\gcd(a^4+2ab^3, a^3+b^3) &= \gcd(a^3+2b^3, a^3+b^3)\\ &= \gcd(b^3, a^3+b^3)\\ &= \gcd(b^3, a^3)\\ &= 1.\end{aligned}$$

3. In the same way as in the previous part, we get that $a^3-b^3 \equiv n^3-(n+1)^3 \equiv n-(n+1) \not\equiv 0 \bmod 3$, so by the properties of the greatest common divisor,

$$\begin{aligned}\gcd(a^4+2ab^3, a^3-b^3) &= \gcd(a^3+2b^3, a^3-b^3)\\ &= \gcd(3b^3, a^3-b^3)\\ &= \gcd(b^3, a^3-b^3)\\ &= \gcd(b^3, a^3)\\ &= 1.\end{aligned}$$

4. Note that, if $2(n+1)n^3=m$, after some algebraic manipulations, we can bound $2(a^4+2ab^3)^3$ from below by repeatedly using that $n+1>n$ and our assumption that n is sufficiently large:

$$\begin{aligned}2(a^4+2ab^3)^3 &= 2(n^4+2n(n+1)^3)^3\\ &\geq 2(n^4+2(n+1)n^3)^3\\ &= 2(n^4+m)^3\\ &= 2n^{12}+6n^8m+6n^4m^2+2m^3\\ &> (n+1)^{12}+3(n+1)^8m+3(n+1)^4m^2+m^3\\ &= ((n+1)^4+m)^3\\ &= ((n+1)^4+2(n+1)n^3)^3\\ &= (b^4+2ba^3)^3,\end{aligned}$$

where the strict inequality is true due to our choice of a sufficiently large n.

Thus, all the claimed pairs (a,b) satisfy the required properties, as desired. $\square$

Lemma 2.4. Let $\frac{a_i}{b_i}$ be fractions with $\gcd(a_i,b_i)=1$ for $1 \leq i \leq n$. Let $M=\operatorname{lcm}(b_1,\dots,b_n)$. Then, the integers $M\frac{a_1}{b_1},\dots,M\frac{a_n}{b_n}$ have a common factor $k>1$ if and only if $a_1,\dots,a_n$ have a common factor $k'>1$.

Proof. If $a_1, \dots, a_n$ have a common factor $k' > 1$, then $M\frac{a_1}{b_1}, \dots, M\frac{a_n}{b_n}$ have a common factor by choosing $k = k'$. Conversely, assume that $M\frac{a_1}{b_1}, \dots, M\frac{a_n}{b_n}$ have a common factor $k > 1$. Let $g = \gcd(k, M)$. We have two cases: either $g < k$ or $g = k$.

1. If $g < k$, consider $k' = \frac{k}{g}$. Then, as $k' \mid k$ and $k \mid M\frac{a_i}{b_i}$, it follows that $k' \mid M\frac{a_i}{b_i}$. This implies that $k'b_i \mid Ma_i$, but as $\gcd(k', M) = 1$, we get that $k' \mid a_i$ for all $1 \le i \le n$. Thus, $k' = \frac{k}{g} > 1$ works as our common factor.

2. Assume that $g = k$. In this case, $k \mid M\frac{a_i}{b_i}$, so $b_i \mid \frac{M}{k}a_i$. As, $\gcd(a_i, b_i) = 1$, we get that $b_i \mid \frac{M}{k}$ for all $1 \le i \le n$. However, $b_i \mid \frac{M}{k}$ for all $1 \le i \le n$ would imply that $\frac{M}{k}$ is a positive integer smaller than $\operatorname{lcm}(n_1, \dots, b_n)$ divisible by $b_1, \dots, b_n$. By definition of least common multiple, this is a contradiction. Thus, this case is not possible, so there is always the desired factor k' dividing $a_1, \dots, a_n$.

Hence, after proving both directions, we are done. $\square$

Theorem 2.5. There are infinitely many primitive solutions to the equation $N = x^3 + y^3 = z^3 + w^3$.

Proof. First, choose a and b as in Lemma 2.3, and let $N = a^3 + b^3$. Consider the curve $C : x^3 + y^3 - N = 0$. The tangent line to C at (a, b) is given by $T : y = -\frac{a^2}{b^2}x + \frac{N}{b^2}$ by Lemma 2.1. This line cuts C again at $(c, d) := \left(-\frac{a^4+2ab^3}{b^3-a^3}, \frac{b^4+2a^3b}{b^3-a^3}\right)$ by Lemma 2.2. Also, by our choice of a, b as in Lemma 2.3, we have that $b > a$. In particular, as $b > a$, we get that $c < 0$ and $d > 0$. If S is the tangent to C at (c, d), by Lemma 2.2 again, we get that this tangent cuts C again at $(e, f) := \left(-\frac{c^4+2cd^3}{d^3-c^3}, \frac{d^4+2c^3d}{d^3-c^3}\right)$. Further, we claim that $e > 0$ and $f < 0$. Indeed, note that

$$f = \frac{d^4 + 2c^3d}{d^3 - c^3} = \frac{d}{d^3 - c^3} \cdot (d^3 + 2c^3) = \frac{d}{d^3 - c^3} \cdot \frac{(b^4 + 2ba^3)^3 - 2(a^4 + 2ab^3)^3}{(b^3 - a^3)^3} < 0$$

by our choice of a, b and Lemma 2.3, and as f is negative but $e^3 + f^3 = N$, e must be positive. Finally, if M is the least common multiple of the denominators of d, f, e, c and $(d', f', e', c') = (Md, -Mf, Me, -Mc)$, we claim that (d', f', e', c') is a primitive solution. Indeed, d', f', e', c' are all positive integers, and

$$d'^3 + f'^3 = (M^3d^3 + M^3c^3) + f'^3 + c'^3 = (Me^3 + Mf^3) + f'^3 + c'^3 = e'^3 + c'^3$$

so it is a solution.

To see it is primitive, assume for the sake of contradiction that they have a common factor $k > 1$. If this were the case, by Lemma 2.4, there would have to be an integer $k' > 1$ dividing the numerators of the reduced fractions of c, d, e, f, so it would have to divide N as well. In particular, as $c = -\frac{a^4+2ab^3}{b^3-a^3}$ is already in reduced form because $\gcd(a^4 + 2ab^3, b^3 - a^3) = 1$, k' would have to divide the numerator $a^4 + 2ab^3$ of c and $N = a^3 + b^3$, but then, putting this together,

$$\begin{aligned} 1 &= \gcd(a^4 + 2ab^3, a^3 + b^3) \\ &= \gcd(a^4 + 2ab^3, N) \\ &\ge k' \\ &> 1 \end{aligned}$$

Thus, we get a contradiction. As a result, the solution is indeed primitive. Hence, we were able to find a primitive solution with $c' \ge c = a^4 + 2ab^3 \ge a$, where a can be as large as we want by Lemma 2.3, so there are indeed infinitely many primitive solutions. $\square$

Remark 1. The process outlined in the proof of Theorem 2.5 can be used as an algorithm to generate infinitely many primitive taxicab solutions.

Remark 2. A similar process would be starting from a known solution (a, b, c, d, N) and, instead of starting by drawing a tangent, we start by drawing the line through $(a, b), (c, d)$. Assume this line cuts the curve for the third time in point (e, f). From (e, f), we draw the tangent until it cuts the curve for the third time in a point (g, h). At least two of the pairs $(a, b), (e, f), (g, h)$ should yield a solution. The only problem with this approach is that it becomes harder to show that the solution is primitive.

3 Density of taxicab numbers

Now that we know infinitely many primitive taxicab numbers exist, we explore the density of taxicab numbers in the integers. In this section, we form a conjecture based on experimentation where we hypothesize that the growth of the number of primitive taxicab numbers is superlinear and that the ratio of primitive taxicab numbers to all taxicab numbers approaches a constant close to 0.3. We use this conjecture to establish a bound on the distance between successive taxicab numbers.

Let us consider the growth of taxicab numbers as it relates to a bound on a, b, c, d. For a positive integer n, we define $F(n)$ to be the number of taxicab solutions (a, b, c, d, N) with $a, b, c, d \leq n$. We define $G(n)$ to be the number of such solutions that are also primitive.

There are $\frac{n(n+1)}{2}$ possible values for $a^3 + b^3$ with $a, b \leq n$. We do not have n^2 possible sums because we consider (a, b) and (b, a) to be the same pair. Consider the set (with size at most $\frac{n(n+1)}{2}$) of all sums of cubes $a^3 + b^3$ with $a, b \leq n$. All sums that appear more than once in this set are taxicab numbers. Therefore, the maximum number of taxicab solutions in our set is $\frac{n(n+1)}{4}$, corresponding to the case where each possible sum occurs exactly twice in our set. This bound on the number of pairs of equivalent sums allows us to bound $F(n)$ from above by $F(n) \leq \frac{n(n+1)}{4}$.

Since we know we can scale taxicab numbers to generate new solutions, we expect $F(n)$ to have at least linear growth. Our experimental results from generating primitive taxicab numbers, however, supports faster than linear growth.

Conjecture 3.2 The function $F(n)$ grows superlinearly. That is, there is no linear function $an + b$ for which $F(n) \leq an + b$ for all n sufficiently large. The ratio $G(n)/F(n)$ approaches a constant value $C \sim 0.3$. In particular, $G(n)$ also grows superlinearly.

Figure 1 below displays the experimental results supporting the conjecture. A computer program was used to generate all taxicab solutions with $a, b, c, d \leq n$ for different values of n, extract those that are primitive, and calculate the ratio of the two. The ratio of primitive taxicab numbers to all taxicab numbers predominantly decreases. The positive concavity in the Figure 1 graph suggests that this ratio is decreasing at a decreasing rate as we increase the bound on a, b, c, d. The ratio appears to asymptotically approach a value of around 0.3. See appendix II for a table of exact experimental data points.

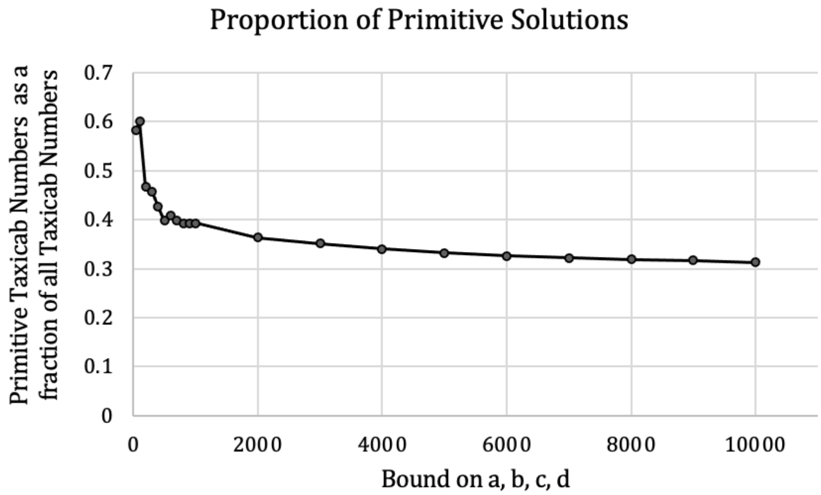

Figure 1. Ratio of primitive solutions to all solutions.

Conjecture 3.2 is a stronger form of Theorem 2.5, as the latter asserts only that $G(n)$ is unbounded. We next give some comments on the relationship between $F(n)$ and $G(n)$. We can write $F(n)$ explicitly in terms of $G(n)$ as

$$F(n) = \sum_{i=1}^{G(n)} \left\lfloor \frac{n}{\max\{a_i, b_i, c_i, d_i\}} \right\rfloor,$$

where a_i, b_i, c_i, and d_i range over the primitive solutions counted by $G(n)$. In particular, we have $F(n) \leq nG(n)$, which means if we can prove $F(n)$ grows superlinearly, then we can conclude that there are infinitely many primitive taxicab solutions. The converse does not hold: knowing that $G(n)$ is unbounded is not enough to conclude that F grows superlinearly. From the lower bound for $F(n)$ given by

$$F(n) = \sum_{i=1}^{G(n)} \left\lfloor \frac{n}{\max\{a_i, b_i, c_i, d_i\}} \right\rfloor \geq G(n) \left\lfloor \frac{n}{\max_i\{a_i, b_i, c_i, c_i\}} \right\rfloor,$$

we see that the behavior of $F(n)$ depends on both $G(n)$ and $\max_i\{a_i, b_i, c_i, d_i\}$. In particular, if it is true that $\frac{G(n)}{F(n)} \sim 0.3$ for large n, then

$$\left\lfloor \frac{n}{\max_i\{a_i, b_i, c_i, c_i\}} \right\rfloor \leq \frac{10}{3}.$$

Thus, there is a primitive taxicab solution between n and $\frac{10}{3}n$ for large n.

4 Bounding the number of solutions to $N = x^3 + y^3$

Now that we have proved there are infinitely many primitive taxicab numbers, let us find some restrictions on the set of possible taxicab solutions. In this section, we bound the number of integer solutions to $N = x^3 + y^3$ based on the prime decomposition of N. The proof uses the factored form of $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$ to reduce the number of possible integer solutions to a finite number of cases, a function of the number of prime factors of N. We prove that primes p and multiples of primes, namely $2p$ and $3p$ under certain restrictions, cannot be written as a sum of two cubes, $x^3 + y^3$, and thus cannot be taxicab numbers.

Proposition 4.1. If (x, y, z, w, N) is a solution, then $x + y \neq z + w$.

Proof. Assume for the sake of contradiction that $x+y=z+w$. Without loss of generality that $x>w$. In this case, let $w=x-i$. This would imply that $z=y+i$. Suppose $(x-i)^3+(y+i)^3=x^3+y^3$. We have

$$\begin{aligned}x^3+y^3&=x^3-3ix^2+3i^2x-i^3+y^3+3i^2y+2yi^2+i^3\\0&=-3ix^2+3i^2x-i^3+3i^2y+2yi^2+i^3\\0&=3i^2(x+y)+3i(y^2-x^2)\\i&=x-y\end{aligned}$$

so $z=y$ and $w=x$ and we have arrived at a contradiction. $\square$

We will now focus on finding solutions to $x^3+y^3=z^3+w^3$ for fixed a $N=x^3+y^3$. We will show that we can bound the number of possible integer solutions to this equation based on the number of prime factors of N. Let us first start with the case where N is prime. We can show that in this case, $N=x^3+y^3$ has no integer solutions (x,y).

Proposition 4.2. Proposition 4.2. For p prime distinct to 2, the equation $p=x^3+y^3$ has no positive integer solutions.

Proof. Recall the identity $p=x^3+y^3=(x+y)(x^2-xy+y^2)$. Since p is prime, the only factorization of p is $p=(1)(p)$. This factorization yields two possible cases for values of the factors in the identity.

$$\begin{cases}x+y=1, x^2-xy+y^2=p\\x+y=p, x^2-xy+y^2=1\end{cases}$$

Since $x\geq 1$ and $y\geq 1$, then $x+y>1$. This means we must have the case where $x+y=p$ and $x^2-xy+y^2=1$.
We solve for y in the first equation to get

$$y=p-x.$$

Then, we substitute this value of y into the other equation, yielding the quadratic equation,

$$\begin{aligned}x^2-x(p-x)+(p-x)^2&=1\\x^2-px+x^2+p^2-2px+x^2&=1\\3x^2-3px+p^2-1&=0.\end{aligned}$$

The discriminant of the above equation is $9p^2-4(3)(p^2-1)=-3p^2+12$. In solving $-3p^2+12\geq 0$, we get that $p\leq 2$. Since p is a positive prime greater than 2, we have arrived at a contradiction. Therefore there are no solutions to $p=x^3+y^3$ for $p>2$ prime. $\square$

After prime numbers, the numbers with the simplest prime factorization are those which are products of two primes. Thus, we will now consider N with two prime factors to see if we can yield any integer solutions in this case.

Proposition 4.3. For $N=pq$ where p,q are primes, $N=x^3+y^3$ has at most four possible integer solutions.

Proof. We again begin with the identity for sums of cubes: $N=pq=(x+y)(x^2-xy+y^2)$.
The different combinations of prime factors p and q yield four cases to consider for this factorization. The four cases are

$$\begin{cases}x+y=1, x^2-xy+y^2=pq\\x+y=pq, x^2-xy+y^2=1\\x+y=p, x^2-xy+y^2=q\\x+y=q, x^2-xy+y^2=p.\end{cases}$$

The first case is impossible, as shown in the proof for prime N. The second case is also impossible as x and y must be equal to 0 and 1 when $x^2 - xy + y^2 = 1$. This leaves the third and fourth cases as possibilities. We just need to analyze one case as we can use a symmetry argument for the other.

Take the third case. We solve for y in the first equation to get

$$y = p - x.$$

Then, we substitute this expression for y into the second equation, and get the quadratic equation,

$$\begin{aligned} x^2 - x(p-x) + (p-x)^2 &= q \\ 3x^2 - 3px + p^2 - q &= 0. \end{aligned}$$

The discriminant of the quadratic equation is $-3p^2 + 12q$. We solve $-3p^2 + 12q > 0$ and get that $p^2 < 4q$, which is true for many values of p and q. The maximum number of solutions from this case is two, given that we have a quadratic. By a symmetry argument for case 4, we get a maximum of two additional solutions. So for $N = pq$, p, q prime, we have at most four solutions to $N = x^3 + y^3$. $\square$

Now, we generalize this method for N with n prime factors in order to bound the number of solutions to $N = x^3 + y^3$ for any integer N.

Theorem 4.4. For N with n prime factors, $N = x^3 + y^3$ has at most $2(\binom{n}{1} + \binom{n}{2} + ...\binom{n}{n-1})$ possible integer solutions.

Proof. Let us list the cases, omitting the ones where either $x + y = 1$ or $x^2 - xy + y^2 = 1$ as we know these yield no solutions. The cases we have left are:

$$\begin{cases} x + y = \text{(one of the factors)}, & x^2 - xy + y^2 = \text{(the remaining } n-1 \text{ factors)} \\ x + y = \text{(two of the factors)}, & x^2 - xy + y^2 = \text{(the remaining } n-2 \text{ factors)} \\ & \vdots \\ x + y = (n-1 \text{ of the factors}), & x^2 - xy + y^2 = \text{(the remaining one factor)} \end{cases}$$

The kth line above represents a total of $\binom{n}{k}$ different cases and each of these is a quadratic with two possible solutions. Therefore, we have $2(\binom{n}{1} + \binom{n}{2} + ...\binom{n}{n-1})$ possible solutions when N has n prime factors.

$\square$

For reference, the maximum number of prime factors of any taxicab number on our list $(a, b, c, d \leq 10,000)$ is 8.

There are some other bounds that can be proven using the fact that $(x + y)^3 = x^3 + y^3 \bmod 3$.

Proposition 4.5. Let N be a positive integer such that no two distinct proper divisors of N that are distinct from 1 are congruent mod 3. Then there is at most one positive integer solution to $x^3 + y^3 = N$.

Proof. Suppose there are two distinct solutions $x_0^3 + y_0^3 = N = x_1^3 + y_1^3$. Since $(x + y)^3 = x^3 + y^3$ mod 3, they must satisfy $x_1 + y_1 = x_0 + y_0$ mod 3. Note that $x_0 + y_0, x_1 + y_1$ are distinct, larger than 1, and divide N by the identity $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$. Since no two distinct factors of N are congruent mod 3, this is a contradiction.

$\square$

A consequence of Proposition 4.5 is that no numbers of the form $3p$ with p a prime distinct from 3, or $2p$ with p a prime congruent to 1 mod 3, are taxicab numbers.

5 Other powers

In section 2, we were able to show there were infinitely many primitive solutions to the equation $N = x^3 + y^3 = z^3 + w^3$. In this section, we explore the equation $N = x^n + y^n = z^n + w^n$ for $n = 2, 4$, and 5 to see if we can arrive to the same conclusion and to get a better understanding of the general case. We prove that there are infinitely many such N for $n = 2$, we propose a method that uses the $n = 2$ case to show this may also be true for the $n = 4$ case, and we provide experimental results that show there are no solutions for the $n = 5$ case that are easily generated by a computer.

First, we consider the case of squares and show that there are an infinite number of solutions to $N = a^2 + b^2 = c^2 + d^2$. In order to prove this, we show a brief lemma that characterizes integers a, b, c, d such that $ab = cd$.

Lemma 5.1. Given $a, b, c, d \in \mathbb{Z}$, $ab = cd$ if and only if there exist $m, n, p, q \in \mathbb{Z}$ such that $a = mn$, $b = pq$, $c = mp$, $d = nq$.

Proof. If $a = mn$, $b = pq$, $c = mp$, $d = nq$, clearly $ab = mnpq = cd$. Conversely, if $ab = cd$, $a \mid cd$, so $\frac{cd}{a}$ is an integer, and thus we can write a as $a = mn$, where $m \mid c$ and $n \mid d$ (say, for example, by taking $m = \gcd(a, c)$ and $n = \frac{a}{\gcd(a,c)}$). Hence, if we let $p = \frac{c}{m}$ and $q = \frac{d}{n}$, we have that $a = mn$, $b = \frac{cd}{a} = \frac{cd}{mn} = \frac{c}{m}\frac{d}{n} = pq$, $c = m \cdot \frac{c}{m} = mp$, and $d = n\frac{d}{n} = nq$. □

With this characterization, we can completely parametrize the solutions to $x^2 + y^2 = z^2 + w^2$ when $n = 2$.

Proposition 5.2. Given $a, b, c, d \in \mathbb{Z}$, $a^2 + b^2 = c^2 + d^2$ if and only if there are integers m, n, p, q such that $a = \frac{pq+mn}{2}$, $b = \frac{nq-mp}{2}$, $c = \frac{pq-mn}{2}$, $d = \frac{nq+mp}{2}$.

Proof. If $a = \frac{pq+mn}{2}$, $b = \frac{nq-mp}{2}$, $c = \frac{pq-mn}{2}$, $d = \frac{nq+mp}{2}$, clearly

$$a^2 + b^2 = \frac{p^2q^2 + m^2n^2 + n^2q^2 + m^2p^2}{4} = c^2 + d^2.$$

Conversely, if $a^2 + b^2 = c^2 + d^2$, then $(a - c)(a + c) = (d - b)(d + b)$, so by Lemma 5.1, there exist $m, n, p, q \in \mathbb{Z}$ such that $a - c = mn$, $a + c = pq$, $d - b = mp$, and $d + b = nq$. Hence, by solving the system of equations, we get that $a = \frac{pq+mn}{2}$, $b = \frac{nq-mp}{2}$, $c = \frac{pq-mn}{2}$, $d = \frac{nq+mp}{2}$, as desired. □

Next, we consider how our method for proving there are an infinite number of solutions to $N = a^n + b^n = c^n + d^n$ for $n = 2$ may help us answer the same question for the case where $n = 4$. By using the above method for squares, the problem for finding solutions to $N = a^4 + b^4 = c^4 + d^4$ reduces to finding $m, n, p, q \in \mathbb{Z}$ such that $\frac{pq+mn}{2}$, $\frac{nq-mp}{2}$, $\frac{pq-mn}{2}$, $\frac{nq+mp}{2}$ are all perfect squares. This method may be helpful in finding an infinite family of solutions for sums of fourth powers.

Our experimental results show that there are 246 solutions to $N = a^4 + b^4 = c^4 + d^4$ for $a, b, c, d \leq 10,000$. The smallest is $635318657 = 59^4 + 158 = 133^4 + 134^4$.

Now let us discuss our experimental findings for the case where $n = 5$. We proved experimentally that there are no solutions to the equation $N = a^5 + b^5 = c^5 + d^5$ for $a, b, c, d \leq 10,000$ (i.e. $N < 2 \cdot 10^{20}$).

We are unsure at this point if any exist. Much less is known about fifth degree polynomials than second or third, and given that 5 is prime, it is unlikely that we can use other results in this paper to draw conclusions about 5th powers.

6 Individual Contributions

We refer to the authors Julianne Flusche, Luis Modes, and Daniel Santiago as authors J, L, and D, respectively. Author J wrote section 1, helped to write section 3 and produced the experimental data that led to Conjecture 3.2. Author J also wrote and proved Theorem 4.4 and Propositions 4.1-4.3. Author L wrote section 2 and proved Theorem 2.5, as well as Lemma 5.1.1 and Proposition 5.1.2. Author D helped to write section 3 and the abstract, and proved and wrote propositions $4.1, 4.5$. All authors contributed to editing and proofreading all sections of the paper.

7 Appendix

I. Table of taxicab numbers

N	(a,b)	(c,d)	prime factors of N
1729	(1, 12)	(9, 10)	[7, 13, 19]
4104	(2, 16)	(9, 15)	[2, 3, 19]
13832	(2, 24)	(18, 20)	[2, 7, 13, 19]
20683	(10, 27)	(19, 24)	[13, 37, 43]
32832	(4, 32)	(18, 30)	[2, 3, 19]
39312	(2, 34)	(15, 33)	[2, 3, 7, 13]
40033	(9, 34)	(16, 33)	[7, 19, 43]
46683	(3, 36)	(27, 30)	[3, 7, 13, 19]
64232	(17, 39)	(26, 36)	[2, 7, 31, 37]
65728	(12, 40)	(31, 33)	[2, 13, 79]
110656	(4, 48)	(36, 40)	[2, 7, 13, 19]
110808	(6, 48)	(27, 45)	[2, 3, 19]
134379	(12, 51)	(38, 43)	[3, 7, 79]
149389	(8, 53)	(29, 50)	[31, 61, 79]
165464	(20, 54)	(38, 48)	[2, 13, 37, 43]
171288	(17, 55)	(24, 54)	[2, 3, 13, 61]
195841	(9, 58)	(22, 57)	[37, 67, 79]
216027	(3, 60)	(22, 59)	[3, 7, 127]
216125	(5, 60)	(45, 50)	[5, 7, 13, 19]
262656	(8, 64)	(36, 60)	[2, 3, 19]
314496	(4, 68)	(30, 66)	[2, 3, 7, 13]
320264	(18, 68)	(32, 66)	[2, 7, 19, 43]
327763	(30, 67)	(51, 58)	[31, 97, 109]
373464	(6, 72)	(54, 60)	[2, 3, 7, 13, 19]
402597	(42, 69)	(56, 61)	[3, 13, 31, 37]
439101	(5, 76)	(48, 69)	[3, 13, 139]
443889	(17, 76)	(38, 73)	[3, 31, 37, 43]
513000	(10, 80)	(45, 75)	[2, 3, 5, 19]
513856	(34, 78)	(52, 72)	[2, 7, 31, 37]
515375	(15, 80)	(54, 71)	[5, 7, 19, 31]
525824	(24, 80)	(62, 66)	[2, 13, 79]
558441	(30, 81)	(57, 72)	[3, 13, 37, 43]
593047	(7, 84)	(63, 70)	[7, 13, 19]
684019	(51, 82)	(64, 75)	[7, 19, 37, 139]
704977	(2, 89)	(41, 86)	[7, 13, 61, 127]
805688	(11, 93)	(30, 92)	[2, 13, 61, 127]
842751	(23, 94)	(63, 84)	[3, 7, 13]
885248	(8, 96)	(72, 80)	[2, 7, 13, 19]
886464	(12, 96)	(54, 90)	[2, 3, 19]
920673	(20, 97)	(33, 96)	[3, 13, 43, 61]
955016	(24, 98)	(63, 89)	[2, 19, 61, 103]
984067	(35, 98)	(59, 92)	[7, 19, 151]
994688	(29, 99)	(60, 92)	[2, 19, 409]
1009736	(50, 96)	(59, 93)	[2, 7, 13, 19, 73]
1016496	(47, 97)	(66, 90)	[2, 3, 13, 181]
1061424	(6, 102)	(45, 99)	[2, 3, 7, 13]
1073375	(23, 102)	(60, 95)	[5, 31, 277]
1075032	(24, 102)	(76, 86)	[2, 3, 7, 79]

N	(a,b)	(c,d)	prime factors of N
1.35019E+12	(7624, 9680)	(8072, 9376)	[2, 3, 7, 13, 43, 103, 727]
1.35029E+12	(7623, 9681)	(8386, 9128)	[2, 3, 7, 19, 67, 103, 139]
1.35053E+12	(8067, 9381)	(8380, 9134)	[2, 3, 7, 139, 727, 8839]
1.35068E+12	(7301, 9870)	(8301, 9200)	[7, 11, 37, 43, 223, 1009]
1.35126E+12	(7266, 9891)	(7560, 9723)	[3, 7, 19, 31, 43, 823]
1.35151E+12	(7286, 9881)	(8234, 9257)	[7, 643, 17167, 17491]
1.35682E+12	(7911, 9516)	(8679, 8892)	[3, 7, 37, 157, 211, 5857]
1.35772E+12	(7980, 9471)	(8743, 8834)	[3, 7, 31, 271, 277]
1.36251E+12	(7638, 9715)	(8059, 9432)	[7, 37, 67, 17491]
1.36255E+12	(7635, 9717)	(8148, 9366)	[2, 3, 7, 61, 139, 241]
1.36269E+12	(7776, 9628)	(8188, 9336)	[2, 13, 19, 229, 337, 1117]
1.36354E+12	(7222, 9956)	(8236, 9302)	[2, 3, 7, 31, 37, 73, 79, 409]
1.37003E+12	(7519, 9813)	(8077, 9447)	[2, 7, 13, 337, 619, 18043]
1.37015E+12	(7518, 9814)	(8337, 9247)	[2, 7, 157, 367, 619]
1.37054E+12	(8070, 9454)	(8329, 9255)	[2, 7, 13, 157, 337, 17791]
1.37471E+12	(7443, 9873)	(8349, 9255)	[2, 3, 7, 13, 37, 163, 859]
1.3756E+12	(7436, 9880)	(7839, 9633)	[2, 3, 7, 13, 37, 1399]
1.37945E+12	(7406, 9910)	(8538, 9114)	[2, 3, 13, 37, 1471, 4513]
1.38537E+12	(7470, 9894)	(7842, 9666)	[2, 3, 7, 31, 1447, 1459]
1.38962E+12	(8006, 9570)	(8565, 9131)	[2, 7, 13, 31, 79, 1153]
1.39581E+12	(7645, 9827)	(8200, 9452)	[2, 3, 7, 13, 43, 421, 1471]
1.40614E+12	(7804, 9764)	(8888, 8896)	[2, 3, 13, 19, 43, 61, 157]
1.40851E+12	(8544, 9224)	(8802, 8990)	[2, 7, 19, 67, 139, 2221]
1.41277E+12	(7945, 9695)	(8040, 9630)	[2, 3, 5, 7, 19, 31, 37]
1.43491E+12	(7689, 9934)	(8793, 9106)	[7, 2557, 4549, 17623]
1.43607E+12	(8504, 9364)	(8798, 9106)	[2, 3, 67, 373, 1489]
1.43684E+12	(7785, 9882)	(8382, 9465)	[3, 7, 13, 31, 151, 661]
1.43853E+12	(8250, 9572)	(8298, 9536)	[2, 7, 19, 31, 37, 67, 73, 241]
1.43868E+12	(8111, 9673)	(8634, 9264)	[2, 3, 13, 19, 157, 171757]
1.45052E+12	(8016, 9780)	(8215, 9641)	[2, 3, 19, 31, 1483]
1.45125E+12	(7734, 9962)	(8436, 9476)	[2, 7, 79, 2239, 9157]
1.4547E+12	(8226, 9648)	(8850, 9132)	[2, 3, 31, 37, 73, 331]
1.46429E+12	(7964, 9862)	(8704, 9302)	[2, 3, 2281, 2971, 3001]
1.46559E+12	(7980, 9856)	(8652, 9352)	[2, 7, 13, 163, 643]
1.4677E+12	(8406, 9560)	(8588, 9414)	[2, 13, 691, 2269, 9001]
1.47735E+12	(8020, 9870)	(8830, 9240)	[2, 5, 13, 139, 457, 1789]
1.4775E+12	(8399, 9601)	(8730, 9330)	[2, 3, 5, 7, 43, 90901]
1.48078E+12	(8524, 9515)	(9011, 9082)	[3, 7, 13, 37, 163, 349, 859]
1.49413E+12	(8182, 9818)	(8400, 9660)	[2, 3, 5, 7, 43, 67]
1.51672E+12	(8040, 9990)	(8325, 9795)	[2, 3, 5, 151, 601, 619]
1.5171E+12	(8124, 9936)	(9117, 9123)	[2, 3, 5, 7, 19, 43, 30703]
1.535E+12	(8711, 9561)	(8882, 9414)	[2, 109, 337, 571, 2287]
1.55733E+12	(8457, 9839)	(9108, 9290)	[2, 19, 487, 2287, 9199]
1.58672E+12	(8400, 9980)	(8988, 9512)	[2, 5, 19, 37, 307, 919]
1.5983E+12	(8446, 9986)	(8880, 9648)	[2, 3, 193, 37441]
1.63354E+12	(8807, 9832)	(8871, 9780)	[3, 19, 37, 109, 127, 6217]
1.63477E+12	(8712, 9911)	(9119, 9570)	[7, 11, 61, 1693, 1699]
1.63567E+12	(8934, 9735)	(9127, 9566)	[3, 7, 31, 43, 67, 109, 127]
1.64087E+12	(8667, 9966)	(8971, 9722)	[3, 7, 31, 67, 673, 6211]
1.6409E+12	(8744, 9907)	(9328, 9395)	[3, 37, 79, 127, 6217]

II. Table of experimental data points

Upper bound for a, b, c, d	**Number of taxicab numbers**	**Number of primitive taxicab numbers**	**Proportion that are primitive**
50	12	7	0.583
100	45	27	0.600
200	135	63	0.467
300	245	116	0.457
400	392	167	0.426
500	563	224	0.398
600	736	300	0.408
700	943	375	0.398
800	1128	443	0.393
900	1362	535	0.393
1000	1590	624	0.392
2000	4358	1583	0.363
3000	7788	2737	0.351
4000	11687	3973	0.340
5000	15,935	5296	0.332
6000	20525	6685	0.326
7000	25,475	8215	0.322
8000	30602	9775	0.319
9000	35967	11392	0.317
10,000	41,563	13027	0.313

References

[1] Mathematica and Wolfram Alpha, *Taxicab Number, MathWorld*. 2024.

[2] Mathematica and Wolfram Alpha, *Bézout's Theorem, MathWorld*. 2024.

[3] Mathematica and Wolfram Alpha, *Vieta's Formulas, MathWorld*. 2024.