Hang Lung Mathematics Awards
Vol. 5 (2012)

Hang Lung Mathematics Awards 2012

Gold Award

Towards Catalan's Conjecture

Team member: Chung Hang Kwan
Teacher: Mr. Yat Ting Tong
School: Sir Ellis Kadoorie Secondary School (West Kowloon)

Hang Lung Mathematics Awards
Vol. 5 (2012), pp. 1–23
© 2012, IMS, CUHK

TOWARDS CATALAN'S CONJECTURE

TEAM MEMBER
Chung Hang Kwan

TEACHER
Mr. Yat Ting Tong

SCHOOL
Sir Ellis Kadoorie Secondary School (West Kowloon)

Abstract. The presented project aims at having an insight on one of the most famous, hard but beautiful problems in number theory—Catalan's Conjecture (This conjecture has become a theorem in 2002). Throughout this project, very advanced techniques and results established are avoided. Most of the results established in this report only require the concepts in elementary number theory, for example: divisibility and congruence. Yet, these techniques can be used delicately to establish a number of particular cases.

1. Introduction

Catalan's Conjecture is a four-variable Diophantine equation. At the first glance, starting to develop approaches from the original statement would be extremely difficult. Also, it is difficult to get a rough idea or insight towards the problem. Under this situation, the first problem-solving strategy should be "looking at simpler cases", i.e. reducing the number of constraints and variables in the original conjecture. However, solving the particular cases is not completely an easy task, a number of techniques have to be developed and they have to be used wisely.

The project is divided into 3 stages, throughout the progress of the stages, stronger and more complicated results are obtained.

Stage 1: Elementary techniques in solving Diophantine equations are used to obtain several weaker starter problems, such as $3^m - 2^n = 1$ and $x^2 - 2^y = 1$

Stage 2: Applying merely the same methods in Stage 1 to the three-variable cases, the desired results cannot be obtained conveniently. Here, several helpful theorems are introduced, such as Zsigmondy's theorem.
Results like $x^r - 1 = p^n$ (p is a prime) are derived.

Stage 3: The discussions on Catalan's Conjecture are extended to generalized form, i.e. $x^r - y^s = n$ (n is a positive integer). The directions of my future studies on this topic are introduced.

2. Investigation Background

Before stating Catalan's Conjecture, perfect power (of natural numbers) is first defined.

Definition 1 (**perfect power**). *A natural number is called a perfect power if it can be written in the form a^b, where a and b are natural numbers with $b \geq 2$.*

The first several terms of perfect power sequence are:

$$1, 4, 8, 9, 16, 25, 27, 32, 36, 49, 64, 81, 100, 121, 125, \ldots$$

Conjecture 2 (**Catalan's Conjecture**). ***The only two consecutive perfect powers are 8 and 9***

Equivalently, it can be stated in the argument of Diophantine euqation:
The only positive integral solution (x, y, r, s) with $x, y, r, s \geq 2$ of the Diophantine equation $x^r - y^s = 1$ is (3,2,2,3)

This was conjectured by French and Belgian mathematician Eugène Charles Catalan in 1844. For the later 160 years, numerous results using the concepts in advanced number theory were established. Some of the remarkable results are:

1. The Diophantine equation $x^p + 1 = y^q$, where p and q are primes with $p \neq 2$ has positive integral solution if and only if both of the following conditions hold[1]:
   (1) $x + 1 = p^{sq-1}y_1^q$, $\frac{x^p+1}{x+1} = py_2^q$, $y = p^s y_1 y_2$, where $\gcd(y_1, y_2) = 1$, $p \nmid y_1 y_2$
   (2) $y - 1 = q^{tp-1}x_1^p$, $\frac{y^p-1}{y-1} = qx_2^p$, $x = q^t x_1 x_2$, where $\gcd(x_1, x_2) = 1$, $q \nmid x_1 x_2$
2. (**KE Zhao**[2], 1965) The Diphantine equation $y^2 - 1 = x^p$, where p is an odd prime, has no positive integral solution if $p \geq 4$.[3]

[1] A result from the Chinese book 'About Indeterminate Equation'(談談不定方程), written by Ke Zhao(柯召) and Sun Qi(孫琦).
[2] Ke Zhao (柯召) (1910-2002) is a Chinese mathematician.
[3] A result from the book 'About Indeterminate Equation'.

3. (**Preda Mihăilescu**[4], 2000-2003) Consider the Diophantine equation $x^p - y^q = 1$ where p and q are odd primes. We have[5]

$$\begin{aligned} p^{q-1} &\equiv 1 \pmod{q^2} \\ q^{p-1} &\equiv 1 \pmod{p^2} \end{aligned} \tag{1}$$

$$\begin{aligned} p &\equiv 1 \pmod{q} \\ q &\equiv 1 \pmod{p} \end{aligned} \tag{2}$$

$$\begin{aligned} p &< 4q^2 \\ q &< 4p^2 \end{aligned} \tag{3}$$

4. Let p and q be odd prime. For $p \leq 41$ or $q \leq 41$, the equation $x^p - y^q = 1$ has no positive integral solution.

Preda Mihăilescu used the results in 3.(1), (2) and (3), as well as 4, to prove that $x^p - y^q = 1$ has only one solution $(x, y, p, q) = (3, 2, 2, 3)$. This leads to the complete proof of Catalan's Conjecture in 2001 (published in 2002). A conjecture has become a theorem!

---

[4]Preda Mihăilescu is a Romanian mathematician (1955-), who gave the complete proof of Catalan's Conjecture in 2002.

[5]3 and 4 are the result stated in the book 'Catalan's Conjecture', written by Rene Schoof and was published in 2008 by Springer (Universitext series). It is a complete historical note on the researches and developments of Catalan's Conjecture.

3. Motivation

When I first saw Catalan's Conjecture, I listed out the perfect powers in ascending order, $1, 4, 8, 9, 16, 25, 27, 32, 36, 49, 64, 81, 100, 121, 125, 128, 144\ldots$ I observed that the difference between the consecutive terms keep increasing and decreasing. I thought if we consider sufficiently large number of perfect powers, there should be a great chance to have two consecutive perfect powers with difference 1 other than (3,2,2,3).

However, the result in 2002 told us that is impossible, and (3,2,2,3) is the unique solution. Therefore, this result is very amazing in my point of view. I decided to carry out investigation in this topic. After reading some parts of the book **Catalan's Conjecture** [8], I knew that this topic is really deep and complicated. Many profound results have been established. But this should not be the end of my attention on this classical problem. I have not seen a number of special cases being proved in the book, as well as the Internet.

As a result, I decided to use my knowledge on elementary number theory to solve some of the particular cases which was developed in my mind in order to gain a deeper understanding on Catalan's Conjecture

I first consider the special solution $3^2 - 2^3 = 1$ and the following observations were made:

1. 2 and 3 are consecutive numbers, i.e. we have consecutive bases and exponents in the particular solution.
2. 2 and 3 are both prime numbers.

These observations were helpful in making a number of special cases, which will be discussed in the following pages.

Notation Note: All of the solution described in each equation are positive integers greater than 1 (without otherwise specified).

4. Stage 1

In this section, the elementary techniques in solving Diophantine equations are employed. The examples of these techniques are factorization, modular arithmetic and inequality bounding. Here, factorization and modular arithmetic are the most useful techniques. They are very effective to establish two-variable special cases.

Using modular arithmetic is an effective way to prove a Diophantine equation has no integral solution or a solution in certain form is not a solution. This can be done by taking congruence of certain natural number to both sides of the equation, showing they are unequal, contradiction arises.

However, only $3^2 - 2^3 = 1$ is a solution, is taking congruence to both sides of the equations (special cases) not suitable? The answer is negative. Although it cannot help reducing the range of all possible solutions, it is helpful to determine the odd-even parity of certain variables.

Theorem 3. *The equation $3^m - 2^n = 1$ has only one positive integral solution (m, n), where both m and n are larger than 1.*

Lemma 4. *If $3^a \equiv 1 \pmod{8}$, then a is a positive even integer.*

Proof. We consider the contrapositive of the statement 'If a is a positive odd integer, then $3^a \not\equiv 1 \pmod{8}$'.
Let $a = 2k+1$, where k is a non-negative integer.
$3^a = 3^{2k+1} = 3(9^k) \equiv 3(1^k) \equiv 3 \not\equiv 1 \pmod{8}$ □

Proof of Theorem 3. Rewrite the equation as $3^m = 1 + 2^n$. When $n \geq 3$, taking (mod 8) to both sides of the equation, we have $3^m \equiv 1 \pmod{8}$. By the above lemma, m is a positive even integer. Let $m = 2k$, where k is a positive integer. Then

$$2^n = 3^m - 1 = 3^{2k} - 1 = (3^k + 1)(3^k - 1)$$

As LHS is a prime factorization, we have

$$\begin{cases} 3^k + 1 = 2^a \\ 3^k - 1 = 2^b \end{cases} \tag{*}$$

where a and b are some positive integers with $n = a + b$ (Note: a and b cannot be zero, otherwise giving non-integral k).

Subtract the two equations, we have $2^a - 2^b = 2$. Clearly, there is only one such pair, i.e. $(a, b) = (2, 1)$. Substituting back into (*), we have $k = 1$. Hence, $m = 2$ and $n = 3$ is a solution of the equation.
Clearly, the case of $n = 2$ gives no positive integral solution.
$(m.n) = (2.3)$ is the only solution to the equation. □

Similarly, we can establish the following theorems:

Theorem 5. *The equation $x^2 - 2^y = 1$ has only one positive integral solution.*

Theorem 6. *The equation $3^x - y^3 = 1$ has only one positive integral solution.*

Proof of Theorem 5. $2^y = x^2 - 1 = (x+1)(x-1)$. Then

$$\begin{cases} x+1 = 2^a \\ x-1 = 2^b \end{cases}$$

where a and b are positive integers with $y = a + b$.
This implies $2^a - 2^b = 2$ again. So, $(a, b) = (2, 1)$ and $(x, y) = (3, 3)$. □

Proof of Theorem 6. $3^x = y^3 + 1 = (y+1)(y^2 - y + 1)$

$$\begin{cases} y+1 = 3^a \\ y^2 - y + 1 = 3^b \end{cases}$$

where a and b are positive integers with $x = a + b$.
Substituting the first equation into the second one, we have

$$(3^a - 1)^2 - (3^a - 1) + 1 = 3^b$$
$$3^{2a} - 3^{a+1} + 3 = 3^b.$$

When $a \geq 1$ and $b \geq 2$, taking (mod 9) to both sides of the equation, we have $3 \equiv 0 \pmod 9$, which is impossible.
So, the case left to check is $(a, b) = (1, 1)$. This gives $y = 2$ and hence $x = 2$
i.e. The only positive integral solution is $(x, y) = (2, 2)$ □

We have seen that the two-variable cases can be established conveniently through elementary number theory, except $x^2 - y^3 = 1$. Although we can factorize $x^2 - 1$ or $y^3 + 1$, we do not have a prime factorization of the remaining expression. My idea towards this equation will be described later.

After discussing the two-variable cases, it is natural for us to extend the above approaches to three-variable cases. However, the situation becomes more complicated and the method of factorization is different.

Theorem 7. *The equation $p^x - y^p = 1$ has only one positive integral solution (p, x, y), where p is a prime and x, y are greater that 1.*

Proof. For $p = 2$, we have $y^2 = 2^x - 1$.
If $x \geq 3$, then $y^2 \equiv -1 \pmod 8$. But $y^2 \equiv 0, 1 \pmod 8$ for any integers y.
Contradiction.
Clearly, $x = 2$ gives no solution.
Therefore, $p = 2$ does not give any solution.

For odd prime p, we have

$$p^x = 1 + y^p = (1+y)(1 - y + y^2 - \ldots + y^{p-1})$$

So, $$\begin{cases} 1+y=p^a \\ 1-y+y^2-\ldots+y^{p-1}=p^b \end{cases}$$
where a and b are positive integers with $x=a+b$

Substituting the first equation into the second one,

$$1-(p^a-1)+(p^a-1)^2-\ldots+(p^a-1)^{p-1}=p^b$$

Expanding LHS and collecting all the 1's, we have

$$p-p^a+(p^{2a}-2p^a)+\ldots+(p^{a(p-1)}-\ldots-(p-1)p^a)=p^b$$

Clearly, LHS is divisible by p. After dividing both sides by p, we have

$$1-p^{a-1}+(p^{2a-1}-2p^{a-1})+\ldots+(p^{a(p-1)-1}-\ldots-(p-1)p^{a-1})=p^{b-1} \qquad (*)$$

If $a\geq 2$ and $b\geq 2$, then LHS is not divisible by p but RHS is divisible by p. Contradiction.

If $a=1$ and $b\geq 2$, then LHS of (*) becomes

$$1-\frac{p(p-1)}{2}+p+\ldots+\Big(p^{p-2}-\ldots+\frac{(p-1)(p-2)}{2}p\Big)$$

Except $1-\frac{p(p-1)}{2}$, all the other terms are divisible by p. Also, RHS is also divisible by p.
So, for odd prime p such that $p\nmid 1-\frac{p(p-1)}{2}$, contradiction occurs.
$p\mid 1-\frac{p(p-1)}{2}$ if and only if $p(2m+p-1)=2$ for some integer m. Then, we must have $p=2$. This contradicts with our assumption that p is an odd prime.
So, the case of $a=1$ and $b\geq 2$ yields no solution.

We claim that $1-y+y^2-\ldots+y^{p-1}\geq 1+y$ (**)

i.e. $b\geq a$
This is equivalent to $\frac{1+y^p}{1+y}\geq 1+y$ or $y^{p-1}-y-2\geq 0$.
Let $f(y)=y^{p-1}-y-2$.
Then $f'(y)=(p-1)y^{p-2}-1$ and $f''(y)=(p-1)(p-2)y^{p-3}$.
$f'(y)=0$ if and only if $y=(\frac{1}{p-1})^{\frac{1}{p-2}}$.

By second derivative test, f attains minimum at $y=(\frac{1}{p-1})^{\frac{1}{p-2}}$. Clearly, $(\frac{1}{p-1})^{\frac{1}{p-2}}<2$. So f is monotonic increasing for $y\geq 2$.
i.e. $f(y)\geq f(2)=2^{p-1}-4\geq 0$, where equality holds if and only if $y=2$ and $p=3$.
The claim is proved. The case of $a\geq 2$ and $b=1$ yields no solution.
The case left is $(a,b)=(1,1)$. This gives the equality of (**). So, $y=2$ and $p=3$.
It is easy to check that $(p,x,y)=(3,2,2)$ is a solution of the equation. □

I have tried to apply the method used in **Theorem 7** to prove several similar, but more general cases of Catalan's Conjecture.

One of the example is
The equation $x^m - 1 = p^n$ has only one positive integral solution (p, x, m, n), where p is a prime, x, m and n are positive integers greater than 1.

Idea
Factorizing LHS first, we have

$$p^n = x^m - 1 = (x-1)(1 + x + x^2 + \ldots + x^{m-1})$$

Then
$$\begin{cases} x - 1 = p^a \\ 1 + x + x^2 + \ldots + x^{m-1} = p^b \end{cases}$$
where a and b are non-negative integers with $a + b = n$
Clearly, $b \neq 0$

If $a \neq 0$, then by the similar arguments in the proof of **Theorem 7**, we can only have $p \mid m$. The odd-even parity of the variables can neither be derived. This result is not sufficient to give a complete proof of this equation. The method in Stage 1 fails for a number of cases of Catalan's Conjecture.
Similarly, the method does not work for $3^n - x^y = 1$.

Therefore under this situation, it is forced to start other approaches to prove more strong and complicated results. [See reviewer's comment (3) and (4)] At some time I believed that there is no simpler method. However, after I start to work on Stage 2, I re-thought about the elementary proofs of some complicated results, and I obtained one of them successfully and a very useful idea. Because the approaches involved are similar to that in this section, I attached the proof here.

Theorem 8. *The equation $x^y - 2^n = 1$ has only one positive integral solution (x, y, n), where x, y and n are positive integers greater than 1.*

Proof. Rewrite and factorize the equation:

$$2^n = x^y - 1 = (x-1)(1 + x + x^2 + \ldots + x^{y-1})$$

If $x - 1 = 1$ or $x = 2$, then $y = 1$ and $n = 0$.
This does not give a desired solution.

If $x - 1 = 2^a$, where a is a positive integer, then both sums in the two brackets of RHS are even. Also x is a positive odd integer.

For the sum in the second brackets of RHS to be even, the number of terms y in the second brackets has to be even.
Let $y = 2k$, where k is a positive integer. Then, we have

$$2^n = x^y - 1 = x^{2k} - 1 = (x^k + 1)(x^k - 1)$$
$$\begin{cases} x^k + 1 = 2^b \\ x^k - 1 = 2^c \end{cases}$$

where b and c are positive integers.
Subtract both of the equations, we have $2^b - 2^c = 2, (b, c) = (2, 1)$.
So, $x^k = 3$. This implies $x = 3$ and $k = 1$. We also have $y = 2$ and $n = 3$.
Therefore, the equation has only one solution $(x, y, n) = (3, 2, 3)$ □

From the proof of **Theorem 3** and **Theorem 8**, proving r is even in the equation $x^r - y^s = 1$, where x, y, r, s can be variables or constants, is a convenient way to give an elementary proof. The same strategy is employed in the proof of **Theorem 9**. In fact, **Theorem 9** is the most tedious case I have encountered when using elementary methods.

Theorem 9. *The equation $(x+1)^y - x^z = 1$ has only one positive integral solution (x, y, z), with x, y, z larger than 1*

Idea
Since $1 = (x + 1)^y - x^z \equiv -(-1)^z \pmod{x + 1}$, z is odd.
Then,

$$(x + 1)^y = 1 + x^z = (1 + x)(1 - x + x^2 - \ldots + x^{z-1})$$
$$(x + 1)^{y-1} = 1 - x + x^2 - \ldots + x^{z-1}$$

If x is odd, then $(x + 1)^{y-1}$ is even, but $1 - x + x^2 - \ldots + x^{z-1}$ is odd.
Contradiction.
Hence, x is even.

Moreover, $x^z = (x + 1)^y - 1 = x(1 + (x + 1) + (x + 1)^2 + \ldots + (x + 1)^{y-1})$

$$x^{z-1} = 1 + (x + 1) + (x + 1)^2 + \ldots + (x + 1)^{y-1}$$

Since x is even, x^{z-1} is even. $1 + (x + 1) + (x + 1)^2 + \ldots + (x + 1)^{y-1}$ is also even.
So, the number of terms y in the sum must be even.
Therefore, we have

$$((x + 1)^k + 1)((x + 1)^k - 1) = x^z \qquad (*)$$

where $k = \frac{1}{2}y$ is a positive integer.

If k is odd, then the first factor in (*) can be further factorized into:

$$(x + 2)[1 - (x + 1) + (x + 1)^2 - \ldots + (x + 1)^{k-1}]$$

Then, we must have $x + 2 \mid x^z$

Originally, I thought there are only small number of cases (of x) for this relation to hold.

Note that $gcd(x, x + 2) = 2$ as x is even. Hence,

$$\begin{cases} x = 2^m p \\ x + 2 = 2q \end{cases} \quad \text{or} \quad \begin{cases} x = 2p \\ x + 2 = 2^m q \end{cases}$$

where m is a positive integer greater than $2, p$ and q are coprime odd positive integers.

Since $x+2 \mid x^z, gcd(x^z, x+2) = x+2$.
For the first case, we have $gcd(x^z, x+2) = 2$. So $x = 0$, which is impossible.
For the second case, we have $gcd(x^z, x+2) = 2^{\min(m,z)}$.
On the other hand, we have $gcd(x^z, x+2) = x+2 = 2^m q$.
Therefore, $q = 1$.

However, I still have not solved the second case, and I found there exist infinitely many x (any x in the form $2^m - 2$ and $z \geq m$) such that $x+2 \mid x^z$. This is the major difficulty I have to overcome. Note that a sound proof will be given in **Stage 2**.

However, I thought the idea of considering the G.C.D. between two factors can be useful, but minute details have to be considered.

I have tried to apply this idea to the equation $x^2 - y^3 = 1$
Rewrite the equation as

$$y^3 = x^2 - 1 = (x+1)(x-1)$$

$gcd(x+1, x-1) = 1$ if x is even
Therefore,
$$\begin{cases} x+1 = p^3 \\ x-1 = q^3 \end{cases}$$
where p and q are positive odd coprime integers.
Then, subtract the two equations and by simple fatorization, the system gives no solution.

$gcd(x+1, x-1) = 2$ if x is odd.
Therefore,
$$\begin{cases} x+1 = 2^{3a+2}p^3 \\ x-1 = 2q^3 \end{cases}$$
where p and q are positive odd coprime integers, a is a non-negative integer.
Then, we have

$$2^{3a+1}p^3 - q^3 = 1$$

Putting $r = 2^a p$, then the equation becomes $2r^3 - q^3 = 1$

Then, a natural question to be asked is that ‘ Is the last equation much easier to be proved than the original equation ?’. I believe the answer is ‘yes’. The RHS of the last equation is a homogeneous polynomial of degree 3, and it can be factorized in domain other than integral one. However, this involved higher knowledge of algebra to do so. After I read some books of abstract algebra, I still cannot give a complete proof.

But it is known that the last equation will not have infinitely many integral solutions. This can be deduced from a result from Axel Thue (1909):

Theorem 10 (Thue's Theorem[6]). *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ be an irreducible polynomials in the integral (i.e. rational) domain and its degree is at least 3. Then consider the Diophantine equation*

$$f(x,y) = a_n x^n + a_{n-1}x^{n-1}y + \ldots + a_1 x y^{n-1} + a_0 y^n = m$$

where m is a non-zero integer.
It has only finitely many integral solutions.

Clearly, $f(x) = 2x^3 - 1$ is irreducible in integral domain, so by Thue's theorem, $2r^3 - q^3 = 1$ has only finitely many solution.

Further investigations are required in the scope of using the factorization in other domains to solve several special cases of Catalan's Conjecture.

[6]based on the book 'An Introduction to Diophantine Equation' (p.103) written by Titu Andreescu and Dorin Andrica, published by Gil Publishing House in 2002.

5. Stage 2

There are still many cases we have not proved yet. Moreover, as we have seen in the previous section, solving the three-variable equations can be painstaking. So, are there any simpler methods that can help us to solve some of the Diophantine equations, especially one with exponential variables? Fortunately, the answer is definite in the special cases of Catalan's Conjecture.

One of the key features in Catalan's Conjecture is that its Diophantine equation can be re-written to involve the expressions $1 \pm x^n$, or more generally $a^n \pm b^n$. These expressions contribute to several important properties in number theory in which they are very useful in solving Diophantine equations.

Definition 11. *We define the function $\upsilon_p(x)$ to be the largest power of prime p in the prime factorization of positive integer x.*
In other words, $a = \upsilon_p(x)$ if $p^a \mid x$ and $p^{a+1} \nmid x$

Basic properties

1. $\upsilon_p(xy) = \upsilon_p(x) + \upsilon_p(y)$
2. $\upsilon_p(x + y) \geq min(\upsilon_p(x), \upsilon_p(y))$
3. $\upsilon_p(x^y) = y\upsilon_p(x)$

Lemma 12 (**Lifting the Exponent (LTE) Lemma**)**.** [7]

(i) *let x and y be integers, n be a positive integer and p be an odd prime such that $p \mid x - y$, but $p \nmid x, y$*
Then $\upsilon_p(x^n - y^n) = \upsilon_p(x - y) + \upsilon_p(n)$

(ii) *Let x and y be integers, n be an odd positive integer and p be an odd prime such that $p \mid x + y$, but $p \nmid x, y$*
Then $\upsilon_p(x^n + y^n) = \upsilon_p(x + y) + \upsilon_p(n)$

(iii) *Let x and y be odd integers such that $4 \mid x - y$, n be a positive integer.*
Then $\upsilon_2(x^n - y^n) = \upsilon_2(x - y) + \upsilon_2(n)$

(iv) *Let x and y be odd integers and n be an even positive integer.*
Then $\upsilon_2(x^n - y^n) = \upsilon_2(x - y) + \upsilon_2(x + y) + \upsilon_2(n) - 1$

[7] There are still some parts of **LTE Lemma** being omitted because they are irrelevant to the main theorems in this project. For more information, please visit http://www.artofproblemsolving.com/Resources/Papers/LTE.pdf, written by Amir Hossein Parvardi. The proof can also be found there. Note that the notation in the **Definition** follows that in the website

LET Lemma (i) and (ii) are the most useful. They can usually reduce the number of possibilities of certain variables greatly.

The idea and the proof of **LET Lemma** is simple and elementary, just using congruence and basic factorization, but its field of application is very wide. Firstly, we return back to **Theorem 7**. A much simpler proof will be given as follow:

Rewrite the equation as: $p^x = 1 + y^p$ (*)

For odd prime p, we have $p^x = 1 + y^p = (1 + y)(1 - y + y^2 - \ldots + y^{p-1})$
Clearly, $p \mid 1 + y$, $p \nmid 1$ and $p \nmid y$ (Otherwise if $p \mid y$, then $0 \equiv 1 \pmod p$, Contradiction)
By LTE Lemma (ii),

$$\upsilon_p(p^x) = \upsilon_p(1 + y^p) = \upsilon_p(1 + y) + \upsilon_p(p) = \upsilon_p(1 + y) + 1$$
$$x = \upsilon_p(1 + y) + 1$$

Hence
$$1 - y + y^2 - \ldots + y^{p-1} = p$$

Or
$$\frac{1 + y^p}{1 + y} = p$$
$$y^p - py + (1-p) = 0$$

Let $f(y) = y^p - py + (1 - p)$. Then $f'(y) = py^{p-1} - p$ and $f''(y) = p(p - 1)y^{p-2}$.
For positive y, $f'(y) = 0$ if and only if $y = 1$.
By second derivative test, f attains minimum at $y = 1$. So, f is monotonic increasing for $y \geq 1$. But y is a positive integer larger than 1.

So, $f(y) \geq f(2) = 2^p - 3p + 1$. Clearly, for $p \geq 3$, the rightmost expression is greater than or equal to 0.
$f(y) = 0$ holds if and only if $p = 3$ and $y = 2$.

Substituting $p = 3$ and $y = 2$ into original equation (*), we have $x = 2$.

The part of $p = 2$ is essentially the same as the proof in Stage 1. **LTE Lemma (iii) and (iv)** cannot be used here. □

However, I have found the other much useful and much simpler lemma to prove much diverse results after I used LTE Lemma for some time. Again, it is based on the expression $a^n \pm b^n$. We can prove that equation does not hold for large number of cases, and we only have to check for the exceptions (bad pairs) in the following lemma:

Lemma 13 (**Zsigmondy's theorem).** [8]

[8] The Zsigmondy's theorem, Pisolve, Art of Problem Solving (mathlinks)
http://www.artofpromblemsolving.com/Forum/download/file/php?id=34511
Zsigmondy's theorem, Wikipedia
http://en.wikipedia.org/wiki/Zsigmondy%27s_theorem

1. *If a, b and n are positive integers, where $a > b$, $gcd(a, b) = 1$ and $n \geq 2$, then $a^n - b^n$ has at least one prime factor which does not divide $a^k - b^k$ for any positive integer $k < n$,*
   with the following exception:
   (i) $a = 2, b = 1$ *and* $n = 6$.
   (ii) $n = 2$ *and* $a + b$ *is a power of 2.*
2. *If a, b and n are positive integers, where $a > b$ and $n \geq 2$, then $a^n + b^n$ has at least one prime factor which does not divide $a^k + b^k$ for any positive integer $k < n$,*
   with the exception: $a = 2$ and $b = 1$

The proof of **Theorem 7** can be much simpler (than using LTE Lemma):

$$p^x = 1 + y^p$$

Note that $x > 1$ and $y > 1$.
By **Zsigmondy's theorem**, there exists one prime factor of $1 + y^p$ which does not divide $1 + y$, with the exception $y = 2$ and $p = 3$.
However, the only prime factor of $1 + y^p$ is p based on the equation. The prime factor not dividing $1 + y$ is p.
So, for the general case[9] (i.e. not the exception case), $1+y = 1$, which is impossible.
So, we are left to consider the case of $y = 2$ and $p = 3$, which gives $x = 2$ □

For the proof of **Theorem 8**, it can be done as follows:
Rewrite the equation as $x^y - 1 = 2^n$.
Clearly, x and 1 are coprime and $y > 1$. Similar to the argument of the proof of **Theorem 7** using **Zsigmondy's theorem**, 2 is the prime factor of $x^y - 1$ not dividing $x - 1$ for the general case. So, $x - 1 = 1$ or $x = 2$. But it is easy to verify that this does not yield desired solution.
For the exception cases,

(i) $x = 2$ and $y = 6$. This pair of (x, y) does not yield integral n.

(ii) $y = 2$ and $x + 1 = 2^m$, where m is a positive integer. This implies

$$(2^m - 1)^2 - 1 = 2^n$$

i.e.

$$2^{m+1}(2^{m-1} - 1) = 2^n$$

$2^{m-1} - 1$ must be a power of 2 (including 1). So, $m = 2$. This gives the solution $(x, y, n) = (3, 2, 3)$.

□

Another advantage of **Zsigmondy's theorem** over the factorization method in **Stage 1** is that we do not have to determine the odd-even parity of the variables, in which these processes can be painstaking.

[9]The term 'general case' will also be used in the later proofs.

Recall that we have not solved the equation $3^n - x^y = 1$ in **Stage 1**. I have failed to use **LTE Lemma** to solve it in the early stage of **Stage 2**. However by using the technique of **Zsigmondy's theorem** the proof becomes easy.

Theorem 14. *The equation $3^n - x^y = 1$ has only one positive integral solution (x, y, n), where x, y and n are greater than 1.*

But in fact, we can generalize the statement of **Theorem 14** very easily, by replacing the '3' in the equation by a prime 'p'. The proof is essentially the same as that of **Theorem 14**.

Theorem 15. *The equation $p^n - x^y = 1$ has only one positive integral solution (p, n, x, y), where x, y, n are positive integers larger than 1 and p is a prime.*

Proof. Rewrite the equation as $p^n = 1 + x^y$.
With the similar arguments as before, p is the prime factor of $1 + x^y$ not dividing $1 + x$. This case is impossible to occur.
The only exception left is $x = 2$ and $y = 3$. Clearly, this yields $p = 3$ and $n = 2$. □

Naturally, we would also want to generalize **Theorem 8**:

Theorem 16. *The equation $x^y - p^n = 1$ has only one positive integral solution (p, n, x, y), where x, y, n are positive integers larger than 1 and p is a prime.*

Proof. Rewrite the equation as $x^y - 1 = p^n$.
By **Zsigmondy's theorem**, p is the prime factor $x^y - 1$ not dividing $x - 1$ for the general case, so $x = 2$. The equation becomes:

$$2^y - 1 = p^n$$

or

$$2^y = 1 + p^n$$

We use **Zsigmondy's theorem** again to get $2 \nmid 1 + p$ for the general case. So, $p = 2$. But this does not yield any desired solution.
Checking the exceptional case $p = 2$ and $n = 3$. This case also gives no solution.

Then, we are left to consider the exceptional case of $x^y - 1$.

(i) $x = 2$ and $y = 6$. This case gives no solution.

(ii) $y = 2$ and $x + 1 = 2^m$, where m is a positive integer.

$$(2^m - 1)^2 - 1 = p^n$$

This is equivalent to

$$2^{m+1}(2^{m-1} - 1) = p^n$$

Clearly, $2 \mid p^n$. This implies $2 \mid p$, i.e. $p = 2$.
The remaining part is essentially the same as that in the proof of **Theorem 8**

using **Zsigmondy's theorem**.
Therefore, $(p, n, x, y) = (2, 3, 3, 2)$. $\square$

We can also prove **Theorem 9** (the most complicated case encountered in **Stage 1**) by using **Zsigmondy's theorem**.

Resrite the equation as

$$(x+1)^y = 1 + x^z$$

For the general case, there exists a prime p such that $p \mid 1 + x^z$ but $p \nmid 1 + x$.
Then RHS is divisible by p but LHS does not. Contradiction.
So, it suffices to consider the case $x = 2$ and $z = 3$.
This gives $y = 2$. $\square$

6. Stage 3

After having a look on the book 'Catalan's Conjecture' [8] and 'Primary cyclotomic' units and a proof of Catalan's conjecture[10], I understand that the proof of four-variable Catalan's equation without any constraints on the variables involves a lot of advanced techniques and numerous deep results have been done before the complete proof came out in 2002.

In fact, in my project, the results $p^n - x^y = 1$ and $x^y - p^n = 1$ were the best results I have established in some sense (four-variable with one constraint)— although they are not the hardest to be proved. In my opinion, further removing the constraint 'p is a prime' or replacing it with a stronger condition will be much difficult to do, possibly beyond the scope of elementary number theory. They cannot be done by using the above methods. For example: though **Zsigmondy's theorem** is a really useful result in my project, it depends much on prime factorization and the expression $a^n \pm b^n$.

However, I was still very interested in this amazing conjecture. I wanted to investigate more on the case on the general Catalan's equation, i.e. replacing the '1' in the Catalan's equation by arbitrary positive integer n. Definitely, it is much harder to solve or even using the techniques in the previous stages to investigate on the special case.

In fact, '1' is a really amazing integer chosen in the original conjecture, as well as it is really important in my research procedures. It is because Catalan's equation can be written to involve the expression $1 + x^n$ or $1 - x^n$. The former one can be factorized for any odd positive integer, while the later one can be factorized for any positive integer. After factoring the expressions, we may consider the prime factorization.

Moreover, for example: the equation $5^n - 3^b = 2$ cannot be solved using the methods in **Stage 1** because we can deduce both a and b are odd easily. (It is well-known that it can be solved by **Pell's equation**[11]).

Of course, we have to add some constraints in the study of this section. The results are relatively weak compared to the previous sections.

The first question to the generalized Catalan's conjecture is 'Are there any n such that the Catalan's equation has infinitely many positive integral solutions (greater than 1) (x, y, r, s) ?'. I believe that it is quite difficult to answer. Perhaps it is related to **abc conjecture**, which is still not yet proven:

---

[10] written by Preda Mihăilescu, on Kuwait Foundation Lecture 30, please visit https://www.dpmms.cam.ac.uk/seminars/Kuwait/abstracts/L30.pdf

[11] The proof can be found on http://www.math.ust.hk/~makyli/190_2010Sp/190_LectNt-rev.pdf p.71

Conjecture 17. ***abc conjecture***[12] ***(Oesterlé-Masser conjecture)*** *For every $\epsilon > 0$, there exist only finitely many triples (of course can be none) (a, b, c) of coprime positive integers such that*

$$a + b = c$$

and

$$c > [rad(abc)]^{1+\epsilon}$$

where $rad(n)$ is defined by:

Definition 18. *Let $p_1^{a_1} \dots p_k^{a_k}$ be the prime factorization of n will all a_i's not equal to 0. Then $rad(n) = p_1 \dots p_k$.*

If abc conjecture is proven to be true, then the answer to the question is 'no'. However, abc conjecture is an extremely comlicated proposition in number theory– although it is highly related to many famous theorems or conjectures[13] in number theory.

The other question is 'Are there any n such that the Catalan's equation has no positive integral solution larger that 1 ?'. Again, it is difficult to answer.

However, considering weaker statements, for example: fixing r and s to be some positive integers, say 2 and 3 respectively. This is a well-known solved problem and the equation is called Mordell's equation[14], i.e. $y^2 - x^3 = n$[15] (fixed positive integer n). It has been proven that Mordell's equation only has **finitely many** integral solutions for any n.

Moreover, is there any n that gives no solution to Mordell's equation ? The answer is also yes and it is well-known that there are various form of n giving no solution and the proofs are elementary, just using the property of congruence and Legendre symbol. For example: $n = (4b - 1)^3 - 4a^2$, where a and b are integers such that a does not have prime factor $p \equiv -1 \pmod 4$[16]

As I have mentioned before, the following theorems established are weak. I was wondering about whether there are n's giving no solution to generalized Catalan's theorem (special case).
For example:

[12]based on the epymt lecture notes (Number Theory and Cryptography) written by Prof. Li Chun Che Charles.

[13]For example: Fermat's Last Theorem, of course Catalan's Conjecture.

[14]With reference to the Chinese book 'About Indeterminate Equation' (談談不定方程), written by Ke Zhao (柯召) and Sun Qi (孫琦) (p.45-61)
and Wikipedia 'Moredell curve' ⟨http://en.wikipedia.org/wiki/Mordell%27s_equation⟩

[15]It is an elliptic curve

[16]'About Indeterminate Equation' (談談不定方程), written by Ke Zhao (柯召) and Sun Qi (孫琦) (p.46)

Theorem 19 (Concerning cubic n in the equation)**.** *The equation $p^n - a^3 = b^3$[17] has no positive integral solution greater than 1 with $p \neq 2, 3$.*

Proof. Here, we employ Fermat's Method of Descent to $p \neq 2, 3$ (so contradiction occurs in later part).
We also use similar strategy to that in **Stage 1**.
Rewrite the equation as

$$p^n = a^3 + b^3 = (a+b)(a^2 - ab + b^2)$$

Then,

$$\begin{cases} a + b = p^c \\ a^2 - ab + b^2 = p^d \end{cases} \tag{*}$$

where c and d are non-negative integers with sum equal to n.
From the second equation,

$$p^d = (a+b)^2 - 3ab = p^{2c} - 3ab \tag{**}$$

As a and b are positibe integers, $p^c = a + b \geq 2$. So, $c \neq 0$.
If $d = 0$, then $c = n$ and $a^2 - ab + b^2 = 1$.
Claim: $a^2 - ab + b^2 = 1$ if and only if $a = b = 1$

Proof.

$$a^2 - ab + b^2 \geq 2ab - ab = ab \geq 1$$

Clearly, equality holds if and only if $a = b = 1$. □

Hence, $p^n = 2$, which implies $p = 2$ and $n = 1$. However in the condition of this theorem, $n \neq 1$. Contradiction. So, $d \neq 0$.

Also, if $d = 2c$, then from (**), we have $a = 0$ or $b = 0$. Contradiction.

So, from (**), we can deduce that $p \mid 3ab$.
If $p \neq 3$, then we have $p \mid a$ or $p \mid b$[18].
Without loss of generality, assume $p \mid a$. Form the first equation of (*), $p \mid a$ also implies $p \mid b$.
Let $a = pa_1$ and $b = pb_1$, where a_1 and b_1 are positive integers.
Then $p^n = a^3 + b^3$ becomes $p^{n-3} = a_1^3 + b_1^3$.
Repeat the above arguments until the power of p becomes 0,1 or 2.
Let k be the number of steps (repeating arguments) required to get to cases in which the power of p becomes $0, 1$ or 2. (k is a positive integer) Denote that 'a' and 'b' by a_k and b_k respectively.

---

[17]A problem form the exercise (Q.2) of the notes 'Diophantine Equation' form math database http://www.mathdb.org/notes_download/elementary/number/ne_N3.pdf. No solution was given in the notes

[18]by Euclid's Lemma

Suppose the power of p becomes 0 and 1. For the first case, we get (a_k, b_k) equal to (1,0) or (0,1). Then, a or b is equal to 0. Contradiction.
For the second case, we can factorize $a_k^3 + b_k^3$. So, p will not be prime except (a_k, b_k) is equal to (1,1). Contradiction. For the case of (1,1), we have $a = b = p^k$, it yields $p = 2$ after putting back into the original equation. Contradiction.

Suppose the power of p becomes 2. $a_k + b_k$ and $a_k^2 - a_k b_k + b_k^2$ both cannot be 1, otherwise for the first case, a or b is equal to 0; for the second case (a_k, b_k) is equal to (1,1), it has been proved that contradiction occurs.
However, the solution of $a_k + b_k = a_k^2 - a_k b_k + b_k^2$ must be (2,2) , (2,1) or (1,2) because

$$a_k + b_k = (a_k + b_k)^2 - 3a_k b_k \geq (a_k + b_k)^2 - 3(\frac{a_k + b_k}{2})^2 = \frac{1}{4}(a_k + b_k)^2$$

So, $4 \geq a_k + b_k$. By checking (a_k, b_k) is equal to (2,2) , (2,1) or (1,2).
(a, b) is equal to $(2p^k, 2p^k), (2p^k, p^k)$ or $(p^k, 2p^k)$. The first one yields $p = 2$ and the later two yield $p = 3$. So, $a_k + b_k = a_k^2 - a_k b_k + b_k^2 = p$ cannot occur. Contradiction.

Therefore, for $p \neq 2, 3$, the equation has no solution. □

Since I have proved this theorem some years ago (at that time I did not know **Zsigmondy's theorem**), I gave a very complicated, but elementray proof. Of course, using Zsigmondy's theorem, we have a delightful and more general proof:

If we have $a = b$, then $p^n = 2a^3$. It is trivial that $p = 2$. So, a is also a power of 2, and let $a = 2^m$, where m is a positive integer.
We have, $(a, b, p, n) = (2^m, 2^m, 2, 3m + 1)$, where m is any integer.

Otherwise, without loss of generality, assume $a > b$.
Then, by **Zsigmondy's theorem**, $p \nmid a + b, a + b = 1$. Contradiction.
It remains to consider the exceptional case, $a = 2$ and $b = 1$. They give $p = 3$ and $n = 2$. □

Theorem 20. [19] *The equation $3^n - x^k = y^k$ has unique positive integral solution, where n, x, k are greater than 1, where x and y are coprime.*

Proof. We have $3^n = x^k + y^k$.
If k is even, then x^k and y^k are both perfect squares. As for any integer m, we have $m^2 \equiv 0, 1 \pmod 3$. So for any positive integers a and b, $3 \mid a^2 + b^2$ if and only if $3 \mid a$ and $3 \mid b$.
So, $gcd(x^{\frac{k}{2}}, y^{\frac{k}{2}}) \geq 3$. This contradicts with the assumption that x and y are coprime. So, k is odd. RHS can be factorized in integral coefficients.

Clearly, $x \neq y$. Then by **Zsigmondy's theorem**, $3 \nmid x + y$.
Therefore, $x + y = 1$. Contradiction.

[19]From the article 'LTE lemma' p.6. However, I gave an alternative proof here.

It remains to consider the special case, (x, y) is equal to $(1, 2)$ or $(2, 1)$, and $k = 3$. Then, we get $n = 2$. $\square$

For the case that x and y are not coprime, let their G.C.D. be d, which is greater than 1. Also, let $x = dx_1$ and $y = dy_1$, where x_1 and y_1 are coprime.
We have

$$d^k(x_1^k + y_1^k) = 3^n$$

So, $d = 3^a$, where a is a positive integer with $n - ka \geq 1$ (otherwise contradiction).

$$x_1^k + y_1^k = 3^{n-ka}$$

Then by using **Theorem 20**, (x_1, y_1) is equal to (1,2) or (2,1), $k = 3$ and $n-ka = 2$. So, (x, y) is equal to $(2(3^a), 3^a)$ or $(3^a, 2(3^a))$, $k = 3$ and $n = 3a + 2$

7. Summary

After finishing the three stages of studying, we establish a number of special cases ranging from two to four-variable equations. Also, a short journey towards generalized Catalan's Conjecture has been taken.

In Stage 1, factorization, congruence and prime factorization were the main tools to study the special cases.

In Stage 2, LTE Lemma and Zsigmondy's theorem were used to develop much stronger theorems.

In Stage 3, the study on Catalan's Conjecture was extended, the '1' in the equation was replaced by any positive integer 'n'. Some further discussions were taken place and two theorems were successfully proved.

8. Conclusions

Although a number of special cases have been successfully established, it is far from the end of my investigation. Firstly, several cases have been thought but I could not complete the proof, only a little idea has been developed. I have to complete the proofs in the future study. Especially, it is important to explore the possibility of using the concept of abstract algebra to solve Diophantine equations. Also, I hope that the constraints of the proved theorems can be weaker so as to step closer to the original Catalan's Conjecture.

Secondly, I hope that I can figure out the elementary proofs of some of the special cases. I believe that we can prove them without using the strong Zsigmondy's theorem. The reason to do so is that the proof of Zsigmondy's theorem requires a little higher concept, for example: properties of cyclotomic polynomials and arithmetic functions.

Thirdly, there is still a large margin in the generalized Catalan's Conjecture in my opinion. Deeper and stronger results should be developed.

REFERENCES

[1] Andreescu, Titu and Andrica, Dorin, *An Introduction to Diophantine Equation*, GIL Publishing House, 2002.
[2] Ke, Zhao (柯召) and Sun, Qi (孫琦), *About Indeterminate Equation* (談談不定方程), Harbin Institute of Techology Press (哈爾濱工業大學), 2011.
[3] Li, Chun Che , *abc conjecture*.
[4] Li, Kin Yin, *Pell's Equation*, http://www.math.ust.hk/~makyli/190_2010Sp/190_LectNt-rev.pdf
[5] Math Database Number Theory Notes Unit 3, *Diophantine Equations*, http://www.mathdb.org/notes_download/elementary/number/ne_N3.pdf
[6] Mihăilescu, Preda, *On Catalan's Conjecture*, https://www.dpmms.cam.ac.uk/Seminars/Kuwait/abstracts/L30.pdf
[7] Parvardi, Amir Hossein, *Lifting the Exponent Lemma (LTE)*, www.artofproblemsolving.com/Resources/Papers/LTE.pdf
[8] Pisolve, *The Zsigmondy's theorem*, http://www.artofproblemsolving.com/Forum/download/file.php?id=34511
[9] Schoof, Rene, *Catalan's Conjecture*, Springer-Verlag, London, 2008.
[10] Wikipedia, *Mordell curve*, http://en.wikipedia.org/wiki/Mordell_curve
[11] Wikipedia, *Zsigmondy's theorem*, http://en.wikipedia.org/wiki/Zsigmondy's_theorem

Reviewer's Comments

This paper is well written. The following is a list of corrections and stylistic suggestions.

1. The reviewer has comments on the wordings, which have been amended in this paper.
2. Punctuation marks are missing in most of the formulas in this paper.
3. "more" should be deleted.
4. "strong and complicated" should be rewritten as "stronger and more complicated".