FERMAT'S CHRISTMAS THEOREM

SACHA DE POYEN-BROWN

Abstract. The purpose of this paper is to provide a casual and conversational exposition of various proofs of Fermat's Christmas theorem. Fermat's Christmas theorem is an accessible result in number theory that can be explained using only concepts familiar to a general audience, but its many proofs use a wide range of approaches likely unfamiliar to the former. For that reason, this paper seeks to use Fermat's Christmas theorem as a jumping off point from the math covered in our REU (apprentice program) to just a few of the varied and interesting techniques of number theory.

Contents

1. History 1
2. Proofs 2
2.1. Girard 2
2.2. Quadratic Character 2
2.3. Zagier 3
3. Conclusion and Further topics 6
Acknowledgments 6
References 6

1. History

Originally formulated by Albert Girard in 1625, Fermat's theorem on the sum of squares, or Christmas theorem as it is colloquially known came to prominence through two letters between the former and to Marianne Mersenne, the first of which was dated December 25, 1865 [1].

Theorem 1.1. *The theorem states that a prime number p can be written as the sum of squares if and only if it is congruent to 1 mod 4.*

Interestingly, Fermat phrased this question rather differently than we do now; he thought of these sums in a much more geometric manner. In his comments on Diophantus' Arithmetica–the book in which Fermat posited his famous "last theorem"–Fermat wrote 48 mathematical statements, often without proof. His annotations were published posthumously by his son [2].

Date: September 23, 2023.

2. Proofs

2.1. **Girard.** The first proof on this topic was due to Girard. For his part, Girard only proved the "only if" direction of this theorem, which holds for all numbers. He didn't show the other implication, which only holds for primes.

Proof. For all integers n, either $n^2 \equiv 0 \pmod 4$ or $n^2 \equiv 1 \pmod 4$ depending on whether n is even or odd. Thus the sum of any two squares is either congruent to $0, 1$, or $2 \pmod 4$ thus if a prime number can be expressed as a sum of two squares, either $p \equiv 1 \pmod 4$ or $p = 2$. □

2.2. **Quadratic Character.** A common proof of Theorem 1.1 uses quadratic character.

Proof. Let p prime such that there exist integers x and y such that $x^2 + y^2 = p$ then under the field $\mathbb{F}_p$,

$$x^2 + y^2 = 0$$

Then we can use algebra to see that:

$$\begin{aligned} x^2 &= -y^2 \\ -1 &= \frac{x^2}{y^2} \\ -1 &= [\frac{x}{y}]^2 \end{aligned}$$

which implies that -1 is a quadratic residue of p. As -1 is a quadratic residue of p we know that there exists an element m of $\mathbb{F}_p$ such that $m^2 = -1$, and thus $m^4 = 1$. From this we know that the subgroup of $\mathbb{F}_p^\times$ generated by m is a subgroup of order 4. By Lagrange's theorem, we know that the cardinality of $\mathbb{F}_p^\times$ is divisible by 4. As $F_p^\times = F_p \setminus \{0\}$, and $|F_p| = p$, we conclude that $4 \mid p - 1$. In other words $p \equiv 1 \pmod 4$.

For the reverse implication, let $p \equiv 1 \pmod 4$ be prime. Then $\mathbb{Z}/p\mathbb{Z} = \mathbb{F}_p$ is a field, and we compute the Legendre symbol of -1 and p as

$$\begin{aligned} \left(\frac{-1}{p}\right) &\equiv (-1)^{\frac{p-1}{2}} \pmod p \\ &\equiv (-1)^{\frac{(4k+1)-1}{2}} \\ &\equiv (-1)^{2k} \\ &\equiv 1 \end{aligned}$$

so -1 is a quadratic residue of $\mathbb{F}_p$, and there exists $z \in \mathbb{F}_p$ such that $z^2 = -1$. Let y be some element in $\mathbb{F}_p$ other than $0, 1$, or -1, and define $x \equiv zy$. Then

$$\begin{aligned} \frac{x}{y} &\equiv z \\ \frac{x^2}{y^2} &\equiv -1 \\ x^2 &\equiv -y^2 \\ x^2 + y^2 &\equiv 0 \pmod{p}. \end{aligned}$$

Then, for some $k \in \mathbb{N}$

$$x^2 + y^2 = kp.$$

Yet $0 < x, y < p$, so

$$\begin{aligned} x^2 + y^2 &< 2p^2 \\ x^2 + y^2 &= p^2. \end{aligned}$$

□

2.3. **Zagier.** Another notable proof of Fermat's Christmas theorem is the one line proof by Don Zagier. It is reproduced in full below:

Proof. The involution on the finite set $S = \{(x, y, z) \in \mathbb{N}^3 : x^2 + 4yz = p\}$ defined by

$$(x, y, z) \to \begin{cases} ((x + 2z, z, y - x - z) \text{ if } x < y - z \\ (2y - x, y, x - y + z) \text{ if } y - z < x < 2y \\ (x - 2y, x - y + z, y) \text{ if } x > 2y \end{cases} \tag{2.1}$$

has exactly one fixed point, so $|S|$ is odd and the involution defined by $(x, y, z) \to (x, z, y)$ also has a fixed point [3]. □

In essence, this proof loosens the condition from $p = x^2 + y^2$, to $p = x^2 + 4yz$, which plainly holds for all $p \equiv 1 \mod 4$, then shows that if there exists such an (x, y, z) then there exists (x', y', y') such that $(x')^2 + (2y')^2 = p$. Before we begin, it might be useful to define the notion of an involution:

Definition 2.2. An involution is a function which is its own inverse.

In other words, a function f is an involution if $f \circ f$ is the identity. One might wonder how it is how we know that there is exactly one fixed point. To see this, we can analyze each of the three cases, $x < y - z$, $y - z < x < 2y$, and $x > 2y$, and identify each of their fixed points.
In the case that $x < y - z$, (x, y, z) is a fixed point only if:

(1) $x = x + 2z$,
(2) $y = z$, and
(3) $z = y - x - z$.

From the first of these equations, we can see that $z = 0$. From that fact, and (2), we know that $y = 0$ as well. Thus by (3), we have $x = 0$. But $(x, y, z) \in \mathbb{N}^3$, yet $(0, 0, 0) \notin \mathbb{N}^3$, so there are no fixed points in the case that $x < y - z$.
When $y - z < x < 2y$, (x, y, z) is a fixed point only if:

(1) $x = 2y - x$,
(2) $y = y$, and
(3) $z = x - y + z$.

It follows from (1) that $y = x$. From (3) we see that z is unconstrained by the choice of x. However, we can further constrain our possible values for x. As $x = y$, we have that

$$\begin{aligned} p &= x^2 + 4xz \\ &= x(x + 4z) \end{aligned}$$

which implies that x divides p. As we know p to be prime, then x can only be 1 or p itself. The latter is impossible as $x^2 + 4xz > x^2 = p^2 > p$, thus we have that any fixed point in this case is of the form $(1, 1, j)$ where $j = \frac{p-1}{4}$.
In the third case, when $x > 2y$, any possible fixed point is constrained by the following equations:

(1) $x = x - 2y$,
(2) $y = x - y + z$, and
(3) $z = y$.

From the first equation, it is clear that $y = 0$. Then by (3) we have $z = 0$, which with (2) gives us that $x = 0$. But $(0, 0, 0)$, is not in the domain of the involution, so it has no fixed points in this case. We conclude the involution only has a fixed point at $(1, 1, j)$.
Zagier goes on to claim this implies that $|S|$ is odd.

Lemma 2.3. *Let $f : S \longrightarrow S$ be an involution which fixes only one point. Then $|S|$ is odd.*

Proof. For all points $x \neq q$, $f(x) \neq x$, and $f(f(x)) = x$. Then we can express f as a permutation consisting of k disjoint 2-cycles and one 1-cycle. Then as the k transpositions are disjoint, no element appears in two transpositions, then there are $2k$ elements which are not fixed by f. Then as f fixes 1 element, we know that f acts on a set of order $2k + 1$ for some natural number k, or equivalently, $|S|$ is odd. □

As (2.1) has exactly one fixed point, $|S| = 2k + 1$ is odd. Then let $g : (x, y, z) \to (x, z, y)$ be an involution on S.

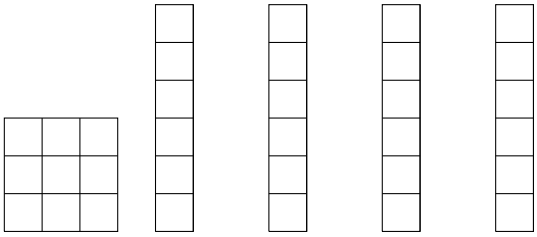
Lemma 2.4. *Let $f : S \longrightarrow S$ be a function, and let $|S|$ be odd. if f has no fixed points, then f isn't an involution.*

Proof. If f is an involution, then f can be written as the product of disjoint 2-cycles, yet each 2-cycle acts on exactly 2 elements of S and fixes all the others, so any composition of n disjoint transpositions acts on $2n$ elements of S. So if f acts on $2k + 1$ elements of S, then it cannot be expressed as the product of disjoint transpositions, and therefore f isn't an involution. □

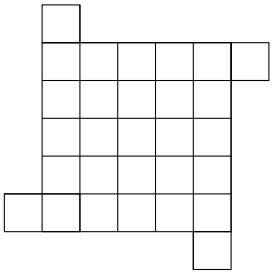
By the contrapositive of (2.4), $(x, y, z) \to (x, z, y)$ has a fixed point, so for all $p \equiv 1 \pmod 4$ there exists an $(x, y, z) \in \mathbb{N}^3$ such that $y = z$ and $x^2 + (2y)^2 = p$.

2.3.1. *Motivation.* While this does explain the mechanics of the proof, it fails to properly motivate it. A reader might wonder "Why did the proof deal with equations of the form $x^2 + 4yz = p$ rather than $x^2 + y^2 = p$?", or perhaps "Where did function (2.1) come from?". To answer the former, it might help to consider the possible values x and y from the original $x^2 + y^2 = p$ equation. As p is an odd prime, then one of x^2 and y^2 must be odd and the other must be even. This without loss of generality we can assume that y^2 is even and thus that y is even and $4|y^2$. Thus we can rewrite this equation as $x^2 + 4(y')^2 = p$. As for the choice of replacing $(y')^2$ with yz, it will be more clear after (2.1) is motivated.

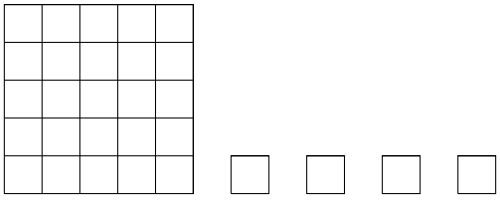
In Zagier's proof, (2.1) is presented without any motivation, so I will present a geometric motivation below [6, 7]. Consider a triple (x, y, z) satisfying the equation of the form $x^2 + 4yz = p$ as a collection of rectangles. For example, $x = 3, y = 5, z = 1$, which satisfies $x^2 + 4yz = 29$, would look like:



Which we can arrange into a windmill shape as follows.

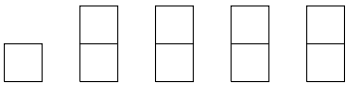


This picture suggests another (x, y, z) triple satisfying $x^2 + 4yz = 29$, namely $x = 5, y = 1, z = 1$. To represent this visually, consider that the same windmill shape could also be constructed by the different collection of rectangles shown below.

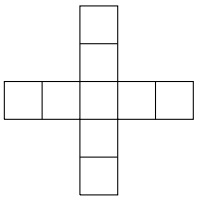


The function in (2.1) does this exact thing, given a triple of the form (x, y, z) such that $x^2 + 4yz = p$, (2.1) gives you a triple (x', y', z') that also satisfies $(x')^2 + 4y'z' = p$. From this visual analogue it is clear that once the value of x is chosen, the values of y and z are fixed (up to $(x, y, z) \to (x, z, y)$). Thus, if there is only one valid choice of x, then we have found a fixed point. No other configuration exists is the case of the $(1, 1, j)$ triple, hence it is the fixed point of

(2.1). To see why this is, lets look at one such case. The simplest such case is $(1, 1, 2)$, which can be represented as:



When we assemble the rectangles into a windmill we get:



When x and y are both 1, it there is no way for to form a square larger than 1 as the "corners" of the square will always be empty. While this "windmill" visualisation is not rigorous, I hope it can provide the reader some motivation for (2.1), as it seems totally arbitrary in Zagier's proof.

3. Conclusion and Further topics

These are just a few of the many interesting proofs of Fermat's Christmas Theorem, there are many more. If the reader is hungry for more, I recommend the papers of Robert Heath-Brown[4], Christian Elsholtz[5], or Euler's original proof of the theorem, which can be found on Wikipedia[1].

Acknowledgments

I would like to thank Ryan Wandsnider for his help and advice throughout the program. I would also like to thank Danil Rudenko and Lazlo Babai for their excellent lectures, as well as J. Peter May and Elizaveta Shuvaeva for their work in organizing the REU.

References

[1] https://en.wikipedia.org/wiki/Fermat%27s_theorem_on_sums_of_two_squares
[2] Kazuya Kato, Nobushige Kurokawa, Takeshi Saito. Translated by Masato Kuwata. Fermat's Dream.
[3] Donald Zagier, A One-Sentence Proof That Every Prime $p \equiv 1 \pmod 4$ Is a Sum of Two Squares. https://people.mpim-bonn.mpg.de/zagier/files/doi/10.2307/2323918/fulltext.pdf
[4] Robert Heath-Brown. https://www.researchgate.net/publication/266218880_Fermat's_two_squares_theorem
[5] Christian Elsholtz. The Liouville Heath-Brown Zagier Proof of the Two Squares Theorem and generalizations. http://www.math.tugraz.at/~elsholtz/WWW/papers/papers30nathanson-new-address3.pdf
[6] Moritz Firsching. Zagier's one-sentence proof of a theorem of Fermat. https://mathoverflow.net/questions/31113/zagiers-one-sentence-proof-of-a-theorem-of-fermat
[7] Alexander Spivak. Sum of Squares. http://mmmf.msu.ru/lect/spivak/summa_sq.pdf