Proof of Fermat's Christmas Theorem

Fermat's Christmas Theorem series

2025-12-25

ITA ENG

Fermat's Christmas Theorem

Let's recap what we have seen so far. Now we are ready to state Fermat's Christmas theorem in terms of congruences.

Theorem 0.1 (Fermat's Christmas Theorem)**.** *An odd prime number p can be written as $p = a^2 + b^2$ with $a, b \in \mathbb{Z}$ if and only if $p \equiv_4 1$.*

In the previous post we proved that if p can be written as the sum of two squares, then it has remainder 1 when divided by 4, i.e. $p \equiv_4 1$. Now we want to prove the converse: if $p \equiv_4 1$, then p can be written as the sum of two squares (this is the hard part).

In this post we develop everything we need in order to give a complete proof of the theorem. First, we shift the problem to the context of Gaussian integers, which are an extension of the integers we know. Then we will see how sums of two squares appear naturally there, and how to recognize the primes that can be written as the sum of two squares.

Gaussian integers

We all know that the equation

$$x^2 + 1 = 0 \tag{1}$$

has no solutions in the integers, since squares are always nonnegative, and if we add 1 to a nonnegative number we can't get zero. Right?

Well, let's add a solution of this equation to the integers: we define

$$i := \sqrt{-1}$$

so that we get

$$i^2 + 1 = \sqrt{-1}^2 + 1 = -1 + 1 = 0$$

We now have a solution to Equation 1.

We define $\mathbb{Z}[i]$ to be the set of integers in which we add this i. The "numbers" in $\mathbb{Z}[i]$ are of the form $a + b \cdot i$, where a and b are usual integers. For example, $1 + 2i$, $5 + 7i$, and $0 + i$ are all Gaussian integers. The special cases of numbers of the form $a + 0 \cdot i$ will be just denoted by a, while the numbers $0 + b \cdot i$ will be denoted just by bi. In particular observe that we have all the integers inside $\mathbb{Z}[i]$: indeed we can recover them as the numbers of the form $a + 0 \cdot i = a$.

But why should we work in Gaussian integers instead of classical integers? We will understand this later, since the existence of a solution for the congruence

$$x^2 + 1 \equiv_p 0$$

is closely related to p being a sum of two squares.

First of all, how do products and sums work in the Gaussian integers?

It is easy: you just have to treat i as a formal variable, and remember that $i^2 = -1$. I will explain it better:

Suppose we want to sum $3 + 2i$ with $5 + i$. Then we do like this:

$$(3 + 2i) + (5 + i) = (3 + 5) + (2 + 1)i = 8 + 3i$$

so we add the numbers that don't have the i with them and the numbers that have the i with them separately. Other examples:

$$(-2 + 4i) + (7 - 3i) = 5 + i, \qquad (3 + 6i) - i = 3 + 7i, \qquad 3 + (-1 + i) = 2 + i.$$

Similarly we do for the products, we just treat i as a formal variable:

$$(3 + 2i) \cdot (5 + i) = 3 \cdot 5 + 3 \cdot i + 2i \cdot 5 + 2i \cdot i = 15 + 3i + 10i + 2i^2 = 15 + 2i^2 + 13i$$

Now remember that $i^2 = -1$ so $2i^2 = -2$, thus

$$15 + 2i^2 + 13i = 15 - 2 + 13i = 13 + 13i$$

Other examples:

$$(-2 + 4i)(7 - 3i) = -2 + 34i, \qquad (3 + 6i) \cdot (-i) = 6 - 3i, \qquad 3 \cdot (-1 + i) = -3 + 3i$$

Now, define the following:

Definition: Let $a + bi$ be a Gaussian integer. We define its conjugate to be $a - bi$.

Example 0.1.

- The conjugate of $3+4i$ is $3-4i$.
- The conjugate of $1-i$ is $1+i$.
- The conjugate of 4 is 4.
- The conjugate of $-6i$ is $6i$.

Now observe this useful thing. If you multiply a Gaussian integer $a+bi$ by its conjugate $a-bi$ you get

$$(a+bi)(a-bi) = a^2 - abi + abi - b^2i^2 = a^2 - b^2i^2$$

and since $i^2 = -1$ we get

$$(a+bi)(a-bi) = a^2 + b^2$$

This is nice, isn't it? Our goal is to understand when a prime is a sum of two squares, and now we have an easy way to produce sums of two squares.

What is a prime in the Gaussian integers?

If you remember, in the second post we defined a prime p as a number such that whenever p divides a product $a \cdot b$, then either p divides a or p divides b. We now do the same in the Gaussian integers. First we define what it means to divide something in the Gaussian integers.

Definition: Let $a+bi$ be a Gaussian integer. We say that $a+bi$ divides $c+di$ if we can write

$$c+di = (a+bi)\cdot(e+fi)$$

for some Gaussian integer $e+fi$.

Definition: Let $a+bi$ be a Gaussian integer. We say that $a+bi$ is prime if whenever $a+bi$ divides a product $(c+di)(e+fi)$ then either $a+bi$ divides $c+di$ or $a+bi$ divides $e+fi$.

We will not prove the following theorem, whose proof is in some way similar to the one made for the integers.

Theorem 0.1 (Unique factorization)**.** *Every Gaussian integer $a+bi$ can be written as the product of Gaussian integer primes, i.e.*

$$a+bi = (c_1+d_1i)\cdots(c_k+d_ki)$$

for some positive integer k. This writing is unique up to multiplying by $-1, i$, and $-i$.

Let's clarify this uniqueness, which is a little bit strange. What we are saying is that if you consider a Gaussian integer like $1+3i$, you can factorize it as

$$1+3i = (1+i)(2+i)$$

(you can check that $1+i$ and $2+i$ are primes by using the definition, but it is maybe a little tedious). Observe now that $i \cdot (-i) = -i^2 = 1$. Therefore we can write

$$1+3i = 1 \cdot (1+i)(2+i) = (i \cdot (-i))(1+i)(2+i) = i \cdot ((-i) \cdot (1+i))(2+i) = i \cdot (1-i)(2+i)$$

this seems to be another factorization of $1+3i$. However, it differs from the other one only because we multiplied by $1 = i \cdot (-i)$. Therefore, we say that up to multiplying by these special numbers (if you think about it you can also do the same by writing $1 = (-1) \cdot (-1)$ and $1 = (-i) \cdot i$) the factorization is unique.

Proof of the theorem (finally!)

Okay, we are almost there. We can now try to understand how the proof will work.

Let p be a prime in the integers. A sketch of the proof is to prove the following statements:

- if $x^2 + 1 \equiv_p 0$ has a solution then $p + 0i$ is not prime in the Gaussian integers;
- if $p + 0i$ is not prime in the Gaussian integers, then p can be written as the sum of two squares;
- if p is the sum of two squares then $p \equiv_4 1$ (we already proved this last time!);
- if $p \equiv_4 1$ then $x^2 + 1 \equiv_p 0$ has a solution.

These implications tell us that if we know that one statement above is true, then all the other three are true. In other words, the statements above are all saying the same thing: they are equivalent, since they imply each other (observe that the first hypothesis is the same as the last conclusion).

Relation between $x^2 + 1 \equiv_p 0$ and Fermat's christmas theorem

The following lemma is the crucial one: through this lemma we will be able to prove Fermat's Christmas Theorem. It gives the right interpretation of what it means for a prime p to be the sum of two squares.

Lemma 0.1. *Let p be a prime number. If the congruence*

$$x^2 + 1 \equiv_p 0$$

has a solution, then $p + 0i$ is not prime in the Gaussian integers.

The meaning of $x^2 + 1 \equiv_p 0$ having a solution is that there exists an integer $x \in \mathbb{Z}$ such that $x^2 + 1 \equiv_p 0$, or equivalently that p divides $x^2 + 1$. Another way to say this is

$$x^2 \equiv_p -1$$

so we are saying that -1 is a square when reduced modulo p.

Proof. Suppose that there exists an $x \in \mathbb{Z}$ such that

$$x^2 + 1 \equiv_p 0$$

This means that p divides $x^2 + 1$. But observe that

$$x^2 + 1 = (x + i)(x - i)$$

in the Gaussian integers. We want to say that $p + 0i$ is not prime in the Gaussian integers.

Suppose that $p + 0i$ were a prime in the Gaussian integers. Since $p + 0i$ divides the product $(x+i)(x-i)$, then it would have to divide one of them. But this is impossible. Indeed, if $p+0i$ divides $x+i$, then we can write $x+i = (p+0i)(a+bi) = pa+pbi$. Confronting the coefficients of i in both sides gives $1 = pb$, which is impossible. The same argument works for $x - i$.

Therefore $p + 0i$ is not prime in the Gaussian integers. □

Primes in the integers which are not primes in the Gaussian integers

Lemma 0.2. *Let p be a prime number. If $p + 0i$ is not prime in $\mathbb{Z}[i]$, then we can write p as a sum of two squares.*

Proof. If $p + 0i$ is not prime, then it can be written as a product of two factors:

$$p + 0i = (a + bi)(c + di)$$

where neither $a + bi$ nor $c + di$ are equal to $1, -1, i, -i$.

Now multiply both sides by their conjugates. On the left we get

$$(p + 0i)(p - 0i) = p^2.$$

On the right we get

$$(a + bi)(a - bi)(c + di)(c - di) = (a^2 + b^2)(c^2 + d^2).$$

Therefore

$$p^2 = (a^2 + b^2)(c^2 + d^2).$$

We have $a^2 + b^2 > 1$ and $c^2 + d^2 > 1$. But p is prime in $\mathbb{Z}$, so the only way two integers bigger than 1 can multiply to p^2 is that

$$a^2 + b^2 = p, \qquad c^2 + d^2 = p.$$

In particular,

$$p = a^2 + b^2,$$

which means that p is a sum of two squares. □

Wilson theorem and solution to $x^2 + 1 \equiv_p 0$

We will use the following theorem to prove what we need.

Theorem 0.1 (Wilson's Theorem)**.** *If p is a prime number, then*

$$(p-1) \cdot (p-2) \cdot (p-3) \cdots 2 \cdot 1 \equiv_p p - 1.$$

i.e. multiplying all numbers from $p-1$ to 1 we get $p-1$ modulo p.

We give an idea of how to prove this theorem.

Idea. Recall that an integer x is invertible modulo p if and only if $\gcd(x, p) = 1$. Since p is prime, it has no common factors with any of the numbers $(p-1), (p-2), \ldots, 2, 1$. This means that every factor on the left hand side is invertible modulo p.

Now, when you list all the invertible residues modulo p, you get exactly $1, 2, \ldots, p-1$. So on the left hand side, every term appears together with its inverse, except for the terms that are their own inverse.

An element a is its own inverse if and only if $a^2 \equiv_p 1$. Modulo a prime, this means $a \equiv_p 1$ or $a \equiv_p -1$. Therefore, everything cancels out in pairs, and only $-1 \equiv_p p - 1$ remains. □

Lemma 0.3. *If p is a prime such that $p \equiv_4 1$ then $x^2 + 1 \equiv_p 0$ has a solution.*

Proof. Since $p \equiv_4 1$, it means that $p - 1$ is divisible by 4, which means that $p - 1 = 4k$ for some integer k. We have then

$$(p-1) \cdot (p-2) \cdot (p-3) \cdots 3 \cdot 2 \cdot 1.$$

Now group the factors in pairs as follows:

$$(p-1) \cdot 1 \cdot (p-2) \cdot 2 \cdots (p-2k) \cdot (2k).$$

Since $p - j \equiv_p -j$, each pair gives

$$(p-j)\cdot j \equiv_p (-j)\cdot j = -j^2.$$

Therefore

$$(p-1)\cdot(p-2)\cdots 2\cdot 1 \equiv_p (-1)^2\cdot(-2)^2\cdots(-(2k))^2 = (1\cdot 2\cdot 3\cdots\cdot(2k-1)\cdot 2k)^2.$$

(Here the minus signs disappear because we are squaring.)

But, by Theorem 0.1 we also know that the left hand side is congruent to $p-1$ modulo p. And since $p-1 \equiv_p -1$ we get that

$$-1 \equiv_p (1\cdot 2\cdots(2k-1)\cdot 2k)^2$$

Define

$$x := 1\cdot 2\cdots(2k-1)\cdot 2k$$

Then we get

$$x^2 + 1 \equiv_p 0.$$

Therefore, we found a solution. □

The theorem is proven, what else can we say?

The theorem is finally proven, but we don't stop asking questions. We come up with two questions.

How many primes can be written as the sum of two squares? Equivalently, how many primes have remainder 1 when divided by 4?

Probably not surprisingly we already expect that there are infinitely many primes that can be written as the sum of two squares. This statement is left as a guided exercise at the end of the post.

An interesting, but non trivial fact is to know also the density of these kinds of primes. Namely, if we consider all the primes up to some integer N, how many of these do we expect to be written as the sum of two squares? The answer is that we expect about half of them to be of this kind. Indeed here you can see a table which could convince you:

N	primes $\leq N$	primes that are a^2+b^2	proportion
10^2	25	12	0.48
10^3	168	81	0.4821...
10^6	78498	39176	0.49907...
10^9	50847534	25423492	0.499994...

This can be proven, however this is far beyond our interests here.

Another natural question now could be

Consider an integer n. When can we write it as the sum of two squares?

So we are trying to generalize from the case of primes to the case of a general integer. This is not difficult to prove knowing Fermat's Christmas Theorem, therefore it is also left as a guided exercise at the end.

Conclusion

I would like to point out that the crucial point in our reasoning was to relate the problem of writing p as the sum of two squares and the solution of the congruence $x^2 + 1$. Indeed, if you think about it, the Gaussian integers were defined by adding the solution of $x^2 + 1 = 0$ to the integers. This is not a coincidence.

There are more general theorems that tell us precisely how the factorization of a prime ideal in an extension of the integers (given by adding something similar to the solutions of a polynomial) relates to the factorization (in our case the solutions) of the polynomial modulo that prime.

In our case we knew that $x^2 + 1 \equiv_p 0$ has a solution, and this is the same as saying that there exists an integer u such that $u^2 \equiv_p -1$. Then modulo p we have the factorization

$$x^2 + 1 \equiv_p (x + u)(x - u),$$

because

$$(x + u)(x - u) = x^2 - u^2 \equiv_p x^2 + 1.$$

This tells us exactly that the prime p can be written as the product of two primes in the Gaussian integers (observe the symmetry: the polynomial splits in two factors exactly as the prime). The theorem which explains this behaviour is known as the Dedekind-Kummer theorem.

Guided exercises

Exercise 1: infinitely many primes that are sums of two squares

Exercise 1

Prove that there are infinitely many primes that can be written as the sum of two squares.

Hint 1

A prime is a sum of two squares if and only if -1 is a square mod p. Suppose there are only finitely many primes such that -1 is a square mod p. What can you say?

Solution to Hint 1

Saying "-1 is a square mod p" means: there exists an integer n such that

$$n^2 \equiv_p -1,$$

i.e. $p \mid (n^2+1)$. So every prime for which -1 is a square mod p divides at least one number of the form n^2+1.

Hint 2

Let k be the product of all primes p such that -1 is a square mod p. Try to imitate Euclid's proof that there are infinitely many primes.

Hint 3

Consider the integer k^2+1. Which primes can divide it?

Solution to Hint 3 (and end of the exercise)

Take a prime q dividing k^2+1. Then

$$k^2 \equiv_q -1,$$

so -1 is a square mod q. By definition of k, this forces $q \mid k$.

But then q divides both k and k^2+1, hence it divides their difference $(k^2+1)-k^2 = 1$, impossible. Contradiction.

Therefore the assumption "only finitely many primes with -1 a square mod p" was false, and there are infinitely many such primes. By Fermat's Christmas theorem, these are exactly the primes that are sums of two squares.

Exercise 2: which integers are sums of two squares?

Exercise 2

Which integers $n \geq 1$ can be written as $n = a^2 + b^2$?

Hint 1

What happens if n is the product of two primes that can be written as the sum of two squares?

Solution to Hint 1

If $p = a^2 + b^2$ and $q = c^2 + d^2$, then pq is again a sum of two squares. Indeed a sum of two squares can be written as a Gaussian integer times its conjugate:

$$a^2 + b^2 = (a + bi)(a - bi), \qquad c^2 + d^2 = (c + di)(c - di).$$

Therefore

$$(a^2 + b^2)(c^2 + d^2) = (a + bi)(a - bi)(c + di)(c - di).$$

The factor $(a + bi)(c + di)$ is a Gaussian integer, and $(a - bi)(c - di)$ is its conjugate. So their product is a sum of two squares.

Hint 2

Consider the number $5 \cdot 7^2$. Can it be written as a sum of two squares?

Solution to Hint 2

Yes:

$$5 \cdot 7^2 = 245 = 14^2 + 7^2.$$

(So the prime $7 \equiv_4 3$ is not "forbidden", but its exponent seems to matter.)

Hint 3

In the previous hint, $5 \equiv_4 1$ and $7 \equiv_4 3$. What theorem do you expect?

Solution to Hint 3 (the statement)

Theorem (sum of two squares). A positive integer n can be written as $n = a^2 + b^2$ if and only if, in the prime factorization of n, every prime $q \equiv_4 3$ appears with an **even** exponent.

Complete proof (guided by the hints)

Step 1: primes $q \equiv_4 3$ must appear an even number of times.

We first prove the key lemma:

Lemma 0.1. *Let q be a prime with $q \equiv_4 3$. If $q \mid (a^2 + b^2)$, then $q \mid a$ and $q \mid b$.*

Proof. Suppose $q \mid (a^2 + b^2)$. If $q \nmid b$, then b has an inverse u modulo q, and

$$(a \cdot u)^2 \equiv_q -1,$$

so -1 is a square mod q. But by Fermat's Christmas theorem, this would force $q \equiv_4 1$, contradiction. Therefore $q \mid b$. Then $q \mid a^2$ and hence $q \mid a$. □

Now assume $n = a^2 + b^2$. Take a prime $q \equiv_4 3$ dividing n. By the lemma, q divides both a and b, so q^2 divides $a^2 + b^2 = n$. This means that in the prime factorization of n, the exponent of q cannot be odd: it must be even.

Step 2: the converse.

Now suppose n has the property that every prime $q \equiv_4 3$ appears with an even exponent. So we can write n as a product of three types of factors:

- a power of 2,
- primes $p \equiv_4 1$ (possibly repeated),
- and squares of primes $q \equiv_4 3$ (possibly repeated).

  Each of these factors is a sum of two squares:

- $2 = 1^2 + 1^2$,
- if $p \equiv_4 1$ then p is a sum of two squares by Fermat's Christmas theorem,
- and if you have a square q^{2m} then $q^{2m} = (q^m)^2 + 0^2$.

  Finally, by Hint 1, a product of sums of two squares (check that the arguments works not only for primes) is again a sum of two squares. Therefore n is a sum of two squares.