

ON THE REPRESENTATION OF THE NATURAL NUMBERS AS SUMS OF TWO SQUARES

Titu Andreescu
Marian Tetiva

1. Introduction. A natural number (nonnegative integer) n can be represented as a sum of two squares if and only if integers x and y exist such that $n = x^2 + y^2$. (Obviously, since $x^2 + y^2 = |x|^2 + |y|^2$, this is equivalent to the existence of some natural numbers x and y such that $n = x^2 + y^2$.) In what follows we will say, in short, that n is *representable* if integers x and y for which $n = x^2 + y^2$ do exist (with the obvious negation *unrepresentable* if this is not the case). For example

$$0 = 0^2 + 0^2, 1 = 0^2 + 1^2, 2 = 1^2 + 1^2, 4 = 0^2 + 2^2, 5 = 1^2 + 2^2,$$

$$8 = 2^2 + 2^2, 9 = 0^2 + 3^2, 10 = 1^2 + 3^2$$

are the representable natural numbers at most equal to 10. The sequence of representable numbers goes on with 13, 16, 17, 18, 20, 25, 26, 29, etc. Clearly, this sequence (which, among many other numbers, contains all the perfect squares) is infinite. On the other hand, one easily checks that the numbers 3, 6, 7, 11, 12, 14, 15, 19, 21, 22, 23, 24, 27, 28 can *not* be expressed as sums of two squares (are unrepresentable) by using a very simple algorithm: if $n - x^2$ is not a perfect square for every natural number $x \leq \sqrt{n}$, then n cannot be represented as a sum of two squares. (Simple, but, unfortunately, not very efficient for n big.) It is well-known that the nonnegative integers congruent to 3 modulo 4 are not representable (because any square is congruent to either 0, or 1 modulo 4, a sum of two squares is never congruent to 3 modulo 4). This shows that there are infinitely many unrepresentable numbers—for example those of the form $4k + 3$, but it would be naïve to believe that all the numbers giving remainder other than 3 when divided by 4 are representable: for instance, 12, 21, or 6 are not representable, and their remainders when divided by 4 are 0, 1, and 2 respectively.

Since the very year of 1625 Girard noticed which of the first natural numbers are representable—and, implicitly, which primes can (or cannot) be expressed as sums of two squares. More precisely, he noticed (of course, in particular cases) that the primes of the form $4k + 1$ are sums of two squares, while those of the form $4k + 3$ are not (for a natural number k ; and 2 is also representable.) Later on Fermat communicated this fact in a letter to Mersenne, in which he also mentioned the number of representations of a prime of the form $4k + 1$ as a sum of two squares. Namely, for a prime $p = 4k + 1$ there are 8 representations $p = x^2 + y^2$ with integers x and y , considering two representations distinct if the pairs (x, y) are so; for example, $5 = (\pm 1)^2 + (\pm 2)^2 = (\pm 2)^2 + (\pm 1)^2$ are the eight representations of 5. Differently speaking, for a prime $p = 4k + 1$ there exists fundamentally one representation $p = x^2 + y^2$ (if we do not take account of the changing of the order, or of the signs of x and y).

Because the letter is dated December 25, 1640, the theorem stating that any prime of the form $4k + 1$ is the sum of two squares is often called *Fermat's Christmas theorem*. Such as in many other cases (the most prominent being that of Fermat's Last Theorem, which states that the Diophantine equation $x^n + y^n = z^n$ has no nontrivial solutions (x, y, z) in integers), Fermat left no argument for his Christmas theorem. Only a century later Euler, in two letters (1747, 1749) to Goldbach provided a proof based on the infinite descent—a method that was also invented by Fermat. Many more proofs are available today—due to Lagrange, Dedekind, Liouville and many more. In what follows, we will deal with three such demonstrations, two of which are elementary (in the sense that they only use rational

integers—one being a simplified form of Euler’s infinite descent, one other based on a lemma of Thue), while the third one relies on the arithmetic properties of the ring of Gauss’s integers.

Theorem 1. (*Fermat’s Christmas theorem*) A prime p is the sum of two squares if and only if $p = 2$, or $p \equiv 1 \pmod{4}$.

2. The first proof of Fermat’s Christmas theorem. Because $2 = 1^2 + 1^2$ is obviously representable, and the primes congruent to 3 modulo 4 are not (actually, *all* the natural numbers congruent to 3 modulo 4 are unrepresentable, as we just have seen), it only remains to prove that any odd prime $p \equiv 1 \pmod{4}$ can be written as a sum of two squares. Many of the proofs of theorem 1 use part (a) from the following result. (Note that, basically, this says that -1 is a quadratic residue for a prime $p \equiv 1 \pmod{4}$ —which is part of one special case of the quadratic reciprocity law.)

Lemma 2. (a) Let $p \equiv 1 \pmod{4}$ be a prime. Then a natural number c exists such that $p \mid c^2 + 1$.

(b) If p is a prime congruent to 1 modulo 4, then there exists a representable multiple mp of p for which m is an integer and $1 \leq m < p$.

Proof. (a) Remember that $(p-1)! \equiv -1 \pmod{p}$, according to Wilson’s theorem (which actually holds for any prime p). On the other hand, we have the obvious congruences

$$-j \equiv p - j \pmod{p}$$

for $j = 1, \dots, \frac{p-1}{2}$, which, multiplied side by side, lead to

$$(-1)^{\frac{p-1}{2}} 1 \cdot 2 \cdot \dots \cdot \frac{p-1}{2} \equiv \frac{p+1}{2} \cdot \frac{p+3}{2} \cdot \dots \cdot (p-1) \pmod{p}.$$

By multiplying both sides of the last congruence with $1 \cdot 2 \cdot \dots \cdot \frac{p-1}{2}$, we find

$$(-1)^{\frac{p-1}{2}} \left(\left(\frac{p-1}{2} \right)! \right)^2 \equiv (p-1)! \pmod{p}.$$

Now, taking into account that $\frac{p-1}{2}$ is even (p being of the form $4k+1$) and by using Wilson’s theorem, the last congruence becomes

$$\left(\left(\frac{p-1}{2} \right)! \right)^2 \equiv -1 \pmod{p},$$

which shows that for $c = \left(\frac{p-1}{2} \right)!$ we have $c^2 + 1 \equiv 0 \pmod{p}$, that is, $p \mid c^2 + 1$.

(b) For the number c from the first part, an integer a exists such that

$$|a| < \frac{p}{2} \quad \text{and} \quad a \equiv c \pmod{p}.$$

Consequently,

$$a^2 + 1 \equiv c^2 + 1 \equiv 0 \pmod{p},$$

meaning that $a^2 + 1 = mp$ for some integer m . As $a^2 + 1 > 0$, we have $m \geq 1$; on the other hand,

$$mp = a^2 + 1 < \left(\frac{p}{2} \right)^2 + 1 < p^2,$$

so we also have $m < p$, as desired. Thus some integer m with $1 \leq m < p$ can be found such that $mp = a^2 + 1 = a^2 + 1^2$ is representable.

Example 3. For instance, with $p = 13$, we have:

$$(6!)^2 = 6! \cdot 6! \equiv 6! \cdot (-7) \cdot (-8) \cdot (-9) \cdot (-10) \cdot (-11) \cdot (-12) = 12! \equiv -1 \pmod{13},$$

where the first congruence relies on $6 \equiv -7$, $5 \equiv -8$, and so on, until $1 \equiv -12$ (all congruences are modulo 13), and the second is Wilson's theorem. In other words, we have that 13 divides $(6!)^2 + 1 = 720^2 + 1$. Check the congruence $720^2 + 1 \equiv 0 \pmod{13}$ without using Wilson's theorem!

As $720 \equiv 5 \pmod{13}$ it follows that $5^2 + 1 \equiv 720^2 + 1 \equiv 0 \pmod{13}$. And, indeed, we have $5^2 + 1 = 2 \cdot 13$, hence the numbers from the lemma can be taken to be $c = 6! = 720$, $a = 5$, and $m = 2$.

Exercise 4. Find the numbers c , a , and m for $p = 17$.

Exercise 5. Prove Wilson's theorem in general, after looking at the particular cases

$$4! + 1 = 1 \cdot 4 \cdot (2 \cdot 3) \equiv 1 \cdot 4 \cdot 1 \equiv -1 \pmod{5},$$

$$6! + 1 = 1 \cdot 6 \cdot (2 \cdot 4) \cdot (3 \cdot 5) \equiv 1 \cdot 6 \cdot 1 \cdot 1 \equiv -1 \pmod{7},$$

or

$$10! = 1 \cdot 10 \cdot (2 \cdot 6) \cdot (3 \cdot 4) \cdot (5 \cdot 9) \cdot (7 \cdot 8) \equiv 1 \cdot 10 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \equiv -1 \pmod{11}.$$

We can see now one proof for theorem 1.

First proof of Fermat's Christmas theorem. We want to show that a prime $p \equiv 1 \pmod{4}$ can be expressed as a sum of two perfect squares. According to the second part of lemma 2, we can find the natural numbers m , a , and b such that $1 \leq m < p$ and $mp = a^2 + b^2$ (the proof actually provides $b = 1$, but this is not important, or useful in what follows). Moreover, we consider further that mp is *the smallest* representable positive multiple of p : therefore we have $1 \leq m < p$ and $mp = a^2 + b^2$ for some natural numbers a and b (not both equal to 0, evidently), but $m'p$ is *not* representable for any $1 \leq m' < m$. Of course, if $m = 1$, then there is nothing left to prove, so we can assume further that $1 < m < p$.

The variant of the division theorem that we used in the proof of the previous lemma ensures the existence of some integers x and y such that

$$x \equiv a \pmod{m}, \quad y \equiv b \pmod{m}$$

and

$$|x| \leq \frac{m}{2}, \quad |y| \leq \frac{m}{2}.$$

We note that x and y cannot be both equal to 0, since then we would have $a \equiv 0 \pmod{m}$ and $b \equiv 0 \pmod{m}$, hence m^2 would divide $a^2 + b^2 = mp$, meaning that this m would divide p —which is clearly impossible for $1 < m < p$. We also have

$$x^2 + y^2 \equiv a^2 + b^2 = mp \equiv 0 \pmod{m},$$

meaning that a natural number n exists such that $x^2 + y^2 = mn$. The previous observation shows that n cannot be 0, and, on the other hand, we have

$$mn = x^2 + y^2 \leq \left(\frac{m}{2}\right)^2 + \left(\frac{m}{2}\right)^2 = \frac{m^2}{2} < m^2,$$

thus $1 \leq n < m$.

Now we have

$$m^2 np = (mp)(mn) = (a^2 + b^2)(x^2 + y^2) = (ax + by)^2 + (ay - bx)^2,$$

where

$$ax + by \equiv a^2 + b^2 = mp \equiv 0 \pmod{m}$$

and

$$ay - bx \equiv ab - ba = 0 \pmod{m}.$$

Consequently, we can find integers s and t for which $ax + by = ms$ and $ay - bx = mt$, therefore the above equality reads

$$m^2 np = (a^2 + b^2)(x^2 + y^2) = (ax + by)^2 + (ay - bx)^2 = m^2(s^2 + t^2),$$

that is, $np = s^2 + t^2$.

We summarize: the assumption that $1 < m$ leads to the existence of an n with $1 \leq n < m$ and such that np is representable. However, this contradicts the minimality of m , and the contradiction shows that we must have $m = 1$, that is, the smallest representable multiple of p is p —which finishes the proof.

Note that the existence of $n < m$ for which np is representable only relies on the hypothesis that mp is representable for some $m > 1$ (it does not use the fact that mp is the smallest such multiple of p). Thus we can start with any $m > 1$ for which mp is the sum of two squares, and we find a positive integer $n < m$ for which np has the same property, then we can go on with the demonstration as follows. If the n that we found is $n = 1$ the proof is done, otherwise the same argument applied to n instead of m leads to some $1 \leq n_1 < n$ for which $n_1 p$ is representable. Then, again, if $1 < n_1$, we can find some $1 \leq n_2 < n_1$ such that $n_2 p$ is representable, and so on. Thus, if each n_j was greater than 1, we would obtain an infinite and strictly decreasing sequence $n = n_0 > n_1 > n_2 > \dots$ of positive integers such that every $n_j p$ ($j = 0, 1, \dots$) is representable. But such a (strictly decreasing) sequence of positive integers clearly does not exist, so we infer that the smallest representable multiple mp of p is obtained for $m = 1$, finishing the proof. Any of the two variants of the proof uses Fermat's method of *infinite descent*, but the second one shows more clearly what the infinite descent really is.

Example 5. We see thus that if we have a representable multiple mp of p (with $1 < m < p$), then another such representable multiple np of p for which $1 \leq n < m$ can be built—and the above proof gives us a method to do that. Namely, if $mp = a^2 + b^2$ for some integers a and b , we first choose integers $x \equiv a \pmod{m}$, $y \equiv b \pmod{m}$ for which $|x| \leq \frac{m}{2}$, $|y| \leq \frac{m}{2}$, and we get $ax + by = ms$ and $ay - bx = mt$ with integers s and t . Then $s^2 + t^2$ is this multiple np .

For instance, we have seen that $2 \cdot 13 = 5^2 + 1^2$, where we have $p = 13$, $m = 2$, $a = 5$ and $b = 1$. Consequently, $x = 1$, $y = 1$, $ax + by = 6 = 2 \cdot 3$ and $ay - bx = 4 = 2 \cdot 2$. It follows that $s = 3$, $t = 2$ and, indeed, $3^2 + 2^2 = 13$ (so we get precisely the representation of 13 as a sum of two squares—which was to be expected, since the only positive multiple of 13 less than $2 \cdot 13$ is 13).

We also have the equality $10 \cdot 17 = 13^2 + 1^2$ (with $p = 17$, $m = 10$, $a = 13$ and $b = 1$). We then find $x = 3$, $y = 1$, $ax + by = 40 = 10 \cdot 4$ and $ay - bx = 10 = 10 \cdot 1$, hence $4^2 + 1^2$ is the representable multiple of 17, lesser than $10 \cdot 17$ (we got again the representation of 17 as a sum of two squares).

We invite the interested reader to check that, starting with the equality $17 \cdot 29 = 22^2 + 3^2$ ($p = 29$, $m = 17$, $a = 22$, $b = 3$, $x = 5$, $y = 3$), we first arrive to $2 \cdot 29 = 7^2 + 3^2$ and then, by repeating the procedure (with $m = 2$, $a = 7$, $b = 3$, $x = 1$, $y = 1$), we get $29 = 5^2 + 2^2$. What do we get if (also for $p = 29$) we start from $13 \cdot 29 = 16^2 + 11^2$? Or from $13 \cdot 29 = 19^2 + 4^2$? Or (for $p = 41$) from $17 \cdot 41 = 24^2 + 11^2$?

3. The second proof of Fermat's Christmas theorem also uses the first part of lemma 2, and one more result known as Thue's lemma.

Lemma 6. (*Thue's lemma*) Let $m \geq 1$ and l be integers, and let A, B be positive real numbers for which $AB \geq m$. Then integers x and y exist such that they are not both zero, $|x| \leq A$, $|y| \leq B$, and $lx \equiv y \pmod{m}$.

Proof. We consider the numbers $lj - k$ for all $j \in \{0, 1, \dots, \lfloor A \rfloor\}$ and $k \in \{0, 1, \dots, \lfloor B \rfloor\}$ (where $\lfloor t \rfloor$ denotes the integral part of the real number t). There are $(\lfloor A \rfloor + 1)(\lfloor B \rfloor + 1) > AB \geq m$ such numbers, therefore (their number being greater than m) two of them yield the same remainder when divided by m . Thus, we can find two distinct pairs (j_1, k_1) and (j_2, k_2) , where $j_1, j_2 \in \{0, 1, \dots, \lfloor A \rfloor\}$, $k_1, k_2 \in \{0, 1, \dots, \lfloor B \rfloor\}$, and

$$lj_1 - k_1 \equiv lj_2 - k_2 \pmod{m} \Leftrightarrow l(j_1 - j_2) \equiv k_1 - k_2 \pmod{m}.$$

Evidently this implies that the numbers $x = j_1 - j_2$ and $y = k_1 - k_2$ fulfill the required conditions (they are integers and at least one of them is nonzero, because the pairs (j_1, k_1) and (j_2, k_2) are distinct; we also have $|x| = |j_1 - j_2| \leq \lfloor A \rfloor \leq A$ and $|y| = |k_1 - k_2| \leq \lfloor B \rfloor \leq B$, and, clearly, $lx \equiv y \pmod{m}$).

The second proof of Fermat's Christmas theorem. Such as in the first proof, we only need to show that any prime $p \equiv 1 \pmod{4}$ is representable. According to lemma 2, there exists a natural number c such that p divides $c^2 + 1$. Thue's lemma (for $m = p$, $l = c$ and $A = B = \sqrt{p}$) ensures the existence of integers x and y , not both zero, such that $|x| \leq \sqrt{p}$, $|y| \leq \sqrt{p}$ and $cx \equiv y \pmod{p}$. The congruence implies $c^2x^2 \equiv y^2 \pmod{p}$, hence it implies that p divides $c^2x^2 - y^2$. We also note that we actually have $|x| < \sqrt{p}$ and $|y| < \sqrt{p}$, because $|x|$ and $|y|$ are integers, while $\sqrt{p}$ is not. Then we get that p divides

$$x^2 + y^2 = (c^2 + 1)x^2 - (c^2x^2 - y^2).$$

Moreover, the inequalities for x and y , and the fact that at least one of them is nonzero yield

$$0 < x^2 + y^2 = |x|^2 + |y|^2 < p + p = 2p.$$

Now we conclude: $x^2 + y^2$ is a multiple of p strictly between 0 and $2p$, meaning that $x^2 + y^2 = p$, therefore that p is representable, as we intended to demonstrate.

We note once again the usefulness of part (a) of lemma 2 in the second proof of theorem 1, as well as the fact that using Thue's lemma considerably simplifies the reasoning, and brings an extra touch of elegance to it (compared to the first proof).

Example 7. For instance, when $p = 29$, we can take $c = 14! \pmod{29} = 12$ for which $p \mid c^2 + 1$. Thue's lemma says that there are integers x and y , not both 0, with absolute values at most $\lfloor \sqrt{p} \rfloor = 5$, and such that $cx \equiv y \pmod{p}$. In fact, we immediately see that $12 \cdot 5 = 60 \equiv 2 \pmod{29}$, so we can choose $x = 5$ and $y = 2$. From the proof above we infer that $29 = 5^2 + 2^2$.

We finish this paragraph with three more exercises for the interested reader: one (pretty obvious) generalization of Thue's lemma, and two more theorems, both similar to the Christmas theorem, and both also stated by Fermat (but, again, we do not know whether Fermat had proofs to these results).

Exercise 8. Let m and n be positive integers, let $a_1, \dots, a_n$ be integers, and let $A_1, \dots, A_n$ be positive real numbers for which $A_1 \cdots A_n \geq m$. Then there exist integers $x_1, \dots, x_n$, not all equal to 0, such that $0 \leq x_i \leq A_i$ for $i = 1, \dots, n$ and $a_1x_1 + \cdots + a_nx_n \equiv 0 \pmod{m}$.

Exercise 9. In this exercise we consider known the fact that, for an odd prime p , a natural number c such that $p \mid c^2 + 2$ exists if and only if p is either 1, or 3 modulo 8. (In other words, -2 is a quadratic residue modulo p if and only if p is congruent to either 1 or 3 modulo 8—one of the special cases of the quadratic reciprocity law.)

Prove that a prime p can be expressed in the form $x^2 + 2y^2$, with integers x and y if and only if $p = 2$ or p is congruent to 1, or to 3 modulo 8.

Exercise 10. The quadratic reciprocity law also implies that for a prime $p > 3$ there exists a natural number c such that $p \mid c^2 + 3$ if and only if p is congruent to 1 modulo 3. (In other words, -3 is a quadratic residue modulo p if and only if p is congruent to 1 modulo 3.)

Prove that a prime p can be expressed in the form $x^2 + 3y^2$, with integers x and y if and only if $p = 3$ or p is congruent to 1 modulo 3.

4. The third proof of Fermat's Christmas theorem uses the ring of *Gaussian integers*, especially its arithmetic properties. This is the ring formed by the set

$$\mathbb{Z}[i] = \{a + bi \mid a, b \in \mathbb{Z}\}$$

together with the usual operations of addition and multiplication of the complex numbers. This ring clearly is a subring of the field of complex numbers (and, on its turn, it contains the ring of usual integers $\mathbb{Z}$ as a subring). Consequently, $\mathbb{Z}[i]$ is commutative and has no nonzero zero divisors (as one says, it is an *integral domain*). Moreover, the *invertible elements* (or *units*) of $\mathbb{Z}[i]$ are 1, -1 , i and $-i$, which can be easily proved by using the *norm* of the Gaussian integers. The norm is a function $N : \mathbb{Z}[i] \rightarrow \mathbb{N}$ that maps every element $z = a + bi$ from $\mathbb{Z}[i]$ to a nonnegative integer $N(z) \in \mathbb{N}$. More precisely, we define

$$N(z) = z\bar{z} \text{ or } N(a + bi) = (a + bi)(a - bi) = a^2 + b^2$$

for each $z = a + bi \in \mathbb{Z}[i]$ (we denote by $\bar{z} = a - bi$ the *complex conjugate* of z). We can also say that the norm $N(z)$ of z is the square of the absolute value of z . One immediately sees that the norms of the Gaussian integers are precisely the representable natural numbers, and that the norm is a *multiplicative* function (such as the absolute value, or the function $z \mapsto \bar{z}$ that associates to each complex number its conjugate). This means that we have $N(wz) = N(w)N(z)$ for all $w, z \in \mathbb{Z}[i]$. Indeed, this equality is easily seen to be equivalent to the *Brahmagupta-Fibonacci identity*:

$$(a^2 + b^2)(c^2 + d^2) = (ac - bd)^2 + (ad + bc)^2,$$

which can be proved by direct computation, and which, of course, holds for any kind of numbers (not necessarily integers) a, b, c, d . (Actually, this identity's first appearance is in Diophantus's book *Arithmetica*. Brahmagupta knew a more general form of the identity, which was brought back to Europe by Fibonacci, via the Arabian mathematicians.) Thus, the multiplicative property of the norm shows that the product of two (or more) representable numbers is also representable. For example, $29 = 5^2 + 2^2$ and $117 = 9^2 + 6^2$ are both representable, hence their product must be so. More precisely, we have (according to the above identity)

$$3393 = 29 \cdot 117 = (5 \cdot 9 - 2 \cdot 6)^2 + (5 \cdot 6 + 2 \cdot 9)^2 = 33^2 + 48^2.$$

The multiplicative property of the norm is useful in many situations. For instance (as said before), it helps to finding the units of $\mathbb{Z}[i]$. Namely, if $z = a + bi \in \mathbb{Z}[i]$ is a unit, then $z' = a' + b'i \in \mathbb{Z}[i]$ exists such that $zz' = 1$, whence we get the equality

$$(a^2 + b^2)(a'^2 + b'^2) = N(z)N(z') = N(zz') = N(1) = 1,$$

in which (a, b, a', b') being integers) $a^2 + b^2$ and $a'^2 + b'^2$ are natural numbers, therefore $a^2 + b^2 = a'^2 + b'^2 = 1$ necessarily follows. As $a^2 + b^2 = 1$ is only possible in the two cases $a = \pm 1$ and $b = 0$, or $a = 0$ and $b = \pm 1$, we conclude that z can only be one of the numbers 1, -1 , i , or $-i$. Conversely, we immediately see that all these are invertible elements in $\mathbb{Z}[i]$, their inverses being 1, -1 , $-i$, and i respectively.

The division theorem works for the Gaussian integers (although there are a few minor differences compared to the division theorem for the usual integers), and it is fundamental in establishing that the arithmetic of $\mathbb{Z}[i]$ is essentially the same as the arithmetic of $\mathbb{Z}$.

Theorem 11. (*The division theorem for the Gaussian integers*) For any $w, z \in \mathbb{Z}[i]$ with $z \neq 0$ there exist $q \in \mathbb{Z}[i]$ (the quotient of w divided to z) and $r \in \mathbb{Z}[i]$ (the remainder of the division) such that

$$w = zq + r \text{ and } N(r) < N(z).$$

Proof. Let

$$\frac{w}{z} = x + yi,$$

where x and y are rational numbers. For these we can find integers s and t such that

$$|x - s| \leq \frac{1}{2} \quad \text{and} \quad |y - t| \leq \frac{1}{2}.$$

Thus we get

$$N((x - s) + (y - t)i) = (x - s)^2 + (y - t)^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

As

$$w = z(x + yi) = z(s + ti) + z((x - s) + (y - t)i),$$

we can choose $q = s + ti$ and $r = z((x - s) + (y - t)i)$ for which the required conditions are fulfilled: we have $q \in \mathbb{Z}[i]$ and $w = zq + r$, hence $r = w - zq \in \mathbb{Z}[i]$ also follows, and

$$N(r) = N(z)N((x - s) + (y - t)i) \leq \frac{N(z)}{2} < N(z).$$

For instance, when $w = 17 + 14i$ and $z = 3 + i$, we have

$$\frac{w}{z} = \frac{13}{2} + \frac{5}{2}i,$$

so we can choose $s = 6$, $t = 3$ that lead to $q = 6 + 3i$ and $r = (3 + i)\left(\frac{1}{2} - \frac{1}{2}i\right) = 2 - i$. We then have, indeed,

$$w = zq + r \Leftrightarrow 17 + 14i = (3 + i)(6 + 3i) + (2 - i)$$

and

$$N(r) = 5 < 10 = N(z).$$

Note that we can also choose (for example) $s = 6$, $t = 2$, yielding the quotient $q = 6 + 2i$ and the remainder $r = 1 + 2i$ (and two more choices are possible), that is (unlike for the division in the ring of integers) the uniqueness of the quotient and the remainder is no longer ensured. However, uniqueness can be achieved if we impose to s and t the slightly more restrictive conditions

$$-\frac{1}{2} \leq x - s < \frac{1}{2} \quad \text{si} \quad -\frac{1}{2} \leq y - t < \frac{1}{2}.$$

We can define divisibility in $\mathbb{Z}[i]$ such as in $\mathbb{Z}$: we say that $z \in \mathbb{Z}[i]$ is *divisible* by $w \in \mathbb{Z}[i]$ (or that w *divides* z ; w is a *divisor* of z , and z is a *multiple* of w) whenever there exists $v \in \mathbb{Z}[i]$ such that $z = wv$. We denote this by $w \mid z$ (as for the usual integers). For example, $7 + 3i$ is divisible by $5 - 2i$ because $7 + 3i = (5 - 2i)(1 + i)$, or $a + bi$ is divisible by c in the ring of Gaussian integers (a, b, c being integers) if and only if both a and b are divisible by c in the ring of integers. Also, one easily checks that $a + bi$ is divisible by $1 + i$ if and only if a and b have the same parity.

Two elements z and w are called *associated in divisibility* or, briefly, *associates* if and only if each of them divides the other. We invite the reader to show that z and w are associates (that is, we have $z \mid w$ and $w \mid z$) if and only if a unit u exists such that $z = uw$ (this holds in any integral domain, in particular in the ring of integers and in the ring of Gaussian integers). Thus, in $\mathbb{Z}[i]$, the associates of $z = a + bi$ are z itself, $-z = -a - bi$, $iz = -b + ai$ and $-iz = b - ai$. (In the ring of integers the associates of x are x and $-x$). The association (in divisibility) is an equivalence relation and it

preserves the properties related to divisibility. For instance, z is divisible by w if and only if any associate of z is divisible by any associate of w .

A nonzero and not invertible element $z \in \mathbb{Z}[i]$ is named *reducible* if it can be expressed as the product of two not invertible (and not associated to z) elements, and *irreducible* otherwise. Thus, $z \in \mathbb{Z}[i]$ is irreducible if an equality $z = z_1 z_2$, with $z_1, z_2 \in \mathbb{Z}[i]$ implies that either z_1 is a unit and z_2 is associate to z , or z_2 is a unit and z_1 is associate to z . For example, $7 + 3i = (5 - 2i)(1 + i)$ is reducible, while $1 + i$ is irreducible (which is easily seen by using the norm). A nonzero and not invertible element $z \in \mathbb{Z}[i]$ is called *prime* if and only if from the fact that z divides a product $z_1 z_2$ it follows that z divides one of the factors z_1 , or z_2 (which immediately extends to any number of factors: if a prime divides a product, then it must divide one of the factors). The elements $0, \pm 1, \pm i$ are considered neither reducible, nor irreducible, neither primes nor not primes (such as are considered $0, 1$, and -1 in $\mathbb{Z}$). We invite the reader to show that any prime element of $\mathbb{Z}[i]$ is irreducible (which is immediate), and, also, to prove that any nonzero and not invertible element of $\mathbb{Z}[i]$ can be expressed as a product of irreducible elements (which can be easily proved either by inducting on the norm of the element, or by contradiction, considering an element that cannot be expressed as product of irreducible elements, and that has a minimal norm among such elements).

We do not intend to expose here the whole theory of divisibility in the ring of Gaussian integers, but we mention two more important results due to which this theory works precisely as in the ring of integers.

Theorem 12. (*Euclid's lemma for the Gaussian integers*) Any irreducible element from $\mathbb{Z}[i]$ is also a prime element. That is, if p is an irreducible element from $\mathbb{Z}[i]$, and $p \mid wz$ for some $w, z \in \mathbb{Z}[i]$, then either $p \mid w$, or $p \mid z$.

Since, as we said above, in $\mathbb{Z}[i]$ every prime element is also irreducible, it follows that (such as for the usual integers) in the ring of the Gaussian integers the prime elements coincide with the irreducible elements (and we will see further which they are). Also note that a more general form of theorem 12 (sometimes called Gauss's lemma) says that, if $x, y, z \in \mathbb{Z}[i]$, z divides xy and is relatively prime to x (that is, if z and x have no common divisor except for the units), then z necessarily divides y . Euclid's lemma is essential in proving that unique factorization holds in $\mathbb{Z}[i]$, which is stated in the following very important result.

Theorem 13. (*Fundamental theorem of arithmetic in $\mathbb{Z}[i]$*) Any nonzero and nonunit (not invertible) element from $\mathbb{Z}[i]$ can be uniquely (apart from the changing of the order of factors, and associations in divisibility) expressed as a product of irreducible (equivalently: prime) elements. In other words, for any $z \in \mathbb{Z}[i] \setminus \{0, 1, -1, i, -i\}$ there exist irreducible (or, equivalently, prime) elements $p_1, \dots, p_k$ in $\mathbb{Z}[i]$ such that $z = p_1 \cdots p_k$ (this is the existence part). Moreover (and here comes the uniqueness part), if $z = q_1 \cdots q_l$ is also a representation of z as a product of irreducible (or prime) elements, then $k = l$ and a permutation α of the set $\{1, \dots, k\}$ exists such that q_j and $p_{\alpha(j)}$ are associates for any $j = 1, \dots, k$.

We invite the interested reader to think how one can define the greatest common divisor and the least common multiple of a family of Gaussian integers, which properties they have, and how are these properties related to the above theorems (basically there is a total analogy with the corresponding notions and properties for the integers). For now, we are ready to see one proof of Fermat's Christmas theorem that relies on the arithmetic of Gaussian integers.

The third proof of Fermat's Christmas theorem. As in the first two proofs, we still need to show that a prime $p \equiv 1 \pmod{4}$ can be written as a sum of two squares. We emphasize that p is a prime *in the ring of integers*, and *not* in the ring of Gaussian integers (we immediately see what is its status from this point of view in $\mathbb{Z}[i]$). We use again part (a) of lemma 2, according to which an integer c exists such that p divides $c^2 + 1$. Evidently, this divisibility still holds in $\mathbb{Z}[i]$, where we have that p divides $c^2 + 1 = (c + i)(c - i)$. If p was a prime in $\mathbb{Z}[i]$, then it would divide one of the factors $c + i, c - i$

(as it divides their product), which is false; for instance, if p divided $c + i$, then integers x, y would exist such that $p(x + yi) = c + i$, whence $py = 1$ would follow, clearly impossible. Consequently p is not a prime, hence it is not irreducible in $\mathbb{Z}[i]$, and this means that we can find $s + ti, u + vi \in \mathbb{Z}[i]$, not invertible (and thus not associates with p), and for which $p = (s + ti)(u + vi)$. For the norms we get the equality

$$p^2 = N(p) = N(s + ti)N(u + vi)$$

in which the norms of $s + ti$ and $u + vi$ are positive integers, and none of them can be 1 (otherwise, $s + ti$ or $u + vi$ would be a unit in $\mathbb{Z}[i]$), hence none of them can be p^2 . Then the only possibility is to have $N(s + ti) = N(u + vi) = p \Leftrightarrow s^2 + t^2 = u^2 + v^2 = p$, and thus p is, indeed, the sum of two squares of integers, as we intended to prove.

Note that we have $(s + ti)(u + vi) = p = s^2 + t^2 = (s + ti)(s - ti)$, therefore $u + vi = s - ti$ is actually the conjugate of $s + ti$ (showing that $p = u^2 + v^2$ and $p = s^2 + t^2$ are not essentially different expressions of p as a sum of two squares). We invite the reader to prove the statements from the following exercise (which, among others, depicts the primes from $\mathbb{Z}[i]$), by using the norm as we did above.

Exercise 14. (a) Show that $1 + i$ is a prime (so an irreducible element, too) in $\mathbb{Z}[i]$.

(b) Show that any prime $q \equiv 3 \pmod{4}$ is also a prime in $\mathbb{Z}[i]$.

(c) Show that, if $p \equiv 1 \pmod{4}$ is expressed as a sum of two squares in the form $p = a^2 + b^2$, then $a + bi$ and $a - bi$ are primes in $\mathbb{Z}[i]$, and they are not associates.

(d) Infer that the primes from $\mathbb{Z}[i]$ are $1 + i$, the (integer) primes $q \equiv 3 \pmod{4}$, the nonunit factors in $\mathbb{Z}[i]$ of the (integer) primes $p \equiv 1 \pmod{4}$, and, of course, all their associates.

(e) Show that if $p = s^2 + t^2 = u^2 + v^2$ are two representations as sums of two squares of a prime $p \equiv 1 \pmod{4}$, then $s + ti$ and $s - ti$ are associated in divisibility, in some order, to $u + vi$ and $u - vi$ respectively, therefore the pair (u, v) coincides with one of the pairs $(\pm s, \pm t)$, or $(\pm t, \pm s)$. In other words, the representation of p as a sum of squares $s^2 + t^2$ is unique if we disregard the changing of the order of summands, or of the signs of s and t .

5. Finally, which numbers are representable? In order to completely describe the natural numbers that can be written as sums of two squares, we need one more helping result.

Lemma 15. Let $p \equiv 3 \pmod{4}$ be a prime, and let a, b be integers such that p divides $a^2 + b^2$. Then p divides a and p divides b .

Proof. Note that for p, a and b as in the statement of the lemma (with p dividing $a^2 + b^2$), we have that p divides a if and only if p divides b . So, if we assume by contradiction that the conclusion is false, that is, we assume that p does not divide at least one of a and b , we actually have that p divides neither a , nor b . Then a is invertible modulo p , meaning that we can find an integer a' such that $aa' \equiv 1 \pmod{p}$. Consequently

$$a^2 + b^2 \equiv 0 \pmod{p} \Rightarrow (aa')^2 + (ba')^2 \equiv 0 \pmod{p} \Rightarrow 1 + c^2 \equiv 0 \pmod{p},$$

or $-1 \equiv c^2 \pmod{p}$ for the integer $c = ba'$, which is relatively prime to p (since so are both b and a'). Now, by hypothesis, $\frac{p-1}{2}$ is odd, and by using Fermat's theorem we get

$$-1 = (-1)^{\frac{p-1}{2}} \equiv c^{p-1} \equiv 1 \pmod{p},$$

which is clearly impossible, p being odd. It remains that the assumption we made is false, hence at least one of a and b is divisible by p . As shown in the beginning of the proof, this means that p divides both a and b , as desired.

We can now provide the answer to the question from the title of this paragraph (which is actually hidden behind the very title of this paper).

Theorem 16. A natural number n can be represented as a sum of two perfect squares if and only if $n = 0$, or $n = 1$, or $n > 1$ and all the prime factors congruent to 3 modulo 4 from the prime factorization of n have, in this factorization, an even exponent.

Proof. For the "if" part, we see that 0 and 1 are representable, so further we can consider that $n > 1$ has a prime factorization as in the statement of the theorem, say this is $n = 2^\alpha p_1^{\beta_1} \cdots p_i^{\beta_i} q_1^{\gamma_1} \cdots q_j^{\gamma_j}$, where $p_1, \dots, p_i$ are distinct primes congruent to 1 modulo 4, $q_1, \dots, q_j$ are distinct primes congruent to 3 modulo 4, and the exponents $\alpha, \beta_1, \dots, \beta_i$ are natural numbers, while $\gamma_1, \dots, \gamma_j$ are even natural numbers. By repeatedly applying the Brahmagupta-Fibonacci identity, one readily sees that any product of representable numbers is also representable. The conclusion follows from this observation and the fact that each of the factors $2, p_1, \dots, p_i$, and $q_1^{\gamma_1}, \dots, q_j^{\gamma_j}$ of n is representable. Indeed, 2 is obviously representable, $p_1, \dots, p_i$ are representable according to Fermat's Christmas theorem, while $q_1^{\gamma_1}, \dots, q_j^{\gamma_j}$ are representable because all of them are squares (as all the exponents $\gamma_1, \dots, \gamma_j$ are even).

Now for the "only if" part we assume, by contradiction, that there exists at least a representable (of course, positive) integer in whose factorization (at least) a prime congruent to 3 modulo 4 appears with an odd exponent. Denote by m the smallest positive integer with these properties (note that we surely have $m > 1$), and let $q \equiv 3 \pmod{4}$ be a prime whose exponent in the factorization of m is odd, say it is $2\gamma + 1$ (where γ is a nonnegative integer), meaning that $q^{2\gamma+1}$ divides m , but $q^{2\gamma+2}$ does not. Moreover, let $m = a^2 + b^2$ be an expression of m as a sum of two squares. Because $q^{2\gamma+1}$ divides m and $2\gamma + 1 \geq 1$, we infer that q divides $a^2 + b^2$. Then, according to lemma 15, q divides a and q divides b , thus there are integers c and d for which $a = qc$ and $b = qd$. Thus $q^{2\gamma+1}$ divides $q^2(c^2 + d^2)$, hence $q^{2\gamma-1}$ divides $m' = c^2 + d^2$. Moreover $q^{2\gamma}$ does not divide m' , otherwise $q^{2\gamma+2}$ would divide $m = q^2 m'$ (which is not the case). Finally, from $m = q^2 m'$ we get $1 < m' < m$. We summarize: our assumption that a representable integer m exists such that in its factorization a prime congruent to 3 modulo 4 appears with an odd exponent led us to the existence of another positive integer $m' < m$ with the same properties ($m' = c^2 + d^2$, and the exponent of q in the factorization of m' is $2\gamma - 1$). But this contradicts the minimality of m , so the assumption is false, and the theorem is proved.

For instance, $585 = 3^2 \cdot 5 \cdot 13$ is representable because it has only one prime congruent to 3 modulo 4 in its factorization, and the exponent of that prime is 2 (585 can be expressed as a sum of two squares in the form $585 = 3^2(1^2 + 8^2) = 3^2 + 24^2$, or in the form $585 = 3^2(4^2 + 7^2) = 12^2 + 21^2$), while $950 = 2 \cdot 5^2 \cdot 19$ is not, because its prime factor 19 (congruent to 3 modulo 4) has the odd exponent 1.

6. We finish with a few more exercises about representable numbers for the interested reader.

Exercise 17. Prove that there are infinitely many triples of consecutive integers such that each member of every such triple is representable.

One idea for the proof is to note that if $(n, n + 1, n + 2)$ is such a triple (in which all numbers $n, n + 1$ and $n + 2$ are representable), then $(4n(n + 1), 4n(n + 1) + 1, 4n(n + 1) + 2)$ is also a triple of representable consecutive natural numbers. Starting with $(8, 9, 10)$ and applying this transform indefinitely solves the problem (inductively). We invite the reader to find two more such triples and, also, to try to solve the problem by finding parameterizations for n such that $n, n + 1$, and $n + 2$ are all sums of two squares. More generally, we have:

Exercise 18. Let h and k be given arbitrary integers. Show that there exist infinitely many natural numbers n such that $n, n + h$ and $n + k$ are all representable.

A parameterization for n would be indicated (although it is harder to find compared to the particular case $h = 1, k = 2$). The inductive way seems not so productive for the general case. Also note that theorem 16 seems not to be useful in solving such a problem.

Exercise 19. Let $s \geq 2$ be an integer. Prove that there exist s consecutive natural numbers, each of which is *not* representable.

This time, theorem 16 can show its power.

Exercise 20. (a) Show that if a natural number n can be written as the sum of the squares of two

rational numbers, then n is also the sum of the squares of two integers.

(b) Show that, if the quotient of two representable positive integers is also an integer, then this quotient is representable, too.

Exercise 21. For a positive integer n let $d_j(n)$ be the number of positive divisors of n which are congruent to j modulo 4, for $j \in \{1, 3\}$. Prove that the number of the representations of n as the sum of the squares of two integers is $4(d_1(n) - d_3(n))$. Here the representations $n = a^2 + b^2 = c^2 + d^2$ are considered distinct whenever the pairs (a, b) and (c, d) differ.

Collaterally, one can prove that the function $n \mapsto f(n) = d_1(n) - d_3(n)$ is multiplicative, that is, we have $f(mn) = f(m)f(n)$ for any relatively prime positive integers m and n . It also follows that $d_1(n) \geq d_3(n)$ for all n , with equality if and only if in the factorization of n at least one prime $q \equiv 3 \pmod{4}$ appears with an odd exponent.

References

1. W. Sierpinski: *Elementary Theory of Numbers*, North Holland, 1988
2. Fermat's Christmas theorem (https://en.wikipedia.org/wiki/Fermat%27s_theorem_on_sums_of_two_squares)