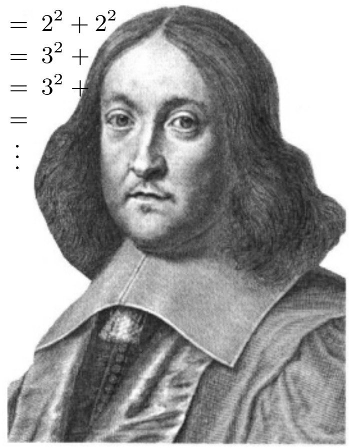
Representing numbers as sums of two squares

Chapter 4

> *Which numbers can be written as sums of two squares?*

$$\begin{aligned}
1 &= 1^2 + 0^2\\
2 &= 1^2 + 1^2\\
3 &=\\
4 &= 2^2 + 0^2\\
5 &= 2^2 + 1^2\\
6 &=\\
7 &=\\
8 &= 2^2 + 2^2\\
9 &= 3^2 +\\
10 &= 3^2 +\\
11 &=\\
&\vdots
\end{aligned}$$



Pierre de Fermat

This question is as old as number theory, and its solution is a classic in the field. The "hard" part of the solution is to see that every prime number of the form $4m + 1$ is a sum of two squares. G. H. Hardy writes that this *two square theorem* of Fermat "is ranked, very justly, as one of the finest in arithmetic." Nevertheless, one of our Book Proofs below is quite recent.

Let's start with some "warm-ups." First, we need to distinguish between the prime $p = 2$, the primes of the form $p = 4m + 1$, and the primes of the form $p = 4m + 3$. Every prime number belongs to exactly one of these three classes. At this point we may note (using a method "à la Euclid") that there are infinitely many primes of the form $4m + 3$. In fact, if there were only finitely many, then we could take p_k to be the largest prime of this form. Setting

$$N_k \;:=\; 2^2 \cdot 3 \cdot 5 \cdots p_k - 1$$

(where $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \ldots$ denotes the sequence of all primes), we find that N_k is congruent to $3 \pmod 4$, so it must have a prime factor of the form $4m + 3$, and this prime factor is larger than p_k — contradiction.

Our first lemma characterizes the primes for which -1 is a square in the field $\mathbb{Z}_p$ (which is reviewed in the box on the next page). It will also give us a quick way to derive that there are infinitely many primes of the form $4m + 1$.

Lemma 1. *For primes $p = 4m + 1$ the equation $s^2 \equiv -1 \pmod p$ has two solutions $s \in \{1, 2, \ldots, p-1\}$, for $p = 2$ there is one such solution, while for primes of the form $p = 4m + 3$ there is no solution.*

■ **Proof.** For $p = 2$ take $s = 1$. For odd p, we construct the equivalence relation on $\{1, 2, \ldots, p-1\}$ that is generated by identifying every element with its additive inverse and with its multiplicative inverse in $\mathbb{Z}_p$. Thus the "general" equivalence classes will contain four elements

$$\{x, -x, \overline{x}, -\overline{x}\}$$

since such a 4-element set contains both inverses for all its elements. However, there are smaller equivalence classes if some of the four numbers are not distinct:

M. Aigner, G.M. Ziegler, *Proofs from THE BOOK*, DOI 10.1007/978-3-662-44205-0_4,

- $x \equiv -x$ is impossible for odd p.
- $x \equiv \overline{x}$ is equivalent to $x^2 \equiv 1$. This has two solutions, namely $x = 1$ and $x = p - 1$, leading to the equivalence class $\{1, p-1\}$ of size 2.
- $x \equiv -\overline{x}$ is equivalent to $x^2 \equiv -1$. This equation may have no solution or two distinct solutions $x_0, p - x_0$: in this case the equivalence class is $\{x_0, p - x_0\}$.

For $p = 11$ the partition is $\{1, 10\}, \{2, 9, 6, 5\}, \{3, 8, 4, 7\}$; for $p = 13$ it is $\{1, 12\}, \{2, 11, 7, 6\}, \{3, 10, 9, 4\}$, $\{5, 8\}$: the pair $\{5, 8\}$ yields the two solutions of $s^2 \equiv -1 \bmod 13$.

The set $\{1, 2, \ldots, p-1\}$ has $p-1$ elements, and we have partitioned it into quadruples (equivalence classes of size 4), plus one or two pairs (equivalence classes of size 2). For $p - 1 = 4m + 2$ we find that there is only the one pair $\{1, p-1\}$, the rest is quadruples, and thus $s^2 \equiv -1 \pmod{p}$ has no solution. For $p - 1 = 4m$ there has to be the second pair, and this contains the two solutions of $s^2 \equiv -1$ that we were looking for. $\square$

Lemma 1 says that every odd prime dividing a number $M^2 + 1$ must be of the form $4m + 1$. This implies that there are infinitely many primes of this form: Otherwise, look at $(2 \cdot 3 \cdot 5 \cdots q_k)^2 + 1$, where q_k is the largest such prime. The same reasoning as above yields a contradiction.

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

Addition and multiplication in $\mathbb{Z}_5$

Prime fields

If p is a prime, then the set $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$ with addition and multiplication defined "modulo p" forms a finite field. We will need the following simple properties:

- For $x \in \mathbb{Z}_p$, $x \neq 0$, the additive inverse (for which we usually write $-x$) is given by $p - x \in \{1, 2, \ldots, p-1\}$. If $p > 2$, then x and $-x$ are different elements of $\mathbb{Z}_p$.
- Each $x \in \mathbb{Z}_p \backslash \{0\}$ has a unique multiplicative inverse $\overline{x} \in \mathbb{Z}_p \backslash \{0\}$, with $x\overline{x} \equiv 1 \pmod{p}$.
  The definition of primes implies that the map $\mathbb{Z}_p \to \mathbb{Z}_p$, $z \mapsto xz$ is injective for $x \neq 0$. Thus on the finite set $\mathbb{Z}_p \backslash \{0\}$ it must be surjective as well, and hence for each x there is a unique $\overline{x} \neq 0$ with $x\overline{x} \equiv 1 \pmod{p}$.
- The squares $0^2, 1^2, 2^2, \ldots, h^2$ define different elements of $\mathbb{Z}_p$, for $h = \lfloor \frac{p}{2} \rfloor$.
  This is since $x^2 \equiv y^2$, or $(x + y)(x - y) \equiv 0$, implies that $x \equiv y$ or that $x \equiv -y$. The $1 + \lfloor \frac{p}{2} \rfloor$ elements $0^2, 1^2, \ldots, h^2$ are called the *squares* in $\mathbb{Z}_p$.

At this point, let us note "on the fly" that for *all* primes there are solutions for $x^2 + y^2 \equiv -1 \pmod{p}$. In fact, there are $\lfloor \frac{p}{2} \rfloor + 1$ distinct squares x^2 in $\mathbb{Z}_p$, and there are $\lfloor \frac{p}{2} \rfloor + 1$ distinct numbers of the form $-(1 + y^2)$. These two sets of numbers are too large to be disjoint, since $\mathbb{Z}_p$ has only p elements, and thus there must exist x and y with $x^2 \equiv -(1 + y^2) \pmod{p}$.

Lemma 2. *No number $n = 4m + 3$ is a sum of two squares.*

■ **Proof.** The square of any even number is $(2k)^2 = 4k^2 \equiv 0 \pmod 4$, while squares of odd numbers yield $(2k+1)^2 = 4(k^2+k)+1 \equiv 1 \pmod 4$. Thus any sum of two squares is congruent to $0, 1$ or $2 \pmod 4$. □

This is enough evidence for us that the primes $p = 4m+3$ are "bad." Thus, we proceed with "good" properties for primes of the form $p = 4m+1$. On the way to the main theorem, the following is the key step.

Proposition. *Every prime of the form $p = 4m+1$ is a sum of two squares, that is, it can be written as $p = x^2+y^2$ for some natural numbers $x, y \in \mathbb{N}$.*

We shall present here two proofs of this result — both of them elegant and surprising. The first proof features a striking application of the "pigeon-hole principle" (which we have already used "on the fly" before Lemma 2; see Chapter 27 for more), as well as a clever move to arguments "modulo p" and back. The idea is due to the Norwegian number theorist Axel Thue.

■ **Proof.** Consider the pairs (x', y') of integers with $0 \le x', y' \le \sqrt{p}$, that is, $x', y' \in \{0, 1, \dots, \lfloor\sqrt{p}\rfloor\}$. There are $(\lfloor\sqrt{p}\rfloor + 1)^2$ such pairs. Using the estimate $\lfloor x\rfloor + 1 > x$ for $x = \sqrt{p}$, we see that we have more than p such pairs of integers. Thus for any $s \in \mathbb{Z}$, it is impossible that all the values $x' - sy'$ produced by the pairs (x', y') are distinct modulo p. That is, for every s there are two distinct pairs

$$(x', y'),\ (x'', y'') \ \in\ \{0, 1, \dots, \lfloor\sqrt{p}\rfloor\}^2$$

with $x' - sy' \equiv x'' - sy'' \pmod p$. Now we take differences: We have $x' - x'' \equiv s(y' - y'') \pmod p$. Thus if we define $x := |x' - x''|$, $y := |y' - y''|$, then we get

$$(x, y) \in \{0, 1, \dots, \lfloor\sqrt{p}\rfloor\}^2 \quad \text{with} \quad x \equiv \pm sy \pmod p.$$

Also we know that not both x and y can be zero, because the pairs (x', y') and (x'', y'') are distinct.

Now let s be a solution of $s^2 \equiv -1 \pmod p$, which exists by Lemma 1. Then $x^2 \equiv s^2y^2 \equiv -y^2 \pmod p$, and so we have produced

$$(x, y) \in \mathbb{Z}^2 \quad \text{with} \quad 0 < x^2 + y^2 < 2p \quad \text{and} \quad x^2 + y^2 \equiv 0 \pmod p.$$

But p is the only number between 0 and $2p$ that is divisible by p. Thus $x^2 + y^2 = p$: done! □

For $p = 13$, $\lfloor\sqrt{p}\rfloor = 3$ we consider $x', y' \in \{0, 1, 2, 3\}$. For $s = 5$, the sum $x' - sy' \pmod{13}$ assumes the following values:

x' \ y'	0	1	2	3
0	0	8	3	11
1	1	9	4	12
2	2	10	5	0
3	3	11	6	1

Our second proof for the proposition — also clearly a Book Proof — was discovered by Roger Heath-Brown in 1971 and appeared in 1984. (A condensed "one-sentence version" was given by Don Zagier.) It is so elementary that we don't even need to use Lemma 1.

Heath-Brown's argument features three linear involutions: a quite obvious one, a hidden one, and a trivial one that gives "the final blow." The second, unexpected, involution corresponds to some hidden structure on the set of integral solutions of the equation $4xy + z^2 = p$.

■ **Proof.** We study the set

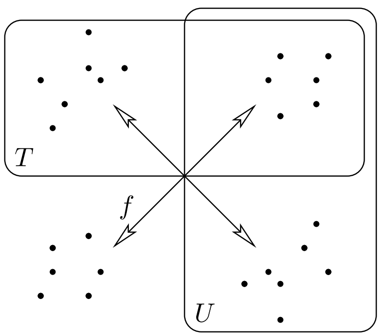
$$S \;:=\; \{(x,y,z) \in \mathbb{Z}^3 : 4xy + z^2 = p, \quad x > 0, \quad y > 0\}.$$

This set is finite. Indeed, $x \geq 1$ and $y \geq 1$ implies $y \leq \frac{p}{4}$ and $x \leq \frac{p}{4}$. So there are only finitely many possible values for x and y, and given x and y, there are at most two values for z.

1. The first linear involution is given by

$$f : S \longrightarrow S, \quad (x,y,z) \longmapsto (y,x,-z),$$

that is, "interchange x and y, and negate z." This clearly maps S to itself, and it is an *involution*: Applied twice, it yields the identity. Also, f has no fixed points, since $z = 0$ would imply $p = 4xy$, which is impossible. Furthermore, f maps the solutions in



$$T \;:=\; \{(x,y,z) \in S : z > 0\}$$

to the solutions in $S \backslash T$, which satisfy $z < 0$. Also, f reverses the signs of $x - y$ and of z, so it maps the solutions in

$$U \;:=\; \{(x,y,z) \in S : (x-y) + z > 0\}$$

to the solutions in $S \backslash U$. For this we have to see that there is no solution with $(x-y)+z = 0$, but there is none since this would give $p = 4xy+z^2 = 4xy + (x-y)^2 = (x+y)^2$.

What do we get from the study of f? The main observation is that since f maps the sets T and U to their complements, it also interchanges the elements in $T \backslash U$ with these in $U \backslash T$. That is, there is the same number of solutions in U that are not in T as there are solutions in T that are not in U — *so T and U have the same cardinality.*

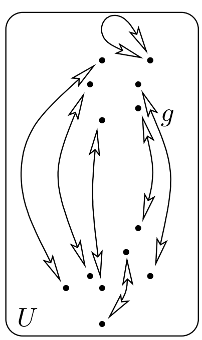
2. The second involution that we study is an involution on the set U:

$$g : U \longrightarrow U, \quad (x,y,z) \longmapsto (x-y+z, y, 2y-z).$$

First we check that indeed this is a well-defined map: If $(x,y,z) \in U$, then $x - y + z > 0, y > 0$ and $4(x-y+z)y + (2y-z)^2 = 4xy + z^2$, so $g(x,y,z) \in S$. By $(x-y+z) - y + (2y-z) = x > 0$ we find that indeed $g(x,y,z) \in U$.

Also g is an involution: $g(x,y,z) = (x-y+z, y, 2y-z)$ is mapped by g to $((x-y+z) - y + (2y-z), y, 2y - (2y-z)) = (x,y,z)$.

And finally g has exactly one fixed point:



$$(x,y,z) \;=\; g(x,y,z) \;=\; (x-y+z, y, 2y-z)$$

implies that $y = z$, but then $p = 4xy + y^2 = (4x+y)y$, which holds only for $y = z = 1$ and $x = \frac{p-1}{4}$.

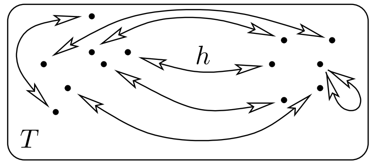
But if g is an involution on U that has exactly one fixed point, then *the cardinality of U is odd.*

3. The third, trivial, involution that we study is the involution on T that interchanges x and y:

$$h : T \longrightarrow T, \quad (x, y, z) \longmapsto (y, x, z).$$

This map is clearly well-defined, and an involution. We combine now our knowledge derived from the other two involutions: The cardinality of T is equal to the cardinality of U, which is odd. But if h is an involution on a finite set of odd cardinality, then it *has a fixed point*: There is a point $(x, y, z) \in T$ with $x = y$, that is, a solution of

$$p = 4x^2 + z^2 = (2x)^2 + z^2. \qquad \square$$



On a finite set of odd cardinality, every involution has at least one fixed point.

Note that this proof yields more — the number of representations of p in the form $p = x^2 + (2y)^2$ is *odd* for all primes of the form $p = 4m+1$. (The representation is actually unique, see [3].) Also note that both proofs are not effective: Try to find x and y for a ten digit prime! Efficient ways to find such representations as sums of two squares are discussed in [1] and [7].

The following theorem completely answers the question which started this chapter.

Theorem. *A natural number n can be represented as a sum of two squares if and only if every prime factor of the form $p = 4m + 3$ appears with an even exponent in the prime decomposition of n.*

■ **Proof.** Call a number n *representable* if it is a sum of two squares, that is, if $n = x^2 + y^2$ for some $x, y \in \mathbb{N}_0$. The theorem is a consequence of the following five facts.

(1) $1 = 1^2 + 0^2$ and $2 = 1^2 + 1^2$ are representable. Every prime of the form $p = 4m + 1$ is representable.

(2) The product of any two representable numbers $n_1 = x_1^2 + y_1^2$ and $n_2 = x_2^2 + y_2^2$ is representable: $n_1 n_2 = (x_1 x_2 + y_1 y_2)^2 + (x_1 y_2 - x_2 y_1)^2$.

(3) If n is representable, $n = x^2 + y^2$, then also nz^2 is representable, by $nz^2 = (xz)^2 + (yz)^2$.

Facts (1), (2) and (3) together yield the “if” part of the theorem.

(4) If $p = 4m + 3$ is a prime that divides a representable number $n = x^2 + y^2$, then p divides both x and y, and thus p^2 divides n. In fact, if we had $x \not\equiv 0 \pmod{p}$, then we could find $\overline{x}$ such that $x\overline{x} \equiv 1 \pmod{p}$, multiply the equation $x^2 + y^2 \equiv 0$ by $\overline{x}^2$, and thus obtain $1 + y^2\overline{x}^2 = 1 + (\overline{x}y)^2 \equiv 0 \pmod{p}$, which is impossible for $p = 4m + 3$ by Lemma 1.

(5) If n is representable, and $p = 4m + 3$ divides n, then p^2 divides n, and n/p^2 is representable. This follows from (4), and completes the proof. $\square$

Two remarks close our discussion:

- If a and b are two natural numbers that are relatively prime, then there are infinitely many primes of the form $am+b$ $(m \in \mathbb{N})$ — this is a famous (and difficult) theorem of Dirichlet. More precisely, one can show that the number of primes $p \leq x$ of the form $p = am + b$ is described very accurately for large x by the function $\frac{1}{\varphi(a)} \frac{x}{\log x}$, where $\varphi(a)$ denotes the number of b with $1 \leq b < a$ that are relatively prime to a. (This is a substantial refinement of the prime number theorem, which we had discussed on page 12.)
- This means that the primes for fixed a and varying b appear essentially at the same rate. Nevertheless, for example for $a = 4$ one can observe a rather subtle, but still noticeable and persistent tendency towards "more" primes of the form $4m+3$. The difference between the counts of primes of the form $4m+3$ and those of the form $4m+1$ changes sign infinitely often. Nevertheless, if you look for a large random x, then chances are that there are more primes $p \leq x$ of the form $p = 4m + 3$ than of the form $p = 4m + 1$. This effect is known as "Chebyshev's bias"; see Riesel [4] and Rubinstein and Sarnak [5].

References

[1] F. W. CLARKE, W. N. EVERITT, L. L. LITTLEJOHN & S. J. R. VORSTER: *H. J. S. Smith and the Fermat Two Squares Theorem,* Amer. Math. Monthly **106** (1999), 652-665.

[2] D. R. HEATH-BROWN: *Fermat's two squares theorem,* Invariant (1984), 2-5.

[3] I. NIVEN & H. S. ZUCKERMAN: *An Introduction to the Theory of Numbers,* Fifth edition, Wiley, New York 1972.

[4] H. RIESEL: *Prime Numbers and Computer Methods for Factorization,* Second edition, Progress in Mathematics **126**, Birkhäuser, Boston MA 1994.

[5] M. RUBINSTEIN & P. SARNAK: *Chebyshev's bias,* Experimental Mathematics **3** (1994), 173-197.

[6] A. THUE: *Et par antydninger til en taltheoretisk metode,* Kra. Vidensk. Selsk. Forh. **7** (1902), 57-75.

[7] S. WAGON: *Editor's corner: The Euclidean algorithm strikes again,* Amer. Math. Monthly **97** (1990), 125-129.

[8] D. ZAGIER: *A one-sentence proof that every prime* $p \equiv 1 \pmod 4$ *is a sum of two squares,* Amer. Math. Monthly **97** (1990), 144.