# Proving mathematical theorems

Mathematics is one of the only areas of knowledge that can objectively be described as "true", because its theorems are derived by using logic. They are not "given" by some authority or common opinion, but they are proven to be correct. Here we consider what proving actually means. We learn how to construct mathematical proofs.

Not all propositions are easy to check. That is because some propositions may involve a large or infinite number of possible cases. As an example, we consider the following claim involving prime numbers. A *prime* is an integer greater than 1 that is divisible only by itself and 1. For example, 2, 3, 5, 7, and 11 are primes. We will return to prime numbers later in this course.

Let us consider the following claim:

$$\textit{For every } n \in \mathbb{N}\textit{, the value of } n^2 + n + 41 \textit{ is prime.}$$

Let us denote

$$f(n) = \text{the value of } n^2 + n + 41 \text{ is prime}$$

If any of the values $f(n)$ for some $n \in \mathbb{N}$ is not prime, then the above proposition if false. We have a python program for this:

```
# Number n is the input
n = 40

# Compute the value f(n) and store it to variable f
f = n**2 + n + 41

# Print the value f(n)
print("f(", n, ") = ", f)

flag = False        # flag is True if we find a number dividing f
i = 2
while i < f:
    if (f % i) == 0:    # test whether i divides f
        flag = True     # we found one, so f cannot be prime
        print(i, "divides", f)  # print the number i
        i = f           # we will exit the loop

    i = i + 1           # we increase i by 1 for the next round

# Output result
if flag:
    print(f, "is not a prime number")
else:
    print(f, "is a prime number")
```

The values are:

| $n$ | $n^2+n+41$ | |
|---|---|---|
| 0 | 41 | (prime) |
| 1 | 43 | (prime) |
| 2 | 47 | (prime) |
| 3 | 53 | (prime) |
| 4 | 61 | (prime) |
| 5 | 71 | (prime) |
| ⋮ | ⋮ | ⋮ |
| 20 | 461 | (prime) |
| 21 | 503 | (prime) |
| ⋮ | ⋮ | ⋮ |
| 39 | 1601 | (prime) |

Interestingly, for $n \in \{0, \ldots, 39\}$, $f(n)$ is a prime! But

$$f(40) = 40^2 + 40 + 41 = 41 \cdot 41$$

is not prime. So it is not true that the expression is prime for all nonnegative integers, and thus the proposition is false!

Mathematical expressions

- "$2 = 1$" and "$0 < 1$"
- "$x^2 + 2x - 4 = 0$" and "$7x < 1$"

differ essentially in the sense that first ones are *closed* with respect to their truth (true or false), but the latter are *open* in the sense that their truth depends on what value the variable $x$ happens to have. Mathematicians have developed notation for dealing with such situations as well.

- **Predicate** $P(x)$: a predicate is the formalization of the mathematical concept of statement. A statement is understood as an assertion that may be true or false, depending on the values of the variables that occur in it. For instance, if $x$ satisfies $P$, then $P(x)$ is **true**. If $x$ does not satisfy $P$, then $P(x)$ is **false**.
- **Existential quantifier** ($\exists$): an existential quantifier is interpreted as "there exists", "there is at least one", or "for some". If $P(x)$ is a predicate, then

  $$(\exists x)\, P(x)$$

  means that *there exists* (at least one) $x$ such that $P(x)$ is true.

- **Universal quantification** ($\forall$): a universal quantification is interpreted as "given any" or "for all". If $P(x)$ is a predicate, then

$$(\forall x)\, P(x)$$

means that $P(x)$ is true for all values of $x$.

**Example 1.** Let $X$ denote the set of all animals of the world. The property that an animal $x$ is a dog can be expressed by a predicate

$$\mathrm{dog}(x).$$

This predicate has the value **true** for all animals $x$ that are dogs and **false** otherwise. Let us consider the following expressions

(a) $(\exists x \in X)\, \mathrm{dog}(x)$ (true; there is at least one dog)

(b) $(\forall x \in X)\, \mathrm{dog}(x)$ (false; all animals are not dogs)

(c) $(\exists x \in X)\, \neg\mathrm{dog}(x)$ (true; there is an animal which is not a dog)

(d) $(\forall x \in X)\, \neg\mathrm{dog}(x)$ (false)

In the following, the "predicate" is a statement involving to numbers:

(e) $(\forall x \in \mathbb{R})\ x^2 \geq 0$ (true)

(f) $(\exists x \in \mathbb{R})\ x^2 + 2x - 3 = 0$ (true: $x = 1$)

(g) $(\forall x, y \in \mathbb{R})\ x < y \Rightarrow x^2 < y^2$ (true)

(h) $(\forall n \in \mathbb{N})\, n^2 + n + 41$ is prime (false)

We have now introduced all logical concepts we need – as well as the notation, and we now begin to look at what the proof of a mathematical theorem means. Let us begin by noting that in some cases, we do not know if a proposition is true or false. For example, the following simple proposition known as Goldbach's Conjecture has been heavily studied since 1742 but we still do not know if it is true.

*Every even integer $n$ greater than 2 is the sum of two primes*

The claim seems to hold at least for small $n \in \mathbb{N}$:

$$\begin{aligned} 4 &= 2+2 \\ 6 &= 3+3 \\ 8 &= 3+5 \\ 10 &= 5+5 \\ 12 &= 5+7 \\ 14 &= 7+7 \\ 16 &= 3+13 \\ 18 &= 5+13 \\ 20 &= 7+13 \\ &\vdots \end{aligned}$$

To show that the claim is not true it would be sufficient to find at least one even number $n > 2$, which is not a sum of two primes. Of course, it has been checked by computer for many values of $n$, but still we do not know the final truth.

In mathematics and logic, a **direct proof** is a way of showing the truth (or falsehood) of a given statement by a straightforward combination of established facts, without making any further assumptions. We begin with two examples example of a direct proof. An integer is **even** if it is of the form $n = 2k$, where $k$ is an integer. The numbers $0, 2, 4, 6, \ldots$ are even. Also the negative numbers $-2, -4, -6, \ldots$ are even. We

**Example 2** (Direct proof)**.** We prove that:

The sum of two even integers is an even integer

Consider two even integers $x$ and $y$. Since they are even, they can be written as

$$x = 2a \quad \text{and} \quad y = 2b$$

for some integers $a$ and $b$. The sum of $x$ and $y$ is:

$$x + y = 2a + 2b = 2(a + b) = 2c,$$

where $c = a + b$. Therefore, $x + y$ has 2 as a factor and therefore is even, so the sum of any two even integers is even.

**Example 3** (Direct proof)**.** We prove that

$$(\forall n \in \mathbb{Z})\, n^3 - n \text{ is divisible by 3.}$$

The steps are:

1. Every integer can be written in form: $3k$, $3k+1$, or $3k+2$.

2. We can write $n^3 - n$ in the form

$$n^3 - n = n(n^2 - 1) = n(n+1)(n-1).$$

3. We consider different cases. If $n = 3k$ for some integer $k$, then

$$n^3 - n = n(n+1)(n-1) = 3k(3k+1)(3k-1).$$

4. If $n = 3k+1$, then

$$n^3 - n = n(n+1)(n-1) = (3k+1)(3k+2)3k.$$

5. If $n = 3k+2$, then

$$n^3 - n = n(n+1)(n-1) = (3k+2)(3k+3)(3k+1) = (3k+2)3(k+1)(3k+1).$$

6. In each case 3–5, $n^3 - n$ is divisible by 3.

Often a mathematical statement is of the form:

| If the assumptions $P_1, \ldots, P_n$ hold, then the conclusion $Q$ holds |
|---|

This can be written as an implication:

$$(P_1 \wedge \cdots \wedge P_n) \Rightarrow Q.$$

Let us denote the proposition $P_1 \wedge \cdots \wedge P_n$ simply by $P$. We need to prove the implication:

$$P \text{ (assumptions)} \Rightarrow Q \text{ (conclusion)}$$

This can be done in two different ways:

**I: Direct proof** We find a chain of implications such that

| $P \Rightarrow V_1, \quad V_1 \Rightarrow V_2, \quad \cdots \quad V_{n-1} \Rightarrow V_n, \quad V_n \Rightarrow Q.$ |
|---|

The proof consists of a sequence of if-then statements. Together these give the desired $P \Rightarrow Q$ implication.

A number $n$ is **odd**, if there is an integer $k$ such that $n = 2k+1$. The integers $1, 3, 5, \ldots$ and $-1, -3, -5, \ldots$ are odd.

**Example 4** (Implication)**.** We prove that

$$(\forall n \in \mathbb{Z})\ \underbrace{n \text{ is odd}}_{\text{assumption}} \Rightarrow \underbrace{n^2 \text{ is odd}}_{\text{conclusion}}$$

This is done in the following steps:

1. Let $n$ be an integer such that $n = 2k + 1$ for some integer $k$.
2. If two real numbers $x$ and $y$ are equal, then $x^2 = x \cdot x = x \cdot y = y \cdot y = y^2$. Thus,
$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1.$$
3. Now $2k^2 + 2k$ is an integer, because $k$ is an integer.
4. Since $n^2 = 2(2k^2 + 2k) + 1$, $n^2$ is odd.

The implication is valid.

**II: Proof by contrapositive** Instead of $P \Rightarrow Q$, we prove its **contrapositive:**

$$\neg Q \Rightarrow \neg P$$

As we have shown $P \Rightarrow Q$ and $\neg Q \Rightarrow \neg P$ are logically equivalent.

**Example 5** (Contrapositive)**.** We prove that

$$(\forall n \in \mathbb{Z})\, n^2 \text{ is even} \Rightarrow n \text{ is even}$$

Let us denote

$$R(n) = \text{“}n^2 \text{ is even”} \quad \text{and} \quad S(n) = \text{“}n \text{ is even”}$$

We need to prove that

$$(\forall n \in \mathbb{Z})\, R(n) \Rightarrow S(n).$$

Now the problem is that this implication is very difficult to prove directly. Namely, if $n^2$ is an even number, then $n^2 = 2k$ for some integer $k$. But how to show that $\sqrt{2k}$ is even? Nor can we verify the correctness of the claim by going through all the values of $n$, since the integers are of infinite amount.

We may prove the claim easily by showing

$$(\forall n \in \mathbb{Z})\, \neg S(n) \Rightarrow \neg R(n).$$

In fact, we have already proved this! Namely:

$$\neg S(n) = \text{“}n \text{ is not even”} = \text{“}n \text{ is odd”}$$

and

$$\neg R(n) = \text{“}n^2 \text{ is not even”} = \text{“}n^2 \text{ is odd”}$$

But this is already proved in Example 4.

Often a mathematical statement is of the form:

$$\boxed{P \text{ is true if and only if } Q \text{ is true}}$$

Note also that "if and only if" if sometimes written as "iff".

We know that $P \Leftrightarrow Q$ is logically equivalent with the proposition $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$. Therefore

$$\boxed{P \Leftrightarrow Q \text{ is proved by showing that } P \Rightarrow Q \text{ and } Q \Rightarrow P}$$

**Example 6.** We prove that

$$(\forall n \in \mathbb{Z})\, n^2 \text{ is odd} \Leftrightarrow n \text{ is odd}$$

We need to prove that

1) $n$ is odd $\Rightarrow n^2$ is odd.

2) $n^2$ is odd $\Rightarrow n$ is odd.

We have already proved that 1) holds in Example 4. Proving 2 as implication is difficult, so we prove its contrapositive:

$$\neg(n \text{ is odd}) \Rightarrow \neg(n^2 \text{ is odd})$$

This can be written in from

$$n \text{ is even} \Rightarrow n^2 \text{ is even}$$

But proving this is very simple:

1. Assume $n$ is even. This means that there exists and integer $k$ such that $n = 2k$.
2. From this is follows that
$$n^2 = (2k)^2 = 4k^2 = 2(2k^2).$$
3. Because $k$ is an integer, also $2k^2$ is an integer. Therefore, $n^2$ is even.

We have now proved both 1) and 2), which means that the equivalence holds.

Note that we have also proved that

$$(\forall n \in \mathbb{Z})\, n^2 \text{ is even} \Leftrightarrow n \text{ is even}$$

Let us now return to the following proposition. If $U$ is a set and $A, B \subseteq U$, then,

$$A \cap B = A \Leftrightarrow A \subseteq B \Leftrightarrow A \cup B = B.$$

We denote by (i)–(iii) the different cases:

(i) $A \cap B = A$,

(ii) $A \subseteq B$,

(iii) $A \cup B = B$.

We show that

$$\text{(i)} \Rightarrow \text{(ii)} \Rightarrow \text{(iii)} \Rightarrow \text{(i)}.$$

This shows that if any of claims (i)–(iii) is true, then they all are.

(i) $\Rightarrow$ (ii): Let us assume that $A \cap B = A$. If $x \in A$, then $x \in A = A \cap B \subseteq B$. Hence, $A \subseteq B$.

(ii) $\Rightarrow$ (iii): Suppose $A \subseteq B$. Clearly, $B \subseteq A \cup B$ holds always. If $x \in A \cup B$, then $x \in A$ or $x \in B$. But because $A \subseteq B$, necessarily $x \in B$ holds. This means that $A \cup B \subseteq B$, and we have shown $A \cup B = B$.

(iii) $\Rightarrow$ (i): Let us assume $A \cup B = B$. Clearly, $A \cap B \subseteq A$ holds always. If $x \in A$, then the assumption $A \cup B = B$ implies $x \in A \subseteq A \cup B = B$. Therefore, $x \in A \cap B$ and $A \subseteq A \cap B$. The proof is finished

In a rectangular triangle, the side opposite the right angle is called **hypotenuse** and two shorter are called **legs**. The **Pythagorean theorem** states that the area of the square whose side is the hypotenuse is equal to the sum of the areas of the squares on the other two legs. This theorem can be written as an equation:

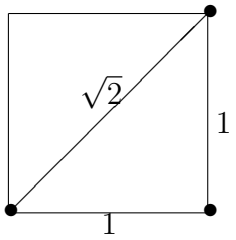
$$a^2 + b^2 = c^2,$$

where $c$ represents the length of the hypotenuse and $a$ and $b$ the lengths of the legs.

Interestingly, the Pythagorean theorem has the result that there are also irrational numbers, that is, numbers that cannot be presented in form

$$\frac{a}{b}$$

The following figure shows a square with side lengths are of length 1. According to Pythagorean theorem, now square the diagonal is a number $c$ such that $c^2 = 2$. Next we show that such a number $c$ cannot be rational. Since the number $c$ cannot be expressed as a rational number, it has an own notation $\sqrt{2}$ (square root).




Often in mathematics we use so-called **indirect proof**, which is also called **proof by contradiction** and **reductio ad absurdum**. The idea is that we assume $\neg P$ to be true and show that a contraction (false) follows. The correctness of $P$ can be proved in a following way:

(1) $P$ is assumed to be false, that is, $\neg P$ is true.

(2) We then show that $\neg P$ implies two mutually contradictory assertions, $Q$ and $\neg Q$.

(3) Since $Q$ and $\neg Q$ cannot both be true, the assumption that $P$ is false must be wrong. Therefore, $P$ must be true.

Since $Q \wedge \neg Q$ if logically equivalence to false **F**, the above method proves the implication

$$\boxed{\neg P \Rightarrow \mathbf{F}}$$

Proof by contradiction is valid by the following truth table showing that $P$ and $\neg P \Rightarrow \mathbf{F}$ are logically equivalent:

| $P$ | $\neg P$ | $\neg P \Rightarrow \mathbf{F}$ |
|---|---|---|
| **T** | **F** | **T** |
| **F** | **T** | **F** |

**Example 7** (Proof by contradiction)**.** We show that $\sqrt{2}$ is not a rational number.

1. Let us assume that $\sqrt{2}$ is rational.
2. From this assumption it follows that there are integers $a$ and $b$ such that $\sqrt{2} = \frac{a}{b}$ We additionally assume that this $\frac{a}{b}$ is simplified to lowest terms, since that can obviously be done with any fraction. Notice that in order for $\frac{a}{b}$ to be in simplest terms, both of $a$ and $b$ cannot be even. One or both must be odd. Otherwise, we could simplify $\frac{a}{b}$ further.
3. As we have noted, $x = y \Rightarrow x^2 = y^2$. This means that $2 = a^2/b^2$.
4. Let us multiply both sides by $b^2$. We get $a^2 = 2b^2$.
5. Because $b$ is an integer, so $b^2$ is an integer and $a^2$ is even.

6. This means that $a$ is even (see Example 5).
7. Therefore, there is an integer $d$ such that $a = 2d$.
8. We obtain $2b^2 = a^2 = (2d)^2 = 4d^2$ and $b^2 = 2d^2$.
9. Because $d$ is an integer, also $d^2$ is an integer. Therefore, $b^2$ is even.
10. We have that both $a$ and $b$ are even.
11. This contradicts with item 2.
12. $\sqrt{2}$ is not rational.

**Example 8** ("Nonproof")**.** Sometimes we can see "proofs" that first appear to be fully correct, but then we notice that there is something wrong in the 'chain' of steps. Here we "prove" that $2 = 1$.

1. Let us assume that there are two real numbers $a$ and $b$ such that $a = b$.
2. We multiply both sides by $a$, which gives the equality $a^2 = ab$.
3. We add to both sides the number $a^2 - 2ab$:
$$a^2 + a^2 - 2ab = ab + a^2 - 2ab.$$
4. The equation can be simplified to the form:
$$2(a^2 - ab) = a^2 - ab.$$
5. We divide the both sides by $a^2 - ab$:
$$2 = 1.$$

The "proof" looks completely harmless, however somewhere in the calculations for an error. This error occurs in step 5, where both sides of the equation are divided by $a^2 - ab$. We assumed first that $a = b$, so dividing by $a^2 - ab$ as the same as dividing by zero. But dividing by zero is an **undefined** operation.

We end this section by giving some general properties how a proof looks like.

A proof begins by leaving a blank line after the proposition to be proved. We start by writing the word *Proof.* The symbol □ is used to indicate the end of the proof. In some texts, the abbreviation Q.E.D. (*quod erat demonstrandum* – Latin: "Thus it has been demonstrated") can be seen to end the proof. In this section, we are used numbering to distinguish the different steps of the proof, but usually such numbering is not used. The following short example illustrates a mathematical proposition and its proof.

**Proposition 9.** *Every odd integer is the difference of two squares.*

*Proof.* Let us assume that $n$ is an odd integer. This means that there is an integer $k$ such that $n = 2k + 1$. Now

$$n = 2k + 1 = \underbrace{(k+1)^2}_{k^2+2k+1} - k^2.$$

This means that $n$ is a difference of $(k + 1)^2$ and $k^2$. □