9 Methods of proof

9.1 What is a proof?

In mathematics one aims to demonstrate that something is "absolutely" true. This way of proceeding is different from the way physics, chemistry or biology is progressing, where one needs to compare the prediction of theories with nature.

So what do we mean when we say that we mathematically prove something?

A proof is *"a convincing argument establishing the truth of a proposition"*. Here *argument* means a sequence of logical rules of inference such as we considered in Chapter 8, which build up a non trivial statement based on some axioms. *Convincing* is rather harder to define, but certainly should mean that the argument should convince other mathematicians!

Constructing proofs can be difficult, as there is no set procedure for doing so. "Any" methods that works will do, so there is plenty of scope for imagination. However, we can indicate some common strategies.

First though, we will illustrate the need for care. Consider the following 'proofs' that $1 = 2$.

Example 9.1.1: Suppose that x and y are non-zero with $x = y$. Then $x^2 = xy$ and so

$$x^2 - y^2 = xy - y^2.$$

Hence

$$(x + y)(x - y) = y(x - y)$$

and so $x + y = y$. Substituting for x we have

$$2y = y$$

and hence $2 = 1$.

Example 9.1.2: Consider

$$x = 1 + \sum_{i \geq 0} (-1)^i = 1 + (1 - 1 + 1 - 1 + \cdots).$$

Now

$$x = 1 + (1 - 1) + (1 - 1) + \cdots = 1.$$

But also

$$x = 1 + 1 + (-1 + 1) + (-1 + 1) + \cdots = 2.$$

Therefore $1 = 2$.

In the first example the error was to divide both sides of an equation by $x - y = 0$. In the second, the error was to assume that the sum $\sum_{i \geq 0} (-1)^i$ is well-defined. These examples illustrate that we must take care that every step in an argument is valid.

There is a difference between an error and a paradox.

An error is a mistake, while a paradox is a result which is surprising or upsets our 'common sense'. Often it demonstrates that our original assumptions were not precise enough. For example, Russell's paradox illustrated the need for a better definition of a set.

We will now consider various general methods of proof.

9.2 Direct proof

A direct proof is a chain of implications

$$p = p_1 \implies p_2 \implies p_3 \implies \cdots \implies p_n = q$$

where p is our hypothesis and q is our conclusion. Each step in the chain is a logical deduction from the preceding step.

Theorem 9.2.1: *If $n \in \mathbb{N}$ is odd then so is n^2.*

Proof: Let

$$\left[\begin{array}{ll} p_1 = & \text{“}n \text{ is odd”} \\ p_2 = & \text{“}n = 2k+1 \text{ for some } k \in \mathbb{N}\text{”} \\ p_3 = & \text{“}n^2 = 2l+1 \text{ for some } l \in \mathbb{N}\text{”} \\ p_4 = & \text{“}n^2 \text{ is odd”} \end{array}\right]$$

As n is odd $n = 2k+1$ $\quad [p_1 \implies p_2]$.

Now

$$\begin{aligned} n^2 = (2k+1)^2 &= 4k^2 + 4k + 1 \\ &= 2(2k^2 + 2k) + 1 \\ &= 2l + 1 \end{aligned}$$

where $l = 2k^2 + 2k \in \mathbb{N}$ $\quad [p_2 \implies p_3]$.

Hence n^2 is odd. $\quad [p_3 \implies p_4]$.

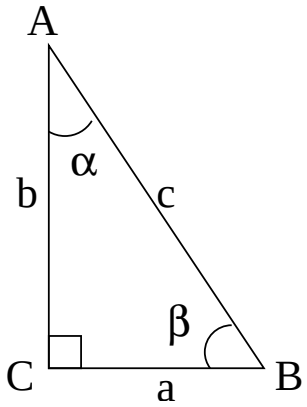
□

Usually we would not write down the steps in square brackets.

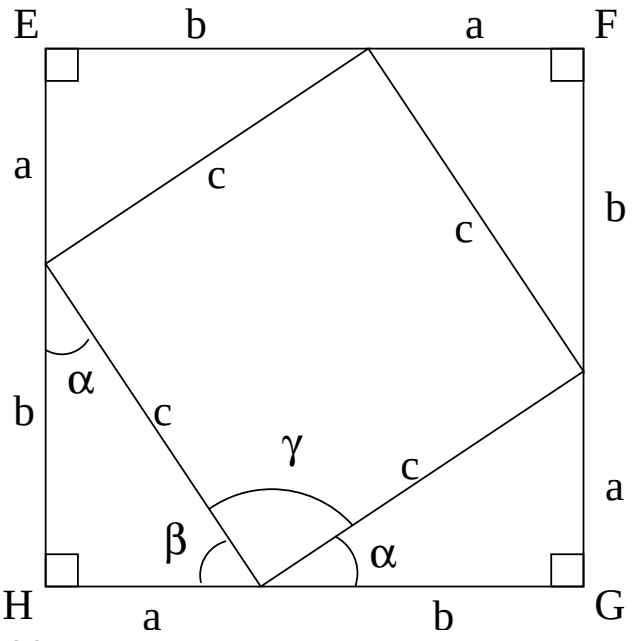
Sometimes we need to be a bit more imaginative.

Theorem 9.2.2: (Pythagoras)
For a right angle triangle we have

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



Proof: Consider the figure



Now

$$\gamma = \pi - (\alpha + \beta) = \pi/2$$

so the interior is a square. Area of *EFGH* is

$$\begin{aligned}&(a+b)^2\\&= 4(\text{Area of } ABC) + c^2\\&= 4(\tfrac{1}{2}ab) + c^2.\end{aligned}$$

So

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

and hence

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

□

Lecture 38

9.3 Indirect proof

Sometimes we cannot go directly from premise to conclusion. There are two basic strategies of indirect proof, by contrapositive and contradiction.

A contrapositive proof uses the fact that $p \longrightarrow q$ is logically equivalent to $(\neg q) \longrightarrow (\neg p)$.

Theorem 9.3.1: *If $n \in \mathbb{N}$ and n^2 is even then n is even.*

Proof: The contrapositive of this statement is *"if n is odd then n^2 is odd"*. We proved this in Theorem 9.2.1. □

Theorem 9.3.2: *If $n \in \mathbb{N}$ and $2^n - 1$ is prime then n is prime.*

Proof: The contrapositive is *"if n is not prime then $2^n - 1$ is not prime"*, so we will prove this.

Let $n = ab$ with $1 < a, b < n$. We have

$$\begin{aligned} 2^n - 1 &= 2^{ab} - 1 \\ &= (2^a)^b - 1 \\ &= (2^a - 1)(2^{a(b-1)} + 2^{a(b-2)} + \cdots + 2^a + 1). \end{aligned}$$

Now $2^a - 1 > 1$ as $a > 1$, and the other factor is also greater than 1 as $b > 1$, hence $2^n - 1 = xy$ for some $1 < x, y < 2^n - 1$. Thus $2^n - 1$ is not prime. □

Note that in both examples we have no idea how to give a direct proof. In the first case because we cannot take square roots of numbers like $n^2 = 2l$ and guarantee that n is even, and in the second because we do not know how to go from the primality of $2^n - 1$ to the primality of n.

Thus, although the contrapositive statement is logically equivalent to the original problem, it turns out to be much easier to work with in each case.

The second method of indirect proof is proof by contradiction. Here, to show that some proposition p is true, we suppose instead that it is false and derive a contradiction (i.e. a logical impossibility).

Theorem 9.3.3: *The square root of 2 is irrational.*

Proof: We will assume that $\sqrt{2}$ is rational, and derive a contradiction.

If $\sqrt{2} \in \mathbb{Q}$ then $\sqrt{2} = \frac{a}{b}$ for some $a, b \in \mathbb{N}$ with $b \neq 0$. We may assume that a and b have no common factors (or else we would divide them out). Then

$$2 = \frac{a^2}{b^2} \quad \text{implies that} \quad a^2 = 2b^2$$

and hence a^2 is even. By Theorem 9.3.1 we deduce that a is even, i.e. $a = 2k$ for some $k \in \mathbb{N}$. Now

$$2 = \frac{4k^2}{b^2} \quad \text{implies that} \quad 2b^2 = 4k^2$$

and hence $b^2 = 2k^2$ is even. But then Theorem 9.3.1 implies that b is even, and hence a and b have common factor 2. This contradicts our choice of a and b. □

Theorem 9.3.4: *Suppose there are $n \geq 2$ people at a party. Then at least two of them have the same number of friends at the party.*

Proof: Suppose no two people have the same number of friends. For each person there are $n-1$ other people, so they can have at most $n-1$ friends at the party. As there are n people each with a different number of friends there must exist for each i with $0 \leq i \leq n-1$ a person with precisely i friends.

This means that some person X has no friends, while another Y has $n-1$ friends. But then Y is friends with everyone at the party, including X. Hence X has a friend, which gives a contradiction. □

Note in this final example that there is no way we could have given a direct proof, as that would have required us to exhibit a particular pair of people with the same number of friends. But we have no idea who is at the party, so this would not be possible.

We have seen four examples of indirect proofs, and these are fairly typical of the general approach. Even when a direct proof is available, it is often easier to give an indirect proof instead.

Lecture 39

9.4 Existence proofs

Often we want to prove that their exists some mathematical object with certain properties. There are two ways to do this: constructive and non-constructive.

A constructive proof is one where we give an actual example that satisfies the conditions.

Consider the polynomial

$$f(n) = n^2 - n + 41.$$

We have $f(1) = 41$, $f(2) = 43$, $f(3) = 47$, $f(4) = 53 \ldots$ and all of these examples are prime. But this does not prove that all $f(n)$ are prime! Checking examples like this will never guarantee that something is always true. In fact we have

Theorem 9.4.1: *There exists $n \in \mathbb{N}$ with $f(n)$ not prime.*

Proof:

$$f(41) = 41^2 - 41 + 41 = 41 \times 41.$$

□

Not all constructive proofs are this easy.

Theorem 9.4.2: *For each $n \in \mathbb{N}$ there exists a sequence of n consecutive composite (i.e. non-prime) numbers.*

Proof: Suppose we are given $n \in \mathbb{N}$. Let $m = (n+1)! + 1$. Then

$$m + 1 = (n+1)! + 2 \text{ is divisible by } 2$$
$$m + 2 = (n+1)! + 3 \text{ is divisible by } 3$$

Indeed

$$m + i = (n+1)! + (i+1) \text{ is divisible by } i + 1$$

for $1 \leq i \leq n$.
So $m + 1, m + 2, \ldots, m + n$ is a sequence of n composite numbers.
□

In both these proofs, the hard part was to guess the right example. Once we had that, the rest was easy.

Unfortunately it is sometimes impossible to find examples in this way. Then we must give a non-constructive proof.

An example of this was the proof of Theorem 9.3.4, where we considered friends at a party. Clearly we could not give the names of two people with exactly the same number of friends without knowing who was at the party!

As another example we have

Theorem 9.4.3: *For all $n \in \mathbb{N}$ there exists a prime greater than n.*

Proof: Given n, let $m = n! + 1$.
Now m must have a prime factor (possibly itself). Let p be such a factor. If $p \leq n$ then p divides $n!$, and hence p does not divide $n! + 1$. Therefore $p > n$ as required. $\square$.

Note that we have not said what p is (and have not claimed that $n! + 1$ is prime). From this we deduce

Corollary 9.4.4: *There are infinitely many primes.*

Proof: Suppose the result is false, so that there exists a largest prime n. But by the Theorem there is a prime greater than n. $\square$.

Sometimes we want to show that there is a unique example. To do this we usually assume there are two such, and then show they must be equal.

As we saw before Theorem 9.4.1, we cannot use examples to prove that a result is always true, although we can use them (as in Theorem 9.4.2) for existence proofs. They can also be used as counterexamples to show that a conjectural result is actually false.

This limitation on the use of examples runs counter to most other disciplines. For example in physics all theories are based on sequences of experimental results, which are just a series of examples.

Of course this work very well — we are happy to assume that the sun will rise again tomorrow on the basis that it has on every other day so far. This method of argument is inductive rather than deductive.

We would like to have a method of induction that is valid in mathematics. This will have to be rather different from that used in the sciences. Next time we will introduce the notion of proof by induction.

Lecture 40

9.5 Induction

Proof by induction can be used when we have a family of propositions $P(n)$ for $n = 1, 2, \ldots$. The idea is to prove that $P(n)$ is true for all values of n by using the following principle:

> If there exists m such that both
> (i) $P(m)$ is true
> (ii) $P(k)$ true implies that $P(k+1)$ is true for all $k \geq m$
> then $P(n)$ is true for all $n \geq m$.

As an application of this principle we will prove

Theorem 9.5.1: *For all $n \geq 1$ we have*

$$\sum_{i=1}^{n} i = \frac{1}{2} n(n+1).$$

Proof: Let $P(n)$ be the statement "$\sum_{i=1}^{n} i = \frac{1}{2} n(n+1)$". Now $P(1)$ is

$$1 = \frac{1}{2} \times 1 \times 2$$

which is true. Now *assume* $P(k)$ is true for some $k \geq 1$. We must show that this implies that $P(k+1)$ is true. $P(k)$ states that

$$1 + 2 + \cdots + k = \frac{1}{2} k(k+1).$$

The left-hand side of $P(k+1)$ is $1 + 2 + \cdots + k + (k+1)$ which by our assumption equals

$$\frac{1}{2} k(k+1) + k + 1 = \frac{1}{2} (k+1)[k+2]$$

which equals the right-hand side of $P(k+1)$. So $P(k)$ true implies that $P(k+1)$ is true, and the result now follows by induction. □

Note that in proof by induction there are two steps. Although the first step is usually very easy it cannot be omitted.

As another example we shall prove

Theorem 9.5.2: *For all $n \geq 1$ we have that 21 divides $4^{n+1} + 5^{2n-1}$.*

Proof: Let $P(n)$ be "21 divides $4^{n+1} + 5^{2n-1}$". Then $P(1)$ says that 21 divides $4^2 + 5$, which is true.

Now assume that $P(k)$ is true for some $k \geq 1$. That is, assume that 21 divides $4^{k+1} + 5^{2k-1}$. We need to show that this implies that 21 divides $4^{k+2} + 5^{2k+1}$. We have

$$\begin{aligned} 4^{k+2} + 5^{2k+1} &= 4 \times 4^{k+1} + 25 \times 5^{2k-1} \\ &= 4(4^{k+1} + 5^{2k-1}) + 21 \times 5^{2k-1}. \end{aligned}$$

Clearly 21 divides $21 \times 5^{2k-1}$, and it divides $4(4^{k+1} + 5^{2k-1})$ by assumption. Hence 21 divides $4^{k+2} + 5^{2k+1}$ as required.

Thus $P(k)$ true implies that $P(k+1)$ true, and the result now follows by induction. □

To see an example where $m \neq 1$, consider

Theorem 9.5.3: *For all $n \geq 4$ we have that $n^2 \leq 2^n$.*

Note that this is false for $n = 3$.

Proof: Let $P(n)$ be "$n^2 \leq 2^n$". Then $P(4)$ is $16 \leq 16$, which is true.

Suppose that $P(k)$ is true for some $k \geq 4$. We want to show that this implies that $(k+1)^2 \leq 2^{k+1}$. We have

$$\begin{aligned}
(k+1)^2 &= k^2 + 2k + 1 \\
&\leq k^2 + 4k \qquad \text{as } k \geq 1 \\
&\leq k^2 + k^2 \qquad \text{as } k \geq 4 \\
&\leq 2(2^k) \qquad \text{as } P(k) \text{ assumed true} \\
&= 2^{k+1}.
\end{aligned}$$

So $P(k)$ true implies that $P(k+1)$ is true, and the result now follows by induction. □

9.6 Conclusions

We have seen various basic examples of some of the most common methods of proof. Discovering proofs is a very powerful technique, and is the basis of all modern mathematics. However, I will conclude with two words of warning.

1. Deductive reasoning depends entirely on the initial hypotheses. If these are incorrect, the conclusions will be useless.

This is particularly important to bear in mind when modelling real world problems. Much of the weakness of mathematical economics is due to the down-playing of this difficulty. More worryingly, we have

2. There are limits to the powers of deductive reasoning.

In the 1930s, Gödel proved (among other things) that there are results about the natural numbers which are true but which can never be proved by deductive reasoning!