Cambridge University Press
978-0-521-29915-2 - The Foundations of Analysis: A Straightforward Introduction
K. G. Binmore
Excerpt
More information

1 PROOFS

1.1 What is a proof?

Everyone knows that theorems require proofs. What is not so widely understood is the nature of the difference between a mathematical proof and the kind of argument considered adequate in everyday life. This difference, however, is an important one. There would be no point, for example, in trying to construct a mathematical theory using the sort of arguments employed by politicians when seeking votes.

The idea of a formal mathematical proof is explained in §1.5. But it is instructive to look first at some plausible types of argument which we shall *not* accept as proofs.

1.2 *Example* We are asked to decide whether or not the expression

$$n^3-4n^2+5n-1$$

is positive for $n=1, 2, 3, \ldots$. One approach would be to construct a table of the expression for as many values of n as patience allows.

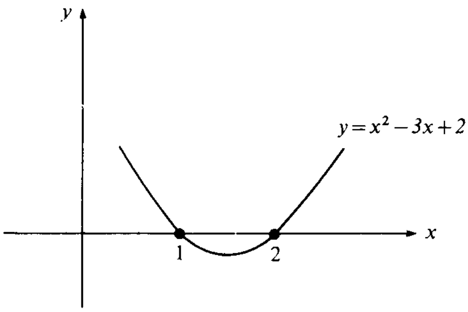
n	n^3-4n^2+5n-1
1	1
2	1
3	5
4	19
5	49
6	101

From the table it seems as though n^3-4n^2+5n-1 simply keeps on getting larger and larger. In particular, it seems reasonable to guess that n^3-4n^2+5n-1 is always positive when $n=1, 2, 3, \ldots$. But few people would maintain that the argument given here is a *proof* of this assertion.

1.3 *Example* In this example the situation is not quite so clear. We are asked to decide whether or not the expression

$$x^2-3x+2$$

is negative for all values of x satisfying $1<x<2$. Since $x^2-3x+2=(x-1)(x-2)$, it is easy to draw a graph of the parabola $y=x^2-3x+2$.



From the diagram it seems quite 'obvious' that $y=x^2-3x+2$ is negative when $1<x<2$ and positive or zero otherwise. But let us examine this question more closely.

How do we know that the graph we have drawn really does represent the behaviour of the equation $y=x^2-3x+2$? School children learn to draw graphs by plotting lots of points and then joining them up. But this amounts to guessing that the graph behaves as we think it should in the gaps between the plotted points. One might counter this criticism by observing that we know from our experience that the use of the graph always leads to correct answers. This would be a clinching argument in the field of physics. But, in mathematics, we are not supposed to accept arguments which are based on our experience of the world.

One might, of course, use a mathematical argument to deduce the properties of the graph, but then the graph would be unnecessary anyway. We are forced (reluctantly) to the conclusion that an appeal to the graph of $y=x^2-3x+2$ cannot be regarded as a *proof* that x^2-3x+2 is negative for $1<x<2$.

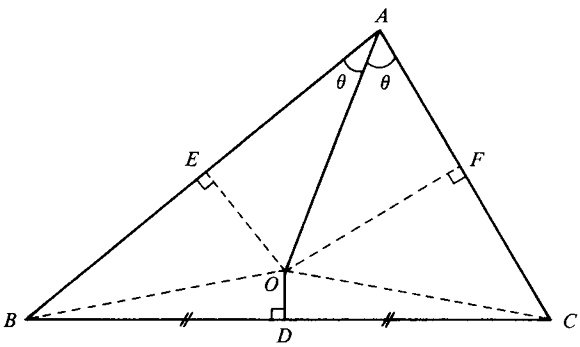
1.4 *Example* In the diagram below, the point O has been chosen as the point of intersection of the bisector of the angle A and the perpendicular bisector of the side BC. The dotted lines are then constructed as shown.

With the help of this diagram we shall show by the methods of elementary geometry that $AB=AC$ – i.e. *all triangles are isosceles*.

The triangles AEO and AFO are congruent and hence

$$AE=AF \quad (1)$$

$$OE=OF. \quad (2)$$



Also the triangles OBD and OCD are congruent. Thus

$$OB = OC. \tag{3}$$

From (2), (3) and Pythagoras' theorem it follows that

$$EB = FC. \tag{4}$$

Finally, from (1) and (4) we obtain that

$$AB = AC.$$

The explanation of this well-known fallacy is that the point O should lie *outside* the triangle ABC. In other words, the diagram does not represent the way things 'really are', and we have been led into error by depending on it. The objection that the diagram was not 'properly drawn' carries no weight since it is clearly not acceptable that a mathematical proof should depend on accurate measurement with ruler and compasses. This means, in particular, that the classical arguments of Euclidean geometry are not acceptable as proofs in modern mathematics because of their dependence on diagrams.

1.5 Mathematical proof

So far we have seen a number of arguments which are not proofs. What then is a proof?

Ideally, the description of a mathematical theory should begin with a list of *symbols*. This should be a finite list and contain all of the symbols which will be used in the theory. For even the simplest theory quite a few symbols will be needed. One will need symbols for variables – e.g. x, y and z. One will need symbols for the logical connectives (and, or, implies, etc.). The highly useful symbols) and (should not be forgotten and for a minimum of mathematical content one should perhaps include the symbols + and =.

Having listed the symbols of the theory (and those mentioned above are just some of the symbols which might appear in the list), it is then necessary

Cambridge University Press
978-0-521-29915-2 - The Foundations of Analysis: A Straightforward Introduction
K. G. Binmore
Excerpt
More information

to specify how these symbols may be put together to make up *formulae* and then how such formulae may be put together to make up *sentences*.

Next, it is necessary to specify which of these sentences are to be called *axioms*.

Finally, we must specify *rules of deduction* which will tell us under what circumstances a sentence may be *deduced* from other sentences.

A mathematical *proof* of a *theorem* S is then defined to be a list of sentences, the last of which is S. Each sentence in the list must be either an *axiom* or else a *deduction* from sentences appearing earlier in the list.

What is more, we demand that all of the processes described above be specified so clearly and unambiguously that even that arch-idiot of intellectuals, the computer, could be programmed to check that a given list of sentences is a proof.

Of course, the ideas set out above only represent an ideal. It is one thing to set a computer to checking a list of several million sentences and quite another to prepare such a list for oneself. Apart from any other consideration, it would be extremely boring.

1.6 *Example* The list of sentences given below shows what a formal proof looks like. It is a proof taken from S. C. Kleene's *Introduction to Metamathematics* (North-Holland, 1967) of the proposition '$a=a$'. This does not happen to be one of his axioms and therefore needs to be proved as a theorem.

(1) $a=b \Rightarrow (a=c \Rightarrow b=c)$
(2) $0=0 \Rightarrow (0=0 \Rightarrow 0=0)$
(3) $\{a=b \Rightarrow (a=c \Rightarrow b=c)\} \Rightarrow \{[0=0 \Rightarrow (0=0 \Rightarrow 0=0)] \Rightarrow [a=b \Rightarrow (a=c \Rightarrow b=c)]\}$
(4) $[0=0 \Rightarrow (0=0 \Rightarrow 0=0)] \Rightarrow [a=b \Rightarrow (a=c \Rightarrow b=c)]$
(5) $[0=0 \Rightarrow (0=0 \Rightarrow 0=0)] \Rightarrow \forall c[a=b \Rightarrow (a=c \Rightarrow b=c)]$
(6) $[0=0 \Rightarrow (0=0 \Rightarrow 0=0)] \Rightarrow \forall b \forall c[a=b \Rightarrow (a=c \Rightarrow b=c)]$
(7) $[0=0 \Rightarrow (0=0 \Rightarrow 0=0)] \Rightarrow \forall a \forall b \forall c[a=b \Rightarrow (a=c \Rightarrow b=c)]$
(8) $\forall a \forall b \forall c[a=b \Rightarrow (a=c \Rightarrow b=c)]$
(9) $\forall a \forall b \forall c[a=b \Rightarrow (a=c \Rightarrow b=c)] \Rightarrow \forall b \forall c[a+0=b \Rightarrow (a+0=c \Rightarrow b=c)]$
(10) $\forall b \forall c[a+0=b \Rightarrow (a+0=c \Rightarrow b=c)]$
(11) $\forall b \forall c[a+0=b \Rightarrow (a+0=c \Rightarrow b=c)] \Rightarrow \forall c[a+0=a \Rightarrow (a+0=c \Rightarrow a=c)]$
(12) $\forall c[a+0=a \Rightarrow (a+0=c \Rightarrow a=c)]$
(13) $\forall c[a+0=a \Rightarrow (a+0=c \Rightarrow a=c)] \Rightarrow [a+0=a \Rightarrow (a+0=a \Rightarrow a=a)]$
(14) $a+0=a \Rightarrow (a+0=a \Rightarrow a=a)$
(15) $a+0=a$
(16) $a+0=a \Rightarrow a=a$
(17) $a=a$.

The above example is given only to illustrate that the formal proofs of even the most trivial propositions are likely to be long and tedious. What is more, although a computer may find formal proofs entirely satisfactory, the human mind needs to have some explanation of the 'idea' behind the proof before it can readily assimilate the details of a formal argument.

What mathematicians do in practice therefore is to write out 'informal proofs' which can 'in principle' be reduced to lists of sentences suitable for computer ingestion. This may not be entirely satisfactory, but neither is the dreadfully boring alternative. In this book our approach will be even less satisfactory from the point of view of those seeking 'absolute certainty', since we shall not even describe in detail the manner in which mathematical assertions can be coded as formal lists of symbols. We shall, however, make a serious effort to remain true to the spirit of a mathematical proof, if not to the letter.

1.7 Obvious

The word 'obvious' is much abused. We shall follow the famous English mathematician G. H. Hardy in interpreting the sentence 'P is obvious' as meaning 'It is easy to think of a proof of P'. This usage accords with what was said in the section above.

A much more common usage is to interpret 'P is obvious' as meaning 'I *cannot* think of a proof of P but I am sure it must be true'. This usage should be avoided.

1.8 The interpretation of a mathematical theory

Observe that in our account of a formal mathematical theory the content has been entirely divorced from 'reality'. This is so that we can be sure, in so far as it is possible to be sure of anything, that the theorems are correct.

But mathematical theories are not made up at random. Often they arise as an attempt to abstract the essential features of a 'real world' situation. One sets up a system of axioms each of which corresponds to a well-established 'real world' fact. The theorems which arise may then be interpreted as predictions about what happens in the 'real world'.

But this viewpoint can be reversed. In many cases it turns out to be very useful when seeking a proof of a theorem to think about the real world situation of which the mathematical theory is an abstraction. This can often suggest an approach which might not otherwise come to mind. It is sometimes useful, for example, to examine theorems in complex analysis in terms of their electrostatic interpretation. In optimisation theory, insight can sometimes be obtained by viewing the theorems in terms of their game-theoretic or economic interpretation.

For our purposes, however, it is the interpretation in terms of geometry that we shall find most useful. One interprets the real numbers as points along an ideal ruler with which we measure distances in Euclidean geometry. This interpretation allows us to draw pictures illustrating propositions in analysis. These pictures then often suggest how the theorem in question may be proved. But it must be emphasised again that these pictures cannot serve as a *substitute* for a proof, since our theorems should be true regardless of whether our geometric interpretation is a good one or a bad one.

Cambridge University Press
978-0-521-29915-2 - The Foundations of Analysis: A Straightforward Introduction
K. G. Binmore
Excerpt
More information

2 LOGIC (I)

2.1 Statements

The purpose of logic is to label sentences either with the symbol T (for *true*) or with the symbol F (for *false*). A sentence which can be labelled in one of these two ways will be called a *statement*.

2.2 *Example* The following are both statements.

(i) Trafalgar Square is in London.
(ii) $2+2=5$.

The first is true and the second is false.

2.3 *Exercise*

Which of the following sentences are statements?

(i) More than 10 000 000 people live in New York City.
(ii) Is Paris bigger than Rome?
(iii) Go jump in a lake!
(iv) The moon is made of green cheese.

2.4 Equivalence

From the point of view of logic, the only thing which really matters about a statement is its *truth value* (i.e. T or F). Thus two statements P and Q are *logically equivalent* and we write

$$P \Leftrightarrow Q$$

if they have the same truth value. If P and Q are both statements, then so is $P \Leftrightarrow Q$ and its truth value may be determined with the aid of the following *truth table*.

P	Q	$P\Leftrightarrow Q$
T	T	T
T	F	F
F	T	F
F	F	T

In this table the right-hand column contains the truth value of '$P\Leftrightarrow Q$' for all possible combinations of the truth values of the statements P and Q.

2.5 *Example* Let P denote the statement 'Katmandu is larger than Timbuktu' and Q denote the statement 'Timbuktu is smaller than Katmandu'. Then P and Q are logically equivalent even though it would be quite difficult in practice to determine what the truth values of P and Q are.

2.6 Not

If P is a statement, the truth value of the statement (not P) may be determined from the following truth table.

P	not P
T	F
F	T

2.7 And, or

If P and Q are statements, the statements 'P and Q' and 'P or Q' are defined by the following truth tables.

P	Q	P and Q
T	T	T
T	F	F
F	T	F
F	F	F

P	Q	P or Q
T	T	T
T	F	T
F	T	T
F	F	F

The English language is somewhat ambiguous in its use of the word 'or'. Sometimes it is used in the sense of 'either/or' and sometimes in the sense of 'and/or'. In mathematics it is always used in the second of these two senses.

2.8 *Example* Let P be the statement 'The Louvre is in Paris' and Q the statement 'The Kremlin is in New York City'. Then 'P and Q' is false but 'P and (not Q)' is true. On the other hand, 'P or Q' and 'P or (not Q)' are both true.

2.9 *Exercise*

(1) If P is a statement, (not P) is its *contradictory*. Show by means of truth tables that 'P or (not P)' is a *tautology* (i.e. that it is true regardless of the truth or falsehood of P). Similarly, show that 'P and (not P)' is a *contradiction* (i.e. that it is false regardless of the truth or falsehood of P).

(2) The following pairs of statements are equivalent regardless of the truth or falsehood of the statements P, Q and R. Show this by means of truth tables in the case of the odd numbered pairs.

(i) P, not (not P)
(ii) P or (Q or R), (P or Q) or R
(iii) P and (Q and R), (P and Q) and R
(iv) P and (Q or R), (P and Q) or (P and R)
(v) P or (Q and R), (P or Q) and (P or R)
(vi) not (P and Q), (not P) or (not Q)
(vii) not (P or Q), (not P) and (not Q).

[*Hint*: For example, the column headings for the truth table in (iii) should read

P	Q	R	P and Q	(P and Q) and R	Q and R	P and (Q and R)

There should be *eight* rows in the table to account for all the possible truth value combinations of P, Q and R.]

(3) From 2(ii) above it follows that it does not matter how brackets are inserted in the expressions 'P or (Q or R)' and '(P or Q) or R' and so we might just as well write 'P or Q or R'. Equally we may write 'P and Q and R' instead of the statements of 2(iii).

Show by truth tables that the statements '(P and Q) or R' and 'P and (Q or R)' need not be equivalent.

(4) Deduce from 2(iv) that the statements 'P and (Q_1 or Q_2 or Q_3)', '(P and Q_1) or (P and Q_2) or (P and Q_3)' are equivalent. Write down similar results which arise from 2(v), 2(vi) and 2(vii). What happens with four or more Qs?

2.10 Implies

Suppose that P and Q are two statements. Then the statement 'P implies Q' (or '$P \Rightarrow Q$') is defined by the following truth table.

CAMBRIDGE

Cambridge University Press
978-0-521-29915-2 - The Foundations of Analysis: A Straightforward Introduction
K. G. Binmore
Excerpt
More information

P	Q	P implies Q
T	T	T
T	F	F
F	T	T
F	F	T

In simple terms the truth of 'P implies Q' means that, from the truth of P, we can deduce the truth of Q. In English this is usually expressed by saying

'*If P, then Q*'

or sometimes

'*P is a sufficient condition for Q*'.

It strikes some people as odd that 'P implies Q' should be defined as true in the case when P is false. This is so that a proposition like '$x>2$ implies $x>1$' may be asserted to be true for all values of x.

Consider next the following truth table.

P	Q	P implies Q	not Q	not P	(not Q) implies (not P)
T	T	T	F	F	T
T	F	F	T	F	F
F	T	T	F	T	T
F	F	T	T	T	T

Observe that the entries in the third and sixth columns are identical. This means that the statements 'P implies Q' and '(not Q) implies (not P)' are either *both* true or *both* false – i.e. they are logically equivalent. Wherever we see 'P implies Q' we might therefore just as well write '(not Q) implies (not P)' since this is an equivalent statement.

The statement '(not Q) implies (not P)' is called the *contrapositive* of 'P implies Q'. In ordinary English it is usually rendered in the form

'*P only if Q*'

or sometimes

'*Q is a necessary condition for P*'.

These last two expressions are therefore two more paraphrases for the simple statement 'P implies Q'.

2.11 *Exercise*

(1) Given that the statements 'P' and '$P \Rightarrow Q$' are both true, deduce that Q is true. [*Hint*: Delete from the truth table for $P \Rightarrow Q$ those rows which do