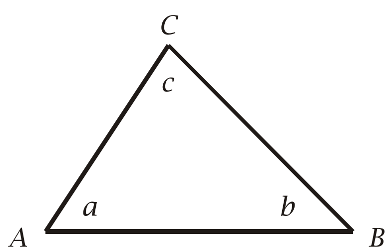
Axiom Systems

Proof

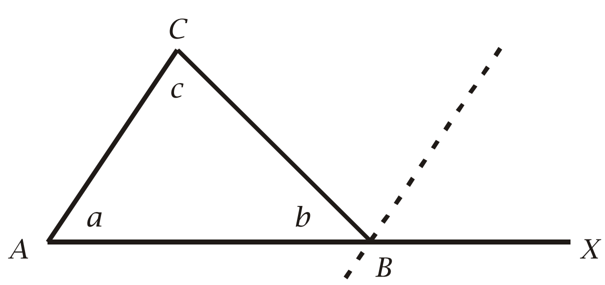
Mathematicians make use of results and theorems. These are statements of general truth that capture some part of our mathematical knowledge. These statements are called *theorems*. (They may also be called *lemmas, results* or *propositions*.) Theorems are established by proofs. For example, the proof that the angle sum of a triangle is 180°.

Theorem
The angle sum of a triangle is 180°.

Proof
Let *ABC* be any triangle, with angles *a, b* and *c* respectively.



Extend the line *AB* to the point *X* as shown in the following diagram and construct a line, *BY*, parallel to *AC* passing through *B*.



Then the angle $C\hat{B}Y$ is equal to the angle *c* since they are alternating angles; and the angle $Y\hat{B}X$ is equal to the angle *a* since they are corresponding angles. However

$$b + C\hat{B}Y + Y\hat{B}X = 180°$$

since these angles add up to the angle subtended by a straight line, which by definition is 180°. Hence $b + c + a = 180°$ and the angle sum of any triangle is 180°.

This is an example of a constructive proof. It illustrates the general principles of a constructive proof. Firstly, the constructive proof builds on definitions and axioms.

<u>Definitions</u>
The angle subtended by a straight-line is 180°.
A triangle is a plane, closed figure of three straight lines.

<u>Axioms</u>
Alternating angles are equal
Corresponding angles are equal

Secondly, the constructive proof begins with the statement, "Let…". This introduces an arbitrary object (or figure). The only property assumed of this figure is that it is a triangle, so whatever is true of it is true of *all* triangles. The constructive proof rests on the idea of taking an argument true of *any* object to be also true of *all* objects of that kind. This is how the generality is created.

The quest for a foundation of knowledge

In logic an argument starts with one or more statements that are assumed or asserted to be true. These statements are called premises. From these premises, a conclusion is drawn. Any argument is said to be *valid* if the premises *force* the conclusion to be true. That is, *if the premises are true, then the conclusion could not possibly be false*. When this is the case we say that the premises *entail* the conclusion, or that the conclusion is *deduced* from the premises. Here is an example of a valid deduction

(1) If Lee Harvey Oswald shot President Kennedy then the last bullet he fired wounded three people (one of whom was President Kennedy).
(2) The last bullet Lee Harvey Oswald fired did not wound three people.
(3) Therefore, Lee Harvey Oswald did not shoot President Kennedy.[1]

[1] This argument is taken from Oliver Stone's film, *JFK*. In that film the statement (2) is derived from another premise : (X) Under the circumstances in which Oswald fired, it is not possible for a single bullet to wound three people.

This argument is *valid*, in the sense that *if the premises are true, then the conclusion could not possibly be false.*[2] Here statement (3) is justified by means of an *inference* from statements (1) and (2). The first statement (1) has the form of a conditional: *if … then …*

If P then Q

where P = "Lee Harvey Oswald shot President Kennedy", and Q = "The last bullet Lee Harvey Oswald fired wounded three people". Here P is called the *antecedent* of the *conditional* and Q is called the *consequent*. In symbols this may be written, $P \to Q$. In the second statement (2), the consequent Q is denied; in other words, it is asserted that Q is not true, in symbols $\neg Q$. The inference is to (3), that P is also not true, in symbols $\neg P$. The whole argument can be written symbolically as

$$\begin{array}{l} P \to Q \\ \underline{\neg Q \quad\;} \\ \neg P \end{array}$$

Proofs are made of inferences, and together these are used in mathematics and logic to *justify* conclusions. The process of justifying mathematical statements cannot be separated from the philosophical question of justifying beliefs in general. Justification creates a chain of deductive inferences. Statements are justified by showing that they follow *logically* from other statements. But this creates a problem of what we call an infinite regress. Suppose statement Z is justified by inference from statement Y, and statement Y is justified by inference from statement X, and so forth. At what point will the process of justifying one statement by deriving it from another stop?

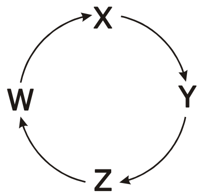
Z is justified by Y
Y is justified by X
X is justified by W
and so on, *ad infinitum*.

Diagrammatically we might represent this situation by

Z ← Y ← X ← W ← ……… ???

The arrows point in the direction of logical inference; but the justification goes the other way – Z is true because it follows logically from Y, and so on. One way we might get around this problem is to propose an alternative structure, Such as justification in a circle.

[2] The argument is controversial. However, the controversy arises from dispute regarding the premises. The argument itself is valid, in the sense that if the premises are tr
ue then the conclusion follows.



This may not seem any better. As we are going around in a circle every statement is justified by appeal to every other – but that does not justify the structure as a whole. Nonetheless, this alternative is proposed in what is known as the *coherence theory of truth*. Yet, if we reject the idea of justifying in a circle (and the coherence theory of truth), then we need to stop the infinite regress in some other way. We need to appeal to some kind of proposition that does not require justification by appeal to other propositions. Such a proposition would be a *self-evident truth*. A self-evident truth is one that justifies itself. The belief that there is a self-evident foundation for knowledge can be called *the axiomatic method*. An axiom in logic or mathematics is a starting proposition from which a theory is developed. In the theory of knowledge whatever is self-evidently true would constitute an axiom of knowledge. Self-evident truths are also called *postulates*.

Euclid's postulates

Historically, the axiomatic method began with the endeavour by Greek mathematicians to find a foundation for geometry. This ran concurrently with the effort by Greek philosophers to find a foundation to knowledge generally. The work in geometry was summarized by Euclid's famous work *The Elements*.

1. It is possible to draw a straight line between any two points.
2. It is possible to produce a finite straight line continuously in a straight line
3. It is possible to describe a circle with any centre and any radius.
4. It is true that all right angles are equal to one another.
5. It is true that, if a straight line falling on two straight lines make the interior angles on the same side less than two right angles, the two straight lines, if produced indefinitely, intersect on that side on which are the angles less than two right angles.

The fifth postulate is also called the parallel postulates. Euclid also advanced five "common notions" that are in effect also postulates.

1. Things which are equal to the same thing are also equal to one another.
2. If equals be added to equals, the whole are equal.
3. If equals be subtracted from equals, the remainders are equal.

4. Things which coincide with one another are equal to one another.
5. The whole is greater than the part.

With the exception of the firth, these postulates were advanced by Euclid as being self-evidently true, and therefore, a foundation on which to construct the whole of geometry as a true body of knowledge. The self-evident nature of the postulates is made by appeal to *geometric intuition*. We directly "see" or "intuit" that geometric objects obey these rules. However, the validity of the fifth postulate (parallel postulate) as a self-evident truth is questionable. In fact, the Greeks also questioned whether it was self-evident and instituted a search for a proof of it from the other postulates, which they did regard as self-evident. That search lasted over 2,000 years, and ended when in the C19th it was shown independently by Lobachevsky (1829), Bolyai (1831) and Gauss (not published) that the fifth postulate cannot be deduced from the other four postulates. In other words, it is *independent* of them. The mathematicians had developed alternative geometries in which the parallel postulate is not valid, and it is now an open question which of these geometries is true of the universe as a whole.

Formal systems

We have seen that the axiomatic method was originally motivated by a quest for self-evident truths, which were called postulates or axioms. A system of axioms was held to embody truths about the real world. The real world was a *model* of the axioms. For example, real space was thought by the Greeks to be a model for Euclid's postulates. The modern approach is to detach the notion of absolute truth from axioms. Axioms become *assumptions*; any structure that conforms to a set of axioms is a *model* for them. From the modern point of view an axiom is a statement that is stipulated to be true for the purpose of constructing a theory in which *theorems* may be derived by its rules of inference. It is a *primitive* statement of a *deductive formal system* and not distinguished from an assumption. Both the axioms and the rules of inference may be given at the outset of the development of the theory as a *formal system*. This means, that the system is given without specific interpretation or model, and the theorems are derived *formally* as if one were playing a game. There is also a philosophy of mathematics known as *formalism*. This asserts that there is no independent source of truth in mathematics and that all mathematics comprises only formal systems. Formalism denies that any axioms are self-evident. Evidently, to be interested in a formal system and to study it does not thereby commit one to a belief in formalism. As the purpose of this chapter is to introduce axiom systems, the philosophical issues will be left to another place. To study mathematics formally, axioms need not be justified, they are just assumed. Further, an axiom system may admit of many models. A set of axioms is *inconsistent* if it is possible to deduce a contradiction from it. An axiom system is *consistent* if it is not inconsistent. One way to demonstrate the consistency of a set of axioms is to show that it has a model.

Peano Postulates for the Natural Numbers

The natural numbers

The natural numbers are

0 1 2 3 4 5 ...

(Here we allow 0 to be a natural number.) The set of natural numbers is the set

$\mathbb{N} = \{0, 1, 2, 3, ...\}$

We seek a set of axioms that embody our intuitions about this chain of numbers. It is a sequence that starts with zero (0) and every member of this chain is followed by another by adding 1. Since zero (0) starts the sequence it cannot follow any number in it. Equality of numbers is captured by the principle that, if $x+1 = y+1$ then $x = y$.

Mathematical induction

Another feature of the natural numbers, is that they permit proof by *mathematical induction*. Proof by mathematical induction is a two-step argument.

> <u>Induction Step</u>
> If the result is true for the *k*th number then the result is true for the *(k+1)*th number.
> <u>Particular Result</u>
> The result is true for $n = 1$ (or for some other starting value).

From which the inference can be drawn:

> <u>Conclusion</u>
> The result is true for all *n* (or for all *n* greater than the starting value).

To illustrate the use of mathematical induction

> <u>Example</u>
> To prove $5^{2n} + 9^n - 2$ is divisible by 8 for all positive integer values *n*.
> Proof by mathematical induction
> Note, we use the symbol $8|N$ to mean 8 divides into the number *N*, or *N* is divisible by 8.
>
> For $n = 1$
> $5^{2n} + 9^n - 2 = 5^2 + 9^1 - 2 = 25 + 9 - 2 = 32 = 8 \times 4$
> Hence

$8 \mid 5^{2n} + 9^n - 2 \qquad \text{for } n = 1$

The induction hypothesis is

$8 \mid 5^{2k} + 9^k - 2$

Now for $n = k + 1$

$$\begin{aligned} 5^{2(k+1)} + 9^{k+1} - 2 &= 5^{2k} \cdot 5^2 + 9^k \cdot 9 - 2 \\ &= 5^{2k}(9 + 16) + 9^k \cdot 9 - 18 + 16 \\ &= 9(5^{2k} + 9^k - 2) + 16 \end{aligned}$$

By the induction hypothesis $8 \mid 5^{2k} + 9^k - 2$, and also $8 \mid 16$, hence

$8 \mid 9(5^{2k} + 9^k - 2) + 16$

and the induction step holds. Therefore, by mathematical induction the result is true for all *n*.

A set of axioms that expresses these primary intuitions about the natural numbers is the *Peano Postulates*.

Peano Axioms

P1 0 is a natural number

P2 If x is a natural number, there is another natural number denoted by x'. It is called the *successor* of x.

P3 $0 \neq x'$ for any natural number x.

P4 If $x' = y'$ then $x = y$

P5 Principle of Induction.
If Q is a property which may or may not hold of natural numbers, and if
(1) 0 has the property, and
(2) whenever a natural number x has the property Q, then x' has the property Q,
then all natural numbers have the property Q.

Relations

Relations

A relation is a pairing of elements of a set according to a rule. We may denote a relation by R or some other symbol.

Examples

1. Let $x, y \in \mathbb{N}$ then xRy iff $x + y = 5$.

This generates the set

$R = \{(0,5),(1,4),(2,3),(3,2),(4,1),(5,0)\}$

2. Let $x, y \in \mathbb{R}$, and xRy iff $xy = 1$. This generates an infinite set of pairs of numbers. Some elements of this set are

$(1,1), \left(2, \frac{1}{2}\right), \left(\sqrt{2}, \frac{1}{\sqrt{2}}\right)$

It is the set of elements $x \in \mathbb{R}$ and there associated reciprocals $\frac{1}{x}$.

Remark

We may denote a relation by xRy or by Rxy. In some cases we adopt a special symbol, for instance, $x \sim y$. Equality of numbers, $x = y$, is a relation.

Equivalence relations

Equality is an example of an equivalence relation. When $x = y$ the two symbols, x and y denote the same number. But objects may be equivalent without being identical. For example, in geometry, going up 1 step and along 1 step is equivalent to going along 1 and up 1, but the two paths are not identical; they are equivalent in the sense that starting from one given point reach the same given point. They are equivalent translations. Equivalence relations are often denoted by the symbol $x \sim y$.

Axioms for equivalence relations

A relation $x \sim y$ is an *equivalence relation* on a set X if it satisfies the following three axioms.

E1 *Reflexiveness axiom*

For all $x \in X$, $x \sim y$

E2 *Symmetry*

For all $x, y \in X$, if $x \sim y$ then $y \sim x$

E3 *Transitivity axiom*

For all $x, y, z \in X$, if $x \sim y$ and $y \sim z$ then $x \sim z$

A relation is an equivalence relation if it is reflexive, symmetric and transitive.

Examples

1 $x, y \in \mathbb{N}$, $x \sim y$ if $x - y$ is divisible by 6.

2. $x, y \in$ set of lines in the Euclidean plane; $x \sim y$ if x is parallel to y.

3. $x = y$ is an equivalence relation.

Order relations

We have seen already that

$\mathbb{N} = \{0, 1, 2, 3, ...\}$

is an ordered set, meaning, every element, after the initial element 0, has an immediate successor. An ordered set is any set on which an *order relation* can be defined. An order relation is a relation $\leq$ on a set X that satisfies the following four axioms

O1 For all $x \in X$, $x \leq x$

O2 For all $x, y, z \in X$, if $x \leq y$ and $y \leq z$ then $x \leq z$

O3 For all $x, y \in X$, if $x \leq y$ and $y \leq x$ then $x = y$

O4 For all $x, y \in X$, either $x \leq y$ or $y \leq x$

The axioms O1 and O2 are the same as the axioms E1 and E3 for an equivalence relation, but an order relation is something quite different from an equivalence relation; symmetry does not hold and O4 expresses the notion of order on the set. An ordered set is also called a *chain*.

Examples

1. The set $\mathbb{N} = \{0, 1, 2, 3, ...\}$ is an ordered set, where $\leq$ is the usual relation of 'greater than or equal' numbers.
2. The set of all words in a dictionary is and ordered set. The order relation is defined by the alphabet.
3. The subset relation, $\subseteq$ between sets is *not* an order relation on any set. Given a set X and its power set (set of all subsets), $P(X)$, then for any two sets $A, B \in P(X)$ we do not have $A \subseteq B$ or $B \subseteq A$. For a specific instance of this, let $X = \{1,2\}$ then $P(X) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$. The null set $\varnothing$ is a subset of every element of $P(X)$, but $\{1\}$ and $\{2\}$ are not ordered. The relation $\subseteq$ on the set $P(X)$ satisfies axioms O1 – O3 above, but not O4. It is called a *partial order*. A set that is partially ordered is called a *poset*.