CHAPTER 1

MATHEMATICS = SET THEORY?

"No one shall drive us out of the paradise that Cantor has created"

David Hilbert

1.1. Set theory

The basic concept upon which the discipline known as *set theory* rests is the notion of *set membership.* A set may be initially thought of simply as a collection of objects, these objects being called *elements* of that collection. *Membership* is the relation that an object bears to a set by dint of its being an element of that set. This relation is symbolised by the Greek letter $\in$ (epsilon). "$x \in A$" means that A is a collection of objects, one of which is x, i.e. x is a *member* (element) of A. When x is not an element of A, this is written $x \notin A$. If $x \in A$, we may also say that x *belongs* to A.

From these fundamental ideas we may build up a catalogue of definitions and constructions that allow us to specify particular sets, and construct new sets from given ones. There are two techniques used here.

(a) *Tabular form:* this consists in specifying a set by explicitly stating all of its elements. A list of these elements is given, enclosed in brackets. Thus

$$\{0, 1, 2, 3\}$$

denotes the collection whose members are all the whole numbers up to 3.

(b) *Set Builder form:* this is a very much more powerful device that specifies a set by stating a property that is possessed by all the elements of the set, and by no other objects. Thus the property of "being a whole number smaller than four" determines the set that was given above in tabular form. The use of properties to define sets is enshrined in the

PRINCIPLE OF COMPREHENSION. *If $\varphi(x)$ is a property or condition pertaining to objects x, then there exists a set whose elements are precisely the objects that have the property (or satisfy the condition) $\varphi(x)$.*

The set corresponding to the property $\varphi(x)$ is denoted

$$\{x: \varphi(x)\}$$

This expression is to be read "the set of all those objects x such that φ is true of x".

EXAMPLE 1. If $\varphi(x)$ is the condition "$x \in A$ and $x \in B$" we obtain the set

$$\{x: x \in A \text{ and } x \in B\}$$

of all objects that belong to both A and B, i.e. the set of objects that A and B have in common. This is known as the *intersection* of the sets A and B, and is denoted briefly by $A \cap B$.

EXAMPLE 2. The condition "$x \in A$ or $x \in B$" yields, by the Comprehension Principle the set

$$\{x: x \in A \text{ or } x \in B\}$$

consisting of all of the elements of A together with all of those of B, and none others. It is called the *union* of A and B, written $A \cup B$.

EXAMPLE 3. The condition "$x \notin A$" determines $-A$, the *complement* of A. Thus

$$-A = \{x: x \notin A\}$$

is the set whose members are precisely those objects that do not belong to A.

These examples all yield new sets from given ones. We may also directly define sets by using conditions that do not refer to any particular sets. Thus from "$x \neq x$" we obtain the set

$$\emptyset = \{x: x \neq x\}$$

of all those objects x such that x is not equal to x. Since no object is distinct from itself, there is nothing that can satisfy the property $x \neq x$, i.e. $\emptyset$ has no members. For this reason $\emptyset$ is known as the *empty* set. Notice that we have already "widened our ontology" from the original conception of a set as something with members to admit as a set something that has no members at all. The notion of an empty collection is often difficult to accept at first. One tends to think initially of sets as objects built up in a rather concrete way out of their constituents (elements). The introduction of $\emptyset$ forces us to contemplate sets as abstract "things-in-themselves". One could think of references to $\emptyset$ as an alternative form of words, e.g. that "$A \cap B = \emptyset$" is a short-hand way of saying "A and B have no elements in common". Familiarity and experience eventually show that the admission of $\emptyset$ as an actual object enhances and simplifies the theory. The justification for calling $\emptyset$ *the* empty set is that there can be only one set with no members. This follows from the definition of equality of sets as embodied in the

PRINCIPLE OF EXTENSIONALITY: *Two sets are equal iff they have the same elements.*

It follows from this principle that if two sets are to be distinct then there must be an object that is a member of one but not the other. Since empty collections have no elements they cannot be so distinguished and so the Extensionality Principle implies that there is only one empty set.

Subsets

The definition of equality of sets can alternatively be conveyed through the notion of *subsets*. A set A is a *subset* of set B, written $A \subseteq B$, if each member of A is also a member of B.

EXAMPLE 1. The set $\{0, 1, 2\}$ is a subset of $\{0, 1, 2, 3\}$, $\{0, 1, 2\} \subseteq \{0, 1, 2, 3\}$.

EXAMPLE 2. For any set A, we have $A \subseteq A$, since each member of A is a member of A.

EXAMPLE 3. For any set A, $\emptyset \subseteq A$, for if $\emptyset$ was not a subset of A, there would be an element of $\emptyset$ that did not belong to A. However $\emptyset$ has no elements at all.

Using this latest concept we can see that, for any sets A and B,

$$A = B \quad \text{iff} \quad A \subseteq B \quad \text{and} \quad B \subseteq A.$$

If $A \subseteq B$ but $A \neq B$, we may write $A \subset B$ (A is a *proper* subset of B).

Russell's Paradox

In stating and using the Comprehension Principle we gave no precise explanation of what a "condition pertaining to objects x" is, nor indeed what sort of entities the letter x is referring to. Do we intend the elements of our sets to be physical objects, like tables, people, or the Eiffel Tower, or are they to be abstract things, like numbers, or other sets themselves? What about the collection

$$V = \{x : x = x\}?$$

All things, being equal to themselves, satisfy the defining condition for this set. Is V then to include everything in the world (itself as well) or should it be restricted to a particular kind of object, a particular universe of discourse?

To demonstrate the significance of these questions we consider the condition "$x \notin x$". It is easy to think of sets that do not belong to themselves. For example the set $\{0, 1\}$ is distinct from its two elements 0 and 1. It is not so easy to think of a collection that includes itself amongst its members. One might contemplate something like "the set of all sets". A somewhat intriguing example derives from the condition

"x is a set derived from the Comprehension Principle by a defining condition expressed in less than 22 words of English".

The sentence in quotation marks has less than 22 words, and so defines a set that satisfies its own defining condition.

Using the Comprehension Principle we form the so-called *Russell set*

$$R = \{x : x \notin x\}.$$

The crunch comes when we ask "Does R itself satisfy the condition $x \notin x$?" Now if $R \notin R$, it does satisfy the condition, so it belongs to the set defined by that condition, which is R, hence $R \in R$. Thus the assumption $R \notin R$ leads to the contradictory conclusion $R \in R$. We must therefore reject this assumption, and accept the alternative $R \in R$. But if $R \in R$, i.e. R is an element of R, it must satisfy the defining condition for R, which is $x \notin x$. Thus $R \notin R$. This time the assumption $R \in R$ has lead to contradiction, so it is rejected in favor of $R \notin R$. So now we have proven both $R \in R$ and $R \notin R$, i.e. R both is, and is not, an element of itself. This is hardly an acceptable situation.

The above argument, known as Russell's Paradox, was discovered by Bertrand Russell in 1901. Set theory itself began a few decades earlier with the work of George Cantor. Cantor's concern was initially with the analysis of the real number system, and his theory, while rapidly becoming of intrinsic interest, was largely intended to give insight into properties of infinite sets of real numbers (e.g. that the set of irrational numbers has "more" elements than the set of rational numbers). During this same period the logician Gottlob Frege made the first attempt to found a definition of "number" and a development of the laws of arithmetic on formal logic and set theory. Frege's system included the Comprehension Principle in a form much as we have given it, and so was shown to be inconsistent (contradictory) by Russell's paradox. The appearance of the later, along with other set-theoretical paradoxes, constituted a crisis in the development of a theoretical basis for mathematical knowledge. Mathematicians were faced with the problem of revising their intuitive ideas about sets and reformulating them in such a way as to avoid inconsistencies. This challenge provided one of the major sources for the burgeoning growth in this century of mathematical logic, a subject which, amongst other things, undertakes a detailed analysis of the axiomatic method itself.

NBG

Set theory now has a rigorous axiomatic formulation – in fact several of them, each offering a particular resolution of the paradoxes.

John von Neuman proposed a solution in the mid-1920's that was later refined and developed by Paul Bernays and Kurt Gödel. The outcome is a group of axioms known as the system NBG. Its central feature is a very simple and yet powerful conceptual distinction between *sets* and *classes*. All entities referred to in NBG are to be thought of as classes, which correspond to our intuitive notion of collections of objects. The word "set" is reserved for those classes that are themselves members of other classes. The statement "x is a set" is then short-hand for "there is a class y such that $x \in y$". Classes that are not sets are called *proper classes*. Intuitively we think of them as "very large" collections. The Comprehension Principle is modified by requiring the objects x referred to there to be sets. Thus from a condition $\varphi(x)$ we can form the *class* of all *sets* (elements of other classes) that satisfy $\varphi(x)$. This is denoted

$$\{x : x \text{ is a set and } \varphi(x)\}.$$

The definition of the Russell class must now be modified to read

$$R = \{x : x \text{ is a set and } x \notin x\}.$$

Looking back at the form of the paradox we see that we now have a way out. In order to derive $R \in R$ we would need the extra assumption that R is a set. If this were true the contradiction would obtain as before, and so we reject it as false. Thus the paradox disappears and the argument becomes nothing more than a proof that R is a *proper class* i.e. a large collection that is not an element of any other collection. In particular $R \notin R$.

Another example of a proper class is V, which we now take to be the class

$$\{x : x \text{ is a set and } x = x\}$$

whose elements are all the sets. In fact NBG has further axioms that imply that $V = R$, i.e. no set is a member of itself.

ZF

A somewhat different and historically prior approach to the paradoxes was proposed by Ernst Zermelo in 1908. This system was later extended by Abraham Fraenkel and is now known as ZF. It can be informally regarded as a theory of "set-building". There is only one kind of entity, the set. All sets are built up from certain simple ones (in fact one can start just with $\emptyset$) by operations like intersection $\cap$, union $\cup$, and complementation $-$. The axioms of ZF legislate as to when such operations can be effected. They can only be applied to sets that have already been constructed, and the result is always a set. Thus proper classes like R are never actually constructed within ZF.

The Comprehension Principle can now only be used relative to a given set, i.e. we cannot collect together all objects satisfying a certain condition, but only those we already know to be members of some previously defined set. In ZF this is known as the

SEPARATION PRINCIPLE. *Given a set A and a condition $\varphi(x)$ there exists a set whose elements are precisely those members of A that satisfy $\varphi(x)$.*

This set is denoted

$$\{x : x \in A \text{ and } \varphi(x)\}.$$

Again we can no longer form the Russell class per se, but only for each set A the set

$$R(A) = \{x : x \in A \text{ and } x \notin x\}.$$

To obtain a contradiction involving the statements $R(A) \in R(A)$ and $R(A) \notin R(A)$ we would need to know that $R(A) \in A$. Our conclusion then is simply that $R(A) \notin A$. In fact in ZF as in NBG no set is an element of itself, so $R(A) = A$. (Note the similarity of this argument to the resolution in NBG – replacing V everywhere by A makes the latter formally identical to the former.)

NBG and ZF then offer some answers to the questions posed earlier. In practical uses of set theory, members of collections may well be physicial objects. In axiomatic presentations of set theory however all objects have a conceptual rather than a material existence. The entities considered are "abstract" collections, whose members are themselves sets. NBG offers a "larger" ontology than ZF. Indeed ZF can be construed as a subsystem of NBG, consisting of the part of NBG that refers only to sets, (i.e. classes that are not proper). We still have not shed any real light on what we mean by a "condition pertaining to objects x" (since sets are never members of themselves, the "less than 22 words" condition mentioned earlier will not be admissible in ZF or NBG). Some clarification of this notion will come later when we consider formal languages and take a closer look at the details of the axioms for systems like ZF.

Consistency

The fact that a particular system avoids Russell's Paradox does not guarantee that it is consistent, i.e. entirely free of contradictions. It is known an inconsistency in either ZF or NBG would imply an inconsistency in the other, and so the two systems stand or fall together. They have been intensively and extensively studied in the last 60 or so years without any contradiction emerging. However there is a real conceptual barrier to the possibility of proving that no such contradiction will ever be found. This was demonstrated by Gödel, around 1930, who showed in effect that any proof of consistency would have to depend on principles whose own consistency was no more certain than that of ZF and NBG themselves. In the decade prior to Gödel's work a group of mathematicians lead by David Hilbert had attempted to establish the consistency of arithmetic and mathematics generally by using only so-called *finitary* methods. These methods are confined to the description of concrete, particular, directly perceivable objects, and principles whose truth is evident by direct inspection. Gödel showed that such methods could never establish the consistency of any system that was powerful enough to develop the arithmetic of ordinary whole numbers. This discovery is regarded as one of the major mathematical events of the 20th century. Its impact on Hilbert's program was devastating, but many people have found in it a source of encouragement, an affirmation of the essentially creative nature of mathematical thought, and evidence against the mechanistic thesis that the mind can be adequately modelled as a physical computing device. As Gödel himself has put it, "either mathematics is too big for the human mind, or the human mind is more than a machine." (cf. Bergamini [65]).

While it would seem there can be no absolute demonstration of the consistency of ZF, there is considerable justification, of an experiential and epistemological nature, for the belief that it contains no contradictions. Certainly if the opposite were the case then, in view of the central role of set theory in contemporary mathematics, a great deal more would be at stake than simply the adequacy of a particular set of postulates.

Which of ZF and NBG is a "better" treatment of set theory? The choice is largely a matter of philosophical taste, together with practical need. ZF seems to enjoy the widest popularity amongst mathematicians generally. Its principle of relativising constructions to particular sets closely reflects the way set theory is actually used in mathematics, where sets are specified within clearly given, mathematically defined contexts (universes). The collection of all sets has not been an object of concern for most working mathematicians. Indeed the sets that they need can generally be obtained within a small fragment of ZF. It is only very recently, with the advent of category theory that a genuine need has arisen amongst mathematicians (other than set-theorists) for a means of handling large collections. These needs are met in a more flexible way by the class-set dichotomy, and have offered a more significant role to NBG and even stronger systems.

The moral to be drawn from these observations is that there is no "correct" way to do set theory. The system a mathematician chooses to work with will depend on what he wishes to achieve.

1.2. Foundations of mathematics

The aim of Foundational studies is to produce a rigorous explication of the nature of mathematical reality. This involves a precise and formal definition, or representation of mathematical concepts, so that their interrelationships can be clarified and their properties better understood. Most approaches to foundations use the axiomatic method. The language to be used is first introduced, generally itself in a precise and formal description. This language then serves for the definition of mathematical notions and the statement of postulates, or axioms, concerning their properties.

The axioms codify ways we regard mathematical objects as actually behaving. The theory of these objects is then developed in the form of statements derived from the axioms by techniques of deduction that are themselves rendered explicit.

It would be somewhat misleading to infer from this that foundational systems act primarily as a basis out of which mathematics is actually created. The artificiality of that view is evident when one reflects that the essential content of mathematics is already there before the basis is made explicit, and does not depend on it for its existence. We may for example think of a real number as an infinite decimal expression, or a point on the number line. Alternatively it could be introduced as an element of a complete ordered field, an equivalence class of Cauchy sequences, or a Dedekind cut. None of these could be said to be *the* correct explanation of what a real number is. Each is an enbodiment of an intuitive notion and we evalute it, not in terms of its correctness, but rather in terms of its effectiveness in explicating the nature of the real number system.

Mathematical discovery is by no means a matter of systematic deductive procedure. It involves insight, imagination, and long explorations along paths that sometimes lead nowhere. Axiomatic presentations serve to describe and communicate the fruits of this activity, often in a different order to that in which they were arrived at. They lend a coherence and unity to their subject matter, an overview of its extent and limitations.

Having clarified our intuitions, the formal framework may then be used for further exploration. It is at this level that the axiomatic method does have a creative role. The systematisation of a particular theory may lead to new internal discoveries, or the recognition of similarities with other theories and their subsequent unification. This however belongs to the "doing" of mathematics. As far as Foundational studies are concerned the role of axiomatics is largely descriptive. A Foundational system serves not so much to prop up the house of mathematics as to clarify the principles and methods by which the house was built in the first place. "Foundations" is a discipline that can be seen as a branch of mathematics standing apart from the rest of the subject in order to describe the world in which the working mathematician lives.

1.3. Mathematics as set theory

The equation of mathematics with set theory can with some justification be seen as a summary of the direction that mathematics has taken in modern times. Many will have heard of the revolution in school curricula called the "New Math". This has largely revolved around the introduction of set theory into elementary education and indicates the preoccupation of the mathematical community with that subject. Of all the foundational frameworks that have been proposed, the set theories have enjoyed the widest acceptance and the most detailed attention. Systems like ZF and NBG provide an elegant formalisation and explanation of the basic notions that the mathematician uses. Paul Cohen, whose work on the independence of the Continuum Hypothesis in 1963 lead to a veritable explosion of set-theoretic activity, has said "by analysing mathematical arguments logicians became convinced that the notion of "set" is the most fundamental concept of mathematics."

Apart from, or perhaps because of, its central role in Foundations, set theory has also dominated the stage of mathematical practise. This is not intended to imply that mathematicians think in set-theoretical concepts, although that is very often the case. Rather the point is that set theory is the basic tool of communication and exposition. It has provided the vehicle for an enormous proliferation of mathematics, both in terms of quantity of knowledge and range of topics and applications. It would be hard to find a recent book on any pure mathematical subject, be it algebra, geometry, analysis, or probability theory, that used no set-theoretical symbolism.

The group of French mathematicians who work under the name of Nicolas Bourbaki undertook in 1935 the formidable task of producing a "fully axiomatised presentation of mathematics in entirety". The result, over 40 years, has been about that many volumes to date, ranging over algebra, analysis and topology. Book 1 of this influential work is devoted to the theory of sets, which provides the framework for the whole enterprise. Bourbaki has said (1949) "... all mathematical theories may be regarded as extensions of the general theory of sets ... on these foundations I state that I can build up the whole of the mathematics of the present day".

The point to be made in this book is that the emergence of category theory has changed the perspectives just described, and that Cohen's statement is no longer even prima facie acceptable. It may be the case that the objects of mathematical study *can* be thought of as sets, but it is not certain that in the future they *will* be so regarded. No doubt the basic language of set theory will continue to be an important tool whenever collections of things are to be dealt with. But the conception of the things themselves as sets has lost some of its prominence through the development of a natural and attractive alternative. It seems indeed very likely that the role of set theory as the lingua universalis for mathematical foundations will be a declining one in the years to come. In case the wrong impression should have been conveyed by the last quotation above, it should be noted that the French mathematicians have been amongst the first to recognise this. René Thom [71] has written that "the old hope of Bourbaki, to see mathematical structures arise naturally from a hierachy of sets, from their subsets, and from their combination, is, doubtless, only an illusion". And in an address given in 1961, Jean Dieudonné made the following prophetic statement:

"In the years between 1920 and 1940 there occurred, as you know, a complete reformation of the classification of different branches of mathematics, necessitated by a new conception of the essence of mathematical thinking itself, which originated from the works of Cantor and Hilbert. From the latter there sprang the *systematic axiomatization* of mathematical science in entirety and the fundamental concept of *mathematical structure*. What you may perhaps be unaware of is that mathematics is about to go through a second revolution at this very moment. This is the one which is in a way completing the work of the first revolution, namely, which is releasing mathematics from the far too narrow conditions by 'set'; it is the theory of *categories and functors*, for which estimation of its range or perception of its consequences is still too early . . . ". (Quoted from Fang [70].)