SET THEORY

John R. Steel
Department of Mathematics, University of California, Berkeley, CA. USA

Keywords: Aronszajn tree, Axiom of Choice, Axiom of Infinity, Borel determinacy, cardinal number, closed unbounded set, Cohen's forcing method, Cohen real, consistency strength, constructible set, Continuum Hypothesis, Covering lemma, countable set, cumulative hierarchy of sets, descriptive set theory, diamond sequence, elementary embedding, fine structure theory, Fodor's lemma, forcing conditions, generic absoluteness, generic extension, generic filter, inner model, iterated ultrapower, large cardinal hypothesis, Lebesgue measure, Levy hierarchy, Levy collapse, Martin's Axiom, Martin's Maximum, measurable cardinal, ordinal number, Powerset Axiom, projective determinacy, random real, relative consistency, Reflection Principle, regular cardinal, set theory, singular cardinal, stationary set, super-compact cardinal, Suslin's Hypothesis, Suslin tree, transfinite induction, transitive model, transitive set, ultrafilter, wellfounded relation, wellorder, Woodin cardinal, Zermelo-Fraenkel Set Theory.

Contents

Summary

All mathematical statements can be expressed in the very simple *language of set theory*, whose quantifiers are understood as ranging over sets, and whose only nonlogical symbol $\in$ stands for the membership relation. All mathematical proofs to date can be carried out granted certain basic axioms about sets. In this sense, the general theory of sets is a foundation for all of mathematics.

The vast majority of mathematical proofs require no more than the axioms of Zermelo-Fraenkel set theory with Choice, or ZFC. Nevertheless, a surprising number of quite basic questions about sets in general are not decided by the axioms of ZFC; moreover many of the more abstract questions of analysis, algebra, and topology are left similarly undecided. Perhaps the most famous of the undecided questions is Cantor's Continuum Problem: what is the cardinality of the set of all real numbers?

In this article we shall describe some of the basic ideas of the general theory of sets, and then move to an exposition of the methods by which one can show that a given statement in the language of set theory is not decided by ZFC. We conclude with a discussion of extensions of ZFC obtained by strengthening the Axiom of Infinity, and show that these suffice to remove the incompleteness of ZFC in one important realm.

1. Introduction

Much of set theory is motivated by the simple question

What are the proper axioms for mathematics?

Mathematicians prove things for a living; what should they take as their common assumptions in these proofs?

In a sense, this question goes back to Euclid, although in his day, merely axiomatizing geometry was a great achievement. It was through Descartes' reduction of geometry to analysis, the reduction of analysis to arithmetic by 19th century mathematicians such as Cauchy and Dedekind, and the reduction of arithmetic to set theory by Gottlob Frege, that the project of axiomatizing all of mathematics became feasible around the turn of the 20th century. These reductions showed that all of the mathematics of the time could be expressed in the very simple *language of set theory*, whose quantifiers are understood as ranging over *pure sets*, and whose only non-logical symbol $\in$ is interpreted as standing for the membership relation. The pure sets are just those built up from the empty set by repeatedly forming sets of objects previously constructed in a process which is iterated into the transfinite. Since we shall not need to consider impure sets (such as the set of knives in my cabinet), henceforth we use *set* to mean pure set.

In the period 1905-1927, Russell, Zermelo, Fraenkel, and Skolem isolated an elegant list of basic statements about sets, expressed in the language of set theory, and showed that from these axioms one could derive all of the mathematics of the time. This system of axioms is now known as Zermelo-Fraenkel Set Theory with Choice, or ZFC. Most of its axioms are set-existence axioms; for example, the Axiom of Infinity asserts that there is an infinite set, and the Powerset Axiom asserts that for every set x, there is a set $P(x)$ whose members are precisely the subsets of x. (See *Formal Logic* for further discussion.) ZFC is a good provisional answer to the question with which we began, and it is still true today that almost everything that mathematicians have proved has been derived from its axioms.

While there is still no hint of a mathematical statement which cannot be expressed in the language of set theory, we have discovered that ZFC is incomplete in important ways. A surprising number of quite basic questions about sets in general are not decided by the axioms of ZFC; moreover, many of the more abstract questions of analysis, algebra, and topology are similarly left undecided. Here is a list of some examples.

The Continuum Hypothesis.

The first theorem of general set theory was Cantor's remarkable discovery, in 1873, that infinite sets come in different sizes. Cantor showed that for any set x, there are strictly more subsets of x than there are elements of x; in modern cardinality notation, that $|x| < |P(x)|$ for all x. He naturally asked whether for infinite x there are any sets of intermediate size. The *Generalized Continuum Hypothesis*, or GCH, asserts that there are not, that is, that for all infinite x, $|P(x)| = |x|^+$, where $|x|^+$ is the least cardinal strictly greater than $|x|$. The special case in which $|x| = \omega$ is the smallest infinite cardinal is called the *Continuum Hypothesis*, or CH, because in this case $|P(x)| = |\mathbb{R}|$. Calling a set *countable* if it is either finite or equinumerous with the set ω of all natural numbers, the Continuum Hypothesis is equivalent to the statement that every uncountable set of real numbers is equinumerous with the set $\mathbb{R}$ of all real numbers.

In 1937, Gödel produced a model of ZFC in which the GCH is true. In 1963, Cohen produced a family of models of ZFC in which the CH is false; there are, as we shall see, many possibilities for the function $|x| \mapsto |P(x)|$ in such models. By itself, ZFC has very little to say about one of the most basic problems in general set theory, the problem of counting the powerset of an infinite set.

Suslin's Hypothesis.

A partial order is a structure $\mathbb{P} = (P, \leq)$, where $\leq$ is a reflexive, antisymmetric, transitive binary relation on P. $\mathbb{P}$ is *linear* if for all $x, y \in P$, $x \leq y$ or $y \leq x$. If $L = (L, \leq)$ is a linear order with no largest or smallest element such that

(1) Every $X \subseteq L$ which is bounded above has a least upper bound, and
(2) there is an $X \subseteq L$ such that $|X| = \omega$ and X is dense in $\mathbb{L}$

(i.e., $\forall a,b \in L(a<b \rightarrow \exists c \in X\,(a<c<b))$,

then $\mathbb{L}$ is isomorphic to $(\mathbb{R},\leq^{\mathbb{R}})$, the real numbers with the standard order on them. In the 1920's, Suslin conjectured that this characterization of $(\mathbb{R},\leq^{\mathbb{R}})$ can be varied by replacing (2) with the ostensibly weaker

(2a) if $\mathcal{I}$ is a family of pairwise disjoint intervals of $\mathbb{L}$, then $\mathcal{I}$ is countable.

A *Suslin line* is a linear order $\mathbb{L}$ without endpoints satisfying (1) and (2a), but not isomorphic to $(\mathbb{R},\leq^{\mathbb{R}})$. *Suslin's Hypothesis* (SH) asserts, confusingly enough, that there are no Suslin lines.

In the late 1960's, Jech used Cohen's techniques to produce a model of ZFC is which there are Suslin lines. Not long afterward, Solovay and Tennenbaum extended Cohen's technique in an important way, and thereby produced a model of ZFC in which there are no Suslin lines.

Scales in $\left(\omega^{\omega},\leq^{*}\right)$.

We define the partial order of eventual domination on $\omega^{\omega}=\{f \mid f:\omega\rightarrow\omega\}$ by: $f\leq^{*} g$ iff $\exists m\forall n\geq m(f(n)\leq g(n))$. A *scale* in $(\omega^{\omega},\leq^{*})$ is a set $S\subseteq\omega^{\omega}$ which is linearly ordered by $\leq^{*}$ and cofinal, in the sense that $\forall f\in\omega^{\omega}\exists g\in S(f\leq^{*} g)$. It is not hard to see that CH implies that there is a scale. There are models of ZFC constructed by Cohen's method in which CH is false and there are no scales, as well as models in which CH is false and nevertheless, there is a scale.

Lebesgue Measure.

In 1902, the analyst Lebesgue defined a natural measure μ on certain sets of reals. If A is an interval, then $\mu(A)$ is just its length. If A is open, then $A=\bigcup_{n<\omega} I_n$ In where the I_n are disjoint intervals, and we set $\mu(A)=\Sigma_{n=1}^{\infty}\mu(I_n)$. If A is closed, then $\mu(A)=\Sigma_{n\in\mathbb{Z}}1-\mu((n,n+1)\setminus A)$. Finally, for arbitrary A, we define the outer measure
$\mu^{+}(A)=\inf(\{\mu(U)\mid U$ is open and $A\subseteq U\})$, and inner measure
$\mu^{-}(A)=\sup(\{\mu(F)\ F$ is closed and $F\subseteq A.\})$.

We say that A is *Lebesgue measurable* iff $\mu^{+}(A)=\mu^{-}(A)$, in which case we write $\mu(A)$ for the common value. Lebesgue's measure leads to a natural and important extension of the Riemann integral to functions which may have infinitely many discontinuities.

Lebesgue showed that the collection of Lebesgue measurable sets is closed under complements and countable unions, and therefore includes the class of Borel subsets of [0,1]. On the other hand, Vitali showed with a heavy use of the Axiom of Choice that there is a non-Lebesgue-measurable set. His construction is the following: for $x, y \in [0,1]$, let $xEy \Leftrightarrow x - y$ is rational. E is an equivalence relation, and by the Axiom of Choice there is a set $A \subseteq [0,1]$ such that A meets each equivalence class of E in exactly one point. Suppose A were Lebesgue measurable, and consider the rational translations $T_q : [0,1] \to [0,2]$ given by

$$T_q(x) = x + q$$

for rational $q \in [0,1]$. Our choice of A guarantees that the images $T_q(A)$ for $q \in \mathbb{Q}$ are pairwise disjoint and cover [1, 2]. But Lebesgue measure is translation invariant, so $\mu(A) = \mu(T_q(A))$ for all q. Since Lebesgue measure is also countably additive, we have a contradiction: if $\mu(A) = 0$, then $\mu([1,2]) = 0$, and if $\mu(A) > 0$, then $\mu([1,2]) = \infty$.

Vitali's example illustrates the fact that *arbitrary* sets of reals may have no geometric content. A dramatic example of this phenomenon is the famous *Banach-Tarski paradox*: there is a decomposition of the unit ball in $\mathbb{R}^3$ into finitely many pieces, which can be moved by translations and rotations to become a decomposition of the unit cube. The pieces in such a decomposition cannot be measurable with respect to the natural extension of Lebesgue measure to $\mathbb{R}^3$.

It is natural to ask whether the sets more familiar to analysts, those given by some construction or definition, can exhibit such pathology, or whether they are all Lebesgue measurable. Gödel showed in 1937 that in his model of ZFC+ GCH, it is also true that there is a simply definable subset of [0, 1] which is not Lebesgue measurable. On the other hand, Solovay in 1965 was able to use Cohen's technique to produce a model of ZFC in which all reasonably definable sets of reals are Lebesgue measurable.

It is also natural to ask whether there is an extension of Lebesgue measure to a measure ν defined on all $A \subseteq [0,1]$A ⊆ [0, 1]. (Such a measure ν cannot be translation-invariant, by Vitali's argument.) In Gödel's model, there is no such extension; on the other hand, Solovay in 1966 produced a model of ZFC in which there is such an extension.

In the rest of this chapter we shall outline the independence techniques of Gödel and Cohen, and show how they apply to these and several other natural problems. The sheer volume of work in this vein can lead to the impression that the set-theorist's sole function delivers to other mathematicians, often other set-theorists, the bad news that they will not be able to solve this or that problem. We hope to lead the reader to a different conclusion: that one of the most interesting and potentially useful global features of the universe V of all sets is that it has within it descriptions of many alternate universes satisfying various natural theories. These alternate universes are not so far from V in some ways; for example, they all have the same natural numbers as does V, so that any statement of number theory true in one of them is true in V. The study of such alternate universes is an inevitable and important part of the global theory of V.

We shall also introduce some plausible reinforcements of the Axiom of Infinity, and show that these strong axioms of infinity (also known as large cardinal hypotheses), decide many of the more concrete statements of mathematical interest which are left undecided by ZFC. This suggests that our answer to "Euclid's question" is still evolving.

2. Some Elementary Tools.

We begin with a quick inventory of some of the more elementary ideas and results of set theory. The reader who would like a gentler and more thorough introduction to this material should see *Formal Logic* .

2.1. Ordinals

Transfinite induction is one of the signature methods of set theory. Indeed, set theory began with Cantor's use of transfinite induction to study trigonometric series.

A *wellfounded relation* is a binary relation *R* on a set *Y* such that every nonempty set $X \subseteq Y$ has an *R*-minimal element, that is, $\exists a \in X \forall b \in Y(\neg bRa)$. Equivalently, *R* is wellfounded iff there is no sequence $\langle a_i \mid i < \omega \rangle$ such that $\forall i(a_{i+1}R_{a_i})$. One can carry out proofs and definitions by transfinite induction along a wellfounded relation; for example, to define a function *f* with domain *Y*, it suffices to define $f(a)$ from $f \mid \{b|bRa\}$. The existence of a unique *f* with domain *Y* satisfying the induction clause at all $a \in Y$ can be proved in ZFC, using the Axiom of Replacement.

A *wellorder* is a linear order whose associated strict order is wellfounded. Any two wellorders $\mathbb{W}_1$ and $\mathbb{W}_2$ can be compared as to length: either $\mathbb{W}_1$ is isomorphic to a proper initial segment of $\mathbb{W}_2$, or $\mathbb{W}_1$ is isomorphic to $\mathbb{W}_2$, or $\mathbb{W}_2$ is isomorphic to a proper initial segment of $\mathbb{W}_1$, moreover these possibilities are mutually exclusive, and the isomorphism is in each case unique. These facts are proved using transfinite induction.

A set *X* is *transitive* if whenever $a \in b$ and $b \in X$,then $a \in X$; that is, *X* is an $\in$-initial segment of the universe of all sets. If $(Y, <^*)$ is a strict wellorder, we define by transfinite induction

$\pi(a) = \{\pi(b) | b <^* a\}$.

The range of π is a transitive set which is strictly wellordered by $\in$;we call such a set an *ordinal*. We have thus shown that every (strict) wellorder is isomorphic to an ordinal. It is easy to show that the ordinal and isomorphism are uniquely determined by the wellorder. If α and β are ordinals, we write $\alpha < \beta$ to mean that α is isomorphic to a proper initial segment of β. It is a pleasant consequence of this choice of representatives for isomorphism types of wellorders (which is due to Von Neumann) that $\alpha < \beta$ iff $\alpha \in \beta$. The relation < wellorders the class of all ordinals, with the least upper bound of a set X being the ordinal $\bigcup X$ .The least ordinal is 0 = $\varnothing$, the order-type of the empty wellorder. If α is an ordinal, then $\alpha+1=\alpha\cup\{\alpha\}$ is the least ordinal strictly greater than α. λ is a *limit ordinal* iff $\lambda>0$ and $\forall\alpha(\lambda\neq\alpha+1)$. The Axiom of Infinity is equivalent to the statement: there is a limit ordinal. The least limit ordinal is called ω, and its members are called natural numbers.

2.2. The Wellordering Theorem

The Axiom of Choice implies that wellorders exist in abundance:

Theorem 2.1 (Zermelo 1905). *Every set admits a wellorder*.

Proof. Let X be given, and let f be a choice function with domain $P(X)$, so that $f(Y)\in Y$ for all nonempty $Y\subseteq X$. We define a function π with domain the ordinals by transfinite induction as follows:

$$\pi(\alpha)=f(X\setminus\{\pi(\beta\mid\beta<\alpha\})\,.$$

That is, we keep listing distinct elements of X, using f to pick the next one. The Axiom of Replacement implies that there is an α such that $\operatorname{ran}(\pi\mid\alpha)=X$. But then the image under π of the wellorder $\in$ on α is a wellorder of X.

This combination of definition by transfinite induction and the Axiom of Choice is quite powerful. Here is another example. A *filter* on a set I is a family $\mathcal{F}\subseteq P(I)$ closed under finite intersections and superset, and not containing $\emptyset$ .One should think of F as a notion of largeness for subsets of I. An *ultrafilter* on I is a filter F on I such that for all $X\subseteq I$, either $X\in\mathcal{F}$ or $I\setminus X\in\mathcal{F}$. Using Zermelo's method, we can show *every filter extends to an ultrafilter*. For let $\langle X_\alpha\mid\alpha<\theta\rangle$ enumerate $P(I)$, and let $\mathcal{F}$ be a given filter on I. We define filters $\mathcal{F}_\alpha$ for $\alpha\leq\theta$ by induction: set $\mathcal{F}_0=\mathcal{F}$,
$\mathcal{F}_{\alpha+1}$ = some filter $\mathcal{G}$ extending $\mathcal{F}_\alpha$ s.t. $X_\alpha\in\mathcal{G}$ or $I\setminus X_\alpha\in\mathcal{G}$,
and

$$\mathcal{F}_\lambda=\bigcup_{\alpha<\lambda}\mathcal{F}_\alpha$$

for λ a limit ordinal. It is not hard to show that a filter $\mathcal{G}$ as above must exist, and that at limit stages, $\mathcal{F}_\lambda$ remains a filter. The desired ultrafilter is then $\mathcal{F}_\theta$.

2.3. The Cumulative Hierarchy; Proper Classes

One further definition by transfinite induction deserves mention, that of the *cumulative hierarchy*. We define

$V_0 = \emptyset,$

$V_{\alpha+1} = V_\alpha \bigcup P(V_\alpha),$

and for λ a limit ordinal,

$V_\lambda = \bigcup_{\alpha<\lambda} V_\alpha.$

Thus V_α consists of those sets which can be built up in $<\alpha$ stages by taking sets of objects previously formed. Each V_α is transitive, and $\alpha \leq \beta \Rightarrow V_\alpha \subseteq V_\beta$. It follows from the Axiom of Foundation of ZFC that every set is in some V_α.We shall also write *V* for the union of all the V_α, the universe of all sets. *V* is not itself a set, but rather what is sometimes called a *proper class*. Another useful proper class is OR, the class of all ordinals. If, as we have claimed, the language of set theory is universal, then our apparent references to proper classes should be eliminable, and indeed they are. Usually, this is because we can replace "$x \in A$" with $\varphi(x)$, where φ is a formula of the language of set theory defining membership in *A*. Thus instead of saying that $V_\alpha \bigcap \mathrm{OR} = \alpha$ (which, incidentally, is true), we could simply say $\forall x(x \in V_\alpha \wedge x \text{ is an ordinal}) \Leftrightarrow x \in \alpha$. Nevertheless, the apparent reference to proper classes can be convenient, so we shall make use of it on occasion.

-
-
-

Biographical Sketch

John R. Steel is a Professor of Mathematics at the University of California, Berkeley. He has worked in all aspects of set theory, and especially in Descriptive Set Theory and the connection between large cardinals and determinacy hypotheses. He has spoken at the International Congress of Mathematicians, and he shared with Donald A. Martin and Hugh Woodin the 1988 Carol Karp Prize of the Association for Symbolic Logic, which is awarded every five years.