Handbook of Set Theory

Matthew Foreman • Akihiro Kanamori
Editors

Handbook of Set Theory

Springer

Matthew Foreman
University of California at Irvine
School of Physical Sciences
Department of Mathematics
Irvine CA 92697-3875
USA
mforeman@math.uci.edu

Akihiro Kanamori
Boston University
College of Arts and Sciences
Department of Mathematics
Boston MA 02215
USA
aki@math.bu.edu

ISBN 978-1-4020-4843-2
e-ISBN 978-1-4020-5764-9
Printed in 3 volumes
DOI 10.1007/978-1-4020-5764-9
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009941207

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Numbers imitate space, which is of such a different nature
—Blaise Pascal

It is fair to date the study of the foundation of mathematics back to the ancient Greeks. The urge to understand and systematize the mathematics of the time led Euclid to postulate axioms in an early attempt to put geometry on a firm footing. With roots in the *Elements*, the distinctive methodology of mathematics has become *proof*. Inevitably two questions arise: *What are proofs?* and *What assumptions are proofs based on?*

The first question, traditionally an internal question of the field of *logic*, was also wrestled with in antiquity. Aristotle gave his famous syllogistic systems, and the Stoics had a nascent propositional logic. This study continued with fits and starts, through Boethius, the Arabs and the medieval logicians in Paris and London. The early germs of logic emerged in the context of philosophy and theology.

The development of analytic geometry, as exemplified by Descartes, illustrated one of the difficulties inherent in founding mathematics. It is classically phrased as the question of *how one reconciles the arithmetic with the geometric*. Are numbers one type of thing and geometric objects another? What are the relationships between these two types of objects? How can they interact? Discovery of new types of mathematical objects, such as imaginary numbers and, much later, formal objects such as free groups and formal power series make the problem of finding a common playing field for all of mathematics importunate.

Several pressures made foundational issues urgent in the 19th century. The development of alternative geometries created doubts about the view that mathematical truth is part of an absolute all-encompassing logic and caused it to evolve towards one in which mathematical propositions follow logically from assumptions that may vary with context.

Mathematical advances involving the understanding of the relationship between the completeness of the real line and the existence of solutions to equations led inevitably to anxieties about the role of infinity in mathematics.

These too had antecedents in ancient history. The Greeks were well aware of the scientific importance of the problems of the infinite which were put forth, not only in the paradoxes of Zeno, but in the work of Eudoxus, Archimedes and others. Venerable concerns about resolving infinitely divisible lines into individual points and what is now called "Archimedes' Axiom" were recapitulated in 19th century mathematics.

In response, various "constructions" of the real numbers were given, such as those using Cauchy sequences and Dedekind cuts, as a way of understanding the relationship between discrete entities, such as the integers or the rationals and the continuum. Even simple operations, such as addition of arbitrary real numbers began to be understood as infinitary operations, defined by some kind of limiting process. The notion of *function* was liberalized beyond those that can be written in closed form. Sequences and series became routine tools for solving equations.

The situation was made acute when Cantor, working on natural problems involving trigonometric series, discovered the existence of different magnitudes of infinity. The widespread use of inherently infinitary techniques, such as the use of the Baire Category Theorem to prove the existence of important objects, became deeply embedded in standard mathematics, making it impossible to simply reject infinity as part of mathematics.

In parallel 19th century developments, through the work of Boole and others, logic became once again a mathematical study. Boole's algebraization of logic made it grist for mathematical analysis and led to a clear understanding of propositional logic. Dually, logicians such as Frege viewed mathematics as a special case of logic. Indeed a very loose interpretation of the work of Frege is that it is an attempt to base mathematics on a broad notion of logic that subsumed all mathematical objects.

With Russell's paradox and the failure of Frege's program, a distinction began to be made between *logic* and *mathematics.* Logic began to be viewed as a formal epistemological mechanism for exploring mathematical truth, barren of mathematical content and in need of assumptions or axioms to make it a useful tool.

Progress in the 19th and 20th centuries led to the understanding of logics involving quantifiers as opposed to propositional logic and to distinctions such as those between first and second-order logic. With the semantics developed by Tarski and the compactness and completeness theorems of Gödel, first-order logic has become widely accepted as a well-understood, unproblematic answer to the question *What is a proof?*

The desirable properties of first-order logic include:

- Proofs and propositions are easily and uncontroversially recognizable.
- There is an appealing semantics that gives a clear understanding of the relationship between a mathematical structure and the formal propositions that hold in it.
- It gives a satisfactory model of what mathematicians actually do: the "rigorous" proofs given by humans seem to correspond exactly to the

> "formal" proofs of first-order logic. Indeed formal proofs seem to provide a normative ideal towards which controversial mathematical claims are driven as part of their verification process.

While there are pockets of resistance to first-order logic, such as constructivism and intuitionism on the one hand and other alternatives such as second-order logic on the other, these seem to have been swept aside, if simply for no other reason than their comparative lack of mathematical fruitfulness.

To summarize, a well-accepted conventional view of foundations of mathematics has evolved that can be caricatured as follows:

$$\text{Mathematical Investigation} = \text{First-Order Logic} + \text{Assumptions}$$

This formulation has the advantage that it segregates the difficulties with the foundations of mathematics into discussions about the underlying assumptions rather than into issues about the nature of reasoning.

So what are the appropriate assumptions for mathematics? It would be very desirable to find assumptions that:

1. involve a simple primitive notion that is easy to understand and can be used to "build" or develop all standard mathematical objects,

2. are *evident*,

3. are *complete* in that they settle all mathematical questions,

4. can be easily recognized as part of a recursive schema.

Unfortunately Gödel's incompleteness theorems make item 3 impossible. *Any* recursive consistent collection $\mathcal{A}$ of mathematical assumptions that are strong enough to encompass the simple arithmetic of the natural numbers will be *incomplete*; in other words there will be mathematical propositions P that cannot be settled on the basis of $\mathcal{A}$. This inherent limitation is what has made the foundations of mathematics a lively and controversial subject.

Item 2 is also difficult to satisfy. To the extent that we understand mathematics, it is a difficult and complex business. The Euclidean example of a collection of axioms that are easily stated and whose content is simple to appreciate is likely to be misleading. Instead of simple, distinctly conceived and obvious axioms, the project seems more analogous to specifying a complicated operating system in machine language. The underlying primitive notions used to develop standard mathematical objects are combined in very complicated ways. The axioms describe the operations necessary for doing this and the test of the axioms becomes how well they code higher level objects as manipulated in ordinary mathematical language so that the results agree with educated mathematicians' sense of correctness.

Having been forced to give up 3 and perhaps 2, one is apparently left with the alternatives:

2′. Find assumptions that are in accord with the intuitions of mathematicians well versed in the appropriate subject matter.

3′. Find assumptions that *describe* mathematics to as large an extent as is possible.

With regard to item 1, there are several choices that could work for the primitive notion for developing mathematics, such as *categories* or *functions*. With no *a priori* reason for choosing one over another, the standard choice of *sets* (or set membership) as the basic notion is largely pragmatic. Taking sets as the primitive, one can easily do the traditional constructions that "build" or "code" the usual mathematical entities: the empty set, the natural numbers, the integers, the rationals, the reals, $\mathbb{C}$, $\mathbb{R}^n$, manifolds, function spaces—all of the common objects of mathematical study.

In the first half of the 20th century a standard set of assumptions evolved, the axiom system called the Zermelo-Fraenkel axioms with the Axiom of Choice (ZFC). It is pragmatic in spirit; it posits sufficient mathematical strength to allow the development of standard mathematics, while explicitly rejecting the type of objects held responsible for the various paradoxes, such as Russell's.

While ZFC is adequate for most of mathematics, there are many mathematical questions that it does not settle. Most prominent among them is the first problem on Hilbert's celebrated list of problems given at the 1900 International Congress of Mathematicians, the *Continuum Hypothesis*.

In the jargon of logic, a question that cannot be settled in a theory T is said to be *independent* of T. Thus, to give a mundane example, the property of being Abelian is independent of the axioms for group theory. It is routine for normal axiomatizations that serve to synopsize an abstract concept internal to mathematics to have independent statements, but more troubling for axiom systems intended to give a definitive description of mathematics itself. However, independence phenomena are now known to arise from many directions; in essentially every area of mathematics with significant infinitary content there are natural examples of statements independent of ZFC.

This conundrum is at the center of most of the chapters in this Handbook. Its investigation has left the province of abstract philosophy or logic and has become a primarily mathematical project. The intent of the Handbook is to provide graduate students and researchers access to much of the recent progress on this project. The chapters range from expositions of relatively well-known material in its mature form to the first complete published proofs of important results. The introduction to the Handbook gives a thorough historical background to set theory and summaries of each chapter, so the comments here will be brief and general.

The chapters can be very roughly sorted into four types. The first type consists of chapters with theorems demonstrating the independence of mathematical statements. Understanding and proving theorems of this type require a thorough understanding of the mathematics surrounding the source of the problem in question, reducing the ambient mathematical constructions to combinatorial statements about sets, and finally using some method (primarily forcing) to show that the combinatorial statements are independent.

A second type of chapter involves delineating the edges of the independence phenomenon, giving proofs in ZFC of statements that on first sight would be suspected of being independent. Proofs of this kind are often extremely subtle and surprising; very similar statements are independent and it is hard to detect the underlying difference.

The last two types of chapters are motivated by the desire to *settle* these independent statements by adding assumptions to ZFC, such as large cardinal axioms. Proposers of augmentations to ZFC carry the burden of marshaling sufficient evidence to convince informed practitioners of the reasonableness, and perhaps truth, of the new assumptions as descriptions of the mathematical universe. (Proposals for axiom systems intended to *replace* ZFC carry additional heavier burdens and appear in other venues than the Handbook.)

One natural way that this burden is discharged is by determining what the supplementary axioms *say*; in other words by investigating the consequences of new axioms. This is a strictly mathematical venture. The theory is assumed and theorems are proved in the ordinary mathematical manner. Having an extensive development of the consequences of a proposed axiom allows researchers to see the overall picture it paints of the set-theoretic universe, to explore analogies and disanalogies with conventional axioms, and judge its relative coherence with our understanding of that universe. Examples of this include chapters that posit the assumption that the Axiom of Determinacy holds in a model of Zermelo-Fraenkel set theory that contains all of the real numbers and proceed to prove deep and difficult results about the structure of definable sets of reals.

Were there an obvious and compelling unique path of axioms that supplement ZFC and settle important independent problems, it is likely that the last type of chapter would be superfluous. However, historically this is not the case. Competing axioms systems have been posited, sometimes with obvious connections, sometimes appearing to have nothing to do with each other.

Thus it becomes important to compare and contrast the competing proposals. The Handbook includes expositions of some stunningly surprising results showing that one axiom system actually implies an apparently unrelated axiom system. By far the most famous example of this are the proofs of determinacy axioms from large cardinal assumptions.

Many axioms or independent propositions are not related by implication, but rather by *relative consistency* results, a crucial idea for the bulk of the Handbook. A remarkable meta-phenomenon has emerged. There appears to be a central spine of axioms to which all independent propositions are comparable in consistency strength. This spine is delineated by large cardinal axioms. There are no known counterexamples to this behavior.

Thus a project initiated to understand the relationships between disparate axiom systems seems to have resulted in an understanding of most known natural axioms as somehow variations on a common theme—at least as far as consistency strength is concerned. This type of unifying deep structure is taken as strong evidence that the axioms proposed reflect some underlying reality and is often cited as a primary reason for accepting the existence of large cardinals.

The methodology for settling the independent statements, such as the Continuum Hypothesis, by looking for evidence is far from the usual deductive paradigm for mathematics and goes against the rational grain of much philosophical discussion of mathematics. This has directed the attention of some members of the philosophical community towards set theory and has been grist for many discussions and message boards. However interpreted, the investigation itself is entirely mathematical and many of the most skilled practitioners work entirely as mathematicians, unconcerned about any philosophical anxieties their work produces.

Thus set theory finds itself at the confluence of the foundations of mathematics, internal mathematical motivations and philosophical speculation. Its explosive growth in scope and mathematical sophistication is testimony to its intellectual health and vitality.

The Handbook project has some serious defects, and does not claim to be a remotely complete survey of set theory; the work of Shelah is not covered to the appropriate extent given its importance and influence and the enormous development of classical descriptive set theory in the last fifteen years is nearly neglected. While the editors regret this, we are consoled that those two topics, at least, are well documented elsewhere. Other parts of set theory are not so lucky and we apologize.

We the editors would like to thank all of the authors for their labors. They have taken months or years out of their lives to contribute to this project. We would especially like to thank the referees, who are the unsung heroes of the story, having silently devoted untold hours to carefully reading the manuscripts simply for the benefit of the subject.

Matthew Foreman
Irvine

Let me express a special gratitude to the Lichtenberg-Kolleg at Göttingen. Awarded an inaugural 2009–2010 fellowship, I was provided with a particularly supportive environment at the Gauss Sternwarte, in the city in which David Hilbert, Ernst Zermelo, and Paul Bernays did their formative work on the foundations of mathematics. Thus favored, I was able to work in peace and with inspiration to complete the final editing and proof-reading of this Handbook.

Akihiro Kanamori
Boston and Göttingen

Contents

Volume 1

Volume 2

Volume 3

List of Contributors

Uri Abraham Department of Mathematics, Ben Gurion University of the Negev, Beer Sheva, Israel, abraham@cs.bgu.ac.il

Tomek Bartoszynski National Science Foundation, Arlington, VA 22230, USA, tomek.bartoszynski@gmail.com

Andreas Blass Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA, ablass@umich.edu

James Cummings Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA 15213, USA, jcumming@andrew.cmu.edu

Patrick Dehornoy Department of Mathematics, Université de Caen, 14 032 Caen cedex, France, dehornoy@math.unicaen.fr

Todd Eisworth Department of Mathematics, Ohio University, Athens, OH 45701, USA, eisworth@math.ohiou.edu

Matthew Foreman Department of Mathematics, University of California at Irvine, Irvine, CA 92697, USA, mforeman@math.uci.edu

Sy D. Friedman Kurt Gödel Research Center for Mathematical Logic, Universität Wien, Wien, Austria, sdf@logic.univie.ac.at

Moti Gitik Department of Mathematics, Tel Aviv University, Tel Aviv, Israel, gitik@math.tau.ac.il

András Hajnal Alfred Renyi Institute of Mathematics, Hungarian Academy of Sciences, Budapest, Hungary, ahajnal@renyi.hu

Greg Hjorth Department of Mathematics and Statistics, University of Melbourne, Melbourne, Australia, greg.hjorth@gmail.com

Steve Jackson Department of Mathematics, University of North Texas, Denton, TX 76203, USA, jackson@unt.edu

Thomas Jech Center for Theoretical Study, Prague 1, Czech Republic, jech@math.cas.cz

Akihiro Kanamori Department of Mathematics, Boston University, Boston, MA 02215, USA, aki@math.bu.edu

Peter Koellner Department of Philosophy, Harvard University, Cambridge, MA, USA, koellner@fas.harvard.edu

Jean A. Larson Department of Mathematics, University of Florida, Gainsville, FL 32611, USA, jal@math.ufl.edu

Paul B. Larson Department of Mathematics and Statistics, Miami University, Oxford, OH 45056, USA, larsonpb@muohio.edu

Menachem Magidor Institute of Mathematics, Hebrew University, Jerusalem, Israel, menachem@math.huji.ac.il

William J. Mitchell Department of Mathematics, University of Florida, Gainsville, FL 32611, USA, mitchell@math.ufl.edu

Itay Neeman Department of Mathematics, University of California at Los Angeles, Los Angeles, CA 90095, USA, ineeman@math.ucla.edu

Ernest Schimmerling Department of Mathematics, Carnegie Mellon University, Pittsburg, PA 15213, USA, eschimme@andrew.cmu.edu

Ralf Schindler Institut für mathematische Logik und Grundlagenforschung, Universität Münster, 48149 Münster, Germany, rds@math.uni-muenster.de

John R. Steel Department of Mathematics, University of California at Berkeley, Berkeley, CA 94720, USA, steel@math.berkeley.edu

Stevo Todorcevic Universite Paris VII – CNRS, 2, Place Jussieu – Case 7012, 75251 Paris Cedex 05, France, stevo@math.jussieu.fr

Philip D. Welch Department of Mathematics, University of Bristol, Bristol BS8 1TW, United Kingdom, p.welch@bristol.ac.uk

W. Hugh Woodin Department of Mathematics, University of California at Berkeley, Berkeley, CA 94720, USA, woodin@math.berkeley.edu

Martin Zeman Department of Mathematics, University of California at Irvine, Irvine, CA 92697, USA, mzeman@math.uci.edu

Introduction

Akihiro Kanamori

Contents

M. Foreman, A. Kanamori (eds.), *Handbook of Set Theory*,
DOI 10.1007/978-1-4020-5764-9_1,

Set theory has entered its prime as an advanced and autonomous research field of mathematics with broad foundational significance, and this Handbook with its expanse and variety amply attests to the fecundity and sophistication of the subject. Indeed, in set theory's further reaches one sees tremendous progress both in its continuing development of its historical heritage, the investigation of the transfinite numbers and of definable sets of reals, as well as its analysis of strong propositions and consistency strength in terms of large cardinal hypotheses and inner models.

This introduction provides a historical and organizational frame for both modern set theory and this Handbook, the chapter summaries at the end being a final elaboration. To the purpose of drawing in the serious, mathematically experienced reader and providing context for the prospective researcher, we initially recapitulate the consequential historical developments leading to modern set theory as a field of mathematics. In the process we affirm basic concepts and terminology, chart out the motivating issues and driving initiatives, and describe the salient features of the field's internal practices. As the narrative proceeds, there will be a natural inversion: Less and less will be said about more and more as one progresses from basic concepts to elaborate structures, from seminal proofs to complex argumentation, from individual moves to collective enterprise. We try to put matters in a succinct yet illuminating manner, but be that as it may, according to one's experience or interest one can skim the all too familiar or too obscure. To the historian this account would not properly be history—it is, rather, a deliberate arrangement, in significant part to lay the ground for the coming chapters. To the seasoned set theorist there may be issues of under-emphasis or over-emphasis, of omissions or commissions. In any case, we take refuge in a wise aphorism: If it's worth doing, it's worth doing badly.

1. Beginnings

1.1. Cantor

Set theory was born on that day in December 1873 when Georg Cantor (1845–1918) established that *the continuum is not countable*—there is no one-to-one correspondence between the real numbers and the natural numbers $0, 1, 2, \ldots$. Given a (countable) sequence of reals, Cantor defined nested intervals so that any real in their intersection will not be in the sequence. In the course of his earlier investigations of trigonometric series Cantor had developed a definition of the reals and had begun to entertain infinite totalities of reals for their own sake. Now with his uncountability result Cantor embarked on a full-fledged investigation that would initiate an expansion of the very concept of number. Articulating cardinality as based on bijection (one-to-one correspondence) Cantor soon established positive results about the existence of bijections between sets of reals, subsets of the plane, and the like. By 1878 his investigations had led him to assert that there are only two infinite cardinalities embedded in the continuum: *Every infinite set of reals is either countable or in bijective correspondence with all the reals.* This was the Continuum Hypothesis (CH) in its nascent context, and the *continuum problem*, to resolve this hypothesis, would become a major motivation for Cantor's large-scale investigations of infinite numbers and sets.

In his magisterial *Grundlagen* of 1883 Cantor developed the *transfinite numbers* and the key concept of *well-ordering*, in large part to take a new, structured approach to infinite cardinality. The transfinite numbers follow the natural numbers $0, 1, 2, \ldots$ and have come to be depicted in his later notation in terms of natural extensions of arithmetical operations:

$$\omega, \omega+1, \omega+2, \ldots \omega+\omega(=\omega{\cdot}2), \\ \ldots \omega{\cdot}3, \ldots \omega{\cdot}\omega(=\omega^2), \ldots \omega^3, \ldots \omega^\omega, \ldots \omega^{\omega^\omega}, \ldots .$$

A *well-ordering* on a set is a linear ordering of it according to which every non-empty subset has a least element. Well-orderings were to carry the sense of sequential counting, and the transfinite numbers served as standards for gauging well-orderings. Cantor developed cardinality by grouping his transfinite numbers into successive number classes, two numbers being in the same class if there is a bijection between them. Cantor then propounded a basic principle: "It is always possible to bring any *well-defined* set into the form of a *well-ordered* set." Sets are to be well-ordered, and they and their cardinalities are to be gauged via the transfinite numbers of his structured conception of the infinite.

The transfinite numbers provided the framework for Cantor's two approaches to the continuum problem, one through cardinality and the other through definable sets of reals, these each to initiate vast research programs. As for the first, Cantor in the *Grundlagen* established results that reduced the continuum problem to showing that the continuum and the countable transfinite numbers have a bijection between them. However, despite several announcements Cantor could never develop a workable correlation, an emerging problem being that he could not *define* a well-ordering of the reals.

As for the approach through definable sets of reals, Cantor formulated the key concept of a *perfect* set of reals (non-empty, closed, and containing no isolated points), observed that perfect sets of reals *are* in bijective correspondence with the continuum, and showed that every closed set of reals is either countable or else have a perfect subset. Thus, Cantor showed that "CH holds for closed sets". The *perfect set property*, being either countable or else having a perfect subset, would become a focal property as more and more definable sets of reals came under purview.

Almost two decades after his initial 1873 result, Cantor in 1891 subsumed it through his celebrated *diagonal* argument. In logical terms this argument turns on the use of the validity $\neg\exists y \forall x(Pxx \longleftrightarrow \neg Pyx)$ for binary predicates P parametrizing unary predicates and became, of course, fundamental to the development of mathematical logic. Cantor stated his new, general result in terms of functions, ushering in totalities of arbitrary functions into mathematics, but his result is cast today in terms of the power set $P(x) = \{y \mid y \subseteq x\}$ of a set x: *For any set x, $P(x)$ has a larger cardinality than x.* Cantor had been extending his notion of set to a level of abstraction beyond sets of reals and the like; this new result showed for the first time that there is a set of a larger cardinality than that of the continuum.

Cantor's *Beiträge* of 1895 and 1897 presented his mature theory of the transfinite, incorporating his concepts of *ordinal number* and *cardinal number*. The former are the transfinite numbers now reconstrued as the "order-types" of well-orderings. As for the latter, Cantor defined the addition, multiplication, and exponentiation of cardinal numbers primordially in terms of set-theoretic operations and functions. Salient was the incorporation of "all" possibilities in the definition of exponentiation: If $\mathfrak{a}$ is the cardinal number of A and $\mathfrak{b}$ is the cardinal number of B then $\mathfrak{a}^{\mathfrak{b}}$ is the cardinal number of the totality, nowadays denoted ${}^{B}A$, of all functions from B into A. As befits the introduction of new numbers Cantor introduced a new notation, one using the Hebrew letter aleph, $\aleph$. $\aleph_0$ is to be the cardinal number of the natural numbers and the successive alephs

$$\aleph_0, \aleph_1, \aleph_2, \ldots, \aleph_\alpha, \ldots$$

indexed by the ordinal numbers are now to be the cardinal numbers of the successive number classes from the *Grundlagen* and thus to exhaust all the infinite cardinal numbers. Cantor pointed out that the exponentiated $2^{\aleph_0}$ is the cardinal number of the continuum, so that CH could now have been stated as

$$2^{\aleph_0} = \aleph_1.$$

However, with CH unresolved Cantor did not even mention the hypothesis in the *Grundlagen*, only in correspondence. Every well-ordered set has an aleph as its cardinal number, but where is $2^{\aleph_0}$ in the aleph sequence?

Cantor's great achievement, accomplished through almost three decades of prodigious effort, was to have brought into being the new subject of set theory as bolstered by the mathematical objectification of the actual infinite and moreover to have articulated a fundamental problem, the continuum problem. Hilbert made this the very first of his famous problems for the 20th Century, and he drew out Cantor's difficulty by suggesting the desirability of "actually giving" a well-ordering of the real numbers.

1.2. Zermelo

Ernst Zermelo (1871–1953), already estimable as an applied mathematician, turned to set theory at Göttingen under the influence of Hilbert. Zermelo analyzed Cantor's well-ordering principle by reducing it to the Axiom of Choice (AC), the abstract existence assertion that every set x has a *choice function*, i.e. a function f with domain x such that for every non-empty $y \in x$,

$f(y) \in y$. Zermelo's 1904 proof of the Well-Ordering Theorem, that with AC *every set can be well-ordered*, would anticipate the argument two decades later for transfinite recursion:

With x a set to be well-ordered, let f be a choice function on the power set $P(x)$. Call $y \subseteq x$ an *f-set* if there is a well-ordering R of y such that for any $a \in y$, $a = f(\{b \in x \mid b \text{ does not } R\text{-precede } a\})$. The well-orderings of f-sets are thus determined by f, and f-sets cohere. It follows that the union of f-sets is again an f-set and must in fact be x itself.

Zermelo's argument provoked open controversy because of its appeal to AC, and the subsequent tilting toward the acceptance of AC amounted to a conceptual shift in mathematics toward arbitrary functions and abstract existence principles. Responding to his critics Zermelo in 1908 published a second proof of the Well-Ordering Theorem and then the first full-fledged axiomatization of set theory, one similar in approach to Hilbert's axiomatization of geometry and incorporating set-theoretic ideas of Richard Dedekind. This axiomatization duly avoided the emerging "paradoxes" like Russell's Paradox, which Zermelo had come to independently, and served to buttress the Well-Ordering Theorem by making explicit its underlying set-existence assumptions. Zermelo's axioms, now formalized, constitute the familiar theory Z, *Zermelo set theory*:

Extensionality (sets are equal if they contain the same members), Empty Set (there is a set having no members), Pairs (for any sets x and y there is a set $\{x, y\}$ consisting exactly of x and y), Union (for any set x there is a set $\bigcup x$ consisting exactly of those sets that are members of some member of x), Power Set (for any set x there is a set $P(x)$ consisting exactly of the subsets of x), Choice (for any set x consisting of non-empty, pairwise disjoint sets, there is a set c such that every member of x has exactly one member in c), Infinity (there is a certain, specified infinite set); and Separation (for any set x and "definite" property P, there is a set consisting exactly of those members of x having the property P).

Extensionality, Empty Set, and Pairs lay the basis for sets. Infinity and Power Set ensure sufficiently rich settings for set-theoretic constructions. Power Set legitimizes "all" for subsets of a given set, and Separation legitimizes "all" for elements of a given set satisfying a property. Finally, Union and Choice (formulated reductively in terms of the existence of a "transversal" set meeting each of a family of sets in one member) complete the encasing of the Well-Ordering Theorem.

Zermelo's axiomatization sought to clarify vague subject matter, and like strangers in a strange land, stalwarts developed a familiarity with sets guided hand-in-hand by the axiomatic framework. Zermelo's own papers, with work of Dedekind as an antecedent, pioneered the reduction of mathematical concepts and arguments to set-theoretic concepts and arguments from axioms. Zermelo's analysis moreover served to draw out what would come to be generally regarded as set-theoretic and combinatorial out of the presumptively logical, with Infinity and Power Set salient and the process being strategically advanced by the segregation of the notion of property to Separation.

Taken together, Zermelo's work in the first decade of the 20th Century initiated a major transmutation of the notion of set after Cantor. With AC Zermelo shifted the notion away from Cantor's inherently well-ordered sets, and with his axiomatization Zermelo ushered in a new abstract, prescriptive view of sets as structured solely by membership and governed and generated by axioms. Through his set-theoretic reductionism Zermelo made evident how his set theory is adequate as a basis for mathematics.

1.3. First Developments

During this period Cantor's two main legacies, the extension of number into the transfinite and the investigation of definable sets of reals, became fully incorporated into mathematics in direct initiatives. The axiomatic tradition would be complemented by another, one that would draw its life more directly from the mathematics.

The French analysts Emile Borel, René Baire, and Henri Lebesgue took on the investigation of definable sets of reals in what would be a typically "constructive" approach. Cantor had established the perfect set property for closed sets and formulated the concept of *content* for a set of reals, but he did not pursue these matters. With these as antecedents the French work would lay the basis for measure theory as well as *descriptive set theory*, the definability theory of the continuum.

Borel, already in 1898, developed a theory of *measure* for sets of reals; the formulation was axiomatic, and at this early stage, bold and imaginative. The sets measurable according to his measure are the now well-known *Borel sets*. Starting with the open intervals (a, b) of reals assigned measure $b-a$, the Borel sets result when closing off under complements and countable unions, measures assigned in a corresponding manner.

Baire in his 1899 thesis classified those real functions obtainable by starting with the continuous functions and closing off under pointwise limits—the *Baire functions*—into classes indexed by the countable ordinal numbers, providing the first transfinite hierarchy after Cantor. Baire's thesis also introduced the now basic concept of *category*. A set of reals is *nowhere dense iff* its closure under limits includes no open set, and a set of reals is *meager* (or *of first category*) *iff* it is a countable union of nowhere dense sets—otherwise, it is *of second category*. Generalizing Cantor's 1873 argument, Baire established the Baire Category Theorem: *Every non-empty open set of reals is of second category.* His work also suggested a basic property: A set of reals A has the *Baire property iff* there is an open set O such that the symmetric difference $(A-O)\cup(O-A)$ is meager. Straightforward arguments show that every Borel set has the Baire property.

Lebesgue's 1902 thesis is fundamental for modern integration theory as the source of his concept of measurability. Lebesgue's concept of measurable set subsumed the Borel sets, and his analytic definition of measurable function subsumed the Baire functions. In simple terms, any *arbitrary* subset of a Borel measure zero set is a Lebesgue measure zero, or *null*, set, and a set is *Lebesgue measurable* if it is the union of a Borel set and a null set, in which case the measure assigned is that of the Borel set. It is this "completion" of Borel measure through the introduction of arbitrary subsets which gives Lebesgue measure its complexity and applicability and draws in wider issues of constructivity. Lebesgue's subsequent 1905 paper was the seminal paper of descriptive set theory: He correlated the Borel sets with the Baire functions, thereby providing a transfinite hierarchy for the Borel sets, and then applied Cantor's diagonalization argument to show both that this hierarchy is proper (new sets appear at each level) and that there is a Lebesgue measurable set which is not Borel.

As descriptive set theory was to develop, a major concern became the extent of the *regularity properties*, those indicative of well-behaved sets of reals, of which prominent examples were Lebesgue measurability, having the Baire property, and having the perfect set property. Significantly, the context was delimited by early explicit uses of AC in the role of providing a well-ordering of the reals: In 1905 Giuseppe Vitali established that there is a non-Lebesgue measurable set, and in 1908 Felix Bernstein established that there is a set without the perfect set property. Thus, Cantor's early contention that the reals are well-orderable precluded the universality of his own perfect set property, and it would be that his new, enumerative approach to the continuum would steadily provide focal examples and counterexamples.

The other, more primal Cantorian legacy, the extension of number into the transfinite, was considerably advanced by Felix Hausdorff, whose work was first to suggest the rich possibilities for a mathematical investigation of the uncountable. A mathematician *par excellence*, he took that sort of mathematical approach to set theory and extensional, set-theoretic approach to mathematics that would come to dominate in the years to come. In a 1908 paper, Hausdorff provided an elegant analysis of scattered linear orders (those having no dense sub-ordering) in a transfinite hierarchy. He first stated the Generalized Continuum Hypothesis (GCH)

$$2^{\aleph_\alpha} = \aleph_{\alpha+1} \quad \text{for every } \alpha.$$

He emphasized cofinality (the *cofinality* $\mathrm{cf}(\kappa)$ of a cardinal number κ is the least cardinal number λ such that a set of cardinality κ is a union of λ sets each of cardinality less than κ) and the distinction between *singular* ($\mathrm{cf}(\kappa) < \kappa$) and *regular* ($\mathrm{cf}(\kappa) = \kappa$) cardinals. And for the first time he broached a "large cardinal" concept, a regular limit cardinal $> \aleph_0$. Hausdorff's work around this time on sets of real functions ordered under eventual domination and having no uncountable "gaps" led to the first plausible mathematical proposition that entailed the denial of CH.

Hausdorff's 1914 text, *Grundzüge der Mengenlehre*, broke the ground for a generation of mathematicians in both set theory and topology. Early on, he defined an ordered pair of sets in terms of (unordered) pairs, formulated functions in terms of ordered pairs, and ordering relations as collections of ordered pairs. He in effect capped efforts of logicians by making these moves in mathematics, completing the set-theoretic reduction of relations and functions. He then presented Cantor's and Zermelo's work systematically, and of particular interest, he used a well-ordering of the reals to provide what is now known as Hausdorff's Paradox. The source of the later and better known Banach-Tarski Paradox, Hausdorff's Paradox provided an implausible decomposition of the sphere and was the first, and a dramatic, synthesis of classical mathematics and the new Zermelian abstract view.

A decade after Lebesgue's seminal 1905 paper, descriptive set theory came into being as a distinct discipline through the efforts of the Russian mathematician Nikolai Luzin. He had become acquainted with the work of the French analysts while in Paris as a student, and in Moscow he began a formative seminar, a major topic of which was the "descriptive theory of functions". The young Pole Wacław Sierpiński was an early participant while he was interned in Moscow in 1915, and undoubtedly this not only kindled the decade-long collaboration between Luzin and Sierpiński but also encouraged the latter's involvement in the development of a Polish school of mathematics and its interest in descriptive set theory. In an early success, Luzin's student Pavel Aleksandrov (and independently, Hausdorff) established the groundbreaking result that *the Borel sets have the perfect set property*, so that "CH holds for the Borel sets".

In the work that really began descriptive set theory, another student of Luzin's, Mikhail Suslin, investigated the *analytic sets* after finding a mistake in Lebesgue's paper. In a brief 1917 note Suslin formulated these sets in terms of an explicit operation $\mathcal{A}$ drawn from Aleksandrov's work and announced two fundamental results: *a set B of reals is Borel iff both B and its complement $\mathbb{R}-B$ are analytic;* and *there is an analytic set which is not Borel.* This was to be his sole publication, for he succumbed to typhus in a Moscow epidemic in 1919 at the age of 25. In an accompanying note Luzin announced that *every analytic set is Lebesgue measurable and has the perfect set property*, the latter result attributed to Suslin. Luzin and Sierpiński in joint papers soon provided proofs, in work that shifted the emphasis to the *co-analytic sets*, complements of analytic sets, and provided for them a basic *tree representation* based on well-foundedness (having no infinite branches) from which the main results of the period flowed.

After this first wave in descriptive set theory had crested, Luzin and Sierpiński in 1925 extended the domain of study to the *projective sets*. For $Y \subseteq \mathbb{R}^{k+1}$, the *projection of Y* is $pY = \{\langle x_1, \ldots, x_k\rangle \mid \exists y(\langle x_1, \ldots, x_k, y\rangle \in Y)\}$. Suslin had essentially noted that *a set of reals is analytic iff it is the projection of a Borel subset of* $\mathbb{R}^2$. Luzin and Sierpiński took the geometric operation of projection to be basic and defined the projective sets as those sets obtainable from the Borel sets by the iterated applications of projection and complementation. The corresponding hierarchy of projective subsets of

$\mathbb{R}^k$ is defined, in modern notation, as follows: For $A \subseteq \mathbb{R}^k$,

$$A \text{ is } \mathbf{\Sigma}^1_1 \quad \textit{iff} \quad A = pY \text{ for some Borel set } Y \subseteq \mathbb{R}^{k+1},$$

A is *analytic* as for $k = 1$, and for $n > 0$,

$$\begin{aligned} A \text{ is } \mathbf{\Pi}^1_n \quad &\textit{iff} \quad \mathbb{R}^k - A \text{ is } \mathbf{\Sigma}^1_n, \\ A \text{ is } \mathbf{\Sigma}^1_{n+1} \quad &\textit{iff} \quad A = pY \text{ for some } \mathbf{\Pi}^1_n \text{ set } Y \subseteq \mathbb{R}^{k+1}, \text{ and} \\ A \text{ is } \mathbf{\Delta}^1_n \quad &\textit{iff} \quad A \text{ is both } \mathbf{\Sigma}^1_n \text{ and } \mathbf{\Pi}^1_n. \end{aligned}$$

($\mathbf{\Sigma}^1_n$ is also written $\underset{\sim}{\Sigma}{}^1_n$; $\mathbf{\Pi}^1_n$ is also written $\underset{\sim}{\Pi}{}^1_n$; and $\mathbf{\Delta}^1_n$ is also written $\underset{\sim}{\Delta}{}^1_n$. One can formulate these concepts with continuous images instead of projections, e.g. A is $\mathbf{\Sigma}^1_{n+1}$ iff A is the continuous image of some $\mathbf{\Pi}^1_n$ set $Y \subseteq \mathbb{R}$. If the basics of continuous functions are in hand, this obviates the need to have different spaces.)

Luzin and Sierpiński recast Lebesgue's use of the Cantor diagonal argument to show that the projective hierarchy is proper, and soon its basic properties were established. However, this investigation encountered obstacles from the beginning. Whether the $\mathbf{\Pi}^1_1$ subsets of $\mathbb{R}$, the co-analytic sets at the bottom of the hierarchy, have the perfect set property and whether the $\mathbf{\Sigma}^1_2$ sets are Lebesgue measurable remained unknown. Besides the regularity properties, the properties of *separation, reduction,* and especially *uniformization* relating sets to others were studied, but there were accomplishments only at the first projective level. The one eventual success and a culminating result of the early period was the Japanese mathematician Motokiti Kondô's 1937 result, the $\mathbf{\Pi}^1_1$ Uniformization Theorem: *Every* $\mathbf{\Pi}^1_1$ *relation can be uniformized by a* $\mathbf{\Pi}^1_1$ *function.* This impasse with respect to the regularity properties would be clarified, surprisingly, by penetrating work of Gödel involving metamathematical methods.

In modern set theory, what has come to be taken for the "reals" is actually *Baire space*, the set of functions from the natural numbers into the natural numbers (with the product topology). Baire space, the "fundamental domain" of a 1930 Luzin monograph, is homeomorphic to the irrational reals and so equivalent for all purposes having to do measure, category, and perfect sets. Already by then it had become evident that a set-theoretic study of the continuum is best cast in terms of Baire space, with geometric intuitions being augmented by combinatorial ones.

During this period AC and CH were explored by the new Polish school, most notably by Sierpiński, Alfred Tarski, and Kazimierz Kuratowski, no longer as underlying axiom and primordial hypothesis but as part of ongoing mathematics. Sierpiński's own earliest publications, culminating in a 1918 survey, not only dealt with specific constructions but also showed how deeply embedded AC was in the informal development of cardinality, measure, and the Borel hierarchy. Even more than AC, Sierpiński investigated CH, and summed up his researches in a 1934 monograph. It became evident how having not only a well-ordering of the reals but one as given by CH whose initial segments are countable led to striking, often initially counter-intuitive, examples in analysis and topology.

1.4. Replacement and Foundation

In the 1920s, fresh initiatives in axiomatics structured the loose Zermelian framework with new features and corresponding axioms, the most consequential moves made by John von Neumann (1903–1957) in his doctoral work, with anticipations by Dmitry Mirimanoff in an informal setting. Von Neumann effected a Counter-Reformation of sorts that led to the incorporation of a new axiom, the Axiom of Replacement: *For any set x and property $P(v,w)$ functional on x (i.e. for any $a \in x$ there is exactly one b such that $P(a,b)$), $\{b \mid P(a,b)$ for some $a \in x\}$ is a set.* The transfinite numbers had been central for Cantor but peripheral to Zermelo; von Neumann reconstrued them as *bona fide* sets, the ordinals, and established their efficacy by formalizing transfinite recursion, the method for defining sets in terms of previously defined sets applied with transfinite indexing.

Ordinals manifest the basic idea of taking precedence in a well-ordering simply to be membership. A set x is *transitive iff* $\bigcup x \subseteq x$, so that x is "closed" under membership, and x is an *ordinal iff* x is transitive and well-ordered by $\in$. Von Neumann, as had Mirimanoff before him, established the key instrumental property of Cantor's ordinal numbers for ordinals: *Every well-ordered set is order-isomorphic to exactly one ordinal with membership.* Von Neumann took the further step of ascribing to the ordinals the role of Cantor's ordinal numbers. To establish the basic ordinal arithmetic results that affirm this role, von Neumann saw the need to establish the Transfinite Recursion Theorem, the theorem that validates definitions by transfinite recursion. The proof was anticipated by the Zermelo 1904 proof, but Replacement was necessary even for the very formulation, let alone the proof, of the theorem. Abraham Fraenkel and Thoralf Skolem had independently proposed Replacement to ensure that a specific collection resulting from a simple recursion be a set, but it was von Neumann's formal incorporation of transfinite recursion as method which brought Replacement into set theory. With the ordinals in place von Neumann completed the restoration of the Cantorian transfinite by defining the *cardinals* as the *initial ordinals*, i.e. those ordinals not in bijective correspondence with any of its predecessors. The infinite initial ordinals are now denoted

$$\omega = \omega_0, \omega_1, \omega_2, \ldots, \omega_\alpha, \ldots,$$

so that ω is to be the set of natural numbers in the ordinal construal. It would henceforth be that we take

$$\omega_\alpha = \aleph_\alpha$$

conflating extension with intension, with the left being a von Neumann ordinal and the right being the Cantorian cardinal concept. Every infinite set x, with AC, is well-orderable and hence in bijective correspondence with a unique initial ordinal ω_α, and the cardinality of x is $|x| = \aleph_\alpha$. It has become customary to use the lower case Greek letters to denote ordinals; $\alpha < \beta$ to denote $\alpha \in \beta$ construed as ordering; On to denote the ordinals; and the middle letters $\kappa, \lambda, \mu, \ldots$ to denote the initial ordinals in their role as the infinite cardinals, with κ^+ denoting the cardinal successor of κ.

Von Neumann provided a new axiomatization of set theory, one that first incorporated what we now call proper classes. A *class* is the totality of all sets that satisfy a specified property, so that membership in the class amounts to satisfying the property, and von Neumann axiomatized the ways to have these properties. Only sets can be members, and so the recourse to possibly proper classes, classes not represented by sets, avoids the contradictions arising from formalizing the known paradoxes. Actually, von Neumann took functions to be primitive in an involved framework, and Paul Bernays in 1930 re-constituted the von Neumann axiomatization with sets and classes as primitive. Classes would not remain a formalized component of modern set theory, but the informal use of classes as objectifications of properties would become increasingly liberal, particularly to convey large-scale issues in set theory.

Von Neumann (and before him Mirimanoff, Fraenkel, and Skolem) also considered the salutary effects of restricting the universe of sets to the *well-founded sets*. The well-founded sets are the sets in the class $\bigcup_\alpha V_\alpha$, where the "ranks" V_α are defined by transfinite recursion:

$$V_0 = \emptyset; \qquad V_{\alpha+1} = P(V_\alpha); \quad \text{and} \quad V_\delta = \textstyle\bigcup_{\alpha<\delta} V_\alpha \quad \text{for limit ordinals } \delta.$$

Von Neumann entertained the Axiom of Foundation: *Every nonempty set x has an $\in$-minimal element, i.e. a $y \in x$ such that $x \cap y$ is empty.* (With AC this is equivalent to having no infinite $\in$-descending sequences.) This axiom amounts to the assertion that the *cumulative hierarchy* exhausts the universe V of sets:

$$V = \textstyle\bigcup_\alpha V_\alpha.$$

In modern terms, the ascribed well-foundedness of $\in$ leads to a ranking function $\rho : V \to \mathrm{On}$ defined recursively by $\rho(x) = \bigcup\{\rho(y) + 1 \mid y \in x\}$, so that $V_\alpha = \{x \mid \rho(x) < \alpha\}$, and one can establish results for all sets by induction on rank.

Zermelo in a 1930 paper offered his final axiomatization of set theory as well as a striking, synthetic view of a procession of models that would have a modern resonance. Proceeding in what we would now call a second-order context, Zermelo amended his 1908 axiomatization Z by adjoining both Replacement and Foundation while leaving out Infinity and AC, the latter being regarded as part of the underlying logic. The now standard axiomatization of set theory

ZFC, *Zermelo-Fraenkel with Choice*,

is recognizable if we inject Infinity and AC, the main difference being that ZFC is a first-order theory (as discussed below). "Fraenkel" acknowledges the early suggestion by Fraenkel to adjoin Replacement; and the Axiom of Choice is explicitly mentioned.

ZF, *Zermelo-Fraenkel*,

is ZFC without AC and is a base theory for the investigation of weak Choice-type propositions as well as propositions that contradict AC.

Zermelo herewith completed his transmutation of the notion of set, his abstract view stabilized by further axioms that structured the universe of sets. Replacement and Foundation focused the notion of set, with the first providing the means for transfinite recursion and induction and the second making possible the application of those means to get results about *all* sets, they appearing in the cumulative hierarchy. Foundation is the one axiom unnecessary for the recasting of mathematics in set-theoretic terms, but the axiom is also the salient feature that distinguishes investigations specific to set theory as a field of mathematics. With Replacement and Foundation in place Zermelo was able to provide natural models of his axioms, each a V_κ where κ is an *inaccessible* cardinal (regular and *strong limit*: if $\lambda < \kappa$, then $2^\lambda < \kappa$), and to establish algebraic isomorphism, initial segment, and embedding results for his models. Finally, Zermelo posited an endless procession of such models, each a set in the next, as natural extensions of their cumulative hierarchies.

Inaccessible cardinals are at the modest beginnings of the theory of *large cardinals*, now a mainstream of modern set theory devoted to the investigation of strong hypotheses and consistency strength. The journal volume containing Zermelo's paper also contained Stanisław Ulam's seminal paper on measurable cardinals, which would become focal among large cardinals. In modern terminology, a *filter over* a set Z is a family of subsets of Z closed under the taking of supersets and of intersections. (Usually excluded from consideration as trivial are $\{X \subseteq Z \mid A \subseteq X\}$ for some set $A \subseteq Z$, the *principal* filters.) An *ultrafilter* U *over* Z is a maximal filter over Z, i.e. for any $X \subseteq Z$, either $X \in U$ or else $Z - X \in U$. For a cardinal λ, a filter is λ-*complete* if it is closed under the taking of intersections of fewer than λ members. Finally, an uncountable cardinal κ is *measurable iff* there is a κ-complete ultrafilter over κ. In a previous, 1929 note Ulam had constructed, using a well-ordering of the reals, an ultrafilter over ω. Measurability thus generalizes a property of ω, and Ulam showed moreover that *measurable cardinals are inaccessible.* In this work, Ulam was motivated by measure-theoretic considerations, and he viewed his work as about $\{0,1\}$-valued measures, the measure 1 sets being the sets in the ultrafilter. To this day, ultrafilters of all sorts in large cardinal theory are also called measures.

A decade later Tarski provided a systematic development of these concepts in terms of ideals. An *ideal over* a set Z is a family of subsets of Z closed under the taking of subsets and of unions. This is the "dual" notion to filters; if I is an ideal (resp. filter) over Z, then $\breve{I} = \{Z - X \mid X \in I\}$ is its dual filter (resp. ideal). An ideal is λ-*complete* if its dual filter is. A more familiar conceptualization in mathematics, Tarski investigated a general notion of ideal on a Boolean algebra in place of the power set algebra $P(Z)$. Although filters and ideals in large cardinal theory are most often said to be *on* a cardinal κ, they are more properly *on* the Boolean algebra $P(\kappa)$. Moreover, the measure-theoretic terminology has persisted: For an ideal $I \subseteq P(Z)$, the I-*measure zero* (negligible) sets are the members of I, the I-*positive measure* (non-negligible) sets are the members of $P(Z) - I$, and the I-*measure one* (all but negligible) sets are the members of the dual filter $\{Z - X \mid X \in I\}$.

Returning to the axiomatic tradition, Zermelo's 1930 paper was in part a response to Skolem's advocacy of the idea of framing Zermelo's 1908 axioms in *first-order logic*, the logic of formal languages based on the quantifiers $\forall$ and $\exists$ interpreted as ranging over the *elements* of a domain of discourse. First-order logic had emerged in 1917 lectures of Hilbert as a delimited system of logic amenable to mathematical investigation. Entering from a different, algebraic tradition, Skolem in 1920 had established a seminal result for semantic methods with the Löwenheim-Skolem Theorem, that a countable collection of first-order sentences, if satisfiable, is satisfiable in a countable domain. For this he introduced what we now call Skolem functions, functions added formally for witnessing $\exists x$ assertions. For set theory Skolem in 1923 proposed formalizing Zermelo's axioms in the first-order language with $\in$ and $=$ as binary predicate symbols. Zermelo's "definite" properties were to be those expressible in this first-order language in terms of given sets, and the Axiom of Separation was to become a *schema* of axioms, one for each first-order formula. As an argument against taking set theory as a foundation for mathematics, Skolem pointed out what has come to be called *Skolem's Paradox*: Zermelo's 1908 axioms cast in first-order logic is a countable collection of sentences, and so if they are satisfiable at all, they are satisfiable in a countable domain. Thus, we have the paradoxical existence of countable models for Zermelo's axioms although they entail the existence of uncountable sets. Zermelo found this antithetical and repugnant. However, strong currents were at work leading to a further, subtler transmutation of the notion of set as based on first-order logic and incorporating its relativism of set-theoretic concepts.

2. New Groundwork

2.1. Gödel

Kurt Gödel (1906–1978) substantially advanced the mathematization of logic by submerging metamathematical methods into mathematics. The main vehicle was the direct coding, "the arithmetization of syntax", in his celebrated 1931 Incompleteness Theorem, which worked dialectically against a program of Hilbert's for establishing the consistency of classical mathematics. But starting an undercurrent, the earlier 1930 Completeness Theorem for first-order logic clarified the distinction between the formal syntax and semantics of first-order logic and secured its key instrumental property with the Compactness Theorem.

Tarski in the early 1930s provided his systematic "definition of truth", exercising philosophers to a surprising extent ever since. Tarski simply schematized truth as a correspondence between formulas of a formal language and set-theoretic assertions about an intended structure interpreting the language and provided a recursive definition of the *satisfaction* relation, when a formula holds in the structure, in set-theoretic terms. The eventual effect of Tarski's mathematical formulation of semantics would be not only to make mathematics out of the informal notion of satisfiability, but also to enrich ongoing mathematics with a systematic method for forming mathematical analogues of several intuitive semantic notions. Tarski would only be explicit much later about satisfaction-in-a-structure for arbitrary structures, this leading to his notion of logical consequence. For coming purposes, the following affirms notation and concepts in connection with Tarski's definition.

For a first-order language, a structure N interpreting that language (i.e. a specification of a domain of discourse as well as interpretations of the function and predicate symbols), a formula $\varphi(v_1, v_2, \ldots, v_n)$ of the language with the (free) variables as displayed, and $a_1, a_2, \ldots, a_n$ in the domain of N,

$$N \models \varphi[a_1, a_2, \ldots, a_n]$$

asserts that the formula φ is satisfied in N according to Tarski's recursive definition when v_i is interpreted as a_i. A subset y of the domain of N is *first-order definable over N iff* there is a $\psi(v_1, v_2, \ldots, v_{n+1})$ and $a_1, a_2, \ldots, a_n$ in the domain of N such that

$$y = \{z \in N \mid N \models \psi[a_1, a_2, \ldots, a_n, z]\}.$$

(The first-order definability of k-ary relations is analogously formulated with v_{n+1} replaced by k variables.)

Through Tarski's recursive definition and an "arithmetization of syntax" whereby formulas are systematically coded by natural numbers, the satisfaction relation $N \models \varphi[a_1, a_2, \ldots, a_n]$ for *sets* N is definable in set theory. On the other hand, by Tarski's result on the "undefinability of truth", the satisfaction relation for V itself is not first-order definable over V.

Set theory was launched as a distinctive field of mathematics by Gödel's construction of the class L leading to the relative consistency of the Axiom of Choice and the Generalized Continuum Hypothesis. In a brief 1939 account Gödel informally presented L essentially as is done today: For any set x let $\mathrm{def}(x)$ denote the collection of subsets of x first-order definable over the structure $\langle x, \in\rangle$ with domain x and the membership relation restricted to it.

Then define:

$$L_0 = \emptyset; \qquad L_{\alpha+1} = \mathrm{def}(L_\alpha), \qquad L_\delta = \bigcup\{L_\alpha \mid \alpha < \delta\} \quad \text{for limit ordinals } \delta;$$

and the *constructible universe*

$$L = \bigcup_\alpha L_\alpha.$$

Gödel pointed out that L "can be defined and its theory developed in the formal systems of set theory themselves". This is actually the central feature of the construction of L. L is definable in ZF via transfinite recursion based on the formalizability of def(x), which was reaffirmed by Tarski's definition of satisfaction. With this, one can formalize the Axiom of Constructibility $V = L$, i.e. $\forall x(x \in L)$. To set a larger context, we affirm the following for a class X: for a set-theoretic formula φ, φ^X denotes φ with its quantifiers restricted to X and this extends to set-theoretic terms t (like $\bigcup x$, $P(x)$, and so forth) through their definitions to yield t^X. X is an *inner model iff* X is a transitive class containing all the ordinals such that φ^X is a theorem of ZF for every axiom φ of ZF. What Gödel did was to show in ZF that L is an inner model which satisfies AC and GCH. He thus established a relative consistency which can be formalized as an assertion: Con(ZF) implies Con(ZFC + GCH).

In the approach via def(x) it is necessary to show that def(x) remains unaltered when applied in L with quantifiers restricted to L. Gödel himself would never establish this *absoluteness of first-order definability* explicitly. In a 1940 monograph, Gödel worked in Bernays' class-set theory and used eight binary operations producing new classes from old to generate L set by set via transfinite recursion. This veritable "Gödel numbering" with ordinals eschewed def(x) and made evident certain aspects of L. Since there is a direct, definable well-ordering of L, choice functions abound in L, and AC holds there. Of the other axioms the crux is where first-order logic impinges, in Separation and Replacement. For this, "algebraic" closure under Gödel's eight operations ensured "logical" Separation for *bounded* formulas, formulas having only quantifiers expressible in terms of $\forall v \in w$, and then the full exercise of Replacement (in V) secured all of the ZF axioms in L.

Gödel's proof that L satisfies GCH consisted of two separate parts. He established the implication $V = L \rightarrow$ GCH, and, in order to apply this implication within L, that $(V = L)^L$. This latter follows from the aforementioned absoluteness of def(x), and in his monograph Gödel gave an alternate proof based on the absoluteness of his eight binary operations.

Gödel's argument for $V = L \rightarrow$ GCH rests, as he himself wrote in his 1939 note, on "a generalization of Skolem's method for constructing enumerable models". This was the first significant use of Skolem functions since Skolem's own to establish the Löwenheim-Skolem theorem, and with it, Skolem's Paradox. Ironically, though Skolem sought through his paradox to discredit set theory based on first-order logic as a foundation for mathematics, Gödel turned paradox into method, one promoting first-order logic. Gödel specifically established his "Fundamental Theorem":

> For infinite γ, every constructible subset of L_γ
>
> belongs to some L_β for a β of the same cardinality as γ.

For infinite α, L_α has the same cardinality as that of α. It follows from the Fundamental Theorem that in the sense of L, the power set of L_{ω_α} is included in $L_{\omega_{\alpha+1}}$, and so GCH follows in L.

The work with L led, further, to the resolution of difficulties in descriptive set theory. Gödel announced, in modern terms: *If* $V = L$, *then* (a) *there is a* $\mathbf{\Delta}^1_2$ *set of reals that is not Lebesgue measurable, and* (b) *there is a* $\mathbf{\Pi}^1_1$ *set of reals without the perfect set property.* Thus, the early descriptive set theorists were confronting an obstacle insurmountable in ZFC! When eventually confirmed and refined, the results were seen to turn on a "good" $\mathbf{\Sigma}^1_2$ well-ordering of the reals in L defined via reals coding well-founded structures and thus connected to the well-founded tree representation of $\mathbf{\Pi}^1_1$ sets. Gödel's results (a) and (b) constitute the first real synthesis of abstract and descriptive set theory, in that the axiomatic framework is incorporated into the investigation of definable sets of reals.

Gödel brought into set theory a method of construction and of argument which affirmed several features of its axiomatic presentation. Most prominently, he showed how first-order definability can be formalized and used to achieve strikingly new mathematical results. This significantly contributed to a lasting ascendancy for first-order logic which, in addition to its sufficiency as a logical framework for mathematics, was seen to have considerable operational efficacy. Moreover, Gödel's work buttressed the incorporation of Replacement and Foundation into set theory, the first immanent in the transfinite recursion and arbitrary extent of the ordinals, and the second as underlying the basic cumulative hierarchy picture that anchors L.

In later years Gödel speculated about the possibility of deciding propositions like CH with large cardinal hypotheses based on the heuristics of *reflection*, and later, *generalization*. In a 1946 address he suggested the consideration of "stronger and stronger axioms of infinity" and reflection down from V: "Any proof of a set-theoretic theorem in the next higher system above set theory (i.e. any proof involving the concept of truth, etc.) is replaceable by a proof from such an axiom of infinity". In a 1947 expository article on the continuum problem Gödel presumed that CH would be shown independent from ZF and speculated more concretely about possibilities with large cardinals. He argued that the axioms of set theory do not "form a system closed in itself" and so the "very concept of set on which they are based suggests their extension by new axioms that assert the existence of still further iterations of the operation of 'set of' ". In an unpublished footnote toward a 1966 revision of the article, Gödel acknowledged "extremely strong axioms of infinity of an entirely new kind", generalizations of properties of ω "supported by strong arguments from analogy". These heuristics would surface anew in the 1960s, when the theory of large cardinals developed a self-fueling momentum of its own, stimulated by the emergence of forcing and inner models.

2.2. Infinite Combinatorics

For decades Gödel's construction of L stood as an isolated monument in the axiomatic tradition, and his methodological advances would only become fully assimilated after the infusion of model-theoretic techniques in the 1950s. In the mean time, the direct investigation of the transfinite as extension of number was advanced, gingerly at first, by the emergence of *infinite combinatorics.*

The 1934 Sierpiński monograph on CH (discussed earlier) having considerably elaborated its consequences, a new angle in the combinatorial investigation of the continuum was soon broached. Hausdorff in 1936 reactivated his early work on gaps in the orderings of functions to show that the reals can be partitioned into $\aleph_1$ Borel sets, answering an early question of Sierpiński. Hausdorff had newly cast his work in terms of functions from ω to ω, the members of Baire space or the "reals", under the ordering of *eventual dominance*: $f \leq^* g$ if $f(n) \leq g(n)$ for all but finitely many $n \in \omega$. Work on this structure and definable sets of reals in the 1930s, and particularly of Fritz Rothberger through the 1940s, isolated what is now called the *dominating number* $\mathfrak{d}$, the least cardinality of a subset of Baire space cofinal in $\leq^*$. $\aleph_1 \leq \mathfrak{d} \leq 2^{\aleph_0}$, but absent CH $\mathfrak{d}$ assumed an independent significance as a pivotal cardinal. Rothberger established incisive results which we now cast as about the relationships to other pivotal cardinals, results which provided new understandings about the structure of the continuum but would become vacuous with the blanket assumption of CH. The investigation of $\mathfrak{d}$ and other "cardinal characteristics (or invariants) of the continuum" would blossom with the advent of forcing.

Taking up another thread, Frank Ramsey in 1930, addressing a problem of formal logic, established a generalization of the pigeonhole principle for finite sets, and in a move transcending purpose and context he also established an infinite version implicitly applying the now familiar Kőnig's Lemma for trees. In modern terms, for ordinals α, β, and δ and $n \in \omega$ the *partition relation*

$$\beta \longrightarrow (\alpha)^n_\delta$$

asserts that for any partition $f : [\beta]^n \to \delta$ of the n-element subsets of β into δ cells, there is an $H \subseteq \beta$ of order type α *homogeneous* for the partition, i.e. all the n-element subsets of H lie in the same cell. Ramsey's theorem for finite sets is: *For any* $n, k, i \in \omega$ *there is an* $r \in \omega$ *such that* $r \longrightarrow (k)^n_i$. The "Ramsey numbers", the least possible r's for various n, k, i, are unknown except in a few basic cases. The (infinite) Ramsey's Theorem is: $\omega \longrightarrow (\omega)^n_i$ *for every* $n, i \in \omega$.

A *tree* is a partially ordered set T such that the predecessors of any element are well-ordered. The *αth level* of T consists of those elements whose predecessors have order-type α, and the *height* of T is the least α such that the αth level of T is empty. A *chain* of T is a linearly ordered subset, and an *antichain* is a subset consisting of pairwise incompatible elements. A *cofinal branch* of T is a chain with elements at every non-empty level of T.