SET THEORY FROM CANTOR TO COHEN

Akihiro Kanamori

Set theory is an autonomous and sophisticated field of mathematics, enormously successful not only at its continuing development of its historical heritage but also at analyzing mathematical propositions and gauging their consistency strength. But set theory is also distinguished by having begun intertwined with pronounced metaphysical attitudes, and these have even been regarded as crucial by some of its great developers. This has encouraged the exaggeration of crises in foundations and of metaphysical doctrines in general. However, set theory has proceeded in the opposite direction, from a web of intensions to a theory of extension *par excellence*, and like other fields of mathematics its vitality and progress have depended on a steadily growing core of mathematical proofs and methods, problems and results. There is also the stronger contention that from the beginning set theory actually developed through a progression of *mathematical* moves, whatever and sometimes in spite of what has been claimed on its behalf.

What follows is an account of the development of set theory from its beginnings through the creation of forcing based on these contentions, with an avowedly Whiggish emphasis on the heritage that has been retained and developed by the current theory. The whole transfinite landscape can be viewed as having been articulated by Cantor in significant part to solve the Continuum Problem. Zermelo's axioms can be construed as clarifying the set existence commitments of a single proof, of his Well-Ordering Theorem. Set theory is a particular case of a field of mathematics in which seminal proofs and pivotal problems actually shaped the basic concepts and forged axiomatizations, these transmuting the very notion of set. There were two main junctures, the first being when Zermelo through his axiomatization shifted the notion of set from Cantor's range of inherently structured sets to sets solely structured by membership and governed and generated by axioms. The second juncture was when the Replacement and Foundation Axioms were adjoined and a first-order setting was established; thus transfinite recursion was incorporated and results about all sets could established through these means, including results about definability and inner models. With the emergence of the cumulative hierarchy picture, set theory can be regarded as becoming a theory of well-foundedness, later to expand to a study of consistency strength. Throughout, the subject has not only been sustained by the axiomatic tradition through Gödel and Cohen but also fueled by Cantor's two legacies, the extension of number into the transfinite as transmuted into the theory of large cardinals and the investigation of definable sets of reals as transmuted into descriptive set theory. All this can be regarded as having a historical and mathematical logic internal to set theory, one that is often misrepresented at critical junctures in textbooks (as will be pointed out). This view, from inside set theory and about itself, serves to shift the focus to

those tensions and strategies familiar to mathematicians as well as to those moves, often made without much fanfare and sometimes merely linguistic, that have led to the crucial advances.

1   CANTOR

1.1   Real Numbers and Countability

Set theory had its beginnings in the great 19th Century transformation of mathematics, a transformation beginning in analysis. Since the creation of the calculus by Newton and Leibniz the function concept had been steadily extended from analytic expressions toward arbitrary correspondences. The first major expansion had been inspired by the explorations of Euler in the 18th Century and featured the infusion of infinite series methods and the analysis of physical phenomena, like the vibrating string. In the 19th Century the stress brought on by the unbridled use of series of functions led first Cauchy and then Weierstrass to articulate convergence and continuity. With infinitesimals replaced by the limit concept and that cast in the ϵ-δ language, a level of deductive rigor was incorporated into mathematics that had been absent for two millenia. Sense for the new functions given in terms of infinite series could only be developed through carefully specified deductive procedures, and proof reemerged as an extension of algebraic calculation and became basic to mathematics in general, promoting new abstractions and generalizations.

Working out of this tradition Georg Cantor[1] (1845–1918) in 1870 established a basic uniqueness theorem for trigonometric series: If such a series converges to zero everywhere, then all of its coefficients are zero. To generalize Cantor [1872] started to allow points at which convergence fails, getting to the following formulation: For a collection P of real numbers, let P' be the collection of limit points of P, and $P^{(n)}$ the result of n iterations of this operation. *If a trigonometric series converges to zero everywhere except on a P where $P^{(n)}$ is empty for some n, then all of its coefficients are zero.*[2]

It was in [1872] that Cantor provided his formulation of the real numbers in terms of fundamental sequences of rational numbers, and significantly, this was for the specific purpose of articulating his proof. With the new results of analysis to be secured by proof and proof in turn to be based on prior principles the regress led in the early 1870s to the appearance of several independent formulations of the real numbers in terms of the rational numbers. It is at first quite striking that the real numbers came to be developed so late, but this can be viewed as part of the expansion of the function concept which shifted the emphasis from the continuum taken as a whole to its extensional construal as a collection of objects. In mathematics objects have been traditionally introduced only with reluctance, but a more arithmetical rather than geometrical approach to the continuum became necessary for the articulation of proofs.

The other well-known formulation of the real numbers is due to Richard Dedekind [1872], through his cuts. Cantor and Dedekind maintained a fruitful correspondence,

---

[1] Dauben [1979], Meschkowski [1983], and Purkert-Ilgauds [1987] are mathematical biographies of Cantor.

[2] See Kechris-Louveau [1987] for recent developments in the Cantorian spirit about uniqueness for trigonometric series converging on definable sets of reals.

especially during the 1870s, in which Cantor aired many of his results and speculations.[3] The formulations of the real numbers advanced three important predispositions for set theory: the consideration of infinite collections, their construal as unitary objects, and the encompassing of arbitrary such possibilities. Dedekind [1871] had in fact made these moves in his creation of ideals, infinite collections of algebraic numbers,[4] and there is an evident similarity between ideals and cuts in the creation of new numbers out of old.[5] The algebraic numbers would soon be the focus of a major breakthrough by Cantor. Although both Cantor and Dedekind carried out an arithmetical reduction of the continuum, they each accommodated its antecedent geometric sense by asserting that each of their real numbers actually corresponds to a point on the line. Neither theft nor honest toil sufficed; Cantor [1872: 128] and Dedekind [1872: III] recognized the need for an *axiom* to this effect, a sort of Church's Thesis of adequacy for the new construal of the continuum as a collection of objects.

Cantor recalled[6] that around this time he was already considering infinite iterations of his P' operation using "symbols of infinity":

$$P^{(\infty)} = \bigcap_n^{\infty} P^{(n)}, P^{(\infty+1)} = P^{(\infty)\prime}, P^{(\infty+2)}, \ldots P^{(\infty \cdot 2)}, \ldots P^{(\infty^2)}, \ldots P^{(\infty^\infty)}, \ldots P^{(\infty^{\infty^\infty})}, \ldots$$

In a crucial conceptual move he began to investigate infinite collections of real numbers and infinitary enumerations for their own sake, and this led first to a basic articulation of size for the continuum and then to a new, encompassing theory of counting. Set theory was born on that December 1873 day when Cantor established that *the real numbers are uncountable*.[7] In the next decades the subject was to blossom through the prodigious progress made by him in the theory of transfinite and cardinal numbers.

The uncountability of the reals was established, of course, via *reductio ad absurdum* as with the irrationality of $\sqrt{2}$. Both impossibility results epitomize how a *reductio* can compel a larger mathematical context allowing for the deniability of hitherto implicit properties. Be that as it may, Cantor the mathematician addressed a specific *problem*, embedded in the mathematics of the time, in his seminal [1874] entitled "On a property of the totality of all real algebraic numbers". After first establishing this property, the countability of the algebraic numbers, Cantor then established: *For any (countable) sequence*

---

[3]The most complete edition of Cantor's correspondence is Meschkowski-Nilson [1991]. Excerpts from the Cantor-Dedekind correspondence from 1872 through 1882 were published in Noether-Cavaillès [1937], and excerpts from the 1899 correspondence were published by Zermelo in the collected works of Cantor [1932]. English translations of the Noether-Cavaillès excerpts were published in Ewald [1996: 843ff.]. An English translation of a Zermelo excerpt (retaining his several errors of transcription) appeared in van Heijenoort [1967: 113ff.]. English translations of Cantor's 1899 correspondence with both Dedekind and Hilbert were published in Ewald [1996: 926ff.].

[4]The algebraic numbers are those real numbers that are the roots of polynomials with integer coefficients.

[5]Dedekind [1872] dated his conception of cuts to 1858, and antecedents to ideals in his work were also entertained around then. For Dedekind and the foundation of mathematics see Dugac [1976] and Ferreirós [2007], who both accord him a crucial role in the development of the framework of set theory.

[6]See his [1880: 358].

[7]A set is *countable* if there is a bijective correspondence between it and the natural numbers $\{0, 1, 2, \ldots\}$. The exact date of birth can be ascertained as December 7. Cantor first gave a proof of the uncountability of the reals in a letter to Dedekind of 7 December 1873 (Ewald [1996: 845ff]), professing that ". . . only today do I believe myself to have finished with the thing . . .".

of reals, every interval contains a real not in the sequence. Cantor appealed to the order completeness of the reals:

Suppose that *s* is a sequence of reals and *I* an interval. Let $a < b$ be the first two reals of *s*, if any, in *I*. Then let $a' < b'$ be the first two reals of *s*, if any, in the open interval (a, b); $a'' < b''$ the first two reals of *s*, if any, in (a', b'); and so forth. Then however long this process continues, the (non-empty) intersection of these nested intervals cannot contain any member of *s*.

By this means Cantor provided a new proof of Joseph Liouville's result [1844, 1851] that there are transcendental numbers (real non-algebraic numbers) and only afterward did Cantor point out the uncountability of the reals altogether. This presentation is suggestive of Cantor's natural caution in overstepping mathematical sense at the time.[8]

Accounts of Cantor's work have mostly reversed the order for deducing the existence of transcendental numbers, establishing first the uncountability of the reals and only then drawing the existence conclusion from the countability of the algebraic numbers.[9] In textbooks the inversion may be inevitable, but this has promoted the misconception that Cantor's arguments are non-constructive.[10] It depends how one takes a proof, and Cantor's arguments have been implemented as algorithms to generate the successive digits of new reals.[11]

1.2   Continuum Hypothesis and Transfinite Numbers

By his next publication [1878] Cantor had shifted the weight to getting bijective correspondences, stipulating that two sets have the same *power* [Mächtigkeit] *iff* there is such a correspondence between them, and established that the reals $\mathbb{R}$ and the *n*-dimensional spaces $\mathbb{R}^n$ all have the same power. Having made the initial breach in [1874] with a negative result about the *lack* of a bijective correspondence, Cantor secured the new ground

---

[8] Dauben [1979: 68ff] suggests that the title and presentation of Cantor [1874] were deliberately chosen to avoid censure by Kronecker, one of the journal editors.

[9] Indeed, this is where Wittgenstein [1956: I, Appendix II, 1-3] located what he took to be the problematic aspects of the talk of uncountability.

[10] A non-constructive proof typically deduces the existence of a mathematical object without providing a means for specifying it. Kac-Ulam [1968: 13] wrote: "The contrast between the methods of Liouville and Cantor is striking, and these methods provide excellent illustrations of two vastly different approaches toward proving the *existence* of mathematical objects. Liouville's is purely *constructive*; Cantor's is purely *existential*." See also Moore [1982: 39]. One exception to the misleading trend is Fraenkel [1930: 237][1953: 75], who from the beginning emphasized the constructive aspect of diagonalization.

The first non-constructive proof widely acknowledged as such was Hilbert's [1890] of his basis theorem. Earlier, Dedekind [1888: §159] had established the equivalence of two notions of being finite with a non-constructive proof that made an implicit use of the Axiom of Choice.

[11] Gray [1994] shows that Cantor's original [1874] argument can be implemented by an algorithm that generates the first *n* digits of a transcendental number with time complexity $O(2^{n^{1/3}})$, and his later diagonal argument, with a tractable algorithm of complexity $O(n^2 \log^2 n \log \log n)$. The original Liouville argument depended on a simple observation about fast convergence, and the digits of the Liouville numbers can be generated much faster. In terms of 2.3 below, the later Baire Category Theorem can be viewed as a direct generalization of Cantor's [1874] result, and the collection of Liouville numbers provides an explicit example of a co-meager yet measure zero set of reals (see Oxtoby [1971: §2]). On the other hand, Gray [1994] shows that every transcendental real is the result of diagonalization applied to *some* enumeration of the algebraic reals.

with a positive investigation of the *possibilities* for having such correspondences.[12] With "sequence" tied traditionally to countability through the indexing, Cantor used "correspondence [Beziehung]". Just as the discovery of the irrational numbers had led to one of the great achievements of Greek mathematics, Eudoxus's theory of geometrical proportions presented in Book V of Euclid's *Elements* and thematically antecedent to Dedekind's [1872] cuts, Cantor began his move toward a full-blown mathematical theory of the infinite.

Although holding the promise of a rewarding investigation Cantor did not come to any powers for infinite sets other than the two as set out in his [1874] proof. Cantor claimed at the end of [1878: 257]:

> Every infinite set of reals either is countable or has the power of the continuum .

This was the *Continuum Hypothesis* (CH) in the nascent context. The conjecture viewed as a primordial question would stimulate Cantor not only to approach the reals *qua* extensionalized continuum in an increasingly arithmetical fashion but also to grapple with fundamental questions of set existence. His triumphs across a new mathematical context would be like a brilliant light to entice others into the study of the infinite, but his inability to establish CH would also cast a long shadow. Set theory had its beginnings not as some abstract foundation for mathematics but rather as a setting for the articulation and solution of the *Continuum Problem*: to determine whether there are more than two powers embedded in the continuum.

In his magisterial *Grundlagen* [1883] Cantor developed the *transfinite numbers* [Anzahlen] and the key concept of *well-ordering*. A *well-ordering* of a set is a linear ordering of it according to which every non-empty subset has a least element. No longer was the infinitary indexing of his trigonometric series investigations mere contrivance. The "symbols of infinity" became autonomous and extended as the transfinite numbers, the emergence signified by the notational switch from the ∞ of potentiality to the ω of completion as the last letter of the Greek alphabet. With this the progression of transfinite numbers could be depicted:

$$0, 1, 2, \ldots \omega, \omega + 1, \omega + 2, \ldots, \omega + \omega (= \omega \cdot 2), \ldots, \omega^2, \ldots, \omega^\omega, \ldots, \omega^{\omega^\omega}, \ldots$$

A corresponding transition from subsets of $\mathbb{R}^n$ to a broader concept of set was signaled by the shift in terminology from "point-manifold [Punktmannigfaltigkeit]" to "set [Menge]". In this new setting well-orderings conveyed the sense of sequential counting and transfinite numbers served as standards for gauging well-orderings.

---

[12]Cantor developed a bijective correspondence between $\mathbb{R}^2$ and $\mathbb{R}$ by essentially interweaving the decimal expansions of a pair of reals to define the associated real, taking care of the countably many exceptional points like $.100\ldots = .099\ldots$ by an *ad hoc* shuffling procedure. Such an argument now seems straightforward, but to have bijectively identified the plane with the line was a stunning accomplishment at the time. In a letter to Dedekind of 29 June 1877 Cantor (Ewald [1996: 860]) wrote, in French in the text, "I see it, but I don't believe it."

Cantor's work inspired a push to establish the "invariance of dimension", that there can be no *continuous* bijection of any $\mathbb{R}^n$ onto $\mathbb{R}^m$ for $m < n$, with Cantor [1879] himself providing an argument. As topology developed, the stress brought on by the lack of firm ground led Brouwer [1911] to definitively establish the invariance of dimension in a seminal paper for algebraic topology.

As Cantor pointed out, every linear ordering of a finite set is already a well-ordering and all such orderings are isomorphic, so that the general sense is only brought out by infinite sets, for which there are non-isomorphic well-orderings. Cantor called the set of natural numbers the first number class (I) and the set of numbers whose predecessors are countable the second number class (II). Cantor conceived of (II) as being bounded above according to a limitation principle and showed that (II) itself is not countable. Proceeding upward, Cantor called the set of numbers whose predecessors are in bijective correspondence with (II) the third number class (III), and so forth. Cantor took a set to be of a *higher power* than another if they are not of the same power yet the latter is of the same power as a subset of the former. Cantor thus conceived of ever higher powers as represented by number classes and moreover took every power to be so represented. With this "free creation" of numbers, Cantor [1883: 550] propounded a basic principle that was to drive the analysis of sets:

> "It is always possible to bring any *well-defined* set into the form of a *well-ordered* set."

He regarded this as a "an especially remarkable law of thought which through its general validity is fundamental and rich in consequences." Sets are to be well-ordered, and thus they and their powers are to be gauged via the transfinite numbers of his structured conception of the infinite.

The well-ordering principle was consistent with Cantor's basic view in the *Grundlagen* that the finite and the transfinite are all of a piece and uniformly comprehendable in mathematics,[13] a view bolstered by his systematic development of the arithmetic of transfinite numbers seamlessly encompassing the finite numbers. Cantor also devoted several sections of the *Grundlagen* to a justificatory philosophy of the infinite, and while this metaphysics can be separated from the mathematical development, one concept was to suggest ultimate delimitations for set theory: Beyond the transfinite was the "Absolute", which Cantor eventually associated mathematically with the collection of all ordinal numbers and metaphysically with the transcendence of God.[14]

The Continuum Problem was never far from this development and could in fact be seen as an underlying motivation. The transfinite numbers were to provide the framework for Cantor's two approaches to the problem, the approach through power and the more direct approach through definable sets of reals, these each to initiate vast research programs.

As for the approach through power, Cantor in the *Grundlagen* established that the second number class (II) is uncountable, yet *any infinite subset of* (II) *is either countable or has the same power as* (II). Hence, (II) has exactly the property that Cantor sought for the reals, and he had reduced CH to the positive assertion that the reals and (II) have the same power. The following in brief is Cantor's argument that (II) is uncountable:

Suppose that s is a (countable) sequence of members of (II), say with initial element a. Let a' be a member of s, if any, such that $a < a'$; let a'' be a member of s, if any, such that $a' < a''$; and so forth. Then however long this process continues, the supremum of these numbers, or its successor, is not a member of s.

---

[13]This is emphasized by Hallett [1984] as Cantor's 'finitism'.

[14]The 'absolute infinite" is a varying but recurring explanatory concept in Cantor's work; see Jané [1995].

This argument was reminiscent of his [1874] argument that the reals are uncountable and suggested a correlation of the reals through their fundamental sequence representation with the members of (II) through associated cofinal sequences.[15] However, despite several announcements Cantor could never develop a workable correlation, an emerging problem in retrospect being that he could not *define* a well-ordering of the reals.

As for the approach through definable sets of reals, this evolved directly from Cantor's work on trigonometric series, the "symbols of infinity" used in the analysis of the P' operation transmuting to the transfinite numbers of the second number class (II).[16] In the *Grundlagen* Cantor studied P' for uncountable P and defined the key concept of a *perfect* set of reals (non-empty, closed, and containing no isolated points). Incorporating an observation of Ivar Bendixson [1883], Cantor showed in the succeeding [1884] that *any uncountable closed set of reals is the union of a perfect set and a countable set.* For a set A of reals, A has the *perfect set property iff* A is countable or else has a perfect subset. Cantor had shown in particular that *closed sets have the perfect set property.*

Since Cantor [1884,1884a] had been able to show that any perfect set has the power of the continuum, he had established that "CH holds for closed sets": every closed set either is countable or has the power of the continuum. Or from his new vantage point, he had reduced the Continuum Problem to determining whether there is a closed set of reals of the power of the second number class. He was unable to do so, but he had initiated a program for attacking the Continuum Problem that was to be vigorously pursued (cf. 2.3 and 2.5).

1.3 Diagonalization and Cardinal Numbers

In the ensuing years, unable to resolve the Continuum Problem through direct correlations with transfinite numbers Cantor approached size and order from a broader perspective that *would* incorporate the continuum. He identified power with *cardinal number*, an autonomous concept beyond being *une façon de parler* about bijective correspondence, and he went beyond well-orderings to the study of linear *order types*. Cantor embraced a structured view of sets, when "well-defined", as being given together with a linear ordering of their members. Order types and cardinal numbers resulted from successive abstraction, from a set M to its order type $\overline{M}$ and then to its cardinality $\overline{\overline{M}}$.

Almost two decades after his [1874] result that the reals are uncountable, Cantor in a short note [1891] subsumed it via his celebrated diagonal argument. With it, he estab-

---

[15] After describing the similarity between ω and $\sqrt{2}$ as limits of sequences, Cantor [1887: 99] interestingly correlated the creation of the transfinite numbers to the creation of the irrational numbers, beyond merely breaking new ground in different number contexts: "The transfinite numbers are in a certain sense *new irrationalities*, and in my opinion the best method of defining the finite irrational numbers [via Cantor's fundamental sequences] is wholly similar to, and I might even say in principle the same as, my method of introducing transfinite numbers. One can say unconditionally: the transfinite numbers *stand or fall* with the finite irrational numbers: they are like each other in their innermost being [Wesen]; for the former like the latter are definite delimited forms or modifications of the actual infinite."

[16] Ferreirós [1995] suggests how the formulation of the second number class as a completed totality with a succeeding transfinite number emerged directly from Cantor's work on the operation P', drawing Cantor's transfinite numbers even closer to his earlier work on trigonometric series.

lished: *For any set L the collection of functions from L into a fixed two-element set has a higher cardinality than that of L.* This result indeed generalized the [1874] result, since the collection of functions from the natural numbers into a fixed two-element set has the same cardinality as the reals. Here is how Cantor gave the argument in general form:[17]

Let M be the totality of all functions from L taking only the values 0 and 1. First, L is in bijective correspondence with a subset of M, through the assignment to each $x_0 \in L$ of the function on L that assigns 1 to x_0 and 0 to all other $x \in L$. However, there cannot be a bijective correspondence between M itself and L. Otherwise, there would be a function $\phi(x, z)$ of two variables such that for every member f of M there would be a $z \in L$ such that $\phi(x, z) = f(x)$ for every $x \in L$. But then, the "diagonalizing" function $g(x) = 1 - \phi(x, x)$ cannot be a member of M since for $z_0 \in L$, $g(z_0) \neq \phi(z_0, z_0)$!

In retrospect the diagonal argument can be drawn out from the [1874] proof.[18] Cantor had been shifting his notion of set to a level of abstraction beyond sets of reals and the like, and the casualness of his [1891] may reflect an underlying cohesion with his [1874]. Whether the new proof is really "different" from the earlier one, through this abstraction Cantor could now dispense with the recursively defined nested sets and limit construction, and he could apply his argument to any set. He had proved for the first time that there is a power higher than that of the continuum and moreover affirmed "the general theorem, that the powers of well-defined sets have no maximum."[19] The diagonal argument, even to its notation, would become method, flowing later into descriptive set theory, the Gödel Incompleteness Theorem, and recursion theory.

Today it goes without saying that a function from L into a two-element set corresponds to a subset of L, so that Cantor's Theorem is usually stated as: *For any set L its power set* $\mathcal{P}(L) = \{X \mid X \subseteq L\}$ *has a higher cardinality than L.* However, it would be an exaggeration to assert that Cantor was working on power sets; rather, he had expanded the 19th Century concept of *function* by ushering in arbitrary functions.[20] In any case, Cantor would now

---

[17]Actually, Cantor took L to be the unit interval of reals presumably to invoke a standard context, but he was clearly aware of the generality.

[18]Moreover, diagonalization as such had already occurred in Paul du Bois-Reymond's theory of growth as early as in his [1869]. An argument is manifest in his [1875: 365ff] for showing that for any sequence of real functions $f_0, f_1, f_2, \ldots$ there is a real function g such that for each n, $f_n(x) < g(x)$ for all sufficiently large reals x.

Diagonalization can be drawn out from Cantor's [1874] as follows: Starting with a sequence s of reals and a half-open interval I_0, instead of successively choosing delimiting *pairs* of reals in the sequence, avoid the members of s *one* at a time: Let I_1 be the left or right half-open subinterval of I_0 demarcated by its midpoint, whichever does not contain the first element of s. Then let I_2 be the left or right half-open subinterval of I_1 demarcated by its midpoint, whichever does not contain the second element of s; and so forth. Again, the nested intersection contains a real not in the sequence s. Abstracting the process in terms of reals in binary expansion, one is just generating the binary digits of the diagonalizing real.

In that letter of Cantor's to Dedekind of 7 December 1873 (Ewald [1996: 845ff]) first establishing the uncountability of the reals, there already appears, quite remarkably, a doubly indexed array of real numbers and a procedure for traversing the array downward and to the right, as in a now common picturing of the diagonal argument.

[19]Remarkably, Cantor had already conjectured in the *Grundlagen* [1883: 590] that the collection of continuous real functions has the same power as the second number class (II), and that the collection of all real functions has the same power as the third number class (III). These are consequences of the later Generalized Continuum Hypothesis and are indicative of the sweep of Cantor's conception.

[20]The 'power' in 'power set' is from 'Potenz' in the German for cardinal exponentiation, while Cantor's

have had to confront, in his function context, a general difficulty starkly abstracted from the Continuum Problem: *From a well-ordering of a set, a well-ordering of its power set is not necessarily definable.* The diagonal argument called into question Cantor's very notion of set: On the one hand, the argument, simple and elegant, should be part of set theory and lead to new sets of ever higher cardinality; on the other hand, these sets do not conform to Cantor's principle that every set comes with a (definable) well-ordering.[21]

Cantor's *Beiträge*, published in two parts [1895] and [1897], presented his mature theory of the transfinite. In the first part he described his post-*Grundlagen* work on cardinal number and the continuum. He quickly posed *Cardinal Comparability*, whether

for cardinal numbers $\mathfrak{a}$ and $\mathfrak{b}$, $\mathfrak{a} = \mathfrak{b}$, $\mathfrak{a} < \mathfrak{b}$, or $\mathfrak{b} < \mathfrak{a}$,

as a property "by no means self-evident" and which will be established later "when we shall have gained a survey over the ascending sequence of transfinite cardinal numbers and an insight into their connection." He went on to define the addition, multiplication, and exponentiation of cardinal numbers primordially in terms of set-theoretic operations and functions. If $\mathfrak{a}$ is the cardinal number of M and $\mathfrak{b}$ is the cardinal number of N, then $\mathfrak{a}^{\mathfrak{b}}$ is the cardinal number of the collection of all functions $: N \to M$, i.e. having domain N and taking values in M. The audacity of considering arbitrary functions from a set N into a set M was encased in a terminology that reflected both its novelty as well as the old view of function as given by an explicit rule.[22] As befits the introduction of new numbers Cantor then introduced a new notation, one using the Hebrew letter aleph, $\aleph$. With $\aleph_0$ the cardinal number of the set of natural numbers Cantor observed that $\aleph_0 \cdot \aleph_0 = \aleph_0$ and that $2^{\aleph_0}$ is the cardinal number of continuum. With this he observed that the [1878] labor of associating the continuum with the plane and so forth could be reduced to a "few strokes

---

'power' is from 'Mächtigkeit'.

[21]This is emphasized in Lavine [1994: IV.2]. Cantor did consider power sets in a letter of 20 September 1898 to Hilbert. In it Cantor entertained a notion of 'completed set', one of the guidelines being that 'the collection of *all subsets* of a completed set M is a completed set." Also, in a letter of 10 October 1898 to Hilbert, Cantor pointed out, in an argument focused on the continuum, that the power set $P(S)$ is in bijective correspondence with the collection of functions from S into $\{0, 1\}$. But in a letter of 9 May 1899 to Hilbert, writing now 'set' for 'completed set', Cantor wrote: "...it is our common conviction that the 'arithmetic continuum' is a 'set' in this sense; the question is whether this truth is provable or whether it is an axiom. I now incline more to the latter alternative, although I would gladly be convinced by you of the former." For the first and third letters in context see Moore [2002: 45] and for the second, Ferreirós [2007: epilogue]; the letters are in Meschkowski-Nilson [1991].

[22]Cantor wrote [1895: 486]: "...by a '*covering* [Belegung] of N with M,' we understand a law by which with every element n of N a definite element of M is bound up, where one and the same element of M can come repeatedly into application. The element of M bound up with n is, in a way, a one-valued function of n, and may be denoted by $f(n)$; it is called a 'covering function [Belegungsfunktion] of n.' The corresponding covering of N will be called $f(N)$." A convoluted description! Arbitrary functions on arbitrary domains are now of course commonplace in mathematics, but several authors at the time referred specifically to Cantor's concept of covering, most notably Zermelo [1904]. Jourdain in his introduction to his English translation of the *Beiträge* wrote (Cantor [1915: 82]): 'The introduction of the concept of 'covering' is the most striking advance in the principles of the theory of transfinite numbers from 1885 to 1895 ....'

With Cantor initially focusing on bijective correspondence [Beziehung] and these not quite construed as functions, Dedekind was the first to entertain an arbitrary function on an arbitrary domain. He [1888: §§21,36] formulated $\phi: S \to Z$, 'a mapping [Abbildung] of a system S in Z', in less convoluted terms, but did not consider the totality of such. He quickly moved to the case $Z = S$ for his theory of chains; see footnote 36.

of the pen" in his new arithmetic:

$$(2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0 \cdot \aleph_0} = 2^{\aleph_0}.$$

Cantor only mentioned

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

these to be the cardinal numbers of the successive number classes from the *Grundlagen* and thus to exhaust all the infinite cardinal numbers.

Cantor went on to present his theory of *order types*, abstractions of linear orderings. He defined an arithmetic of order types and characterized the order type η of the rationals as the countable dense linear order without endpoints, introducing the "forth" part of the now familiar back-and-forth argument of model theory.[23] He also characterized the order type θ of the reals as the perfect linear order with a countable dense set; whether a realist or not, Cantor the mathematician was able to provide a characterization of the continuum.

The second *Beiträge* developed the *Grundlagen* ideas by focusing on well-orderings and construing their order types as the *ordinal numbers*. Here at last was the general proof via order comparison of well-ordered sets that ordinal numbers are comparable. Cantor went on to describe ordinal arithmetic as a special case of the arithmetic of order types and after giving the basic properties of the second number class defined $\aleph_1$ as its cardinal number. The last sections were given over to a later preoccupation, the study of ordinal exponentiation in the second number class. The operation was defined via a transfinite recursion and used to establish a normal form, and the pivotal ϵ-numbers satisfying $\epsilon = \omega^\epsilon$ were analyzed.

The two parts of the *Beiträge* are not only distinct by subject matter, cardinal number and the continuum vs. ordinal number and well-ordering, but between them there developed a wide, irreconcilable breach. In the first part nowhere is the [1891] result $\mathfrak{a} < 2^{\mathfrak{a}}$ stated even in a special case; rather, it is made clear [1895: 495] that the procession of transfinite cardinal numbers is to be secured through their construal as the alephs. However, the second *Beiträge* does not mention any aleph beyond $\aleph_1$, nor does it mention CH, which could now have been stated as

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

(Cantor did state this in an 1895 letter.[24]) Ordinal comparability was secured, but cardinal comparability was not reduced to it. Every well-ordered set has an aleph as its cardinal number, but where is $2^{\aleph_0}$ in the aleph sequence?

Cantor's initial [1874] proof led to the Continuum Problem. That problem was embedded in the very interstices of the early development of set theory, and in fact the structures that Cantor built, while now of intrinsic interest, emerged in significant part out of efforts to articulate and solve the problem. Cantor's [1891] diagonal argument, arguably a transmutation of his initial [1874] proof, exacerbated a growing tension between having well-orderings and admitting sets of arbitrary functions (or power sets). David Hilbert, when

---

[23] See Plotkin [1993] for an analysis of the emergence of the back-and-forth argument.
[24] See Moore [1989: 99].

he presented his famous list of problems at the 1900 International Congress of Mathematicians at Paris, made the Continuum Problem the very first problem and intimated Cantor's difficulty by suggesting the desirability of "actually giving" a well-ordering of the reals.

The next, 1904 International Congress of Mathematicians at Heidelberg was to be a generational turning point for the development of set theory. Julius Kőnig delivered a lecture in which he provided a detailed argument that purportedly established that $2^{\aleph_0}$ is not an aleph, i.e. that the continuum is not well-orderable. The argument combined the now familiar inequality $\aleph_\alpha < \aleph_\alpha^{\aleph_0}$ for α of cofinality ω with a result from Felix Bernstein's Göttingen dissertation [1901: 49] which alas does not universally hold.[25] Cantor was understandably upset with the prospect that the continuum would simply escape the number context that he had devised for its analysis.

Accounts differ on how the issue was resolved. Although one has Zermelo finding an error within a day of the lecture, the weight of evidence is for Hausdorff having found the error.[26] Whatever the resolution, the torch had passed from Cantor to the next generation. Zermelo would go on to formulate his Well-Ordering Theorem and axiomatize set theory, and Hausdorff, to develop the higher transfinite in his study of order types and cofinalities.[27]

2   MATHEMATIZATION

2.1   Axiom of Choice and Axiomatization

Ernst Zermelo[28] (1871–1953), born when Cantor was establishing his trigonometric series results, had begun to investigate Cantorian set theory at Göttingen under the influence of Hilbert. In just over a month after the Heidelberg congress, Zermelo [1904] formulated what he soon called the *Axiom of Choice* (AC) and with it, established his Well-Ordering Theorem:

> Every set can be well-ordered .

Zermelo thereby shifted the notion of set away from the implicit assumption of Cantor's principle that every well-defined set is well-ordered and replaced that principle by an

---

[25]The *cofinality* of an ordinal number α is the least ordinal number β such that there is a set of form $\{\gamma_\xi \mid \xi < \beta\}$ unbounded in α, i.e. for any $\eta < \alpha$ there is an $\xi < \beta$ such that $\eta \leq \gamma_\xi < \alpha$. α is *regular* if its cofinality is itself, and otherwise α is *singular*. There concepts were not clarified until the work of Hausdorff, brought together in his [1908], discussed in 2.6.

Kőnig applied Bernstein's equality $\aleph_\alpha^{\aleph_0} = \aleph_\alpha \cdot 2^{\aleph_0}$ as follows: If $2^{\aleph_0}$ were an aleph, say $\aleph_\beta$, then by Bernstein's equality $\aleph_{\beta+\omega}^{\aleph_0} = \aleph_{\beta+\omega} \cdot 2^{\aleph_0} = \aleph_{\beta+\omega}$, contradicting Kőnig's inequality. However, Bernstein's equality fails when α has cofinality ω and $2^{\aleph_0} < \aleph_\alpha$. Kőnig's published account [1905] acknowledged the gap.

[26]See Grattan-Guinness [2000: 334] and Purkert [2002].

[27]And as with many incorrect proofs, there would be positive residues: Zermelo soon generalized Kőnig's inequality to the fundamental Zermelo-Kőnig inequality for cardinal exponentiation, which implies that the cofinality of $2^{\aleph_\alpha}$ is larger than α, and Hausdorff [1904: 571] published his recursion formula $\aleph_{\beta+1}^{\aleph_\alpha} = \aleph_{\beta+1} \cdot \aleph_\beta^{\aleph_\alpha}$, in form like Bernstein's result.

[28]Ebbinghaus [2007] is a substantive biography of Zermelo. See Kanamori [1997,2004] for Zermelo's work in set theory.

explicit axiom about a wider notion of set, incipiently unstructured but soon to be given form by axioms.

In retrospect, Zermelo's *argument* for his Well-Ordering Theorem can be viewed as pivotal for the development of set theory. To summarize the argument, suppose that x is a set to be well-ordered, and through Zermelo's Axiom-of-Choice hypothesis assume that the power set $\mathcal{P}(x) = \{y \mid y \subseteq x\}$ has a choice function, i.e. a function γ such that for every non-empty member y of $\mathcal{P}(x)$, $\gamma(y) \in y$. Call a subset y of x a γ-*set* if there is a well-ordering R of y such that for each $a \in y$,

$$\gamma(\{z \mid z \notin y \text{ or } z \, R \, a \text{ fails}\}) = a \,.$$

That is, each member of y is what γ "chooses" from what does not already precede that member according to R. The main observation is that γ-sets cohere in the following sense: If y is a γ-set with well-ordering R and z is a γ-set with well-ordering S, then $y \subseteq z$ and S is a prolongation of R, or vice versa. With this, let w be the union of all the γ-sets, i.e. all the γ-sets put together. Then w too is a γ-set, and by its maximality it must be all of x and hence x is well-ordered.

The converse to this result is immediate in that if x is well-ordered, then the power set $\mathcal{P}(x)$ has a choice function.[29] Not only did Zermelo's argument analyze the connection between having well-orderings and having choice functions on power sets, it anticipated in its defining of approximations and taking of a union the proof procedure for von Neumann's Transfinite Recursion Theorem (cf. 3.1).[30]

Zermelo [1904: 516] noted without much ado that his result implies that every infinite cardinal number is an aleph and satisfies $\mathfrak{m}^2 = \mathfrak{m}$, and that it secured Cardinal Comparability — so that the main issues raised by Cantor's *Beiträge* are at once resolved. Zermelo maintained that the Axiom of Choice, to the effect that *every* set has a choice function, is a "logical principle" which "is applied without hesitation everywhere in mathematical deduction", and this is reflected in the Well-Ordering Theorem being regarded as a theorem. The axiom is consistent with Cantor's view of the finite and transfinite as unitary, in that it posits for infinite sets an unproblematic feature of finite sets. On the other hand, the Well-Ordering Theorem shifted the weight from Cantor's well-orderings with their residually temporal aspect of numbering through *successive* choices to the use of a function for making *simultaneous* choices.[31] Cantor's work had served to exacerbate a growing discord among mathematicians with respect to two related issues: whether infinite collections can be mathematically investigated at all, and how far the function concept is to be extended. The positive use of an arbitrary function operating on arbitrary subsets of a set having been made explicit, there was open controversy after the appearance of Zermelo's proof. This can be viewed as a turning point for mathematics, with the subsequent tilt-

---

[29]Namely, with $\prec$ a well-ordering of x, for each non-empty member y of $\mathcal{P}(x)$, let $\gamma(y)$ be the the $\prec$-least member of y.

[30]See Kanamori [1997] for more on the significance of Zermelo's argument, in particular as a fixed point argument.

[31]Zermelo himself stressed the importance of simultaneous choices over successive choices in criticism of an argument of Cantor's for the Well-Ordering Theorem in 1899 correspondence with Dedekind, discussed in 2.2. See Cantor [1932: 451] or van Heijenoort [1967: 117].

ing toward the acceptance of the Axiom of Choice symptomatic of a conceptual shift in mathematics.

In response to his critics Zermelo published a second proof [1908] of his Well-Ordering Theorem, and with axiomatization assuming a general methodological role in mathematics he also published the first full-fledged axiomatization [1908a] of set theory. But as with Cantor's work this was no idle structure building but a response to pressure for a new mathematical context. In this case it was not for the formulation and solution of a *problem* like the Continuum Problem, but rather to clarify a specific *proof*. In addition to codifying generative set-theoretic principles, a substantial motive for Zermelo's axiomatizing set theory was to buttress his Well-Ordering Theorem by making explicit its underlying set existence assumptions.[32] Initiating the first major transmutation of the notion of set after Cantor, Zermelo thereby ushered in a new abstract, prescriptive view of sets as structured solely by membership and governed and generated by axioms, a view that would soon come to dominate. Thus, proof played a crucial role by stimulating an axiomatization of a field of study and a corresponding transmutation of its underlying notions.

The objections raised against Zermelo's first proof [1904] mainly played on the ambiguities of a γ-set's well-ordering being only implicit, as for Cantor's sets, and on the definition of the well-ordering being impredicative — defined as a γ-set and so drawn from a collection of which it is already a member. Largely to preclude these objections Zermelo in his second [1908] proof resorted to a rendition of orderings in terms of segments and inclusion first used by Gerhard Hessenberg [1906: 674ff] and a closure approach with roots in Dedekind [1888]. Instead of extending initial segments toward the desired well-ordering, Zermelo got at the collection of its final segments by taking an *intersection* in a larger setting.[33]

With his [1908a] axiomatization, Zermelo "started from set theory as it is historically given" to seek out principles sufficiently restrictive "to exclude all contradictions" and sufficiently wide "to retain all that is valuable". However, he would transform set theory by making explicit *new* existence principles and promoting a generative point of view. Zermelo had begun working out an axiomatization as early as 1905, addressing issues raised by his [1904] proof.[34] The mature presentation is a precipitation of seven axioms, and these do not just reflect "set theory as it is historically given", but explicitly buttress his proof(s) of the Well-Ordering Theorem.

Zermelo's seven set axioms, now formalized, constitute the familiar theory Z, Zermelo set theory: Extensionality, Elementary Sets ($\emptyset$, $\{a\}$, $\{a, b\}$), Separation, Power Set, Union, Choice, and Infinity. His setting allowed for urelements, objects without members yet distinct from each other. But Zermelo focused on sets, and his Axiom of Extensional-

---

[32] Moore [1982: 155ff] supports this contention using items from Zermelo's *Nachlass*.

[33] To well-order a set M using a choice function φ on $\mathcal{P}(M)$, Zermelo defined a Θ-*chain* to be a collection Θ of subsets of M such that: (a) $M \in \Theta$; (b) if $A \in \Theta$, then $A - \{\varphi(A)\} \in \Theta$; and (c) if $Z \subseteq \Theta$, then $\bigcap Z \in \Theta$. He then took the intersection I of all Θ-chains, and observed that I is again a Θ-chain. Finally, he showed that I provides a well-ordering of M given by: $a \prec b$ *iff* there is an $A \in I$ such that $a \notin A$ and $b \in A$. I thus consists of the final segments of the same well-ordering as provided by the [1904] proof. Note that this second proof is less parsimonious than the [1904] proof, as it uses the power set of the power set of M.

[34] This is documented by Moore [1982: 155ff] with items from Zermelo's *Nachlass*.

ity announced the espousal of an extensional viewpoint. In line with this AC, a "logical principle" in [1904] expressed in terms of an informal choice function, was framed less instrumentally: It posited for a set consisting of non-empty, pairwise disjoint sets the existence of a *set* that meets each one in a unique element.[35] However, Separation retained an intensional aspect with its "separating out" of a new set from a given set using a *definite* property, where a property is "*definite [definit]* if the fundamental relations of the domain, by means of the axioms and the universally valid laws of logic, determine without arbitrariness whether it holds or not." But with no underlying logic formalized, the ambiguity of definite property would become a major issue. With Infinity and Power Set Zermelo provided for sufficiently rich settings for set-theoretic constructions. Tempering the logicians' extravagant and problematic "all" the Power Set axiom provided the provenance for "all" for subsets of a given set, just as Separation served to capture "all" for elements of a given set satisfying a property. Finally, Union and Choice completed the encasing of Zermelo's proof(s) of his Well-Ordering Theorem in the necessary set existence principles. Notably, Zermelo's recursive [1904] argumentation also brought him in proximity of the Transfinite Recursion Theorem and thus of Replacement, the next axiom to be adjoined in the subsequent development of set theory (cf. 3.1).

Fully two decades earlier Dedekind [1888] had provided an incisive analysis of the natural numbers and their arithmetic in terms of sets [Systeme], and several overlapping aspects can serve as points of departure for Zermelo's axiomatization.[36] The most immediate is how Dedekind's argumentation extends to Zermelo's [1908] proof of the Well-Ordering Theorem, which in the transfinite setting brings out the role of AC. Both Dedekind and Zermelo set down rules for sets in large part to articulate arguments involving simple set operations like "set of", union, and intersection. In particular, both had to argue for the equality of sets resulting after involved manipulations, and extensionality became operationally necessary. However vague the initial descriptions of sets, sets are to be determined solely by their elements, and the membership question is to be determinate.[37] The looseness of Dedekind's description of sets allowed him [1888: §66] the latitude to "prove" the existence of infinite sets, but Zermelo just stated the Axiom of Infinity as a set existence principle.

The main point of departure has to do with the larger issue of the role of proof for articulating sets. By Dedekind's time proof had become basic for mathematics, and indeed

---

[35]Russell [1906] had previously arrived at this form, his Multiplicative Axiom. The elimination of the 'pairwise disjoint' by going to a choice function formulation can be established with the Union Axiom, and this is the only use of that axiom in the second, [1908] proof of the Well-Ordering Theorem.

[36]In current terminology, Dedekind [1888] considered arbitrary sets S and mappings $\phi\colon S \to S$ and defined a *chain* [Kette] to be a $K \subseteq S$ such that $\phi``K \subseteq K$. For $A \subseteq S$, the *chain of A* is the intersection of all chains $K \supseteq A$. A set $N$ is *simply infinite iff* there is an injective $\phi\colon N \to N$ such that $N - \phi``N \neq \emptyset$. Letting 1 be a distinguished element of $N - \phi``N \neq \emptyset$ Dedekind considered the chain of {1}, the chain of {$\phi(1)$}, and so forth. Having stated an inherent induction principle, he proceeded to show that these sets have all the ordering and arithmetical properties of the natural numbers (that are established nowadays in texts for the (von Neumann) finite ordinals).

[37]Dedekind [1888: §2] begins a footnote to his statement about extensional determination with: "In what manner this determination is brought about, and whether we know a way of deciding upon it, is a matter of indifference for all that follows; the general laws to be developed in no way depend upon it; they hold under all circumstances."

his work did a great deal to enshrine proof as the vehicle for algebraic abstraction and generalization.[38] Like algebraic constructs, sets were new to mathematics and would be incorporated by setting down the rules for their proofs. Just as calculations are part of the sense of numbers, so proofs would be part of the sense of sets, as their "calculations". Just as Euclid's axioms for geometry had set out the permissible geometric constructions, the axioms of set theory would set out the specific rules for set generation and manipulation. But unlike the emergence of mathematics from marketplace arithmetic and Greek geometry, sets and transfinite numbers were neither laden nor buttressed with substantial antecedents. Like strangers in a strange land stalwarts developed a familiarity with them guided hand in hand by their axiomatic framework. For Dedekind [1888] it had sufficed to work with sets by merely giving a few definitions and properties, those foreshadowing Extensionality, Union, and Infinity. Zermelo [1908a] provided more rules: Separation, Power Set, and Choice.

Zermelo [1908], with its rendition of orderings in terms of segments and inclusion, and Zermelo [1908a], which at the end cast Cantor's theory of cardinality in terms of functions cast as set constructs, brought out Zermelo's *set-theoretic reductionism*. Zermelo pioneered the reduction of mathematical concepts and arguments to set-theoretic concepts and arguments from axioms, based on sets doing the work of mathematical objects. Zermelo's analyses moreover served to draw out what would come to be generally regarded as set-theoretic out of the presumptively logical. This would be particularly salient for Infinity and Power Set and was strategically advanced by the relegation of property considerations to Separation.

Zermelo's axiomatization also shifted the focus away from the transfinite numbers to an abstract view of sets structured solely by $\in$ and simple operations. For Cantor the transfinite numbers had become central to his investigation of definable sets of reals and the Continuum Problem, and sets had emerged not only equipped with orderings but only as the developing context dictated, with the "set of" operation never iterated more than three or four times. For Zermelo his second, [1908] proof of the Well-Ordering Theorem served to eliminate any residual role that the transfinite numbers may have played in the first proof and highlighted the set-theoretic operations. This approach to (linear) ordering was to preoccupy his followers for some time, and through this period the elimination of the use of transfinite numbers where possible, like ideal numbers, was regarded as salutary.[39] Hence, Zermelo rather than Cantor should be regarded as the creator of abstract set theory.

---

[38]cf. the first sentence of the preface to Dedekind [1888]: 'In science nothing capable of proof ought to be accepted without proof."

[39]Some notable examples: Lindelöf [1905] proved the Cantor-Bendixson result, that every uncountable closed set is the union of a perfect set and a countable set, without using transfinite numbers. Suslin's [1917], discussed in 2.5, had the unassuming title, 'On a definition of the Borel sets without transfinite numbers", hardly indicative of its results, so fundamental for descriptive set theory. And Kuratowski [1922] showed, pursuing the approach of Zermelo [1908], that inclusion chains defined via transfinite recursion with intersections taken at limits can also be defined without transfinite numbers. Kuratowski [1922] essentially formulated Zorn's Lemma, and this was the main success of the push away from explicit well-orderings. Especially after the appearance of Zorn [1935] this recasting of AC came to dominate in algebra and topology.

Outgrowing Zermelo's pragmatic purposes axiomatic set theory could not long fore-stall the Cantorian initiative, as even $2^{\aleph_0} = \aleph_1$ could not be asserted directly, and in the 1920s John von Neumann was to fully incorporate the transfinite using Replacement (cf. 3.1).[40] On the other hand, Zermelo's axioms had the advantages of schematic sim-plicity and open-endedness. The generative set formation axioms, especially Power Set and Union, were to lead to Zermelo's [1930] cumulative hierarchy picture of sets, and the vagueness of the *definit* property in the Separation Axiom was to invite Thoralf Skolem's [1923] proposal to base it on first-order logic, enforcing extensionalization (cf. 3.2).

2.2   Logic and Paradox

At this point, the incursions of a looming tradition can no longer be ignored. Gottlob Frege is regarded as the greatest philosopher of logic since Aristotle for developing quantifica-tional logic in his *Begriffsschrift* [1879], establishing a logical foundation for arithmetic in his *Grundlagen* [1884], and generally stimulating the *analytic tradition* in philosophy. The architect of that tradition was Bertrand Russell who in his earlier years, influenced by Frege and Giuseppe Peano, wanted to found all of mathematics on the certainty of logic. But from a logical point of view Russell [1903] became exercised with paradox. He had arrived at Russell's Paradox in late 1901 by analyzing Cantor's diagonal argument applied to the class of all classes,[41] a version of which is now known as Cantor's Paradox of the largest cardinal number. Russell [1903: §301] also refocused the Burali-Forti Para-dox of the largest ordinal number, after reading Cesare Burali-Forti's [1897].[42] Russell's Paradox famously led to the tottering of Frege's mature formal system, the *Grundgesetze* [1893, 1903].[43]

Russell's own reaction was to build a complex logical structure, one used later to de-velop mathematics in Whitehead and Russell's 1910-3 *Principia Mathematica*. Russell's *ramified theory of types* is a scheme of logical definitions based on *orders* and *types* indexed by the natural numbers. Russell proceeded "intensionally"; he conceived this scheme as a classification of propositions based on the notion of *propositional function*, a notion not reducible to membership (extensionality). Proceeding in modern fashion, we may say that the universe of the *Principia* consists of *objects* stratified into disjoint types T_n, where T_0 consists of the *individuals*, $T_{n+1} \subseteq \{Y \mid Y \subseteq T_n\}$, and the types T_n for $n > 0$ are further ramified into orders O_n^i with $T_n = \bigcup_i O_n^i$. An object in O_n^i is to be defined either in terms of individuals or of objects in some fixed O_m^j for some $j < i$ and $m \leq n$, the definitions allowing for quantification only over O_m^j. This precludes Russell's Paradox and other "vicious circles", as objects consist only of previous objects and are built up

---

[40]Textbooks usually establish the Well-Ordering Theorem by first introducing Replacement, formalizing transfinite recursion, and only then defining the well-ordering using (von Neumann) ordinals; this amounts to another historical misrepresentation, but one that resonates with how acceptance of Zermelo's proof broke the ground for formal transfinite recursion.

[41]Grattan-Guinness [1974], Coffa [1979], Moore [1988], and Garciadiego [1992] describe the evolution of Russell's Paradox.

[42]Moore-Garciadiego [1981] and Garciadiego [1992] describe the evolution of the Burali-Forti Paradox.

[43]See the exchange of letters between Russell and Frege in van Heijenoort [1967: 124ff]. Russell's Paradox showed that Frege's Basic Law V is inconsistent.

through definitions referring only to previous stages. However, in this system it is impossible to quantify over all objects in a type T_n, and this makes the formulation of numerous mathematical propositions at best cumbersome and at worst impossible. Russell was led to introduce his *Axiom of Reducibility*, which asserts that *for each object there is a predicative object consisting of exactly the same objects*, where an object is *predicative* if its order is the least greater than that of its constituents. This axiom reduced consideration to individuals, predicative objects consisting of individuals, predicative objects consisting of predicative objects consisting of individuals, and so on—the *simple theory of types*. In traumatic reaction to his paradox Russell had built a complex system of orders and types only to collapse it with his Axiom of Reducibility, a fearful symmetry imposed by an artful dodger.

The mathematicians did not imbue the paradoxes with such potency. Unlike Russell who wanted to get at everything but found that he could not, they started with what could be got at and peered beyond. And as with the invention of the irrational numbers, the outward push eventually led to the positive subsumption of the paradoxes.

Cantor in 1899 correspondence with Dedekind considered the collection Ω of all ordinal numbers as in the Burali-Forti Paradox, but he used it *positively* to give mathematical expression to his Absolute.[44] First, he distinguished between two kinds of multiplicities (Vielheiten): There are multiplicities such that when taken as a unity (Einheit) lead to a contradiction; such multiplicities he called "*absolutely infinite or inconsistent multiplicities*" and noted that the "totality of everything thinkable" is such a multiplicity. A multiplicity that can be thought of without contradiction as "being together" he called a "*consistent multiplicity* or a 'set [Menge]'". Cantor then used the Burali-Forti Paradox argument to point out that the class Ω of all ordinal numbers is an inconsistent multiplicity. He proceeded to argue that every *set* can be well-ordered through a presumably recursive procedure whereby a well-ordering is defined through successive choices. The set must get well-ordered, else all of Ω would be injectible into it, so that the set would have been an inconsistent multiplicity instead.[45]

Zermelo found Russell's Paradox independently and probably in 1902,[46] but like Cantor, he did not regard the emergence of the paradoxes so much as a crisis as an overall delimitation for sets. In the Zermelian generative view [1908: 118], "... if in set theory we confine ourselves to a number of established principles such as those that constitute the basis of our proof — principles that enable us to form initial sets and to derive new sets from given ones – then all such contradictions can be avoided." For the first theorem of his axiomatic theory Zermelo [1908a] subsumed Russell's Paradox, putting it to use as is done now to establish that for any set x there is a $y \subseteq x$ such that $y \notin x$, and hence that there is no universal set.[47]

---

[44] See footnote 3 for more about the 1899 correspondence. Purkert [1989: 57ff] argues that Cantor had already arrived at the Burali-Forti Paradox around the time of the *Grundlagen* [1883]. On the interpretations supported in the text *all* of the logical paradoxes grew out of Cantor's work —with Russell shifting the weight to paradox.

[45] G.H. Hardy [1903] and Philip Jourdain [1904, 1905] also gave arguments involving the injection of Ω, but such an approach would only get codified at a later stage in the development of set theory in the work of von Neumann [1925] (cf. 3.1).

[46] See Kanamori [2004: §1].

[47] In 2.6 Hartogs's Theorem is construed as a positive subsumption of that other, the Burali-Forti Paradox.

The differing concerns of Frege-Russell logic and the emerging set theory are further brought out by the analysis of the function concept as discussed below in 2.4, and those issues are here rehearsed with respect to the existence of the null class, or empty set.[48] Frege in his *Grundlagen* [1884] eschewed the terms "set [Menge]" and "class [Klasse]", but in any case the extension of the concept "not identical with itself" was key to his definition of zero as a logical object. Ernst Schröder, in the first volume [1890] of his major work on the algebra of logic, held a traditional view that a class is merely a collection of objects, without the { } so to speak. In his review [1895] of Schröder's [1890], Frege argued that Schröder cannot both maintain this view of classes and assert that there is a null class, since the null class contains no objects. For Frege, logic enters in giving unity to a class as the extension of a *concept* and thus makes the null class viable.

It is among the set theorists that the null class, *qua* empty set, emerged to the fore as an elementary concept and a basic building block. Cantor himself did not dwell on the empty set. At one point he did write [1880: 355] that "the identity of two pointsets P and Q will be expressed by the formula $P \equiv Q$"; defined disjoint sets as "*lacking intersection*"; and then wrote [1880: 356] "for the absence of points . . . we choose the letter O; $P \equiv O$ indicates that the set P contains *no single point*." (So, "$\equiv O$" is arguably more like a predication for being empty at this stage.)

Dedekind [1888: §2] deliberately excluded the empty set [Nullsystem] "for certain reasons", though he saw its possible usefulness in other contexts. Zermelo [1908a] wrote in his Axiom II: "There exists a (improper [uneigentliche]) set, the *null set* [*Nullmenge*] 0, that contains no element at all." Something of intension remained in the "(improper [uneigentliche])", though he did point out that because of his Axiom I, the Axiom of Extensionality, there is a single empty set. Finally, Hausdorff [1914] unequivocally opted for the empty set [Nullmenge]. However, a hint of predication remained when he wrote [1914: 3]: ". . . the equation $A = 0$ means that the set A has no element, vanishes [verschwindet], is empty." The use to which Hausdorff put "0" is much as "∅" is used in modern mathematics, particularly to indicate the extension of the conjunction of mutually exclusive properties.

The set theorists, unencumbered by philosophical motivations or traditions, attributed little significance to the empty set beyond its usefulness. Although embracing both extensionality and the null class may engender philosophical difficulties for the logic of classes, the empty set became commonplace in mathematics simply through use, like its intimate, zero.

2.3   Measure, Category, and Borel Hierarchy

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

The French analysts Emile Borel, Ren´e Baire, and Henri Lebesgue took on the investigation of definable sets of reals in what was to be a paradigmatically constructive

---

[48]For more on the empty set, see Kanamori [2003a].

approach. Cantor [1884] 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.[49]

Soon after completing his thesis Borel [1898: 46ff] considered for his theory of measure those sets of reals obtainable by starting with the intervals and closing off under complementation and countable union. The formulation was axiomatic and in effect impredicative, and seen in this light, bold and imaginative; the sets are now known as the *Borel sets* and quite well-understood.

Baire in his thesis [1899] took on a dictum of Lejeune Dirichlet's that a real function is any arbitrary assignment of reals, and diverging from the 19th Century preoccupation with pathological examples, sought a constructive approach via pointwise limits. His *Baire class 0* consists of the continuous real functions, and for countable ordinal numbers $\alpha > 0$, *Baire class α* consists of those functions f not in any previous class yet obtainable as pointwise limits of sequences f_0, f_1, f_2, ... of functions in previous classes, i.e. $f(x) = \lim_{n \to \infty} f_n(x)$ for every real x. The functions in these classes are now known as the *Baire* functions, and this was the first stratification into a transfinite hierarchy after Cantor.[50]

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*. 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 has the *Baire property iff* it has a meager symmetric difference with some open set. Straightforward arguments show that every Borel set has the Baire property.

Lebesgue's thesis [1902] is fundamental for modern integration theory as the source of his concept of measurability. Inspired in part by Borel's ideas but notably containing non-constructive aspects, Lebesgue's concept of measurable set through its closure under countable unions subsumed the Borel sets, and his analytic definition of measurable function through its closure under pointwise limits subsumed the Baire functions. Category and measure are quite different; there is a co-meager (complement of a meager) set of reals that has Lebesgue measure zero.[51] Lebesgue's first major work in a distinctive direction would be the seminal paper in descriptive set theory:

In the memoir [1905] Lebesgue investigated the Baire functions, stressing that they are exactly the functions definable via analytic expressions (in a sense made precise). He first established a correlation with the Borel sets by showing that they are exactly the preimages of open intervals via Baire functions. With this he introduced the first hierarchy for the Borel sets, his *open sets of class α* not being in any previous class yet being preimages of some open interval via some Baire class α function. After verifying various

---

[49]See Kanamori [1995] for more on the emergence of descriptive set theory. See Moschovakis [1980] or Kanamori [2003] for the mathematical development.

[50]Baire mainly studied the finite levels, particularly the classes 1 and 2. He [1898] pointed out that Dirichlet's function that assigns 1 to rationals and 0 to irrationals is in class 2 and also observed with a non-constructive appeal to Cantor's cardinality argument that there are real functions that are not Baire.

[51]See footnote 11. See Hawkins [1975] for more on the development of Lebesgue measurability. See Oxtoby [1971] for an account of category and measure in juxtaposition.

closure properties and providing characterizations for these classes Lebesgue established two main results. The first demonstrated the necessity of exhausting the countable ordinal numbers: *The Baire hierarchy is proper, i.e. for every countable α there is a Baire function of class α; correspondingly the hierarchy for the Borel sets is analogously proper.* The second established transcendence beyond countable closure for his concept of measurability: *There is a Lebesgue measurable function which is not in any Baire class; correspondingly there is a Lebesgue measurable set which is not a Borel set.*

The first result was the first of all hierarchy results, and a precursor of fundamental work in mathematical logic in that it applied Cantor's enumeration and diagonalization argument to achieve a transcendence to a next level. Lebesgue's second result was also remarkable in that he actually provided an explicitly defined set, one that was later seen to be the first example of a non-Borel analytic set (cf. 2.5). For this purpose, the reals were for the first time regarded as encoding something else, namely countable well-orderings, and this not only further embedded the transfinite into the investigation of sets of reals, but foreshadowed the later coding results of mathematical logic.

Lebesgue's results, along with the later work in descriptive set theory, can be viewed as pushing the mathematical frontier of the actual infinite past $\aleph_0$, which arguably had achieved a mathematical domesticity through increasing use in the late 19th Century, through Cantor's second number class to $\aleph_1$. It is somewhat ironic but also revealing, then, that this grew out of work by analysts with a definite constructive bent. Baire [1899: 36] viewed the infinite ordinal numbers and hence his function hierarchy as merely *une façon de parler*, and continued to view infinite concepts only in potentiality. Borel [1898] took a pragmatic approach and seemed to accept the countable ordinal numbers. Lebesgue was more equivocal but still accepting; recalling Cantor's early attitude Lebesgue regarded the ordinal numbers as an indexing system, "symbols" for classes, but nonetheless he worked out their basic properties, even providing a formulation [1905: 149] of proof by transfinite induction. All three analysts expressed misgivings about AC and its use in Zermelo's proof.[52]

As descriptive set theory was to develop, a major concern became the extent of the *regularity properties*, those properties indicative of well-behaved sets of reals of which Lebesgue measurability, the Baire property, and the perfect set property are the prominent examples. These properties seemed to get at basic features of the extensional construal of the continuum, yet resisted inductive approaches. Early explicit uses of AC through its role in providing a well-ordering of the reals showed how it allowed for new constructions: Giuseppe Vitali [1905] established that there is a non-Lebesgue measurable set of reals, and Felix Bernstein [1908], that there is a set of reals without the perfect set property. Soon it was seen that neither of these examples have the Baire property. Thus, that the reals are well-orderable, an early contention of Cantor's, permitted constructions that precluded the universality of the regularity properties, in particular his own approach to the Continuum Problem through the perfect set property.

---

[52]See Moore [1982: 2.3].

2.4 Hausdorff and Functions

Felix Hausdorff was the first developer of the transfinite after Cantor, the one whose work first suggested the rich possibilities for a mathematical investigation of the higher transfinite. A mathematician *par excellence*, Hausdorff took that sort of mathematical approach to set theory and extensional, set-theoretic approach to mathematics that would dominate in the years to come. While the web of 19th Century intension in Cantor's work, especially his approach toward functions, now seems rather remote, Hausdorff's work seems familiar as part of the modern language of mathematics.

In [1908] Hausdorff brought together his extensive work on *uncountable* order types.[53] Deploring all the fuss being made over foundations by his contemporaries (p.436) and with Cantor having taken the Continuum Problem as far as seemed possible, Hausdorff proceeded to venture beyond the second number class with vigor. He provided an elegant analysis of scattered linear order types (those having no dense subtype) in a transfinite hierarchy, and constructed the η_α sets, prototypes for saturated model theory. He first stated the *Generalized Continuum Hypothesis* (GCH), that $2^{\aleph_\alpha} = \aleph_{\alpha+1}$ for every α, clarified the significance of cofinality, and first considered (p.443) the possibility of an uncountable regular limit cardinal, the first *large cardinal.*

Large cardinal hypotheses posit cardinals with properties that entail their transcendence over smaller cardinals, and as it has turned out, provide a superstructure of hypotheses for the analysis of strong propositions in terms of consistency. Hausdorff observed that uncountable regular limit cardinals, also known now as *weakly inaccessible cardinals*, are a natural closure point for cardinal limit processes. In penetrating work of only a few years later Paul Mahlo [1911,1912,1913] investigated hierarchies of such cardinals based on higher fixed-point phenomena, the *Mahlo cardinals.* The theory of large cardinals was to become a mainstream of set theory.[54]

Hausdorff's classic text, *Grundzüge der Mengenlehre* [1914] dedicated to Cantor, broke the ground for a generation of mathematicians in both set theory and topology. A compendium of a wealth of results, it emphasized mathematical approaches and procedures that would eventually take firm root.[55] After giving a clear account of Zermelo's first, [1904] proof of the Well-Ordering Theorem, Hausdorff (p.140ff) emphasized its maximality aspect by giving synoptic versions of Zorn's Lemma two decades before Zorn [1935], one of them now known as Hausdorff's Maximality Principle.[56] Also, Hausdorff (p.304) provided the now standard account of the Borel hierarchy of sets, with the still persistent F_σ and G_δ notation. Of particular interest, Hausdorff (p.469ff, and also in [1914a]) used AC to provide what is now known as Hausdorff's Paradox, an implausible decomposition of the sphere and the source of the better known Banach-Tarski Paradox

---

[53] See Plotkin [2005] for translations and careful analyses of Hausdorff's work on ordered sets.

[54] See Kanamori [2003] for more on large cardinals.

[55] Hausdorff's mathematical attitude is reflected in a remark following his explanation of cardinal number in a revised edition [1937:§5] of [1914]: 'This formal explanation says what the cardinal numbers are supposed to do, not what they are. More precise definitions have been attempted, but they are unsatisfactory and unnecessary. Relations between cardinal numbers are merely a more convenient way of expressing relations between sets; we must leave the determination of the 'essence' of the cardinal number to philosophy.'

[56] Hausdorff's Maximality Principle states that if *A* is a partially ordered set and *B* is a linearly ordered subset, then there is a $\subseteq$-maximal linearly ordered subset of *A* including *B*.

from Stefan Banach and Alfred Tarski's [1924].[57] Hausdorff's Paradox was the first, and a dramatic, synthesis of classical mathematics and the Zermelian abstract view.

Hausdorff's reduction of functions through a defined ordered pair highlights the differing concerns of the earlier Frege-Russell logic and the emerging set theory.[58] Frege [1891] had two fundamental categories, *function* and *object*, with a function being "unsaturated" and supplemented by objects as arguments. A *concept* is a function with two possible values, the True and the False, and a *relation* is a concept that takes two arguments. The extension of a concept is its graph or course-of-values [Werthverlauf], which is an object, and Frege [1893: §36] devised an iterated or double course-of-values [Doppelwerthverlauf] for the extension of a relation. In these involved ways Frege assimilated relations to functions. As for the ordered pair, Frege in his *Grundgesetze* [1893: §144] provided the extravagant definition that the ordered pair of *x* and *y* is that class to which all and only the extensions of relations to which *x* stands to *y* belong.[59]

On the other hand, Peirce [1883], Schröder [1895], and Peano [1897] essentially regarded a relation from the outset as just a collection of ordered pairs. Whereas Frege was attempting an analysis of thought, Peano was mainly concerned with recasting ongoing mathematics in economical and flexible symbolism and made many reductions, e.g. construing a *sequence* in analysis as a *function* on the natural numbers. Peano from his earliest logical writings had used "(x, y)" to indicate the ordered pair in formula and function substitutions and extensions. In [1897] he explicitly formulated the ordered pair using "$(x; y)$" and moreover raised the two main points about the ordered pair: First, equation 18 of his Definitions stated the instrumental property which is all that is required of the ordered pair:

$$(*) \qquad\qquad \langle x, y \rangle = \langle a, b \rangle \ \textit{iff} \ \ x = a \text{ and } y = b \,.$$

Second, he broached the possibility of reducibility, writing: "The idea of a pair is fundamental, i.e. we do not know how to express it using the preceding symbols."

In Whitehead and Russell's *Principia Mathematica* [1910-3], relations distinguished in intension and in extension were derived from "propositional" functions taken as fundamental and other "descriptive" functions derived from relations. They [1910: ∗55] like Frege defined an ordered pair derivatively, in their case in terms of classes and relations, and also for a specific purpose.[60] Previously Russell [1903: §27] had criticized Peirce

---

[57]Hausdorff's Paradox states that a sphere can be decomposed into four pieces Q, A, B, C with Q countable and A, B, C, and $B \cup C$ all pairwise congruent. Even more implausibly, the Banach-Tarski Paradox states that a ball can be decomposed into finitely many pieces that can be rearranged by rigid motions to form two balls of the same size as the original ball. Raphael Robinson [1947] later showed that there is such a decomposition into just five pieces with one of them containing a single point, and moreover that five is the minimal number. See Wagon [1985] for more on these and similar results; they stimulated interesting developments in measure theory that, rather than casting doubt on AC, embedded it further into mathematical practice (cf. 2.6).

[58]For more on the ordered pair, see Kanamori [2003a].

[59]This definition, which recalls the Whitehead–Russell definition of the cardinal number 2, depended on Frege's famously inconsistent Basic Law V. See Heck [1995] for more on Frege's definition and use of his ordered pair.

[60]Whitehead and Russell had first defined a cartesian product by other means, and only then defined their ordered pair $x{\downarrow}y$ as $\{x\} \times \{y\}$, a remarkable inversion from the current point of view. They [1910: ∗56] used their ordered pair initially to define the ordinal number 2.

and Schröder for regarding a relation "essentially as a class of couples," although he did not mention this shortcoming in Peano.[61] Commenting obliviously on *Principia* Peano [1911,1913] simply reaffirmed an ordered pair as basic, defined a relation as a class of ordered pairs, and a function extensionally as a kind of relation, referring to the final version of his *Formulario Mathematico* [1905-8: 73ff.] as the source.

Capping this to and fro Norbert Wiener [1914] provided a definition of the ordered pair in terms of unordered pairs of classes only, thereby reducing relations to classes. Working in Russell's theory of types, Wiener defined the ordered pair $\langle x, y \rangle$ as

$$\{\{\{x\}, \Lambda\}, \{\{y\}\}\}$$

when x and y are of the same type and Λ is the null class (of the next type), and pointed out that this definition satisfies the instrumental property (∗) above. Wiener used this to eliminate from the system of *Principia* the Axiom of Reducibility for propositional functions of two variables; he had written a doctoral thesis comparing the logics of Schröder and Russell.[62] Although Russell praised Sheffer's stroke, the logical connective not-both, he was not impressed by Wiener's reduction. Indeed, Russell would not have been able to accept it as a genuine analysis. Unlike Russell, Willard V.O. Quine in a major philosophical work *Word and Object* [1960: §53] regarded the reduction of the ordered pair as a paradigm for philosophical analysis.

Making no intensional distinctions Hausdorff [1914: 32ff,70ff] defined an ordered pair in terms of unordered pairs, formulated functions in terms of ordered pairs, and the ordering relations as collections of ordered pairs.[63] Hausdorff thus made both the Peano [1911,1913] and Wiener [1914] moves *in* mathematical practice, completing the reduction of functions to sets.[64] This may have been congenial to Peano, but not to Frege nor Russell, they having emphasized the primacy of functions. Following the pioneering work of Dedekind and Cantor Hausdorff was at the crest of a major shift in mathematics of which the transition from an intensional, rule-governed conception of function to an extensional, arbitrary one was a large part, and of which the eventual acceptance of the Power Set Axiom and the Axiom of Choice was symptomatic.

In his informal setting Hausdorff took the ordered pair of x and y to be

$$\{\{x, 1\}, \{y, 2\}\}$$

---

[61] In a letter accepting Russell's [1901] on the logic of relations for publication in his journal *Rivista*, Peano had pointedly written 'The classes of couples correspond to relations" (see Kennedy [1975: 214]) so that relations are extensionally assimilated to classes. Russell [1903: §98] argued that the ordered pair cannot be basic and would itself have to be given sense, which would be a circular or an inadequate exercise, and 'It seems therefore more correct to take an intensional view of relations . . .".

[62] See Grattan-Guinness [1975] for more on Wiener's work and his interaction with Russell.

[63] He did not so define arbitrary relations, for which there was then no mathematical use, but he was the first to consider general *partial* orderings, as in his maximality principle. Before Hausdorff and going beyond Cantor, Dedekind was first to consider non-linear orderings, e.g. in his remarkably early, axiomatic study [1900] of lattices.

[64] As to historical priority, Wiener's note was communicated to the Cambridge Philosophical Society, presented on 23 February 1914, while the preface to Hausdorff's book is dated 15 March 1914. Given the pace of book publication then, it is arguable that Hausdorff came up with his reduction first.

where 1 and 2 were intended to be distinct objects alien to the situation.[65] In any case, the now-standard definition is the more intrinsic

$$\{\{x\}, \{x, y\}\}$$

due to Kazimierz Kuratowski [1921: 171]. Notably, Kuratowski's definition is a by-product of his analysis of Zermelo's [1908] proof of the Well-Ordering Theorem.[66]

2.5 *Analytic and Projective Sets*

A decade after Lebesgue's seminal paper [1905], descriptive set theory emerged 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 had addressed Baire's functions with a intriguing use of CH. What is now known as a *Luzin set* is an uncountable set of reals whose intersection with any meager set is countable, and Luzin established: CH *implies that there is a Luzin set*.[67] This would become a paradigmatic use of CH, in that a recursive construction was carried out in $\aleph_1$ steps where at each state only countable many conditions have to be attended to, in this case by applying the Baire Category Theorem. Luzin showed that the characteristic function of his set escaped Baire's function classification, and Luzin sets have since become pivotal examples of "special sets" of reals.

In Moscow Luzin began an important seminar, and from the beginning a major topic 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 a 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.

Of the three regularity properties, Lebesgue measurability, the Baire property, and the perfect set property (cf. 2.3), the first two were immediate for the Borel sets. However, nothing had been known about the perfect set property beyond Cantor's own result that the closed sets have it and Bernstein's that with a well-ordering of the reals there is a set not having the property. Luzin's student Pavel Aleksandrov [1916] established the

---

[65]It should be pointed out that the definition works even when x or y is 1 or 2 to maintain the instrumental property ($*$) of ordered pairs.

[66]The general adoption of the Kuratowski pair proceeded through the major developments of mathematical logic: Von Neumann initially took the ordered pair as primitive but later noted (von Neumann [1925:VI]; [1928: 338];[1929: 227]) the reduction via the Kuratowski definition. Gödel in his incompleteness paper [1931: 176] also pointed out the reduction. In his footnote 18, Gödel blandly remarked: 'Every proposition about relations that is provable in [*Principia Mathematica*] is provable also when treated in this manner, as is readily seen." This stands in stark contrast to Russell's labors in *Principia* and his antipathy to Wiener's reduction of the ordered pair. Tarski [1931: n.3] pointed out the reduction and acknowledged his compatriot Kuratowski. In his recasting of von Neumann's system, Bernays [1937: 68] also acknowledged Kuratowski [1921] and began with its definition for the ordered pair. It is remarkable that Nicolas Bourbaki in his treatise [1954] on set theory still took the ordered pair as primitive, only later providing the Kuratowski reduction in the [1970] edition.

[67]Mahlo [1913a] also established this result.

groundbreaking result that *the Borel sets have the perfect set property*, so that "CH holds for the Borel sets".[68]

In the work that really began descriptive set theory another student of Luzin's, Mikhail Suslin, investigated the *analytic sets* following a mistake he had found in Lebesgue's paper.[69] Suslin [1917] formulated these sets in terms of an explicit operation $\mathcal{A}$[70] and announced two fundamental results: *a set B of reals is Borel iff both B and $\mathbb{R} - B$ are analytic;* and *there is an analytic set which is not Borel.*[71] 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 [1917] announced the regularity properties: *Every analytic set is Lebesgue measurable, has the Baire property, and has the perfect set property*, the last result attributed to Suslin.

Luzin and Sierpiński in their [1918] and [1923] provided proofs, and the latter paper was instrumental in shifting the emphasis toward the co-analytic sets, i.e. sets of reals X such that $\mathbb{R} - X$ is analytic. They used well-founded relations to provide a basic *tree representation* of co-analytic sets, one from which the main results of the period flowed, and it is here that well-founded relations entered mathematical practice.[72]

After the first wave in descriptive set theory brought about by Suslin [1917] and Luzin [1917] had crested, Luzin [1925a] and Sierpiński [1925] extended the domain of study to the *projective sets*. For $Y \subseteq \mathbb{R}^{k+1}$ and with ordered k-tuples defined from the ordered pair, the *projection of Y* is

$$pY = \{\langle x_1, ..., x_k \rangle \mid \exists y(\langle x_1, ..., x_k, y \rangle \in Y)\}.$$

Suslin [1917] had essentially noted that *a set of reals is analytic iff it is the projection of a Borel subset of $\mathbb{R}^2$*.[73] Luzin and Sierpiński took the geometric operation of projection to

[68]After getting a partial result [1914: 465ff], Hausdorff [1916] also showed, in essence, that the Borel sets have the perfect set property.

[69]Sierpiński [1950: 28ff] describes Suslin's discovery of the mistake.

[70]A *defining system* is a family $\{X_s\}_s$ of sets indexed by finite sequences s of natural numbers. The result of the Operation $\mathcal{A}$ on such a system is that set $\mathcal{A}(\{X_s\}_s)$ defined by:

$$x \in \mathcal{A}(\{X_s\}_s) \ \ iff \ \ (\exists f\colon \omega \to \omega)(\forall n \in \omega)(x \in X_{f|n})$$

where $f|n$ denotes that sequence determined by the first n values of f. For a set X of reals, X is *analytic iff* $X = \mathcal{A}(\{X_s\}_s)$ for some defining system $\{X_s\}_s$ consisting of closed sets of reals.

[71]Luzin [1925] traced the term "analytic" back to Lebesgue [1905] and pointed out how the original example of a non-Borel Lebesgue measurable set there was in fact the first example of a non-Borel analytic set.

[72]Building on the penultimate footnote, suppose that Y is a co-analytic set of reals, i.e. $Y = \mathbb{R} - X$ with $X = \mathcal{A}(\{X_s\}_s)$ for some closed sets X_s, so that for reals x,

$$x \in Y \ iff \ x \notin X \ iff \ (\forall f\colon \omega \to \omega)(\exists n \in \omega)(x \notin X_{f|n}).$$

For finite sequences s_1 and s_2 define: $s_1 \prec s_2$ *iff* s_2 is a proper initial segment of s_1. For a real x define: $T_x = \{s \mid x \in X_t \text{ for every initial segment } t \text{ of } s\}$. Then:

$$x \in Y \ iff \ \prec \ \text{on } T_x \text{ is a well-founded relation} ,$$

i.e. there is no infinite descending sequence $\ldots \prec s_2 \prec s_1 \prec s_0$. T_x is a tree (cf. 3.5). Well-founded relations were explicitly defined much later in Zermelo [1935]. Constructions recognizable as via recursion along a well-founded relation had already occurred in the proofs that the Borel have the perfect set property in Aleksandrov [1916] and Hausdorff [1916].

[73]Borel subsets of $\mathbb{R}^k$ are defined analogously to those of $\mathbb{R}$.

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 $A \subseteq \mathbb{R}^k$,

$$A \text{ is } \boldsymbol{\Sigma_1^1} \quad \textit{iff} \quad A = pY \text{ for some Borel set } Y \subseteq \mathbb{R}^{k+1} \text{ ,}$$

i.e. A is analytic[74] and for integers $n > 0$,

$$A \text{ is } \boldsymbol{\Pi_n^1} \quad \textit{iff} \quad \mathbb{R}^k - A \text{ is } \boldsymbol{\Sigma_n^1} \text{ ,}$$

$$A \text{ is } \boldsymbol{\Sigma_{n+1}^1} \quad \textit{iff} \quad A = pY \text{ for some } \boldsymbol{\Pi_n^1} \text{ set } Y \subseteq \mathbb{R}^{k+1} \text{ , and}$$

$$A \text{ is } \boldsymbol{\Delta_n^1} \quad \textit{iff} \quad A \text{ is both } \boldsymbol{\Sigma_n^1} \text{ and } \boldsymbol{\Pi_n^1} \text{ .}$$

Luzin [1925a] and Sierpiński [1925] 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 basic obstacles from the beginning. Luzin [1925a] emphasized that whether the $\boldsymbol{\Pi_1^1}$ sets, the co-analytic sets at the bottom of the hierarchy, have the perfect set property was a major question. In a confident and remarkably prophetic passage he declared that his efforts towards its resolution led him to a conclusion "totally unexpected", that "one does not know *and one will never know*" of the family of projective sets, although it has cardinality $2^{\aleph_0}$ and consists of "effective sets", whether every member has cardinality $2^{\aleph_0}$ if uncountable, has the Baire property, or is even Lebesgue measurable. Luzin [1925b] pointed out the specific problem of establishing whether the $\boldsymbol{\Sigma_2^1}$ sets are Lebesgue measurable. Both these difficulties were also pointed out by Sierpiński [1925]. This basic impasse in descriptive set theory was to remain for over a decade, to be surprisingly resolved by penetrating work of Gödel involving metamathematical methods (cf. 3.4).

2.6   *Equivalences and Consequences*

In this period AC and CH began to be explored no longer as underlying axiom and primordial hypothesis but as part of mathematics. Consequences were drawn and even equivalences established, and this mathematization, like the development of non-Euclidean geometry, led eventually to a deflating of metaphysical attitudes and attendant concerns about truth and existence.

Friedrich Hartogs [1915] established an equivalence result for AC, and this was the first substantial use of Zermelo's axiomatization after its appearance. The axiomatization had initially drawn ambivalent response among commentators,[75] especially those exercised by the paradoxes, and its assimilation by structuring sets and clarifying arguments began with such uses.

As noted in 1.3, Cardinal Comparability had become a concern for Cantor by the time of his *Beiträge* [1895]; Hartogs showed in Zermelo's system *sans* AC that *Cardinal Comparability implies that every set can be well-ordered.* Thus, an evident consequence of

---

[74] Analytic subsets of $\mathbb{R}^k$ are defined as for the case $k = 1$ in terms of a defining system consisting of closed subsets of $\mathbb{R}^k$.

[75] See Moore [1982: 3.3].

every set being well-orderable also implied that well-ordering principle, and this first "reverse mathematics" result established the equivalence of the well-ordering principle, Cardinal Comparability, and AC over the base theory.

Hartogs actually established without AC what is now called *Hartogs's Theorem*: *For any set M, there is a well-orderable set E not injectible into M.* Cardinal Comparability would then imply that M is injectible into E and hence is well-orderable. For the proof Hartogs first worked out a theory of ordering relations in Zermelo's system in terms of inclusion chains as in Zermelo's [1908] proof.[76] He then used Power Set and Separation to get the set M_W of well-orderable subsets of M and the set E of equivalence classes partitioning M_W according to order-isomorphism. Finally, he showed that E itself has an inherited well-ordering and is not injectible into M.[77] Reminiscent of Zermelo's subsumption of Russell's Paradox in the denial of a universal set, Hartogs's Theorem can be viewed as a subsumption of the Burali-Forti Paradox into the Zermelian setting.

The first explicit uses of AC mostly amounted to appeals to a well-ordering of the reals, Cantor's preoccupation. Those of Vitali [1905] and Bernstein [1908] were mentioned in 2.3, and Hausdorff's Paradox [1914,1914a], in 2.4. Georg Hamel [1905] constructed by transfinite recursion a basis for the reals as a vector space over the rationals; cited by Zermelo [1908, 114], this provided a useful basis for later work in analysis and algebra. These various results, jarring at first, broached how a well-ordering allows for a new kind of *arithmetical* approach to the continuum.

The full exercise of AC in ongoing mathematics first occurred in the pioneering work of Ernst Steinitz [1910] on abstract fields. This was the first instance of an emerging phenomenon in algebra and topology: the study of axiomatically given structures with the range of possibilities implicitly including the transfinite. Steinitz studied algebraic closures of fields and even had an explicit transfinite parameter in the transcendence degree, the number of indeterminates necessary for closure. Typical of the generality in the years to come was Hausdorff's [1932] result using well-orderings that *every vector space has a basis.* As algebra and topology developed however, such results as these came to be based on the maximal principles that Hausdorff had first broached (cf. 2.4) and began to dominate after the appearance of Zorn's Lemma [1935]. Explicit well-orderings seemed out of place at this level of organization, and Zorn's Lemma had the remarkable feature that its hypothesis was easily checked in most applications.

Poland since its reunification in 1918 featured an active school of mathematics establishing foundational results in mathematical logic, topology, and analysis, and at Warsaw Tarski and Kuratowski together with Sierpiński were making crucial contributions to set theory and the elucidation of its role in mathematics. The Polish school of mathematics carried out a penetrating investigation of the role of AC in set theory and analysis. Sierpiński's earliest publications, culminating in his survey [1918], not only dealt with specific constructions but showed how deeply embedded AC was in the informal development of cardinality, measure, and the Borel hierarchy (cf. 2.3), supporting Zermelo's

---

[76]This is better done in Kuratowski [1921]. The Hausdorff [1914] approach with an ordered pair could have been taken, but that only became standard later when more general relations were considered.

[77]As with Zermelo's Well-Ordering Theorem, textbooks usually establish Hartogs's Theorem after first introducing Replacement and (von Neumann) ordinals, and this amounts to a historical misrepresentation.

contention [1904: 516] that the axiom is applied "everywhere in mathematical deduction". Tarski [1924], explicitly building his work on Zermelo's system, provided several propositions of cardinal arithmetic equivalent to AC, most notably that $\mathfrak{m}^2 = \mathfrak{m}$ for every infinite cardinal $\mathfrak{m}$. Adolf Lindenbaum and Tarski in their [1926] gave further cardinal equivalents, some related to the Hartogs [1915] result, and announced that GCH, in the form that $\mathfrak{m} < \mathfrak{n} < 2^{\mathfrak{m}}$ holds for no infinite cardinals $\mathfrak{m}$ and $\mathfrak{n}$, implies AC. This study of consequences led to other choice principles, further implications and sometimes converses in a continuing cottage industry.[78]

The early mathematical study of AC extended to the issue of its independence. Abraham Fraenkel's first investigations [1922] directly addressed Zermelo's axioms, pointing out the need for the Replacement Axiom and attempting an axiomatization of the *definit* property for the Separation Axiom (cf. 3.1). The latter was motivated in part by the need to better articulate independence proofs for the various axioms. Fraenkel [1922a] came to the fecund idea of starting with urelements and some initial sets closing off under set-theoretic operations to get a model. For the independence of AC he started with urelements $a_n, \overline{a_n}$ for $n \in \omega$ and the set $A = \{\{a_n, \overline{a_n}\} \mid n \in \omega\}$ of unordered pairs and argued that for any set M in the resulting model there is a co-finite $A_M \subseteq A$ such that M is invariant if members of any $\{a_n, \overline{a_n}\} \in A_M$ are permuted. This immediately implies that there is no choice function for A in the model. Finally, Fraenkel argued that the model satisfies the other Zermelo axioms, except Extensionality because of the urelements.

Fraenkel's early model building emphasized the Zermelian generative framework, anticipated well-founded recursion, and foreshadowed the later play with models of set theory. That Extensionality was not to be had precluded settling the matter, but just as for the early models of non-Euclidean or finite geometries Fraenkel's achievement lay in stimulating interest in mathematical constructions despite relaxing some basic tenet. Fraenkel tried to develop his approach from time to time, but it needed the articulation that would come with the full espousal of the satisfaction relation. In the latter 1930s Lindenbaum and Andrzej Mostowski so cast and extended Fraenkel's work. Mostowski [1939] forged a method according to post-Gödelian sensibilities, bringing out the importance of groups of permutations leaving various urelements fixed, and the resulting models as well as later versions are now known as the *Fraenkel-Mostowski models.*

Even more than AC, Sierpí nski investigated CH, and summed up his researches in a monograph [1934]. He provided several notable equivalences to CH, e.g. (p.11) the plane $\mathbb{R}^2$ is the union of countably many curves, where a *curve* is a set of form $\{\langle x, y \rangle \mid y = f(x)\}$ or $\{\langle x, y \rangle \mid x = f(y)\}$ with f a real function.

Moreover, Sierpí nski presented numerous consequences of CH from the literature, one in particular implying a host of others: Mahlo [1913a] and Luzin [1914] had shown that CH *implies that there is a Luzin set*, an uncountable set of reals whose intersection with any meager set is countable (cf. 2.5). To state one consequence, say that a set X of reals has *strong measure zero iff* for any sequence $\epsilon_0, \epsilon_1, \epsilon_2, \ldots$ of positive reals there is a sequence of intervals $I_0, I_1, I_2, \ldots$ such that the length of I_n is less than ϵ_n for each n and $X \subseteq \bigcup_n I_n$. Borel [1919] conjectured that such sets are countable. However, Sierpí nski [1928] showed that *a Luzin set has strong measure zero.* Analogous to a Luzin set, a

---

[78] See Moore [1982], especially its 5.1, for other choice principles.

Sierpiński set is an uncountable set of reals whose intersection with any Lebesgue measure zero set is countable. Sierpi´nski [1924] showed that CH *implies that there is a Sierpiński set*, and emphasized [1934] an emerging duality between measure and category.

The subsequent work of Fritz Rothberger would have formative implications for the Continuum Problem. He [1938] observed that if both Luzin and Sierpi´nski sets exist, then they have cardinality $\aleph_1$, so that the joint existence of such sets of the cardinality of the continuum implies CH. Then in penetrating analyses of the work of Sierpinski and Hausdorff on gaps (cf. 2.1) Rothberger [1939,1948] considered other sets and implications between cardinal properties of the continuum *independent* of whether CH holds. It became newly clarified that absent CH one can still isolate uncountable cardinals $\leq 2^{\aleph_0}$ that gauge and delimit various recursive constructions, and this approach was to blossom half a century later in the study of *cardinal characteristics (or invariants)* of the continuum.[79]

These results cast CH in a new light, as a construction principle. Conclusions had been drawn from having a well-ordering of the reals, but one given by CH allowed for recursive constructions where at any stage only countably many conditions corresponding to as many reals had to be handled. The construction of a Luzin set was a simple recursive application of the Baire Category Theorem, and later constructions took advantage of the possibility of diagonalization at each stage. However, whereas the new constructions using AC, though jarring at first, were eventually subsumed as concomitant with the acceptance of the axiom and as expressions of the richness of possibility, constructions from CH clashed with that very sense of richness for the continuum. It was the *mathematical* investigation of CH that increasingly raised doubts about its truth and certainly its provability (cf. end of 3.4).

3   CONSOLIDATION

3.1   Ordinals and Replacement

In the 1920s fresh initiatives structured the loose Zermelian framework with new features and corresponding developments in axiomatics: von Neumann's work with ordinals and Replacement; the focusing on well-founded sets and the cumulative hierarchy; and extensionalization in first-order logic. Von Neumann effected a counter-reformation of sorts: The transfinite numbers had been central for Cantor but peripheral to Zermelo; von Neumann reconstrued them as *bona fide* sets, now called simply the *ordinals*, and established their efficacy by formalizing transfinite recursion.

Von Neumann [1923,1928], and before him Dimitry Mirimanoff [1917,1917a] and Zermelo in unpublished 1915 work,[80] isolated the now familiar concept of ordinal, with the basic idea of taking precedence in a well-ordering simply to be membership. Appealing to forms of Replacement Mirimanoff and Von Neumann then established the key instrumental property of Cantor's ordinal numbers for ordinals: *Every well-ordered set is order-*

---

[79]See Miller [1984] for more on special sets of reals and van Douwen [1984] as a trend setting paper for cardinal characteristics of the continuum. See Blass [2008] and Bartoszyński [2008] for recent work on cardinal characteristics.

[80]See Hallett [1984: 8.1].

isomorphic to exactly one ordinal with membership. Von Neumann in his own axiomatic presentation took the further step of ascribing to the ordinals the role of Cantor's ordinal numbers. Thus, like Kepler's laws by Newton's, Cantor's principles of generation for his ordinal numbers would be subsumed by the Zermelian framework. For this reconstrual of ordinal numbers and already to define the arithmetic of ordinals 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. With the ordinals in place von Neumann completed the restoration of the Cantorian transfinite by defining the *cardinals* as the *initial ordinals*, those ordinals not in bijective correspondence with any of its predecessors. Now, the infinite initial ordinals are denoted

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

so that ω is to be the set of natural numbers in the ordinal construal, and the identification of different intensions is signaled by

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

with the left being a von Neumann ordinal and the right being the Cantorian cardinal number.

Replacement has been latterly regarded as somehow less necessary or crucial than the other axioms, the purported effect of the axiom being only on large-cardinality sets. Initially, Abraham Fraenkel [1921,1922] and Thoralf Skolem [1923] had independently proposed adjoining Replacement to ensure that $E(a) = \{a, \mathcal{P}(a), \mathcal{P}(\mathcal{P}(a)), \dots\}$ be a set when a is the particular infinite set $Z_0 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots\}$ posited by Zermelo's Axiom of Infinity, since, as they pointed out, Zermelo's axioms cannot establish this. However, even $E(\emptyset)$ cannot be proved to be a set from Zermelo's axioms,[81] and if his axiom of Infinity were reformulated to accommodate $E(\emptyset)$, there would still be many finite sets a such that $E(a)$ cannot be proved to be a set.[82] Replacement serves to rectify the situation by admitting new infinite sets defined by "replacing" members of the one infinite set given by the Axiom of Infinity. In any case, the full exercise of Replacement is part and parcel of transfinite recursion, which is now used everywhere in modern set theory, and it was von Neumann's formal incorporation of this method into set theory, as necessitated by his proofs, that brought in Replacement.

That Replacement became central for von Neumann was intertwined with his taking of function, in its full extensional sense, instead of set as primitive and his establishing of a context for handling *classes*, collections not necessarily sets. He [1925,1928a] formalized the idea that a class is *proper*, i.e. not a set, exactly when it is in bijective correspondence with the entire universe, and this exactly when it is not an element of any class. This thus brought in another move from Cantor's 1899 correspondence with Dedekind (cf. 2.2). However, von Neumann's axiomatization [1925,1928] of function was complicated, and

---

[81] The union of $E(Z_0)$, with membership restricted to it, models Zermelo's axioms yet does not have $E(\emptyset)$ as a member.

[82] See Mathias [2001].

reverting to sets as primitive Paul Bernays (cf. his [1976]) recast and simplified von Neumann's system. Still, the formal incorporation of proper classes introduced a superstructure of objects and results distant from mathematical practice. What was to be inherited was a predisposition to entertain proper classes in the mathematical development of set theory, a willingness that would have crucial ramifications (cf. 3.6).

3.2   Well-Foundedness and the Cumulative Hierarchy

With ordinals and Replacement, set theory continued its shift away from pretensions of a general foundation toward a theory of a more definite subject matter, a process fueled by the incorporation of well-foundedness. Mirimanoff [1917] was the first to study the well-founded sets, and the later hierarchical analysis is distinctly anticipated in his work. But interestingly enough well-founded relations next occurred in the direct definability tradition from Cantor, descriptive set theory (cf. 2.5).

In the axiomatic tradition Fraenkel [1922], Skolem [1923] and von Neumann [1925] considered the salutary effects of restricting the universe of sets to the well-founded sets. Von Neumann [1929: 231,236ff] formulated in his functional terms the Axiom of Foundation, that every set is well-founded,[83] and defined the resulting hierarchy of sets in his system via transfinite recursion: In modern notation, the axiom, as is well-known, entails that the universe V of sets is stratified into cumulative ranks V_α, where

$$V_0 = \emptyset; \ V_{\alpha+1} = \mathcal{P}(V_\alpha); \ V_\delta = \bigcup_{\alpha<\delta} V_\alpha \text{ for limit ordinals } \delta \, ;$$

and

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

Von Neumann used this, *the cumulative hierarchy*, to establish the first relative consistency result in set theory via "inner models"; his argumentation in particular established the consistency of Foundation relative to Zermelo's axioms plus Replacement.

During this period mathematical logic gained new currency, and a tussle based on the different approaches of first- and second-order logic to set theory would lead to a substantial axiomatic development.[84] The prescient Skolem [1923] made the proposal of using for Zermelo's definite properties for the Separation Axiom those properties expressible in first-order logic with $\in$ as a binary relation symbol. After Leopold Löwenheim [1915] had broken the ground for model theory with his result about the satisfiability of a first-order sentence, Skolem [1920,1923] had located the result solidly in first-order logic and generalized it to the Löwenheim-Skolem Theorem: *If a countable collection of first-order sentences is satisfiable, then it is satisfiable in a countable domain.* That Skolem intended for set theory to be a first-order system without a privileged interpretation for $\in$ becomes evident in the initial application of the Löwenheim-Skolem Theorem to get

---

[83]$\forall x(x \neq \emptyset \longrightarrow \exists y \in x(x \cap y = \emptyset))$. This is von Neumann's Axiom VI4 in terms of sets. The term 'Foundation [Fundierung]" itself comes from Zermelo [1930].

[84]First-order logic is the logic of formal languages consisting of formulas built up from specified function and relation symbols using logical connectives and first-order quantifiers $\forall$ and $\exists$, quantifiers to be interpreted as ranging over the *elements* of a domain of discourse. Second-order logic has quantifiers to be interpreted as ranging over arbitrary subsets of a domain.

Skolem's Paradox: In first-order logic Zermelo's axioms are countable, Separation having become a *schema*, a schematic collection of axioms, one for each first-order formula; the theorem then implies the existence of countable models of the axioms although they entail the existence of uncountable sets. Skolem intended by this means to deflate the possibility of set theory becoming a foundation for mathematics. Exercised by this relativism and by the recent work of Fraenkel and von Neumann, Zermelo [1929] in his first publication in set theory in two decades proposed an axiomatization of his *definit* property in second-order terms. In direct response Skolem [1930] pointed out possible difficulties with this approach and reaffirmed his first-order formulation, completing the backdrop for a new axiomatic synthesis.

Zermelo in his remarkable [1930] offered his final axiomatization of set theory as well as a striking view of a procession of natural models that would have a modern resonance. While ostensibly a response to Skolem [1930], the dramatically new picture of sets in Zermelo [1930] reflects gained experience and the germination of ideas over a prolonged period. The main axiomatization incorporated Replacement but also the Axiom of Foundation. In contrast to Zermelo [1908a], while urelements continued to be allowed, Infinity was eschewed and Choice was regarded as part of the underlying logic. Concerning Separation and Replacement it becomes evident from how Zermelo proceeded that he regarded their applicability in a fully second-order context.

As described in above, Foundation in modern set theory ranks the universe of sets into the cumulative hierarchy $V = \bigcup_\alpha V_\alpha$. Zermelo substantially advanced this schematic generative picture with his inclusion of Foundation in an axiomatization. Replacement and Foundation focused the notion of set, with the first making possible the means of transfinite recursion and induction, and the second making possible the application of those means to get results about *all* sets. It is now almost banal that Foundation is the one axiom unnecessary for the recasting of mathematics in set-theoretic terms, but the axiom ascribes to membership the salient feature that distinguishes investigations specific to set theory as an autonomous field of mathematics. Indeed, it can be fairly said that modern set theory is at base a study couched in well-foundedness, the Cantorian well-ordering doctrines adapted to the Zermelian generative conception of sets.

In [1930] Zermelo described a range of models for set theory, each an initial segment of a cumulative hierarchy built on an initial set of urelements. Zermelo then established a categoricity of sorts for his axioms, one made possible by his second-order context. He showed that his models are characterized up to isomorphism by two cardinals, the number of their urelements and the height of their ordinals. Moreover, he established that if two models have the same number of urelements yet different heights, then one is isomorphic to an initial segment of the other's cumulative hierarchy. Grappling with Power Set and Replacement he characterized the heights of his models ("Grenzzahlen") as $\aleph_0$ or the *(strongly) inaccessible cardinals*, those uncountable regular cardinals κ that are strong limit, i.e. if $\lambda < \kappa$, then $2^\lambda < \kappa$.

Zermelo posited an endless procession of models, each a set in a next, advocating a dynamic view of sets that was a marked departure from Cantor's (and later, Gödel's) realist presumption of a fixed universe of sets. In synthesizing the sense of progression inherent in the new cumulative hierarchy picture and the sense of completion in the limit

numbers, the inaccessible cardinals, he promoted the crucial idea of internal models of set theory. The open-endedness of Zermelo's original [1908a] axiomatization had been structured by Replacement and Foundation, but he advanced a new open-endedness with an eternal return of models approaching Cantor's Absolute.

In the process, inaccessible cardinals became structurally relevant. Sierpiński-Tarski [1930] had formulated these cardinals arithmetically as those uncountable cardinals that are not the product of fewer cardinals each of smaller power and observed that they are weakly inaccessible — the first large cardinal concept, from Hausdorff [1908: 443] (cf. 2.4). Be that as it may, in the early model-theoretic investigations of set theory the inaccessible cardinals provided the natural models as envisioned by Zermelo. Moreover, strong large cardinal hypotheses emerging in the 1960s were to be formulated in terms of these initial segments of the cumulative hierarchy.[85]

The journal volume containing Zermelo's paper also contained Stanisław Ulam's seminal paper [1930] on *measurable cardinals*, the most important of all large cardinals. For a set S, U is a (non-principal) *ultrafilter over S iff* U is a collection of subsets of S containing no singletons, closed under the taking of supersets and finite intersections, and such that for any $X \subseteq S$, either $X \in U$ or $S - X \in U$. For a cardinal λ, an ultrafilter U is λ-*complete iff* for any $D \subseteq U$ of cardinality less than λ, $\bigcap D \in U$. Finally, an uncountable cardinal κ is *measurable iff* there is a κ-complete ultrafilter over κ. Thus, a measurable cardinal is a cardinal whose power set is structured with a two-valued measure having strong closure property.

Measurability embodied the first large cardinal confluence of Cantor's two legacies, the investigation of definable sets of reals and the extension of number into the transfinite: Distilled from measure-theoretic considerations related to Lebesgue measure, the concept *also* entailed inaccessibility in the transfinite. Moreover, the initial airing generated a problem that was to keep the spark of large cardinals alive for the next three decades: *Can the least inaccessible cardinal be measurable?* In the 1960s consequences of, and a structural characterization of, measurability were established that became fundamental in the setting structured by the new Zermelian emphasis on well-foundedness (cf. 3.6).

3.3   *First-Order Logic and Extensionalization*

The final structuring of set theory before it was to sail forth on its independent course as a distinctive field of mathematics was its full extensionalization in first-order logic.[86] However influential Zermelo's [1930] and despite his subsequent advocacy [1931,1935] of infinitary logic, his efforts to forestall Skolem were not to succeed, as stronger currents were at work in the direction of first-order formalization.

Hilbert effected a basic shift in the development of mathematical logic when he took Whitehead and Russell's *Principia Mathematica*, viewed it as an uninterpreted formalism, and made it an object of mathematical inquiry. The book [1928][87] by Hilbert and Wilhelm

---

[85] See Kanamori [2003: chap.5].

[86] See Goldfarb [1979] and Moore [1988a] for more on the emergence of first-order logic.

[87] The historical development is clarified by the fact that while this book was published in light of the developments of the 1920s, it has a large overlap with unpublished lecture notes for a 1917-8 course given by Hilbert

Ackermann reads remarkably like a recent text. In marked contrast to the formidable works of Frege and Russell with their forbidding notation and all-inclusive approach, it proceeded pragmatically and upward to probe the extent of structure, making those moves emphasizing forms and axiomatics typical of modern mathematics. After a complete analysis of sentential logic it distinguished and focused on first-order logic ("functional calculus", and later "(restricted) predicate calculus") as already the source of significant problems. Thus, while Frege and Russell never separated out first-order logic, Hilbert through his mathematical investigations established it as a subject in its own right.

Hilbert in the 1920s developed proof theory, i.e. metamathematics, and proposed his program of establishing the consistency of classical mathematics. The issues here gained currency because of Hilbert's preeminence, just as mathematics in the large had been expanded in the earlier years of the century by his reliance on non-constructive proofs and transcendental methods and his advocacy of new contexts. Through this expansion the full exercise of AC had become a mathematical necessity (cf. 2.6) and arbitrary functions, and so Power Set, had become implicitly accepted in the extensive investigation of higher function spaces.

Hilbert-Ackermann [1928: 65ff,72ff] raised two crucial questions directed at the further possibilities for first-order logic: the completeness of its axioms and the Decision Problem [Entscheidungsproblem]. These as well as Hilbert's program for securing consistency were to be decisively informed by penetrating work that for set theory eventually led to its first sophisticated metamathematical result, the relative consistency of AC and GCH.

Kurt Gödel (1906–1978), born when Zermelo was devising his proofs of the Well-Ordering Theorem, virtually completed the mathematization of logic by submerging meta-mathematical methods into mathematics. The main vehicle was of course the direct coding, "the arithmetization of syntax", in his celebrated Incompleteness Theorem [1931]. Establishing a fundamental distinction between what is *true* about the natural numbers and what is *provable*, this theorem transformed Hilbert's consistency program and led to the undecidability of the Decision Problem from Hilbert–Ackermann [1928] and the development of recursion theory. Gödel's work showed in particular that for a (schematically definable) collection of axioms A, its *consistency*, that from A one cannot prove a contradiction, has a formal counterpart in an arithmetical formula Con(A) about natural numbers. Gödel's "second" theorem asserts that if A is consistent and subsumes some elementary arithmetic of the natural numbers, then Con(A) cannot be proved from A. But starting an undercurrent, the earlier Completeness Theorem [1930] from his thesis answered affirmatively a Hilbert–Ackermann [1928] question about semantic completeness, clarified the distinction between the formal syntax and semantics of first-order logic, and secured its key instrumental property with the Compactness Theorem.

Tarski [1933,1935] then completed the mathematization of logic by providing his definition of truth, exercising philosophers to a surprising extent ever since. Through Hilbert-Ackermann [1928] and Gödel [1930] the satisfaction relation had been informal, and in that sense completeness could be said to have remained inadequately articulated. Tarski simply extensionalized truth in formal languages and provided a formal, *recursive* definition of the satisfaction relation in set-theoretic terms. This new response to a growing

_____

at Göttingen.

need for a mathematical framework became the basis for model theory, but thus cast into mathematics truth would leave behind any semantics in the real meaning of the word. Tarski's [1933] was written around the same time as his [1931], a seminal paper that highlights the thrust of his initiative. In [1931] Tarski gave a precise mathematical (that is, set-theoretic) formulation of the informal concept of a (first-order) definable set of reals, thus infusing the intuitive notion of definability into ongoing mathematics. This mathematization of intuitive or logical notions was accentuated by Kuratowski-Tarski [1931], where second-order quantification over the reals was correlated with the geometric operation of projection, beginning the process of explicitly wedding descriptive set theory to mathematical logic. The eventual effect of Tarski's [1933] mathematical formulation of (so-called) 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.[88]

In this process of extensionalization first-order logic came to be accepted as the canonical language because of its mathematical possibilities as epitomized by the Compactness Theorem, and higher-order logics became downgraded as the workings of the power set operation in disguise. Skolem's early suggestion for set theory was thus taken up generally, and again the ways of paradox were positively subsumed, as the negative intent of Skolem's Paradox gave way to the extensive, internal use of Skolem functions from the Löwenheim-Skolem Theorem in set-theoretic constructions.

3.4 Relative Consistency

Set theory was launched on an independent course as a distinctive field of mathematics by Gödel's construction of L [1938,1939] leading to the relative consistency of the Axiom of Choice and the Generalized Continuum Hypothesis. Synthesizing all that came before, Gödel built on the von Neumann ordinals as sustained by Replacement to formulate a relative Zermelian universe of sets based on logical definability, a universe imbued with a Cantorian sense of enumerative order.

Gödel's advances in set theory can be seen as part of a steady intellectual development. In a lecture [1933] on the foundations of mathematics Gödel propounded the axiomatic set theory "as presented by Zermelo, Fraenkel and von Neumann" as "a natural generalization of [Russell's simple] theory of types, or rather, what becomes of the theory of types if certain superfluous restrictions are removed." First, the types can be taken to be cumulative, and second, the process can be continued into the transfinite. As for how far this cumulative hierarchy of sets is to continue, "the first two or three [infinite] types already suffice to define very large [Cantorian ordinal numbers]" which can then serve to index the process, and so on. Implicitly referring to his incompleteness result Gödel noted that for a formal system S based on the theory of types a number-theoretic proposition can be constructed which is unprovable in S but becomes provable if to S is adjoined "the next higher type and the axioms concerning it."[89] Thus, although he never mentioned

---

[88]Incidentally, Tarski [1931] stated a result whose proof led to Tarski's well-known theorem [1951] that the elementary theory of real closed fields is decidable via the elimination of quantifiers.

[89]Gödel was evidently referring to propositions like Con(S). In a prescient footnote, 48a, to his incomplete-

Zermelo [1930], Gödel was entertaining its cumulative hierarchies but as motivated by the theory of types.

It is to this initiative, separately fueled by Zermelo and Gödel, that one can date how the formation of sets out of sets iterated into the transfinite as embodied by the cumulative hierarchy can be regarded as a motivation for the subject matter of set theory. In a notable inversion, what has come to be regarded as the *iterative conception* became a heuristic for motivating the axioms of set theory generally.[90] The iterative conception of sets, like Tarski's definition of truth, has exercised philosophers to a surprising extent with respect to extrinsic justifications. This has opened the door to a metaphysical appropriation in the following sense: It is as if there is some notion of set that is "there", in terms of which the axioms must find some further justification. But set theory has no particular obligations to mirror some prior notion of set arrived at *a posteriori*. Replacement and Choice for example do not quite "fit" the iterative conception,[91] but if need be, Replacement can be "justified" in terms of achieving algebraic closure of the axioms, a strong motivation in the work of Fraenkel and the later Zermelo, and choice can be "justified" in terms of Cantorian well-ordering doctrines or as a logical principle as Zermelo did.

In his first announcement [1938] about L Gödel described it as a hierarchy "which can be obtained by Russell's ramified hierarchy of types, if extended to include transfinite orders." Indeed, with L Gödel had refined the cumulative hierarchy of sets to a cumulative hierarchy of *definable* sets which is analogous to the orders of Russell's *ramified* theory. Gödel's further innovation was to continue the indexing of the hierarchy through *all* the ordinals. Von Neumann's canonical well-orderings would be the spine for a thin hierarchy of sets, and this would be the key to both the AC and CH results.

In a brief account [1939] 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 $\langle x, \in \rangle$.[92] Then define:

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

and the *constructible universe*

$$L = \bigcup\{L_\alpha \mid \alpha \text{ is an ordinal}\}.$$

Gödel pointed out that L "can be defined and its theory developed in the formal systems of set theory themselves." This is a remarkable understatement of arguably the central

---

ness paper [1931] Gödel had already written: "...the true reason for the incompleteness inherent in all formal systems of mathematics is that the formation of ever higher types can be continued into the transfinite ...while in any formal system at most denumerably many of them are available. For it can be shown that the undecidable propositions constructed here become decided whenever appropriate higher types are added (for example, the type ω to the system P [Peano Arithmetic]). An analogous situation prevails for the axiom system of set theory."

[90] Shoenfield [1967: 238ff][1977], Wang [1974a], Boolos [1971], and Scott [1974] motivate the axioms of set theory in terms of an iterative concept of set based on stages of construction. Parsons [1977] raises issues about this approach.

[91] See Boolos [1971] for Replacement and Scott [1974: 214] for Choice.

[92] For a first-order formula $\varphi(v_1, \ldots, v_n)$ in $\in$, $\varphi^x(x_1, \ldots, x_n)$ is the restriction of the formula to x, i.e. each $\forall y$ is replaced by $\forall y \in x$ and each $\exists y$ is replaced by $\exists y \in x$ (with these abbreviations having the expected formal articulation). A set $y \subseteq x$ is *first-order definable over* $\langle x, \in \rangle$ if there is a first-order formula $\varphi(v, v_1, \ldots, v_n)$ and $a_1, \ldots, a_n$ all in x such that $y = \{z \in x \mid \varphi^x(z, a_1, \ldots, a_n)\}$.

feature of the construction of L. L is a class definable in set theory via a transfinite recursion that could be based on the formalizability of $\text{def}(x)$, the definability of definability, which was later reaffirmed by Tarski's systematic definition of the satisfaction relation in set-theoretic terms. With this, one can formalize the *Axiom of Constructibility* $V = L$, i.e. $\forall x(x \in L)$. In modern parlance, an *inner model* is a transitive class[93] containing all the ordinals such that, with membership and quantification restricted to it, the class satisfies each axiom of ZF. In summary terms, what Gödel did was to show in ZF that L is an inner model, and moreover that L satisfies AC and CH. He thus established the relative consistency Con(ZF) implies Con(ZFC + GCH).

In the approach via $\text{def}(x)$ it is necessary to show that $\text{def}(x)$ remains unaltered when applied in L with quantifiers restricted to L. Gödel himself would never establish this *absoluteness of first-order definability* explicitly, preferring in his one rigorous published exposition of L to take an approach that avoids $\text{def}(x)$ altogether.

In his monograph [1940], based on 1938 lectures, Gödel provided a specific, formal presentation of L in a class-set theory emanating from that of Paul Bernays (cf. [1976]), a theory based in turn on a theory of von Neumann [1925]. Using eight binary operations producing new classes from old, Gödel generated L set by set via transfinite recursion. This veritable "Gödel numbering" with ordinals bypassed $\text{def}(x)$ 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,[94] 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 $V = L \to$ GCH, and, in order to apply this implication within L, that L as defined within L with quantifiers restricted to L is again L itself. This latter follows from the aforementioned absoluteness of $\text{def}(x)$, and in [1940] Gödel gave an alternate proof based on the absoluteness of his eight binary operations.

Gödel's argument for $V = L \to$ GCH rests, as he himself wrote in [1939], 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 [1939] specifically established:

(∗)   For infinite α, every constructible subset of L_α
                  belongs to some L_β for a β of the same cardinality as α.

It is straightforward to show that for infinite α, L_α has the same cardinality as that of α. It follows from (∗) that in the sense of L, the power set of $L_{\aleph_\alpha}$ is included in $L_{\aleph_{\alpha+1}}$, and so GCH follows in L. To establish (∗), Gödel actually iterated the Skolem closure procedure,

---

[93] A class C is *transitive* if members of members of C are themselves members of C, so that C is "closed under membership".

[94] That is, those first-order formulas in which all the quantifiers can be rendered as $\forall x \in y$ and $\exists x \in y$.

and made the first use of the now familiar Mostowski collapse (cf. 3.6). In an incisive 1939 lecture Gödel announced the version of $(*)$ for countable α as the crux of the consistency proof of CH and asserted that "this fundamental theorem constitutes the corrected core of the so-called Russellian axiom of reducibility."[95] Thus, Gödel established another connection between L and Russell's ramified theory of types. But while Russell had to *postulate* his ill-fated Axiom of Reducibility for his finite orders, Gödel was able to *derive*, with an important use of Replacement, an analogous form for his transfinite hierarchy that asserts that the types are delimited in the hierarchy of orders.

The synthesis at L extended to the resolution of difficulties in descriptive set theory (cf. end of 2.5). Gödel [1938] announced, in modern terms: *If $V = L$, then (a) there is a* Δ^1_2 *set of reals that is not Lebesgue measurable, and (b) there is a a* Π^1_1 *set of reals without the perfect set property.* Thus, the descriptive set theorists were confronting an obstacle insurmountable in ZFC! Gödel [1938] listed each of these impossibility results on an equal footing with his AC and GCH results. Unexpected, they were the first instances of metamathematical methods resolving outstanding mathematical problems that exhibited no prior connection to such methods. When eventually confirmed and refined, the results were seen to turn on a Σ^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 a Π^1_1 set (cf. 2.5).[96]

Set theory had progressed to the point of establishing, in addition to a consistent resolution of CH, a consistent possibility for a definable well-ordering of the reals as Cantor had wanted, one that synthesizes the two historical sources of well-foundedness. Put into a broader historical context, formal definability was brought into descriptive set theory by Tarski [1931], and by Kuratowski-Tarski [1931] and Kuratowski [1931] which pursued the basic connection between existential number quantifiers and countable unions and between existential real quantifiers and projection and used these "logical symbols" to aid in the classification of sets in the Borel and projective hierarchies. Gödel's results (a) and (b) constitute the first real synthesis of abstract and descriptive set theory, in that the axiomatic framework is brought to bear on the investigation of definable sets of reals.

Gödel brought into set theory a method of construction and argument which affirmed several features of its axiomatic presentation. Most prominently, Gödel showed how first-order definability can be formalized and used in a transfinite recursive construction to establish striking new mathematical results. This significantly contributed to a lasting ascendancy for first-order logic which beyond its *sufficiency* as a logical framework for

---

[95]See Gödel [1939a: 141].

[96]When every real is in L, this Σ^1_2 well-ordering is also Δ^1_2 and does not satisfy Fubini's Theorem for Lebesgue measurable subsets of the plane, and this is one way to confirm (a). What may have been Gödel's original argument for (b) is given in Kanamori [2003: 170].

Texts establish (b) indirectly via the Kondô Π^1_1 Uniformization Theorem, and this leads to a historical point about Gödel the working mathematician. As 1938 correspondence with von Neumann makes evident, Gödel was working on one-to-one continuous images of Π^1_1 sets, and his [1938] actually states the results (a) and (b) in these terms. In a 1939 letter, von Neumann informed Gödel of Kondô [1939], the paper containing the uniformization result, from which it is immediate that the Σ^1_2 sets are exactly the one-to-one continuous images of Π^1_1 sets. In a replying letter to von Neumann of 20 March 1939 Gödel wrote: 'The result of Kondô is of great interest to me and will definitely allow an important simplification in the consistency proof of [(a)] and [(b)] of the attached offprint.' See Gödel [2003] for the Gödel–von Neumann correspondence.

mathematics was seen to have considerable *operational efficacy*. Moreover, Gödel's construction buttressed the incorporation of Replacement and Foundation into set theory. Replacement was immanent in the arbitrary extent of the ordinals for the indexing of L and in its formal definition via transfinite recursion. As for Foundation, underlying the construction was the well-foundedness of sets. Gödel in a footnote to his 1939 note wrote: "In order to give A [the axiom $V = L$] an intuitive meaning, one has to understand by 'sets' all objects obtained by building up the simplified hierarchy of types on an empty set of individuals (including types of arbitrary transfinite orders)."

How Gödel transformed set theory can be broadly cast as follows: On the larger stage, from the time of Cantor, sets began making their way into topology, algebra, and analysis so that by the time of Gödel, they were fairly entrenched in the structure and language of mathematics. But how were sets viewed among set *theorists*, those investigating sets as such? Before Gödel, the main concerns were what sets *are* and how sets and their axioms can serve as a reductive basis for mathematics. Even today, those preoccupied with ontology, questions of mathematical existence, focus mostly upon the set theory of the early period. After Gödel, the main concerns became what sets *do* and how set theory is to advance as an autonomous field of mathematics. The cumulative hierarchy picture was in place as subject matter, and the metamathematical methods of first-order logic mediated the subject. There was a decided shift toward epistemological questions, e.g. what can be proved about sets and on what basis.

As a pivotal figure, what was Gödel's own stance? What he *said* would align him more with his predecessors, but what he *did* would lead to the development of methods and models. In a 1944 article on Russell's mathematical logic, in a 1947 article on Cantor's continuum problem (and in a 1964 revision), and in subsequent lectures and correspondence, Gödel articulated his philosophy of "conceptual realism" about mathematics. He espoused a staunchly objective "concept of set" according to which the axioms of set theory are true and are descriptive of an objective reality schematized by the cumulative hierarchy. Be that as it may, his actual mathematical work laid the groundwork for the development of a range of models and axioms for set theory. Already in the early 1940s Gödel worked out for himself a possible model for the negation of AC, and in a 1946 address he described a new inner model, the class of ordinal definable sets.

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*. Already in that 1946 address he suggested[97] the consideration of "stronger and stronger axioms of infinity," and reflection as follows: "Any proof of a set-theoretic theorem in the next higher system above set theory (i.e. any proof involving the concept of truth ...) is replaceable by a proof from such an axiom of infinity." This ties in with the class of all ordinal numbers cast as Cantor's Absolute: A largeness property ascribable to the class might be used to derive some set-theoretic proposition; but any such property confronts the antithetical contention that the class is mathematically incomprehendable, fostering the synthetic move to a large cardinal posited with the property.

In the expository article [1947] on the Continuum Problem Gödel presumed that CH would be shown independent from ZF and speculated more concretely about possibilities

---

[97] See Gödel [1990: 151].

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' ", citing Zermelo [1930] and echoing its theme. In an unpublished footnote 20 toward a 1966 revision of [1947] Gödel was to acknowledge[98] "extremely strong axioms of infinity of an entirely new kind", generalizations of properties of ω "supported by strong arguments from analogy". This heuristic of generalization ties in with Cantor's view of the finite and transfinite as unitary, with properties like inaccessibility and measurability technically satisfied by ω being too accidental were they not also ascribable to higher cardinals through the uniformity of the set-theoretic universe.[99]

Gödel [1947] at the end actually argued against CH by drawing on the work of Sierpinski and others (cf. 2.6) to exhibit six "paradoxical" consequences. One of them is the existence of a Luzin set of cardinality of the continuum, and three others actually follows from the existence of such a set. This brought to the fore Gödel's stance about what is true in set theory. Whether CH is proved consistent or independent of ZFC, he believed in a "truth of the matter" both from the point of view of intuitions about the continuum and from his philosophical standpoint. That CH is implausible because it led to various implausible conclusions became a prominent attitude, one that would stay with set theory through its subsequent development.

3.5   *Combinatorics*

Gödel's construction of *L* was both a culmination in all major respects of the early period in set theory and a source for much that was to follow. But for quite some time it was to remain an isolated monument in the axiomatic tradition. No doubt the intervening years of war were a prominent factor, but there was a continuing difficulty in handling definability within set theory and a stultifying lack of means for constructing models of set theory to settle issues of consistency and independence. It would take a new generation versed in emerging model-theoretic methods to set the stage for the next major methodological advances.

In the mean time, the direct investigation of the transfinite as extension of number was advanced, gingerly at first, by a new initiative. The seminal results of *infinite combinatorics* were established beginning in the 1930s. As for algebra and topology, it was natural to extend concepts over the transfinite, and significantly, the combinatorics that would have the most bearing there had their roots in the mathematization of logic.

Frank Ramsey [1930] established a special case of the Decision Problem of Hilbert-Ackermann [1928], the decidability of validity for the $\exists\forall$ formulas with identity. For this purpose he established a basic generalization of the pigeonhole principle. In a move that transcended purpose and context he also established an infinite version implicitly applying the now familiar Kőnig's Lemma for trees. Stated more generally for graphs by Dénes Kőnig [1927: 121] the lemma had also figured implicitly in Löwenheim [1915]. In

---

[98] See Gödel [1990: 260ff].

[99] See Wang [1974: §§1,4] for more on Gödel's view on heuristics as well as the criteria of *intrinsic necessity* and *pragmatic success* for accepting new axioms.

what follows we affirm the general terminology for the formulation of Ramsey's results and then Kőnig's Lemma, anticipating extensions into the transfinite.

For ordinals α, β, and δ and $n \in \omega$ the *partition property*

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

is the assertion 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 showed that for any k, n, and r all in ω, there is a $m \in \omega$ such that $m \longrightarrow (k)^n_r$. Skolem [1933] sharpened Ramsey's argument and thereby lowered the possibilities for the m's, but to this day the least such m's, the *Ramsey numbers*, have not been determined except in the simplest cases. Ramsey's infinite version is: $\omega \longrightarrow (\omega)^n_r$ for every $n, r \in \omega$. This partition property and its variants have been adapted to a variety of situations, and today *Ramsey theory* is a thriving field of combinatorics.[100]

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 incomparable elements. A *branch* of T is a maximal chain, and a *cofinal branch* of T is a branch with elements at every non-empty level of T. Finally, for a cardinal κ, a κ-*tree* is a tree of height κ each of whose levels has cardinality less than κ, and

κ has the *tree property iff* every κ-tree has a cofinal branch .

Finite trees of course are quite basic to current graph theory and computer science. With infinite trees the concerns are rather different, typically involving cofinal branches. Kőnig's Lemma asserts that *ω has the tree property.*

The first systematic study of transfinite trees was carried out in Đuro Kurepa's thesis [1935], and several properties emerging from his investigations, particularly for ω_1-trees, would later become focal in the combinatorial study of the transfinite.

An *Aronszajn tree* is an ω_1-tree without a cofinal branch ,

i.e. a counterexample to the tree property for ω_1. Kurepa [1935: §9,thm 6] gave Nachman Aronszajn's result that *there is an Aronszajn tree.*

A *Suslin tree* is an ω_1-tree with no uncountable chains or antichains .

Kurepa [1935: appendix] reduced a hypothesis growing out of a problem of Suslin [1920] about the characterizability of the order type of the reals to a combinatorial property of ω_1 as follows: *Suslin's Hypothesis holds iff there are no Suslin trees.*

A *Kurepa tree* is an ω_1-tree with at least ω_2 cofinal branches ,

and *Kurepa's Hypothesis* deriving from Kurepa [1942: 143], is the assertion that such trees exist. Much of this would be rediscovered, and both Suslin's Hypothesis and Kurepa's Hypothesis would be resolved three decades later with the advent of forcing, several of the resolutions in terms of large cardinal hypotheses.[101]Kurepa's work also anticipated

---

[100]See the text Graham-Rothschild-Spencer [1990] and the compendium Nešetřil-R̆odl [1990] for the recent work on Ramsey Theory.

[101]See Todorčević [1984] for a wide-ranging account of transfinite trees.

another development from a different quarter:

Paul Erdős, although an itinerant mathematician for most of his life, was the prominent figure of a strong Hungarian tradition in combinatorics, and through some seminal results he introduced major initiatives into the detailed combinatorial study of the transfinite. Erdős and his collaborators simply viewed the transfinite numbers as a combinatorially rich source of intrinsically interesting problems, the concrete questions about graphs and mappings having a natural appeal through their immediacy. One of the earliest advances was Erdős-Tarski [1943] which concluded enticingly with an intriguing list of six combinatorial problems, the positive solution to any, as it was to turn out, amounting to the existence of a large cardinal. In a footnote various implications were noted, one of them being essentially that *for inaccessible κ, the tree property for κ implies $\kappa \longrightarrow (\kappa)^2_2$*, generalizing Ramsey's $\omega \longrightarrow (\omega)^2_2$ and making explicit the Kőnig Lemma property needed. While Kurepa was investigating distinctive properties of uncountable trees, Erdős-Tarski [1943] was evidently motivated by strong properties of ω to formulate direct combinatorial generalizations to inaccessible cardinals by analogy.[102] The situation would become considerably clarified, but only two decades later.[103]

The detailed investigation of partition properties began in earnest in the 1950s, with Erdős and Richard Rado's [1956] being representative. For a cardinal κ, let κ^+ denote its successor cardinal and set $\exp_0(\kappa) = \kappa$ and $\exp_{n+1}(\kappa) = 2^{\exp_n(\kappa)}$. What became known as *the* Erdős-Rado Theorem asserts: *For any infinite cardinal κ and $n \in \omega$,*

$$\exp_n(\kappa)^+ \longrightarrow (\kappa^+)^{n+1}_\kappa .$$

This was established using the basic tree argument underlying Ramsey's results, whereby a homogeneous set is not constructed recursively, but a tree is constructed such that its branches provide homogeneous sets, and a counting argument ensures that there must be a homogeneous set of sufficient cardinality. Kurepa [1937,1939] in effect had actually established the case $n = 1$ and shown that $\exp_1(\kappa)^+$ was the least possible. The $\exp_n(\kappa)^+$ was also shown to be the least possible in the general case, and so unlike for the Ramsey numbers in the finite case an exact analysis was quickly achieved in the transfinite. This was to be a recurring phenomenon, that the gross features of transfinite cardinality make its combinatorics actually easier than in the analogous finite situation. And notably, iterated cardinal exponentiation figured prominently, so that shedding deeper concerns the power set operation became further domesticated in the arithmetic of combinatorics. In fact, assuming GCH simplified results and formulations, and this was often done, as in Erdős, András Hajnal, and Rado's [1965], representative of the 1960s. Increasingly, a myriad of versions have been investigated in the larger terrain without GCH.[104]

Still among the Hungarians, Géza Fodor [1956] established the now familiar *regressive*

---

[102]On the other hand, Kurepa [1935: §10.3] did ask whether every inaccessible cardinal has the tree property, a question only resolved by work of Hanf (cf. 3.6).

[103]The details of implications asserted at the end of Erdős-Tarski [1943] were worked out in an influential seminar conducted by Tarski and Mostowski at Berkeley in 1958-9, and appeared in Erdős-Tarski [1961].

[104]The results of Erdős-Hajnal-Rado [1965] were extended in Byzantine detail to the general situation without GCH by the book Erdős-Hajnal-Máté-Rado [1984]. See Hajnal–Larson [2008] for recent work on partition relations.

function lemma for stationary sets: *If λ regular and uncountable, S is stationary in λ,*[105] *and $f: S \rightarrow \lambda$ is regressive (i.e. $f(\xi) < \xi$ for $\xi \in S$), then there is an $\alpha < \lambda$ such that $\{\xi \in S \mid f(\xi) = \alpha\}$ is stationary in λ.* It is a basic fact and a simple exercise now, but then it was the culmination of a progression of results beginning with a special case established in Aleksandrov-Urysohn [1929] and getting to the right largeness notion of stationarity. The contrast with how the lemma's earlier precursors were considered difficult and even paradoxical is striking, indicative of both the novelty of uncountable cofinality and the great leap forward that set theory has made.

3.6 Model-Theoretic Methods

Model theory began in earnest with the method of diagrams of Abraham Robinson's thesis [1951] and the related method of constants from Leon Henkin's thesis which gave a new proof [1949] of the Gödel Completeness Theorem. Tarski had set the stage with his definition of truth and more generally his casting of formal languages and structures in set-theoretic terms, and with him established at the University of California at Berkeley a large part of the development in the 1950s and 1960s would take place there. The construction of models freely used transfinite methods and soon led to new questions in set theory, but also set theory was to be decisively advanced by the infusion of model-theoretic methods.

The first relevant result was a generalization accreditable to Mostowski [1949] of the Mirimanoff–von Neumann result that every well-ordered set is order-isomorphic to exactly one ordinal with membership. A binary relation R is *extensional on X iff* for any $x \neq y$ both in X there is a $z \in X$ such that $z R x$ iff $\neg z R y$. Recall that x is *transitive iff* members of members of x are themselves members of x, so that x is "closed under membership". *If R is a well-founded relation on a set X and extensional on X, there is a unique isomorphism of $\langle X, R \rangle$ onto $\langle T, \in \rangle$ where T is transitive, i.e. a bijection $\pi: X \rightarrow T$ such that for any $x, y \in X$, $x R y$ iff $\pi(x) \in \pi(y)$.* T is the *transitive collapse* of X, and π the *collapsing isomorphism.* Thus, the linearity of well-orderings has been relaxed to well-foundedness and an analogue of the Axiom of Extensionality, and the transitive sets become canonical representatives as ordinals are for well-orderings. Gödel [1939: 222] had made the first substantial use of the transitive collapse; Mostowski [1949: 147] established the general result much later; and John Shepherdson [1951: 171] in a structured setting that brought out a further necessary hypothesis for classes X: R is *set-like,* i.e. for any $x \in X$, $\{y \mid y R x\}$ is a set. The initial applications in Mostowski [1949] and Shepherdson [1953] were to establish the independence of the assertion that there is a transitive set M which with $\in$ restricted to it is a model of set theory. While the Mirimanoff–von Neumann result was basic to the analysis of number in the transfinite, the transitive collapse result grew in significance from specific applications and came to epitomize how well-foundedness made possible a coherent theory of models of set theory.

The relationship between ZFC and Bernays-Gödel (BG), the class-set theory brought

---

[105] A set $C \subseteq \lambda$ is *closed unbounded in λ iff* C contains its limit points, i.e. those $0 < \alpha < \lambda$ such that $C \cap \alpha = \alpha$, and is cofinal, i.e. $\bigcup C = \lambda$. A set $S \subseteq \lambda$ is *stationary in λ iff* for any C closed unbounded in λ, $S \cap C$ is not empty.

into prominence by its use in Gödel [1940], was clarified during this period. As analyzed in Hao Wang [1949], BG can be construed as an extension of ZFC via the introduction of class variables intended to range over subcollections of *V* and correlative axioms, together with a comprehension principle asserting for each formula φ with just one set variable *v* free and no class variables quantified that there is a corresponding class $\{v \mid \varphi\}$. Ilse Novak [1950] and Barkley Rosser and Wang in their [1950] established that *if* ZFC *is consistent, then so is* BG by providing model-theoretic interpretations of BG relative to ZFC. Then Mostowski [1950] showed that BG is a conservative extension of ZF, i.e. any sentence σ without class variables provable in BG is already provable in ZFC. Subsequently, Joseph Shoenfield [1954] showed how to convert directly a proof of such a σ in BG into a proof in ZFC. These results reinforced the impression that, as far as the axiomatic tradition from Zermelo through Gödel is concerned, there is essentially one set theory, and one might as well work in the parsimonious ZFC.

Shepherdson [1951,1952,1953] studied "inner" models of set theory, with [1952] giving a rigorous first-order account of the results of Zermelo [1930]. The term is now reserved for the case mentioned in 3.4: A transitive class containing all the ordinals such that, with membership and quantification restricted to it, the class satisfies each axiom of ZF. The archetypal inner model is Gödel's *L*, and $L \subseteq M$ for any inner model *M* since the construction of *L* carried out in *M* is again *L*. Because of this Shepherdson [1953] observed that the relative consistency of hypotheses like the negation of CH cannot be established via inner models.

Hajnal [1956,1961] and Azriel Levy [1957,1960] developed generalizations of *L* that were to become basic in a richer setting. For a set *A*, Hajnal formulated the *constructible closure L(A) of A*, i.e. the smallest inner model *M* such that $A \in M$, and Levy formulated the *class L[A] of sets constructible relative to A*, i.e. the smallest inner model *M* such that for every $x \in M$, $A \cap x \in M$.[106] *L(A)* realizes the algebraic idea of building up a model starting from a set of generators, and *L[A]* the idea of building up a model using *A* construed as a predicate. *L(A)* may not satisfy AC since e.g. it may not have a well-ordering of *A*, yet *L[A]* always satisfies that axiom. This distinction was only to surface later, as both Hajnal and Levy took *A* to be a set of ordinals, when *L(A)* = *L[A]*. Hajnal and Levy (and also Shoenfield [1959], who formulated a special version of Levy's construction) used these models to establish conditional independence results of the sort: if the failure of CH is consistent, then so is that failure together with $2^\lambda = \lambda^+$ for sufficiently large cardinals λ.

After Richard Montague [1956,1961] applied reflection phenomena to investigate finite axiomatizability for set theory, Levy [1960a,1960b] also formulated reflection principles and established their broader significance. The *Reflection Principle* for ZF, drawn from Montague [1961: 99] and from Levy [1960a: 234], asserts: *For any (first-order) formula $\varphi(v_1, \ldots, v_n)$ and any ordinal β, there is a limit ordinal $\alpha > \beta$ such that for any $x_1, \ldots, x_n \in$*

---

[106]To formulate *L(A)*, define: $I_0(A)$ = the smallest transitive set $\supseteq \{A\}$ (to ensure that the resulting class is transitive); $L_{\alpha+1} = \text{def}(L_\alpha(A))$ (where def is as in 3.4); $L_\delta = \bigcup_{\alpha<\delta} L_\alpha(A)$ for limit $\delta > 0$; and finally $L(A) = \bigcup_\alpha L_\alpha(A)$. To formulate *L[A]*, first let $\text{def}^A(x)$ denote the collection of subsets of *x* first-order definable over $\langle x, \in, A \cap x \rangle$, i.e. $A \cap x$ is now allowed as a predicate in the definitions. Then define: $L_0[A] = \emptyset$; $L_{\alpha+1}[A] = \text{def}^A(L_\alpha[A])$; $L_\delta[A] = \bigcup_{\alpha<\delta} L_\alpha[A]$ for limit $\delta > 0$; and finally $L[A] = \bigcup_\alpha L_\alpha[A]$.

V_α,

$$\varphi[x_1, \ldots, x_n] \text{ iff } \varphi^{V_\alpha}[x_1, \ldots, x_n],$$

i.e. the formula holds exactly when it holds with all the quantifiers restricted to V_α. Levy showed that this schema is equivalent to the conjunction of the Replacement schema together with Infinity in the presence of the other axioms of ZF. Moreover, he formulated reflection principles in local form that characterized cardinals in the Mahlo hierarchy (2.4), conceptually the least large cardinals after the inaccessible cardinals. Then William Hanf and Dana Scott in their [1961] posited analogous reflection principles for higher-order formulas, leading to what are now called the *indescribable cardinals*, and eventually Levy [1971] carried out a systematic study of the sizes of these cardinals.[107] The model-theoretic reflection idea thus provided a coherent scheme for viewing the bottom of an emerging hierarchy of large cardinals as a generalization of Replacement and Infinity, one that resonates with the procession of models in Zermelo [1930]. The heuristic of reflection had been broached in 1946 remarks by Gödel (cf. 3.4), and another point of contact is the formulation of the concept of *ordinal definable set* in those remarks. With the class of ordinal definable sets formalized by OD $= \bigcup_\alpha \text{def}(V_\alpha)$, the adequacy of this definition is based on some form of the Reflection Principle for ZF. With tc(y) denoting the smallest transitive set $\supseteq y$, let HOD $= \{x \mid \text{tc}(\{x\}) \subseteq \text{OD}\}$. the class of *hereditarily ordinal definable sets*. As adumbrated by Gödel, HOD is an inner model in which AC, though not necessarily CH, holds. The basic results about this inner model were to be rediscovered several times.[108] In these several ways reflection phenomena both as heuristic and as principle became incorporated into set theory, bringing to the forefront what was to become a basic feature of the study of well-foundedness.

The set-theoretic generalization of first-order logic allowing transfinitely indexed logical operations was to lead to the solution of the problem of whether the least inaccessible cardinal can be measurable (cf. 3.2). Extending familiarity by abstracting to a new domain Tarski [1962] defined the *strongly compact* and *weakly compact* cardinals by ascribing natural generalizations of the key compactness property of first-order logic to the corresponding infinitary languages. These cardinals had figured in Erdős-Tarski [1943] (cf. 3.5) in combinatorial formulations that was later seen to imply that *a strongly compact cardinal is measurable,* and *a measurable cardinal is weakly compact.* Tarski [1962] pointed out that his student William Hanf (cf. [1964]) established, using the satisfaction relation for infinitary languages, that *there are many inaccessible cardinals (and Mahlo cardinals) below a weakly compact cardinal. A fortiori, the least inaccessible cardinal is not measurable.* This breakthrough was the first result about the size of measurable cardinals since Ulam's original paper [1930] and was greeted as a spectacular success for metamathematical methods. Hanf's work radically altered size intuitions about problems coming to be understood in terms of large cardinals and ushered in model-theoretic methods into the study of large cardinals beyond the Mahlo cardinals.[109]

Weak compactness was soon seen to have a variety of characterizations; most notably in terms of 3.5, κ is weakly compact *iff $\kappa \to (\kappa)^2_2$ iff $\kappa \to (\kappa)^n_\lambda$ for every $n \in \omega$ and $\lambda < \kappa$ iff*

---

[107] See Kanamori [2003: §6].

[108] See Myhill-Scott [1971], especially p. 278.

[109] See Kanamori [2003: §4] for these results about strongly and weakly compact cardinals.

κ is inaccessible and has the tree property. Erdős and Hajnal [1962] noted that the study of stronger partition properties had progressed to the point where a combinatorial proof that the least inaccessible cardinal is not measurable could have been given before Hanf came to his argument. However, model-theoretic methods quickly led to far stronger conclusions, particularly through the connection that had been made in Ehrenfeucht-Mostowski [1956] between partition properties and *sets of indiscernibles*.[110]

The concurrent emergence of the *ultraproduct construction* in model theory set the stage for the development of the modern theory of large cardinals in set theory. With a precursor in Skolem's [1933a,1934] construction of a non-standard model of arithmetic the ultraproduct construction was brought to the forefront by Tarski and his students after Jerzy Łoś's [1955] adumbration of its fundamental theorem. This new method of constructing concrete models brought set theory and model theory even closer together in a surge of results and a lasting interest in ultrafilters. Measurable cardinals had been formulated (cf. 3.2) in terms of ultrafilters construed as two-valued measures; Jerome Keisler [1962] struck on the idea of taking the ultrapower of a measurable cardinal κ by a κ-complete ultrafilter over κ to give a new proof of Hanf's result, seeing the crucial point that the completeness property led to a well-founded, and so in his case well-ordered, structure.

Then Scott [1961] made the further, crucial move of taking the ultrapower of the universe V itself by such an ultrafilter. The full exercise of the transitive collapse as a generalization of the correlation of ordinals to well-ordered sets now led to an inner model $M \neq V$ and an elementary embedding $j: V \to M$.[111] With this Scott established: *If there is a measurable cardinal, then $V \neq L$.* Large cardinal hypotheses thus assumed a new significance as a means for maximizing possibilities away from Gödel's delimitative construction. Also, the Cantor-Gödel realist view of a fixed set-theoretic universe notwithstanding, Scott's construction fostered the manipulative use of inner models in set theory. The construction provided one direction and Keisler [1962a] the other of a new characterization that established a central structural role for measurable cardinals: *There is an elementary embedding $j: V \to M$ for some inner model $M \neq V$ iff there is a measurable cardinal.* This result is not formalizable in ZFC because of the use of the satisfaction relation and the existential assertion of a proper class, but technical versions are. Despite the lack of formalizability such existential assertions have been widely entertained since, and with this set theory in practice could be said to have overleaped the bounds of ZFC. On the other hand, that the existence of a class elementary embedding is equivalent to the existence of a certain set, the witnessing ultrafilter for a measurable cardinal, can be considered a means of formalization in ZFC, one that would be paradigmatic for such reductions.

Work of Petr Vopěnka, who started the active Prague seminar in set theory in the spring of 1963, would be closely connected to that of Scott. Aware of the limitations of inner

---

[110]See Kanamori [2003: §§7, 8, 9] for more on partition relations and sets of indiscernibles, particularly their role in the formulation the set of natural numbers $0^{\#}$ and its role of transcendence over L.

[111]That is, for any formula $\varphi(v_1, \ldots, v_n)$ and sets $x_1, \ldots, x_n$, $\varphi(x_1, \ldots, x_n) \longleftrightarrow \varphi^M(j(x_1), \ldots, j(x_n))$, i.e. the formula holds of the x_is exactly when it holds of the $j(x_i)$s with the quantifiers restricted to M. Thus elementary embeddings are just the extension of algebraic monomorphisms to the preservation of logical properties.

models for establishing independence results Vopěnka (cf. [1965]) embarked on a systematic study of (mostly ill-founded) class models of Bernays-Gödel set theory using ultrapower and direct limit constructions. Vopěnka not only established [1962] Scott's result on the incompatibility of measurability and constructibility via different means, but he and his student Karel Hrb´aček in their [1966] soon established a global generalization for inner models $L(A)$: *If there is a strongly compact cardinal, then $V \neq L(A)$ for any set A.*

Through model-theoretic methods set theory was brought to the point of entertaining elementary embeddings into well-founded models,[112] soon to be transfigured by a new method for getting well-founded *extensions* of well-founded models.

4 INDEPENDENCE

4.1 Forcing

Paul Cohen (1934–2007), born just before Gödel established his relative consistency results, established the independence of AC from ZF and the independence of CH from ZFC [1963,1964]. That is, Cohen established that Con(ZF) implies Con(ZF + ¬AC) and Con(ZF) implies Con(ZFC + ¬CH). These results delimited ZF and ZFC in terms of the two fundamental issues at the beginnings of set theory. But beyond that, Cohen's proofs were soon to flow into method, becoming the inaugural examples of *forcing*, a remarkably general and flexible method for extending models of set theory. Forcing has strong intuitive underpinnings and reinforces the notion of set as given by the first-order ZF axioms with conspicuous uses of Replacement and Foundation. If Gödel's construction of L had launched set theory as a distinctive field of mathematics, then Cohen's method of forcing began its transformation into a modern, sophisticated one.[113]

Cohen's approach was to start with a model M of ZF and adjoin a set G, one that would exhibit some desired new property. He realized that this had to be done in a minimal fashion in order that the resulting structure also model ZF, and so imposed restrictive conditions on both M and G. He took M to be a countable standard model, i.e. a countable transitive set that together with the membership relation restricted to it is a model of ZF.[114] The ordinals of M would then coincide with the predecessors of some ordinal ρ, and M would be the cumulative hierarchy $M = \bigcup_{\alpha<\rho} V_\alpha \cap M$. Cohen then established a system of terms to denote members of the new model, finding it convenient to use a ramified language: For each $x \in M$ let $\dot{x}$ be a corresponding constant; let $\dot{G}$ be a new constant; and for each $\alpha < \rho$ introduce quantifiers $\forall_\alpha$ and $\exists_\alpha$. Then develop a hierarchy of terms as follows: $\dot{M}_0 = \{\dot{G}\}$, and for limit ordinals $\delta < \rho$, $\dot{M}_\delta = \bigcup_{\alpha<\delta} \dot{M}_\alpha$. At the successor stage, let $\dot{M}_{\alpha+1}$ be the collection of terms $\dot{x}$ for $x \in V_\alpha \cap M$ and "abstraction" terms

---

[112]See Keisler-Tarski [1964] for a comprehensive account of the theory of large cardinals through the use of ultrapowers in the early 1960s.

[113]According to Scott (Bell [1985: ix]): 'Set theory could never be the same after Cohen, and there is simply no comparison whatsoever in the sophistication of our knowledge about models of set theory today as contrasted to the pre-Cohen era."

[114]The existence of such a model is an avoidable assumption in formal relative consistency proofs via forcing.

corresponding to formulas allowing parameters from $\dot{M}_\alpha$ and quantifiers $\forall_\alpha$ and $\exists_\alpha$. It is crucial that this ramified language with abstraction terms be entirely formalizable in M, through a systematic coding of symbols. Once a set G is provided from the outside, a model $M[G] = \bigcup_{\alpha<\rho} M_\alpha[G]$ would be determined by the terms, where each $\dot{x}$ is to be interpreted by x for $x \in M$ and $\dot{G}$ is to be interpreted by G, so that: $M_0[G] = \{G\}$; for limit ordinals $\delta < \rho$, $M_\delta[G] = \bigcup_{\alpha<\delta} M_\alpha[G]$; and $M_{\alpha+1}[G]$ consists of the sets in $V_\alpha \cap M$ together with sets interpreting the abstraction terms as the corresponding definable subsets of $M_\alpha[G]$ with $\forall_\alpha$ and $\exists_\alpha$ ranging over this domain.

But what properties can be imposed on G to ensure that $M[G]$ be a model of ZF? Cohen's key idea was to tie G closely to M through a system of sets in M called *conditions* that would approximate G. While G may not be a member of M, G is to be a subset of some $Y \in M$ (with $Y = \omega$ a basic case), and these conditions would "force" some assertions about the eventual $M[G]$ e.g. by deciding some of the membership questions, whether $x \in G$ or not, for $x \in Y$. The assertions are to be just those expressible in the ramified language, and Cohen developed a corresponding *forcing relation* $p \Vdash \varphi$, "p forces φ", between conditions p and formulas φ, a relation with properties reflecting his approximation idea. For example, if $p \Vdash \varphi$ and $p \Vdash \psi$, then $p \Vdash \varphi \,\&\, \psi$. The conditions are ordered according to the constraints they impose on the eventual G, so that if $p \Vdash \varphi$, and q is a stronger condition, then $q \Vdash \varphi$. Scott actually provided the now common forcing symbol $\Vdash$, and with Cohen having worked with prenex formulas, Scott showed how to proceed generally by separating out negation with: $p \Vdash \neg\varphi$ *iff* for no stronger condition q does $q \Vdash \varphi$. It was crucial to Cohen's approach that the forcing relation, like the ramified language, be definable in M.

The final ingredient is that the whole scaffolding is given life by incorporating a certain kind of set G. Stepping out of M and making the only use of its countability, Cohen enumerated the formulas of the ramified language in a countable sequence and required that G be completely determined by a countable sequence of stronger and stronger conditions $p_0, p_1, p_2, \ldots$ such that for every formula φ of the ramified language exactly one of φ or $\neg\varphi$ is forced by some p_n. Such a G is called a *generic* set. Cohen was able to show that the resulting $M[G]$ does indeed satisfy the axioms of ZF: Every assertion about $M[G]$ is already forced by some condition; the forcing relation is definable in M; and so the ZF axioms, holding in M, most crucially Power Set and Replacement, can be applied to derive corresponding forcing assertions about ZF axioms holding in $M[G]$.

The foregoing outline in its main features reflects how forcing was viewed by July 1963 and presented by Cohen himself in a course in Spring 1965.[115] He first described the case when $G \subseteq \omega$ and the conditions p are functions from some finite subset of ω into $\{0, 1\}$ and $p \Vdash \dot{n} \in \dot{G}$ if $p(n) = 1$ and $p \Vdash \dot{n} \notin \dot{G}$ if $p(n) = 0$. Today, a G so adjoined to M is called a *Cohen real over M*. Cohen established the independence of CH by adjoining a set which can be construed as a sequence of many Cohen reals. He established the independence of AC by a version of the above scheme where in addition to $\dot{G}$ there are also new constants $\dot{G}_i$ for $i \in \omega$, with $\dot{G}$ to be interpreted by a set X of Cohen reals, each an interpretation of some $\dot{G}_i$. The point is that X is not well-orderable in the extension.

The appeal to a countable model in Cohen's approach is a notable positive subsumption

---

[115]See Cohen [1966].

of Skolem's Paradox (cf. 3.2) into a new method. Remarkably, Skolem [1923: 229] had entertained the possibility of adjoining a new subset of the natural numbers to a countable model of Zermelo's system and getting a new model, adding in a footnote that "it is quite probable" that the Continuum Hypothesis is not decided by Zermelo's axioms. Just as starting with a countable standard model is not formally necessary for relative consistency results, other features of Cohen's argument would soon be reformulated, reorganized, and generalized, but the main thrust of his constructive approach through definability and genericity would remain. Cohen's particular achievement lies in devising a concrete procedure for extending well-founded models of set theory in a minimal fashion to well-founded models of set theory with new properties but without altering the ordinals.[116] Set theory had undergone a sea-change, and beyond how the subject was enriched, it is difficult to convey the strangeness of it.

The creation of forcing is a singular phenomenon in the development of set theory not only since it raised the level of the subject dramatically but also since it could well have occurred decades earlier. But however epochal Cohen's advance there was a line of development for which it did provide at least a semblance of continuity: Interest in independence results for weak versions of AC had been on the rise from the mid-1950's, with more and more sophisticated Fraenkel-Mostowski models being constructed.[117] Solomon Feferman, who was one of the first to whom Cohen explained his ideas for the independence proofs in the process of their development, was the first after him to establish results by forcing; Levy soon followed; and among their first results were new independences from ZF for weak versions of AC (Feferman-Levy [1963], Feferman [1965]). Cohen [1965: 40] moreover acknowledged the similarities between his AC independence result and the previous Fraenkel-Mostowski models. In fact, consistencies first established via Fraenkel-Mostowski models were soon "transferred" to consequence of ZF via forcing by correlating urelements with generic sets.[118]

After an initial result by Feferman [1963], Levy [1963,1965,1970] also probed the limits of ZFC definability, establishing consistency results about definable sets of reals and well-orderings and in descriptive set theory. Intriguingly, inaccessible cardinals were brought in to overcome a technical hurdle in this study; Levy [1963: IV] applied the defining properties of such a cardinal to devise its "collapse" to $\aleph_1$ by making every smaller ordinal countable, and this forcing is now known as the *Levy collapse*.

Forcing was quickly generalized and applied to achieve wide-ranging results, particularly by Robert Solovay. He above all epitomized this period of great expansion in set theory with his mathematical sophistication and fundamental results about and with forcing, and in the areas of large cardinals and descriptive set theory. Just weeks after Cohen's breakthrough Solovay [1963,1965] elaborated the independence of CH by characterizing the possibilities for the size of 2^κ for regular κ and made the first exploration

---

[116]Scott continued (Bell [1985: ix]): 'I knew almost all the set-theoreticians of the day, and I think I can say that no one could have guessed that the proof would have gone in just this way. Model-theoretic methods had shown us how many *non-standard* models there were; but Cohen, starting from very primitive first principles, found the way to keep the models *standard* (that is, with a well-ordered collection of ordinals)."

[117]See Moore [1982: 5.1].

[118]See Felgner [1971] and Jech [1973] for more on the independence of weak versions of AC and transfers, and Pincus [1972] for a strong transfer theorem.

of a spectrum of cardinals. Then William Easton [1964,1970] established the definitive result for powers of regular cardinals: *Suppose that* GCH *holds and F is a class function from the class of regular cardinals to cardinals such that for κ ≤ λ, F(κ) ≤ F(λ) and the cofinality of F(κ) is greater than κ. Then there is a forcing extension preserving cofinalities in which* $2^\kappa = F(\kappa)$ *for every regular κ.* Thus, as Solovay had seen locally, the *only* restriction beyond monotonicity on the power function for regular cardinals is that given by the Zermelo–Kőnig inequality.[119] Easton's result vitally infused forcing not only with the introduction of proper classes of forcing conditions but the now basic idea of a product analysis and the now familiar concept of *Easton support.* Through its reduction Easton's result focused interest on the possibilities for powers of *singular* cardinals, and this *Singular Cardinals Problem* together with the *Singular Cardinals Hypothesis* would stimulate the further development of set theory much as the Continuum Problem and the Continuum Hypothesis had stimulated its early development.[120]

In the Spring of 1964 Solovay [1965b,1970] established a result remarkable for its mathematical depth and revelatory of what standard of argument was possible with forcing: *If there is an inaccessible cardinal, then in an inner model of a forcing extension every set of reals is Lebesgue measurable, has the Baire property, and has the perfect set property.* Like Cohen's results, this contextually decided issues dating back to the turn of the century and before; as described in 2.3 the regularity properties of sets of reals was a major concern of the early descriptive set theorists. Classical counterexamples show that Solovay's inner model cannot have a well-ordering of the reals, and so AC fails there. However, he established that the model satisfies the Principle of Dependent Choices, a principle sufficient for the formal rendition of the traditional theory of measure and category. Thus, Solovay's work vindicated the early descriptive set theorists in the sense that the regularity properties can consistently hold for all sets of reals in a *bona fide* model for mathematical analysis. For his result Solovay applied the Levy collapse of an inaccessible cardinal and built on its definability properties as first exploited by Levy [1963: IV]; for the Lebesgue measurability he introduced a new kind of forcing beyond Cohen's direct ways of adjoining new sets of ordinals or collapsing cardinals, that of adding a *random real.* Solovay's work not only opened the door to a wealth of different forcing arguments, but to this day his original definability arguments remain vital to descriptive set theory.

The perfect set property, central to Cantor's direct approach to the Continuum Problem through definability (1.2,2.3,2.5), led to the first acknowledged instance of a new phenomenon in set theory: the derivation of *equiconsistency* results with large cardinal hypotheses based on the complementary methods of forcing and inner models. A large cardinal hypothesis is typically transformed into a proposition about sets of reals by forcing that "collapses" that cardinal to $\aleph_1$ or "enlarges" the power of the continuum to that cardinal. Conversely, the proposition entails the same large cardinal hypothesis in the clarity of an inner model. Solovay's result provided the forcing direction from an inaccessible cardinal to the proposition that every set of reals has the perfect set property (and $\aleph_1$ is regular). But Ernst Specker [1957: 210] had in effect established that if this obtains,

---

[119]See 1.3, especially footnote 27.

[120]The Singular Cardinal Hypothesis asserts that 2^λ for singular λ is the least possible with respect to the powers 2^μ for $\mu < \lambda$, as given by monotonicity and the Zermelo-Kőnig inequality.

then $\aleph_1$ (of V) is inaccessible in L. Thus, Solovay's use of an inaccessible cardinal was actually necessary, and its collapse to $\aleph_1$ complemented Specker's observation. Other propositions figuring in the initial applications of inaccessibility to forcing turned out to require inaccessibility, further integrating it into the interstices of set theory.[121] The emergence of such equiconsistency results is a subtle transformation of earlier hopes of Gödel (cf. 3.4): Propositions can be positively subsumed if there are enough ordinals, how many being specified by positing a large cardinal.[122] On the other hand, forcing quickly led to the conclusion that there could be no direct implication for CH: Levy and Solovay (Levy [1964], Solovay [1965a], Levy-Solovay [1967]) established that measurable cardinals neither imply nor refute CH, with an argument generalizable to most inaccessible large cardinals. Rather, the subsumption for many other propositions would be in terms of consistency, the methods of forcing and inner models being the operative modes of argument. In a new synthesis of the two Cantorian legacies, hypotheses of length concerning the extent of the transfinite are correlated with hypotheses of width concerning sets of reals.

It was the incisive work of Scott and Solovay through this early period that turned Cohen's breakthrough into a general method of wide applicability. Scott simplified Cohen's original formulation as noted above; Solovay made the important move to general partial orders and generic filters; and they together developed, with vicissitudes, the formulation in terms of Boolean-valued models.[123] These models forcibly showed how to avoid Cohen's ramified language as well as his dependence on a countable model. With their elegant algebraic trappings and seemingly more complete information they held the promise of being the right approach to independence results. But Shoenfield [1971] showed that forcing with partial orders can get at the gist of the Boolean approach in a straightforward manner. Moreover, Boolean-valued models were soon found to be too abstract and unintuitive for establishing *new* consistency results, so that within a few years set theorists were generally working with partial orders. It is a testament to Cohen's concrete approach that in this return from abstraction even the use of ramified languages has played an essential role in careful forcing arguments at the interface of recursion theory and set theory.

---

[121]The original application of the Levy collapse in Levy [1963: IV] also turned out to require an inaccessible cardinal (Levy [1970: 131ff]) – a remarkable turn of events for an apparently technical artifact at the beginning of forcing.

Many years later, Saharon Shelah [1980,1984] was able to establish the necessity of Solovay's inaccessible for the proposition that every set of reals is Lebesgue measurable; on the other hand, Shelah also showed that the inaccessible is not necessary for the proposition that every set of reals has the Baire property.

[122]There is a telling antecedent in the result of Gerhard Gentzen [1936,1943] that the consistency strength of arithmetic can be exactly gauged by an ordinal ε_0, i.e. transfinite induction up to that ordinal in a formal system of notations. Although Hilbert's program of establishing consistency by finitary means could not be realized, Gentzen provided an exact analysis in terms of ordinal length. Proof theory blossomed in the 1960s with the analysis of other theories in terms of such lengths, the proof theoretic ordinals.

[123]Vopěnka had developed a similar concept in a reworking [1964] of the independence of CH. The concept was generalized and simplified in a series of papers on the so-called ∇-models from the active Prague seminar founded by Vopěnka (see Hájek [1971: 78]), culminating in the exposition Vopěnka [1967]. However, the earlier papers did not have much impact, partly because of an involved formalism in which formulas were valued in a complete lattice rather than Boolean algebra.

4.2 Envoi

Building on his Lebesgue measurability result Solovay soon reactivated the classical descriptive set theory program (cf. 2.5) of investigating the extent of the regularity properties by providing characterizations for the Σ^1_2 sets, the level at which Gödel established from $V = L$ the failure of the properties (cf. 3.4), and showed in particular that the regularity properties for these sets follow from the existence of a measurable cardinal. Thus, although measurable cardinals do not decide CH, they do establish the perfect set property for Σ^1_2 sets (Solovay [1969]) so that "CH holds for the Σ^1_2 sets" – a vindication of Gödel's hopes for large cardinals through a direct implication. Donald Martin and Solovay in their [1969] then applied large cardinal hypotheses weaker than measurability to push forward the old tree representation ideas of the classical descriptive set theorists, with the hypotheses cast in the new role of securing well-foundedness in this context.

The method of forcing as part of the axiomatic tradition together with the transmutations of Cantor's two legacies, large cardinals furthering the extension of number into the transfinite and descriptive set theory investigating definable sets of reals, established set theory as a sophisticated field of mathematics, a study of well-foundedness expanded into one of consistency strength. With the further development of forcing through increasingly sophisticated iteration techniques questions raised in combinatorics and over a broad landscape would be resolved in terms of consistency, sometimes with equiconsistencies in terms of large cardinals. The theory of large cardinals would itself be much advanced with the heuristics of reflection and generalization and sustained through increasing use in the study of consistency strength. In the most distinctive and intriguing development of contemporary set theory, the investigation of the determinacy of games, large cardinals would be further integrated into descriptive set theory. They were not only used to literally incorporate well-foundedness of inner models into the study of tree representations, historically first context involving well-foundedness, but also to provide the exact hypotheses, with Woodin cardinals, for gauging consistency strength.[124]

Stepping back to gaze at modern set theory, the thrust of mathematical research should deflate various possible metaphysical appropriations with an onrush of new models, hypotheses, and results. Shedding much of its foundational burden, set theory has become an intriguing field of mathematics where formalized versions of truth and consistency have become matters for manipulation as in algebra. As a study couched in well-foundedness ZFC together with the spectrum of large cardinals serves as a court of adjudication, in terms of relative consistency, for mathematical statements that can be informatively contextualized in set theory by letting their variables range over the set-theoretic universe. Thus, set theory is more of an open-ended framework for mathematics rather than an elucidating foundation. It is as a field *of* mathematics that both proceeds with its own internal questions and is capable of contextualizing over a broad range which makes of set theory an intriguing and highly distinctive subject.

---

[124]See Kanamori [2003] for these recent developments.

ACKNOWLEDGEMENTS

BIBLIOGRAPHY

[Addison *et al.*, 1965] J. W. Addison Jr., L. Henkin, and A. Tarski, eds. *The Theory of Models*, Proceedings of the 1963 International Symposium at Berkeley, North-Holland, Amsterdam 1965.

[Aleksandrov, 1916] P. S. Aleksandrov. Sur la puissance des ensembles mesurables *B*, *Comptes Rendus Hebdomadaires des S´eances de l'Acad´emie des Sciences, Paris* 162 (1916), 323–325.

[Aleksandrov and Urysohn, 1929] P. S. Aleksandrov and P. S. Urysohn. Mémoire sur les espaces topologiques compacts, *Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen Tweed Reeks. Afdeling Natuurkunde* 14 (1929), 1-96.

[Baire, 1898] R. Baire. Sur les fonctions discontinues qui se rattachent aux fonctions continues, *Comptes Rendus Hebdomadaires des S´eances de l'Acad´emie des Sciences, Paris* 129 (1898), 1621–1623.

[Baire, 1899] R. Baire. Sur les fonctions de variables réelles, *Annali di Matematica Pura ed Applicata* (3)3 (1899), 1-123.

[Banach, 1967] S. Banach. In S. Hartman and E. Marczewski, eds., *Oeuvres*, vol. 1, Państwowe Wydawnictwo Naukowe, Warsaw 1967.

[Banach and Tarski, 1924] S. Banach and A. Tarski. Sur la décomposition des ensembles de points en parties respectivement congruentes, *Fundamenta Mathematicae* 6 (1924), 244–277. Reprinted in Banach [1967], 118–148. Reprinted in Tarski [1986] vol. 1, 121–154.

[Bar-Hillel *et al.*, 1961] Y. Bar-Hillel, E. I. J. Poznanski, M. O. Rabin, and A. Robinson, eds. *Essays on the Foundations of Mathematics*, Magnes Press, Jerusalem 1961.

[Bartoszyński, 2008] T. Bartoszyński. Invariants of Meaure and Category, in: Foreman–Kanamori [2008].

[Bell, 1985] J. L. Bell. *Boolean-Valued Models and Independence Proofs in Set Theory*, second edition, Oxford Logic Guides #12, Oxford University Press, Oxford 1985.

[Benacerraf and Putnam, 1964] P. Benacerraf and H. Putnam, eds. *Philosophy of Mathematics. Selected Readings.* Prentice Hall, Englewood Cliffs, N.J. 1964.

[Benacerraf and Putnam, 1983] P. Benacerraf and H. Putnam. Second edition of [1964]. Cambridge University Press, Cambridge 1983.

[Bendixson, 1883] I. Bendixson. Quelques théorèmes de la théorie des ensembles de points, *Acta Mathematica* 2 (1883), 415–429.

[Bernays, 1937] P. Bernays. A system of axiomatic set theory, Part I, *The Journal of Symbolic Logic* 2 (1937), 65–77.

[Bernays, 1976] P. Bernays. A system of axiomatic set theory, in Gert H. M¨uller (editor), *Sets and Classes. On the Work of Paul Bernays*, North-Holland, Amsterdam 1976, 1–119. Individually appeared in various parts in *The Journal of Symbolic Logic* 2 (1937), 65–77; 6 (1941), 1–17; 7 (1942), 65–89 and 133–145; 8 (1943), 89–106; 3 (1948), 65–79; and 19 (1954), 81–96.

[Bernstein, 1908] F. Bernstein. Zur Theorie der trigonometrischen Reihen, *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematische-Physische Klasse* 60 (1908), 325–338.

[Blass, 1984] A. R. Blass. Existence of bases implies the Axiom of Choice, in Baumgartner, James E., Donald A. Martin, and Saharon Shelah (editors), *Axiomatic Set Theory*, Contemporary Mathematics vol. 31, American Mathematical Society, Providence 1984, 31–33.

[Blass, 2008] A. R. Blass. Combinatorial Cardinal Charateristics of the Continuum, in: Foreman–Kanamori [2008].

[Boolos, 1971] G. S. Boolos. The iterative concept of set, *Journal of Philosophy* 68 (1971), 215–231. Reprinted in Benacerraf-Putnam [1983], 486–502, and in [1998] below, 13–29.

[Boolos, 1998] G. S. Boolos. *Logic, Logic, and Logic*, Harvard University Press, Cambridge 1998.

[Borel, 1898] E. Borel. *Le¸cons sur la th´eorie des fonctions*, Gauthier-Villars, Paris 1898.

[Borel, 1919] E. Borel. Sur la classification des ensembles de mesure nulle, *Bulletin de la Soci´et´e Math´ematique de France* 47 (1919), 97–125.

[Bourbaki, 1954] N. Bourbaki. *El´ements de Math´ematique. I. Th´eorie des Ensembles*, Hermann, Paris 1954.

[Bourbaki, 1970] N. Bourbaki. *El´ements de Mathematique. I. Th´eorie des Ensembles*, combined edition, Hermann, Paris 1970.

[Brouwer, 1911] L. E. J. Brouwer. Beweis der Invarianz der Dimensionenzahl, *Mathematische Annalen* 70 (1911), 161–165. Reprinted in [1976] below, 430–434.

[Brouwer, 1976] L. E. J. Brouwer. Freudenthal, Hans (editor), *Collected Works*, vol. 2, North-Holland, Amsterdam 1976.

[Burali-Forti, 1897] C. Burali-Forti. Una questione sui numeri transfiniti, *Rendiconti del Circolo Matematico di Palermo* 11 (1897), 154–164. Translated in van Heijenoort [1967], 104–111.

[Cantor, 1872] G. Cantor. Über die Ausdehnung eines Satzes aus der Theorie der trignometrischen Reihen, *Mathematische Annalen* 5 (1872), 123–132. Reprinted in [1932] below, 92–102.

[Cantor, 1874] G. Cantor. Über eine Eigenschaft des Inbegriffes aller reellen algebraischen Zahlen, *Journal für die reine und angewandte Mathematik* 77 (1874), 258–262. Reprinted in [1932] below, 115–118. Translated with commentary in Ewald [1996], 839–843.

[Cantor, 1878] G. Cantor. Ein Beitrag zur Mannigfaltigkeitslehre, *Journal für die reine und angewandte Mathematik* 84 (1878), 242–258. Reprinted in [1932] below, 119–133.

[Cantor, 1879] G. Cantor. Über einen Satz aus der Theorie der stetigen Mannigfaltigkeiten, *Nachrichten von der Königlichen Gesellschaft der Wissenschaften und der Georg-Augusts-Universität zu Göttingen* (1879), 127–135. Reprinted in [1932] below, 134–138.

[Cantor, 1880] G. Cantor. Über unendliche, lineare Punktmannigfaltigkeiten. II, *Mathematische Annalen* 17 (1880), 355–358. Reprinted in [1932] below, 145–148, but excluding the referenced remark.

[Cantor, 1883] G. Cantor. Über unendliche, lineare Punktmannigfaltigkeiten. V. *Mathematische Annalen* 21 (1883), 545–591. Published separately as *Grundlagen einer allgemeinen Mannigfaltigkeitslehre. Ein mathematisch-philosophischer Versuch in der Lehre des Unendlichen*, B.G. Teubner, Leipzig 1883. Reprinted in [1932] below, 165–209. Translated with commentary in Ewald [1996], 878–920.

[Cantor, 1884] G. Cantor. Über unendliche, lineare Punktmannigfaltigkeiten. VI, *Mathematische Annalen* 23 (1884), 453–488. Reprinted in [1932] below, 210–246.

[Cantor, 1884a] G. Cantor. De la puissance des ensembles parfaits de points, *Acta Mathematica* 4 (1884), 381–392. Reprinted in [1932] below, 252–260.

[Cantor, 1887] G. Cantor. Mitteilungen zur Lehre vom Transfiniten, *Zeitschrift für Philosophie und philosophische Kritik* 91 (1887), 81–125 and 252–270. Reprinted in [1932] below, 378–439.

[Cantor, 1891] G. Cantor. Über eine elementare Frage der Mannigfaltigkeitslehre, *Jahresbericht der Deutschen Mathematiker-Vereinigung* I (1891), 75–78. Reprinted in [1932] below, 278–280. Translated with commentary in Ewald [1996], 920–922.

[Cantor, 1895] G. Cantor. Beitr¨age zur Begr¨undung der transfiniten Mengenlehre. I, *Mathematische Annalen* 46 (1895), 481–512. Translated in [1915] below. Reprinted in [1932] below, 282–311.

[Cantor, 1897] G. Cantor. Beitr¨age zur Begr¨undung der transfiniten Mengenlehre. II, *Mathematische Annalen* 49 (1897), 207–246. Translated in [1915] below. Reprinted in [1932] below, 312–351.

[Cantor, 1915] G. Cantor. *Contributions to the Founding of the Theory of Transfinite Numbers.* Translation of [1895] and [1897] above with introduction and notes, by Philip E.B. Jourdain, Open Court, Chicago 1915. Reprinted Dover, New York 1965.

[Cantor, 1932] G. Cantor. E. Zermelo, ed., *Gesammelte Abhandlungen mathematicschen und philosophischen Inhalts*, Julius Springer, Berlin 1932. Reprinted in Springer, Berlin 1980.

[Cavaillès, 1962] J. Cavaillès. *Philosophie Math´ematique.* Paris, Hermann 1962. Includes French translation of Noether-Cavaillès [1937].

[Coffa, 1979] J. A. Coffa. The humble origins of Russell's paradox, *Russell* 33–34 (1979), 31–37.

[Cohen, 1963] P. J. Cohen. The independence of the Continuum Hypothesis. I, *Proceedings of the National Academy of Sciences U.S.A.* 50 (1963), 1143–1148.

[Cohen, 1964] P. J. Cohen. The independence of the Continuum Hypothesis. II, *Proceedings of the National Academy of Sciences U.S.A.* 51 (1964), 105–110.

[Cohen, 1965] P. J. Cohen. Independence results in set theory, in Addison-Henkin-Tarski [1965], 39-54.

[Dauben, 1979] J. W. Dauben. *Georg Cantor. His Mathematics and Philosophy of the Infinite,* Harvard University Press, Cambridge 1979. Paperback edition 1990.

[Dedekind, 1871] R. Dedekind. Über die Komposition der bin¨aren quadratischen Formen, supplement X to Dirichlet's *Vorlesungen über Zahlentheorie*, second edition, F. Vieweg, Braunschweig 1871. Reprinted in [1932] below, vol. 3, 399–407. Excerpt translated with commentary in Ewald [1996], 762-765.

[Dedekind, 1872] R. Dedekind. *Stetigkeit und irrationale Zahlen,* F. Vieweg, Braunschweig 1872. Fifth, 1927 edition reprinted in [1932] below, vol. 3, 315–334. Translated in [1963] below, 1-27. Translated with commentary in Ewald [1996], 765–779.

[Dedekind, 1888] R. Dedekind. *Was sind und was sollen die Zahlen?*, F. Vieweg, Braunschweig 1888. Sixth, 1930 edition reprinted in [1932] below, vol. 3, 335–390. Second, 1893 edition translated in [1963] below, 29-115. Third, 1911 edition translated with commentary in Ewald [1996], 787–833.

[Dedekind, 1900] R. Dedekind. Über die von drei Moduln erzeugte Dualgruppe, *Mathematische Annalen* 53 (1900), 371–403. Reprinted in [1932] below, vol. 2, 236–271.

[Dedekind, 1932] R. Dedekind. Fricke, Robert, Emmy Noether, and Öystein Ore (editors), *Gesammelte mathematische Werke*, F. Vieweg, Braunschweig 1932. Reprinted by Chelsea, New York 1969.

[Dedekind, 1963] R. Dedekind. *Essays on the Theory of Numbers*. Translations by Wooster W. Beman, Dover, New York 1963. (Reprint of original edition, Open Court, Chicago 1901).

[du Bois-Reymond, 1869] P. du Bois-Reymond. Bemerkungen über die verschidenen Werthe, welche eine Function zweier reellen Variabeln erhält, wenn man diese Variabeln entweder nach einander oder gewissen Beziehungen gemäss gleichzeitig verschwinden lässt, *Journal für die reine und angewandte Mathematik* 70 (1869), 10–45.

[du Bois-Reymond, 1875] P. du Bois-Reymond. Über asymptotische Werthe, infinitäre Approximationen und infinitäre Auflösung von Gleichungen, *Mathematische Annalen* 8 (1875), 363–414.

[Dugac, 1976] P. Dugac. *Richard Dedekind et les fondements des mathématiques,* Collection des travaux de l'Académie Internationale d'Histoire des Sciences #24, Vrin, Paris 1976.

[Easton, 1964] W. B. Easton. Powers of regular cardinals. Ph.D. thesis, Princeton University 1964. Abstracted as: Proper classes of generic sets, *Notices of the American Mathematical Society* 11 (1964), 205. Published in abridged form as [1970] below.

[Easton, 1970] W. B. Easton. Powers of regular cardinals, *Annals of Mathematical Logic* 1 (1970), 139–178.

[Ebbinghaus, 2007] H.-D. Ebbinghaus. *Ernst Zermelo. An Approach to his Life and Work*, Springer, Berlin 2007.

[Ehrenfeucht and Mostowski, 1956] A. Ehrenfeucht and A. M. Mostowski. Models of axiomatic theories admitting automorphisms, *Fundamenta Mathematicae* 43 (1956), 50–68. Reprinted in Mostowski [1979], 494–512.

[Erdős and Hajnal, 1962] P. Erdős and A. Hajnal. Some remarks concerning our paper 'On the structure of set mappings", *Acta Mathematica Academiae Scientiarum Hungaricae* 13 (1962), 223–226.

[Erdős *et al.*, 1984] P. Erdős, A. Hajnal, A. Máté, and R. Rado. *Combinatorial Set theory: Partition Relations for Cardinals*, North-Holland, Amsterdam 1984.

[Erdős *et al.*, 1965] P. Erdős, A. Hajnal, and R. Rado. Partition relations for cardinal numbers, *Acta Mathematica Academiae Scientiarum Hungaricae* 16 (1965), 93–196.

[Erdős and Rado, 1956] P. Erdős and R. Rado. A partition calculus in set theory, *Bulletin of the American Mathematical Society* 62 (1956), 427–489. Reprinted in Gessel-Rota [1987], 179–241.

[Erdős and Tarski, 1943] P. Erdős and A. Tarski On families of mutually exclusive sets, *Annals of Mathematics* 44 (1943), 315–329. Reprinted in Tarski [1986] vol. 2, 591–605.

[Erdős and Tarski, 1961] P. Erdős and A. Tarski. On some problems involving inaccessible cardinals, in: Bar-Hillel–Poznanski–Rabin–Robinson [1961], 50–82. Reprinted in Tarski [1986] vol. 4, 79–111.

[Ewald, 1996] W. Ewald, ed. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, two volumes, Clarendon Press, Oxford 1996.

[Feferman, 1963] S. Feferman. Some applications of the notions of forcing and generic sets (summary), in Addison–Henkin–Tarski [1965], 89–95.

[Feferman, 1965] S. Feferman. Some applications of the notions of forcing and generic sets, *Fundamenta Mathematicae* 56 (1965), 325-345.

[Feferman and Levy, 1963] S. Feferman and A. Levy. Independence results in set theory by Cohen's method. II (abstract), *Notices of the American Mathematical Society* 10 (1963), 593.

[Felgner, 1971] U. Felgner. *Models of ZF-Set Theory,* Lecture Notes in Mathematics #223, Springer, Berlin 1971.

[Ferreirós, 1995] J. Ferreirós. 'What fermented in me for years": Cantor's discovery of transfinite numbers, *Historia Mathematica* 22 (1995), 33–42.

[Ferreirós, 2007] J. Ferreirós. *Labyrinth of Thought: A History of Set Theory and its Role in Modern Mathematics*, second edition, Birkhauser, Basel 2007.

[Fodor, 1956] G. Fodor. Eine Bemerking zur Theorie der regressiven Funktionen, *Acta Scientiarum Mathematicarum (Szeged)* 17 (1956), 139–142.

[Foreman and Kanamori, 2008] M. Foreman and A. Kanamori. *The Handbook of Set Theory*, Springer, Berlin 2008.

[Fraenkel, 1921] A. A. Fraenkel. Über die Zermelosche Begründung der Mengenlehre (abstract), *Jahresbericht der Deutschen Mathematiker-Vereinigung* 30II (1921), 97–98.

[Fraenkel, 1922] A. A. Fraenkel. Zu den Grundlagen der Cantor-Zermeloschen Mengenlehre, *Mathematische Annalen* 86 (1922), 230–237.

[Fraenkel, 1922a] A. A. Fraenkel. Über den Begriff 'definit' und die Unabhängigkeit des Auswahlaxioms, *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse* (1922), 253–257. Translated in van Heijenoort [1967], 284–289.

[Fraenkel, 1930] A. A. Fraenkel. Georg Cantor, *Jahresbericht der Deutschen Mathematiker-Vereinigung* 39 (1930), 189–266. Published separately as *Georg Cantor*, B.G. Teubner, Leipzig 1930. Published in abridged form in Cantor [1932], 452–483.

[Fraenkel, 1953] A. A. Fraenkel. *Abstract Set Theory,* North-Holland, Amsterdam 1953.

[Frege, 1879] G. Frege. *Begriffsschrift, eine der arithmetischen nachgëbildete Formelsprache des reinen Denkens,* Nebert, Halle 1879. Reprinted Hildesheim, Olms 1964. Translated in van Heijenoort [1967], 1–82.

[Frege, 1884] G. Frege. *Die Grundlagen der Arithmetik, eine logisch-mathematische Untersuchung über den Begriff der Zahl,* Wilhelm Köbner, Breslau 1884. Translated with German text by John L. Austin, as *The Foundations of Arithmetic, A logico-mathematical enquiry into the concept of number*; Blackwell, Oxford 1950. Later editions without German text, Harper, New York.

[Frege, 1891] G. Frege. Function und Begriff, Hermann Pohle, Jena 1891. Translated in [1952] below, 21–41.

[Frege, 1893] G. Frege. *Grundgesetze der Arithmetik, Begriffsschriftlich abgeleitet*. Vol. 1, Hermann Pohle, Jena 1893. Reprinted Hildesheim, Olms 1962. Partially translated in [1964] below.

[Frege, 1895] G. Frege. Kritische Beleuchtung einiger Punkte in E. Schröders *Vorlesungen über die Algebra der Logik, Archiv für systematische Philosophie* 1 (1895), 433–456. Translated in [1952] below, 86–106.

[Frege, 1903] G. Frege. Vol. 2 of [1893]. Reprinted Hildesheim, Olms 1962.

[Frege, 1952] G. Frege. P. Geach and M. Black, trans. and eds., *Translations from the Philosophical Writings of Gottlob Frege*, Blackwell, Oxford 1952. Second, revised edition 1960. Latest edition, Rowland & Littlewood, Totowa 1980.

[Frege, 1964] G. Frege. M. Furth, trans. and ed., *The Basic Laws of Arithmetic. Exposition of the System*, University of California Press, Berkeley 1964.

[Garciadiego, 1992] A. Garciadiego. *Bertrand Russell and the Origins of the Set-theoretic "Paradoxes",* Birkhäuser, Boston 1992.

[Gentzen, 1936] G. Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie, *Mathematicshe Annalen* 112 (1936), 493–565. Translated in [1969] below, 132–213.

[Gentzen, 1943] G. Gentzen. Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie, *Mathematicshe Annalen* 119 (1943), 140–161. Translated in [1969] below, 287–308.

[Gentzen, 1969] G. Gentzen. M. E. Szabo, ed., *The Collected Papers of Gerhard Gentzen*, North-Holland, Amsterdam 1969.

[Gessel and Rota, 1987] I. Gessel and G.-C. Rota, eds. *Classic Papers in Combinatorics,* Birkhäuser, Boston 1987.

[Gödel, 1930] K. F. Gödel. Die Vollständigkeit der Axiome des logischen Funktionenkalküls, *Monatshefte für Mathematik und Physik* 37 (1930), 349–360. Reprinted and translated in [1986] below, 102–123.

[Gödel, 1931] K. F. Gödel. Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I, *Monatshefte für Mathematik und Physik* 38 (1931), 173–198. Reprinted and translated with minor emendations by the author in [1986] below, 144–195.

[Gödel, 1933] K. F. Gödel. The present situation in the foundations of mathematics, in [1995] below, 45–53, and the page references are to these.

[Gödel, 1938] K. F. Gödel. The consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis, *Proceedings of the National Academy of Sciences U.S.A.* 24 (1938), 556–557. Reprinted in [1990] below, 26–27.

[Gödel, 1939] K. F. Gödel. Consistency-proof for the Generalized Continuum-Hypothesis, *Proceedings of the National Academy of Sciences U.S.A.* 25 (1939), 220–224. Reprinted in [1990] below, 28–32.

[Gödel, 1939a] K. F. Gödel. Vortrag Göttingen, text and translation in [1995] below, 126–155, and the page references are to these.

[Gödel, 1940] K. F. Gödel. *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory,* Annals of Mathematics Studies #3, Princeton University Press, Princeton 1940. Reprinted in [1990] below, 33–101.

[Gödel, 1947] K. F. Gödel. What is Cantor's Continuum Problem?, *American Mathematical Monthly* 54 (1947), 515–525. Errata 55 (1948), 151. Reprinted in [1990] below, 176–187. Revised and expanded version in Benacerraf-Putnam [1964], 258–273. This version reprinted with minor emendations by the author in [1990] below, 254–270.

[Gödel, 1986] K. F. Gödel. S. Feferman, *et al.*, eds., *Kurt Gödel, Collected Works*, vol. I, Oxford University Press, Oxford 1986.

[Gödel, 1990] K. F. Gödel. S. Feferman, *et al.*, eds., *Kurt Gödel, Collected Works*, vol. II, Oxford University Press, New York 1990.

[Gödel, 1995] K. F. Gödel. S. Feferman, *et al.*, eds., *Kurt Gödel, Collected Works*, vol. III, Oxford University Press, New York 1995.

[Gödel, 2003] K. F. Gödel. S. Feferman and J. W. Dawson, eds., *Kurt Gödel, Collected Works*, vol. V, Clarendon Press, Oxford 2003.

[Goldfarb, 1979] W. D. Goldfarb. Logic in the Twenties: the nature of the quantifier, *The Journal of Symbolic Logic* 44 (1979), 351–368.

[Graham *et al.*, 1990] R. L. Graham, B. L. Rothschild, and J. H. Spencer. *Ramsey Theory*, second edition, Wiley & Sons, New York 1990.

[Grattan-Guinness, 1978] I. Grattan-Guinness. Wiener on the logics of Russell and Schröder. An account of his doctoral thesis, and of his subsequent discussion of it with Russell, *Annals of Science* 32 (1978), 103–132.

[Grattan-Guinness, 1978a] I. Grattan-Guinness. How Bertrand Russell discovered his paradox, *Historia Mathematica* 5 (1978), 127–137.

[Grattan-Guinness, 2000] I. Grattan-Guinness. *The Search for Mathematical Roots 1870–1940: Logics, Set Theories and the Foundations of Mathematics from Cantor through Russell and Gödel*, Princeton University Press, Princeton 2000.

[Gray, 1994] R. Gray. Georg Cantor and Transcendental Numbers, *American Mathematical Monthly* 101 (1994), 819–832.

[Hájek, 1971] P. Hájek. Sets, semisets, models, in: Scott [1971], 67-81.

[Hajnal, 1956] A. Hajnal. On a consistency theorem connected with the generalized continuum problem, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 2 (1956), 131–136.

[Hajnal, 1961] A. Hajnal. On a consistency theorem connected with the generalized continuum problem, *Acta Mathematica Academiae Scientiarum Hungaricae* 12 (1961), 321–376.

[Hajnal and Larson, 2008] A. Hajnal and J. A. Larson. Partition relations, in: Foreman–Kanamori [2008].

[Hallett, 1984] M. Hallett. *Cantorian Set Theory and Limitation of Size*, Oxford Logic Guides #10, Clarendon Press, Oxford 1984.

[Hamel, 1905] G. Hamel. Eine Basis aller Zahlen und die unstetigen Lösungen der Funktional-gleichung: $f(x+y) = f(x) + f(y)$, *Mathematische Annalen* 60 (1905), 459–462.

[Hanf, 1964] W. P. Hanf. Incompactness in languages with infinitely long expressions, *Fundamenta Mathematicae* 53 (1964), 309–324.

[Hartogs, 1915] F. Hartogs. Über das Problem der Wohlordnung, *Mathematische Annalen* 76 (1915), 436–443.

[Hardy, 1903] G. H. Hardy. A theorem concerning the infinite cardinal numbers, *The Quarterly Journal of Pure and Applied Mathematics* 35 (1903), 87–94; reprinted in [1979] below, vol. 7, 427–434.

[Hardy, 1979] G. H. Hardy. I. W. Busbridge and R. A. Rankin, eds., *Collected Papers of G.H. Hardy*, Clarendon, Oxford 1979.

[Hausdorff, 1908] F. Hausdorff. Grundzüge einer Theorie der geordneten Mengen, *Mathematische Annalen* 76, 436–443, 1908.

[Hausdorff, 1914] F. Hausdorff. *Grundzüge der Mengenlehre*, de Gruyter, Leipzig 1914. Reprinted by Chelsea, New York 1965.

[Hausdorff, 1914a] F. Hausdorff. Bemerkung über den Inhalt von Punktmengen, *Mathematische Annalen* 75 (1914), 428–433.

[Hausdorff, 1916] F. Hausdorff. Die Mächtigkeit der Borelschen Mengen, *Mathematische Annalen* 77 (1916), 430–437, 1916.

[Hausdorff, 1932] F. Hausdorff. Zur Theorie der linearen metrischen Räume, *Journal für die reine und angewandte Mathematik* 167 (1932), 294–311.

[Hausdorff, 1937] F. Hausdorff. *Mengenlehre*, third, revised edition of [1914]. Translated by John R. Auman as *Set Theory*, Chelsea, New York 1962.

[Hawkins, 1975] T. W. Hawkins. *Lebesgue's Theory of Integration. Its Origins and Development,* second edition, Chelsea, New York 1975.

[Heck, 1995] R. G. Heck, Jr. Definition by induction in Frege's *Grundgesetze der Arithmetik*, in: William Demopoulos (editor) *Frege's Philosophy of Mathematics*, Harvard University Press, Cambridge 1995, 295–233.

[Henkin, 1949] L. Henkin. The completeness of the first-order functional calculus, *The Journal of Symbolic Logic* 14 (1949), 159-166.

[Hessenberg, 1906] G. Hessenberg. *Grundbegriffe der Mengenlehre,* Vandenhoeck & Ruprecht, Göttingen 1906. Reprinted in *Abhandlungen der Fries'schen Schule, Neue Folge* 1 (1906), 479–706.

[Hilbert, 1890] D. Hilbert. Über die Theorie der algebraischen Formen, *Mathematische Annalen* 36 (1890), 473-534. Translated in [1978] below, 143–224.

[Hilbert, 1900] D. Hilbert. Mathematische Probleme. Vortrag, gehalten auf dem internationalem Mathematiker-Kongress zu Paris. 1900. *Nachrichten von der Königlichen Gesellschaften der Wissenschaften zu Göttingen* (1900), 253–297. Translated in the *Bulletin of the American Mathematical Society* 8 (1902), 437–479.

[Hilbert, 1926] D. Hilbert. Über das Unendliche, *Mathematische Annalen* 95 (1926), 161–190. Translated in van Heijenoort [1967], 367–392. Partially translated in Benacerraf-Putnam [1983], 183–201.

[Hilbert, 1978] D. Hilbert. M. Ackerman, trans., *Hilbert's Invariant Theory Papers*, Math Sci Press, Brookline.

[Hilbert and Ackermann, 1928] D. Hilbert and W. Ackermann. *Grundzüge der theoretischen Logik*, Julius Springer, Berlin 1928. Second edition, 1938; third edition, 1949. Second edition was translated by Hammond, Lewis M., George G. Leckie, and F. Steinhardt as *Principles of Mathematical Logic*, Chelsea, New York 1950.

[Howard and Rubin, 1998] P. Howard and J. E. Rubin. *Consequences of the Axiom of Choice*, Mathematical Surveys and Monographs, vol. 59, American Mathematical Society, Providence, 1998.

[Jané, 1995] I. Jané. The role of the absolute infinite in Cantor's conception of set, *Erkenntnis* 42 (1995), 375–402.

[Jech, 1973] T. J. Jech. *The Axiom of Choice*, North-Holland, Amsterdam 1973.

[Jech, 2003] T. J. Jech. *Set Theory*, the third millennium edition, revised and expanded, Springer, Berlin 2003.

[Jourdain, 1904] P. E. B. Jourdain. On the transfinite cardinal numbers of well-ordered aggregates, *Philosophical Magazine* 7 (1904), 61–75.

[Jourdain, 1905] P. E. B. Jourdain. On a proof that every aggregate can be well-ordered, *Mathematische Annalen* 60, 465–470.

[Kac and Ulam, 1968] M. Kac and S. M. Ulam. *Mathematics and Logic*, Praeger, New York 1968.

[Kanamori, 1995] A. Kanamori. The emergence of descriptive set theory, in: Hintikka, Jaakko (editor) *Essays on the Development of the Foundations of Mathematics*, Kluwer, Dordrecht 1995, 241–266.

[Kanamori, 1997] A. Kanamori. The mathematical import of Zermelo's well-ordering theorem, *The Bulletin of Symbolic Logic* 3 (1997), 281–311.

[Kanamori, 2003] A. Kanamori. *The Higher Infinite. Large Cardinals in Set Theory from their Beginnings*, second edition, Springer, Berlin 2003.

[Kanamori, 2003a] A. Kanamori. The empty set, the singleton, and the ordered pair, *The Bulletin of Symbolic Logic* 9 (2003), 273–298.

[Kanamori, 2004] A. Kanamori. Zermelo and set theory, *The Bulletin of Symbolic Logic* 10 (2004), 487–553.

[Kanamori, 2007] A. Kanamori. G¨odel and set theory, *The Bulletin of Symbolic Logic* 13 (2007), 153–188.

[Kechris and Louveau, 1987] A. S. Kechris and A. Louveau. *Descriptive Set Theory and the Structure of Sets of Uniqueness*, London Mathematical Society Lecture Note Series #128, Cambridge University Press, Cambridge 1987.

[Keisler, 1962] H. J. Keisler. Some applications of the theory of models to set theory, in: Nagel, Ernest, Patrick Suppes, and Alfred Tarski (editors), *Logic, Methodology and Philosophy of Science*, Proceedings of the 1960 [and first] International Congress (Stanford, California), Stanford University Press, Stanford 1962, 80–86.

[Keisler, 1962a] H. J. Keisler. The equivalence of certain problems in set theory with problems in the theory of models (abstract), *Notices of the American Mathematical Society* 9 (1962), 339-340.

[Keisler and Tarski, 1964] H. J. Keisler and A. Tarski. From accessible to inaccessible cardinals, *Fundamenta Mathematicae* 53 (1964), 225–308. Corrections 57 (1965), 119. Reprinted in Tarski [1986] vol. 4, 129–213.

[Kennedy, 1975] H. C. Kennedy. Nine letters from Giuseppe Peano to Bertrand Russell, *Journal of the History of Philosophy* 13 (1975), 205–220.

[Kőnig, 1927] D. Kőnig. Über eine Schlussweise aus dem Endlichen ins Unendliche: Punktmengen. Kartenf¨arben. Verwandtschaftsbeziehungen. Schachspiel. *Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae, sectio scientiarum mathematicarum* 3 (1927), 121–130.

[Kőnig, 1905] J. (G.) Kőnig. Zum Kontinuum-Problem, in: Krazer, A. (editor) *Verhandlungen des Dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904.* B.G. Teubner, Leipzig 1905, 144–147. Reprinted in *Mathematische Annalen* 60 (1905), 177–180 and Berichtigung, 462.

[Kunen and Vaughan, 1984] K. Kunen and J. E. Vaughan, eds. *Handbook of Set-Theoretic Topology*, North-Holland, Amsterdam 1984.

[Kuratowski, 1921] K. Kuratowski. Sur la notion de l'ordre dans la théorie des ensembles, *Fundamenta Mathematicae* 2 (1921), 161–171. Reprinted in [1988] below, 1–11.

[Kuratowski, 1922] K. Kuratowski. Une méthode d'élimination des nombres transfinis des raisonnements mathématiques, *Fundamenta Mathematicae* 3 (1922), 76–108.

[Kuratowski, 1931] K. Kuratowski. Evaluation de la class Borélienne ou projective d'un ensemble de points à l'aide des symboles logiques, *Fundamenta Mathematicae* 17 (1931), 249–272.

[Kuratowski, 1988] K. Kuratowski. K. Borsuk *et al.*, eds., *Selected Papers*, Państwowe Wydawnictwo Naukowe, Warsaw 1988.

[Kuratowski and Tarski, 1931] K. Kuratowski and A. Tarski. Les opérations logiques et les ensembles projectifs, *Fundamenta Mathematicae* 17 (1931), 240–248. Reprinted in Tarski [1986] vol. 1, 551–559, and in Kuratowski [1988], 367–375. Translated in Tarski [1983], 143–151.

[Kurepa, 1935] Đ. R. Kurepa. *Ensembles ordonnés et ramifiés*, Thèse, Paris. Published as: *Publications mathématiques de l'Université de Belgrade* 4 (1935), 1-138.

[Kurepa, 1942] Đ. R. Kurepa. propos d'une généralisation de la notion d'ensembles bien ordonnés, *Acta Mathematica* 75 (1942), 139–150.

[Kurepa, 1959] Đ. R. Kurepa. On the cardinal number of ordered sets and of symmetrical structures in dependence on the cardinal numbers of its chains and antichains, *Glasnik Matematičko-fizički i astronomski, Periodicum mathematico-physicum et astronomicum* 14 (1959), 183–203.

[Lavine, 1994] S. M. Lavine. *Understanding the Infinite*, Harvard University Press, Cambridge 1994.

[Lebesgue, 1902] H. Lebesgue. Intégrale, longueur, aire, *Annali di Matematica Pura ed Applicata* (3)7 (1902), 231–359. Reprinted in [1972] below, vol. 1, 203–331.

[Lebesgue, 1905] H. Lebesgue. Sur les fonctions représentables analytiquement, *Journal de mathématiques pures et appliquées* (6)1 (1905), 139-216. Reprinted in [1972] below, vol. 3, 103–180.

[Lebesgue, 1972] H. Lebesgue. *Oeuvres Scientifiques*, Kundig, Geneva 1972.

[Levy, 1957] A. Levy. Indépendance conditionelle de $V = L$ et d'axiomes qui se rattachent au système de M. Gödel, *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris* 245 (1957), 1582–1583.

[Levy, 1960] A. Levy. Axiom schemata of strong infinity in axiomatic set theory, *Pacific Journal of Mathematics* 10 (1960), 223–238.

[Levy, 1960a] A. Levy. Principles of reflection in axiomatic set theory, *Fundamenta Mathematicae* 49 (1960), 1–10.

[Levy, 1963] A. Levy. Independence results in set theory by Cohen's method. I, III, IV (abstracts), *Notices of the American Mathematical Society* 10 (1963), 592–593.

[Levy, 1964] A. Levy. Measurable cardinals and the continuum hypothesis (abstract), *Notices of the American Mathematical Society* 11 (1964), 769-770.

[Levy, 1965] A. Levy. Definability in axiomatic set theory I, in: Bar-Hillel, Yehoshua (editor), *Logic, Methodology and Philosophy of Science*, Proceedings of the 1964 International Congress, Jerusalem. North-Holland, Amsterdam 1965, 127–151.

[Levy, 1970] A. Levy. Definability in axiomatic set theory II, in: Bar-Hillel, Yehoshua (editor), *Mathematical Logic and Foundations of Set Theory*, North-Holland, Amsterdam 1970, 129–145.

[Levy and Solovay, 1967] A. Levy and R. M. Solovay. Measurable cardinals and the Continuum Hypothesis, *Israel Journal of Mathematics* 5 (1967), 234–248.

[Lindelöf, 1905] E. Lindelöf. Remarques sur un théorème fondamental de la théorie des ensembles, *Acta Mathematica* 29 (1905), 183-190.

[Lindenbaum and Tarski, 1926] A. Lindenbaum and A. Tarski. Communication sur les recherches de la théorie des ensembles, *Sprawozdana z Posiedzeń Towarzystwa Naukowego Warszawskiego, Wydział III, Nauk Matematyczno-fizycznych (Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie, Classe III, Sciences Mathématiques et Physiques)* 19 (1926), 299–330. Reprinted in Tarski [1986] vol. 1, 171–204.

[Liouville, 1844] J. Liouville. Des remarques relatives 1° à des classes très-étendues de quantités dont la valeur n'est ni rationelle ni même réductible à des irrationnelles algébriques; 2° à un passage du livre des *Principes* où Newton calcule l'action exercée par une sphère sur un point extérieur, *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris* 18 (1844), 883–885.

[Liouville, 1851] J. Liouville. Sur des classes très-étendues de quantités dont la valeur n'est ni algébrique ni même réductible à des irrationnelles algébriques, *Journal de mathématiques pures et appliquées* 16 (1851), 133–142.

[Łoś, 1955] J. Łoś. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres, in: Skolem, Thoralf, *et al.* (editors), *Mathematical Interpretation of Formal Systems*, North-Holland, Amsterdam 1955, 98–113.

[Löwenheim, 1915] L. Löwenheim.Über Möglichkeiten im Relativkalkul, *Mathematische Annalen* 76 (1915), 447–470. Translated in van Heijenoort [1967], 228–251.

[Luzin, 1914] N. N. Luzin. Sur un problème de M. Baire, *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris* 158 (1914), 1258–1261.

[Luzin, 1917] N. N. Luzin. Sur la classification de M. Baire, *Comptes Rendus Hebdomadaires des S´eances de l'Acad´emie des Sciences, Paris* 164 (1917), 91–94.

[Luzin, 1925] N. N. Luzin. Sur un problème de M. Emile Borel et les ensembles projectifs de M. Henri Lebesgue; les ensembles analytiques, *Comptes Rendus Hebdomadaires des S´eances de l'Acad´emie des Sciences, Paris* 180 (1925), 1318–1320.

[Luzin, 1925a] N. N. Luzin. Sur les ensembles projectifs de M. Henri Lebesgue, *Comptes Rendus Hebdomadaires des S´eances de l'Acad´emie des Sciences, Paris* 180 (1925), 1572–1574.

[Luzin, 1925b] N. N. Luzin. Les propriétés des ensembles projectifs, *Comptes Rendus Hebdomadaires des S´eances de l'Acad´emie des Sciences, Paris* 180 (1925), 1817–1819.

[Luzin and Sierpiński, 1918] N. N. Luzin and W. Sierpiński. Sur quelques propriétés des ensembles (A), *Bulletin de l'Acad´emie des Sciences de Cracovie, Classe des Sciences Math´ematiques et Naturelles, S´erie A* (1918), 35–48. Reprinted in Sierpiński [1975], 192–204.

[Luzin and Sierpiński, 1923] N. N. Luzin and W. Sierpiński. Sur un ensemble non mesurable *B*, *Journal de Math´ematiques Pures et Appliqu´ees* (9)2 (1923), 53-72. Reprinted in Sierpiński [1975], 504-519.

[Mahlo, 1911] P. Mahlo. Über lineare transfinite Mengen, *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematische-Physische Klasse* 63 (1911), 187–225.

[Mahlo, 1912] P. Mahlo. Zur Theorie und Anwendung der ρ_0-Zahlen, *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematische-Physische Klasse* 64 (1912), 108–112.

[Mahlo, 1913] P. Mahlo. Zur Theorie und Anwendung der ρ_0-Zahlen II, *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematische-Physische Klasse* 65 (1913), 268–282.

[Mahlo, 1913a] P. Mahlo. Über Teilmengen des Kontinuums von dessen M¨achtigkeit, *Berichte über die Verhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematische-Physische Klasse* 65 (1913), 283–315.

[Martin and Solovay, 1969] D. A. Martin and R. M. Solovay. A basis theorem for Σ_3^1 sets of reals, *Annals of Mathematics* 89 (1969), 138-159.

[Mathias, 2001] A. R. D. Mathias. Slim models of Zermelo set theory, *The Journal of Symbolic Logic* 66 (2001), 487–496.

[Meschkowski, 1983] H. Meschkowski. *Georg Cantor. Leben. Werk und Wirkung*, Bibliographisches Institut, Mannheim 1983.

[Meschkowski and Nilson, 1991] H. Meschkowski and W. Nilson. *Briefe / Georg Cantor*, Springer, Berlin 1991.

[Miller, 1984] A. W. Miller. Special subsets of the real line, in: Kunen-Vaughan [1984], 201–233.

[Mirimanoff, 1917] D. Mirimanoff. Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles, *L'Enseignment Math´ematique* 19 (1917), 37–52.

[Mirimanoff, 1917a] D. Mirimanoff. Remarques sur la théorie des ensembles et les antinomies Cantoriennes. I, *L'Enseignment Math´ematique* 19 (1917), 209–217.

[Montague, 1956] R. M. Montague. Zermelo-Fraenkel set theory is not a finite extension of Zermelo set theory (abstract), *Bulletin of the American Mathematical Society* 62 (1956), 260.

[Montague, 1961] R. M. Montague. Fraenkel's addition to the axioms of Zermelo, in: Bar-Hillel–Poznanski–Rabin–Robinson [1961], 91–114.

[Moore, 1982] G. H. Moore. *Zermelo's Axiom of Choice. Its Origins, Development and Influence*, Springer, New York 1982.

[Moore, 1988] G. H. Moore. The roots of Russell's paradox, *Russell* 8 (new series) (1988), 46–56.

[Moore, 1988a] G. H. Moore. The emergence of first-order logic, in: Aspray, William, and Philip Kitcher (editors), *History and Philosophy of Modern Mathematics*, Minnesota Studies in the Philosophy of Science, vol. 11, University of Minnesota Press, Minneapolis 1988, 95–135.

[Moore, 1989] G. H. Moore. Towards a history of Cantor's Continuum Problem, in: Rowe, David E., and John McCleary (editors), *The History of Modern Mathematics*, vol. 1: *Ideas and their Reception*, Academic Press, Boston 1989, 79–121.

[Moore, 2002] G. H. Moore. Hilbert on the infinite: The role of set theory in the evolution of Hilbert's thought, *Historia Mathematics* 29 (2002), 40–64.

[Moore and Garciadiego, 1981] G. H. Moore and A. R. Garciadiego. Burali-Forti's paradox: A reappraisal of its origins, *Historia Mathematica* 8 (1981), 319–350.

[Moschovakis, 1980] Y. Moschovakis. *Descriptive Set Theory*, North-Holland, Amsterdam 1980.

[Mostowski, 1939] A. M. Mostowski. Über die Unabh¨angigkeit des Wohlordnungssatzes vom Ordnungsprinzip, *Fundamenta Mathematicae* 32 (1939), 201–252. Translated in [1979] below, vol. 1, 290–338.

[Mostowski, 1949] A. M. Mostowski. An undecidable arithmetical statement, *Fundamenta Mathematicae* 36 (1949), 143–164. Reprinted in [1979] below, vol. 1, 531–552.

[Mostowski, 1950] A. M. Mostowski. Some impredicative definitions in the axiomatic set theory, *Fundamenta Mathematicae* 37 (1950), 111–124; correction in 38 (1952), 238.

[Mostowski, 1979] A. M. Mostowski. *Foundational Studies. Selected Works,* North-Holland, Amsterdam 1979.

[Myhill and Scott, 1971] J. R. Myhill and D. S. Scott. Ordinal definability, in: Scott [1971], 271–278.

[Nešetřil and Rödl, 1990] J. Nešetřil and V. Rödl. *Mathematics of Ramsey Theory*, Springer, Berlin 1990.

[Noether and Cavaillès, 1937] E. Noether and J. Cavaillès, eds. *Briefwechsel Cantor-Dedekind*, Hermann, Paris 1937. Translated into French in Cavaillès [1962].

[Novak, 1950] I. L. Novak. A construction for models of consistent systems, *Fundamenta Mathematicae* 37 (1950), 87–110.

[Oxtoby, 1971] J. C. Oxtoby. *Measure and Category. A Survey of the Analogies between Topological and Measure Spaces*, Springer, New York 1971.

[Parsons, 1977] C. Parsons. What is the iterative concept of set? in: Butts, Robert E., and Jaakko Hintikka (editors), *Logic, Foundations of Mathematics and Computability Theory.* Proceedings of the Fifth International Congress of Logic, Methodology, and the Philosophy of Science (London, Ontario 1975), The University of Western Ontario Series in Philosophy and Science, vol. 9; D. Reidel, Dordrecht 1977, 335–367. Reprinted in Benacerraf-Putnam [1983], 503–529.

[Peano, 1897] G. Peano. Studii di logica matematica, *Atti della Accademia delle scienze di Torino, Classe di scienze fisiche, matematiche e naturali* 32 (1897), 565–583. Reprinted in [1958] below, 201–217.

[Peano, 1905-8] G. Peano. *Formulario Mathematico*, Bocca, Torino 1905-8. Reprinted Edizioni Cremonese, Rome 1960.

[Peano, 1911] G. Peano. Sulla definizione di funzione. *Atti della Accademia nazionale dei Lincei, Rendiconti, Classe di scienze fisiche, matematiche e naturali* 20-I (1911), 3–5.

[Peano, 1913] G. Peano. Review of: A.N. Whitehead and B. Russell, *Principia Mathematica*, vols. I,II, *Bollettino di bibliografia e storia delle scienze matematiche Loria* 15 (1913), 47–53, 75–81. Reprinted in [1958] below, 389–401.

[Peano, 1958] G. Peano. *Opere Scelte*, vol. 2, Edizioni Cremonese, Rome 1958.

[Peckhaus, 1990] V. Peckhaus. "Ich habe mich wohl gehütet alle Patronen auf einmal zu verschiessen." Ernst Zermelo in Göttingen, *History and Philosophy of Logic* 11 (1990), 19–58.

[Peirce, 1883] C. S. Peirce. A theory of probable inference. Note B. The logic of relatives, *Studies in Logic by Members of the John Hopkins University*. Boston 1883, 187–203. Reprinted in: Hartshorne, Charles, and Paul Weiss (editors), *Collected Papers of Charles Sanders Peirce*, vol. 3; Harvard University Press, Cambridge 195–209.

[Pincus, 1972] D. Pincus. Zermelo-Fraenkel consistency results by Fraenkel-Mostowski methods, *The Journal of Symbolic Logic* 37 (1972), 721–743.

[Plotkin, 1993] J. M. Plotkin. Who put the "Back" in "Back-and-Forth?" in: Crossley, Remmel, Shore and Sweedler (editors), *Logical Methods, In Honor of Anil Nerode's Sixtieth Birthday*, Birkhäuser, Boston 1993.

[Plotkin, 2005] J. M. Plotkin. *Hausdorff on Ordered Sets*, American Mathematical Society, Providence 2005.

[Purket, 1989] W. Purkert. Cantor's views on the foundations of mathematics, in: Rowe, David E., and John McCleary (editors), *The History of Modern Mathematics*, vol. 1: *Ideas and their Reception*, Academic Press, Boston 1989, 49–65.

[Purket, 2002] W. Purket. Grundzüge der Mengenlehre – Historische Einführung, in: Felix Hausdorff, *Gesammelte Werke*, vol. 2, 1–89.

[Purket and Ilgauds, 1987] W. Purkert and H. J. Ilgauds. *Georg Cantor: 1845–1918*, Birkhäuser, Basel 1987.

[Quine, 1960] W. V. O. Quine. *Word and Object*, MIT Press, Cambridge 1960.

[Ramsey, 1930] F. P. Ramsey. On a problem of formal logic, *Proceedings of the London Mathematical Society* (2)30 (1930), 264–286. Reprinted in Gessel-Rota [1987], 2–24.

[Robinson, 1947] R. M. Robinson. On the decomposition of spheres, *Fundamenta Mathematicae* 34 (1947), 246–270.

[Robinson, 1951] A. Robinson. *On the Metamathematics of Algebra*, North-Holland, Amsterdam 1951.

[Rosser and Wang, 1950] J. B. Rosser and H. Wang. Non-standard models for formal logics, *The Journal of Symbolic Logic* 15 (1950), 113–129.

[Rothberger, 1938] F. Rothberger. Eine Äquivalenz zwischen der Kontinuumhypothese und der Existenz der Lusinschen und Sierpińskischen Mengen, *Fundamenta Mathematicae* 30 (1938), 215–217.

[Rothberger, 1939] F. Rothberger. Sur un ensemble toujours de première categorie qui est depourvu de la propreté λ, *Fundamenta Mathematicae* 32 (1939), 294–300.

[Rothberger, 1948] F. Rothberger. On some problems of Hausdorff and Sierpiński, *Fundamenta Mathematicae* 35 (1948), 29–46.

[Rubin and Rubin, 1985] H. Rubin and J. E. Rubin. *Equivalents of the Axiom of Choice, II.* Amsterdam, North-Holland 1985. Revised, expanded version of their *Equivalents of the Axiom of Choice*, North-Holland, Amsterdam 1963.

[Russell, 1901] B. Russell. Sur la logique des relations avec des applications à la théorie des séries, *Revue de math ́ematiques (Rivista di matematica) 7*, 115–148; partly reprinted in Gregory H. Moore (editor), *The Collected Papers of Bertrand Russell*, vol. 3, Routledge, London 1993, 613–627; translated in same, 310–349.

[Russell, 1903] B. Russell. *The Principles of Mathematics*, Cambridge University Press, Cambridge 1903. Later editions, George Allen & Unwin, London.

[Russell, 1906] B. Russell. On some difficulties in the theory of transfinite numbers and order types, *Proceedings of the London Mathematical Society* (2)4 (1906), 29–53. Reprinted in [1973] below, 135–164.

[Russell, 1959] B. Russell. *My Philosophical Development*, George Allen & Unwin, London 1959.

[Russell, 1973] B. Russell. D. Lackey, ed., *Essays in Analysis*, George Braziller, New York 1973.

[Schr ̈oder, 1890] E. Schr ̈oder. *Vorlesungen über die Algebra der Logik (exakte Logik).* Vol. 1. Leipzig, B.G. Teubner 1890. Reprinted in [1966] below.

[Schr ̈oder, 1895] E. Schr ̈oder. *Vorlesungen über die Algebra der Logik (exakte Logik).* Vol. 3: *Algebra und Logik der Relative.* Leipzig, B.G. Teubner 1895. Reprinted in [1966] below.

[Schr ̈oder, 1966] E. Schr ̈oder. *Vorlesungen über die Algebra der Logik* (three volumes), Chelsea, New York 1966.

[Scott, 1961] D. S. Scott. Measurable cardinals and constructible sets, *Bulletin de l'Acad ́emie Polonaise des Sciences, S ́erie des Sciences Math ́ematiques, Astronomiques et Physiques* 9 (1961), 521–524.

[Scott, 1971] D. S. Scott. *Axiomatic Set Theory*, Proceedings of Symposia in Pure Mathematics vol. 13, part 1, American Mathematical Society, Providence 1971.

[Scott, 1974] D. S. Scott. Axiomatizing set theory, in: Jech, Thomas J. (editor) *Axiomatic Set Theory.* Proceedings of Symposia in Pure Mathematics vol. 13, part 2, American Mathematical Society, Providence 1974, 207–214.

[Shelah, 1980] S. Shelah. Going to Canossa (abstract). *Abstracts of papers presented to the American Mathematical Society* 1 (1980), 630.

[Shelah, 1984] S. Shelah. Can you take Solovay's inaccessible away? *Israel Journal of Mathematics* 48 (1984), 1–47.

[Shepherdson, 1951] J. C. Shepherdson. Inner models for set theory —Part I, *The Journal of Symbolic Logic* 16 (1951), 161–190.

[Shepherdson, 1952] J. C. Shepherdson. Inner models for set theory —Part II, *The Journal of Symbolic Logic* 17 (1952), 225–237.

[Shepherdson, 1953] J. C. Shepherdson. Inner models for set theory —Part III, *The Journal of Symbolic Logic* 18 (1953), 145–167.

[Shoenfield, 1954] J. R. Shoenfield. A relative consistency proof, *The Journal of Symbolic Logic* 19 (1954), 21–28.

[Shoenfield, 1959] J. R. Shoenfield. On the independence of the axiom of constructibility, *American Journal of Mathematics* 81 (1959), 537–540.

[Shoenfield, 1967] J. R. Shoenfield. *Mathematical Logic*, Addison-Wesley, Reading 1967.

[Shoenfield, 1971] J. R. Shoenfield. Unramified forcing, in: Scott [1971], 357–381.

[Shoenfield, 1977] J. R. Shoenfield. Axioms of set theory, in: Barwise, K. Jon (editor), *Handbook of Mathematical Logic*, North-Holland, Amsterdam 1977, 321-344.

[Sierpiński, 1918] W. Sierpiński. L'axiome de M. Zermelo et son rôle dans la Théorie des Ensembles et l'Analyse, *Bulletin de l'Acad ́emie des Sciences de Cracovie, Classe des Sciences Math ́ematiques et Naturelles, S ́erie A* (1918), 97–152. Reprinted in [1975] below, 208–255.

[Sierpiński, 1924] W. Sierpiński. Sur l'hypothèse du continu ($2^{\aleph_0} = \aleph_1$), *Fundamenta Mathematicae* 5 (1924), 177–187. Reprinted in [1975] below, 527–536.

[Sierpiński, 1925] W. Sierpiński. Sur une class d'ensembles, *Fundamenta Mathematicae* 7 (1925), 237–243. Reprinted in [1975] below, 571–576.

[Sierpiński, 1928] W. Sierpiński. Sur un ensemble non dénombrable, dont toute image continue est de mesure nulle, *Fundamenta Mathematicae* 11 (1928), 302–304. Reprinted in [1975] below, 702–704.

[Sierpiński, 1934] W. Sierpiński. *Hypothèse du Continu*, Monografie Matematyczne vol. 4, Warsaw 1934. Second, revised edtion, Chelsea, New York 1956.

[Sierpiński, 1950] W. Sierpiński. *Les ensembles projectifs et analytiques*, Mémorial des Sciences Mathématiques #112, Gauthier-Villars, Paris 1950.

[Sierpiński, 1975] W. Sierpiński. S. Hartman *et al.*, eds., *Oeuvres Choisies*. vol. 2, Warsaw, Państwowe Wydawnictwo Naukowe 1975.

[Sierpiński, 1976] W. Sierpiński. S. Hartman *et al.*, eds., *Ouevres Choisies*. vol. 3, Państwowe Wydawnictwo Naukowe, Warsaw 1976.

[Sierpiński and Tarski, 1930] W. Sierpiński and A. Tarski. Sur une propriété caractéristique des nombres inaccessibles, *Fundamenta Mathematicae* 15 (1930), 292–300. Reprinted in Sierpiński [1976], 29-35, and in Tarski [1986] vol. 1, 289–297.

[Skolem, 1920] T. Skolem. Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze nebst einem Theoreme über dichte Mengen, *Videnskaps-selskapets Skrifter, I. Matematisk-Naturvidenskabelig Klass* (1920, #4), 1-36. Reprinted in [1970] below, 103–136. Partially translated in van Heijenoort [1967], 252–263.

[Skolem, 1923] T. Skolem. Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre, in: *Matematikerkongressen i Helsingfors den 4–7 Juli 1922, Den femte skandinaviska matematikerkongressen, Redogörelse,* Akademiska-Bokhandeln, Helsinki 1923, 217-232. Reprinted in [1970] below, 137–152. Translated in van Heijenoort [1967], 290–301.

[Skolem, 1930] T. Skolem. Einige Bemerkungen zu der Abhandlung von E. Zermelo: "Über die Definitheit in der Axiomatik", *Fundamenta Mathematicae* 15 (1930), 337–341. Reprinted in [1970] below, 275–279.

[Skolem, 1933] T. Skolem. Ein kombinatorischer Satz mit Anwendung auf ein logisches Entscheidungsproblem, *Fundamenta Mathematicae* 20 (1933), 254–261. Reprinted in [1970] below, 337–344.

[Skolem, 1933a] T. Skolem. Über die Unmöglichkeit einer vollständigen Charakterisierung der Zahlenreihe mittels eines endlichen Axiomensystems, *Norsk Matematisk Forenings Skrifter* 2(#10) (1933), 73–82. Reprinted in [1970] below, 345–354.

[Skolem, 1934] T. Skolem. Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Assagen mit ausschliesslich Zahlenvariablen, *Fundamenta Mathematicae* 23 (1934), 150–161. Reprinted in [1970] below, 355–366.

[Skolem, 1970] T. Skolem. J. E. Fenstad, ed., *Selected Works in Logic*, Univesitetsforlaget, Oslo 1970.

[Solovay, 1963] R. M. Solovay. Independence results in the theory of cardinals. I, II (abstracts), *Notices of the American Mathematical Society* 10 (1963), 595.

[Solovay, 1965] R. M. Solovay. $2^{\aleph_0}$ can be anything it ought to be (abstract), in: Addison-Henkin-Tarski [1965], 435.

[Solovay, 1965a] R. M. Solovay. Measurable cardinals and the continuum hypothesis (abstract), *Notices of the American Mathematical Society* 12 (1965), 132.

[Solovay, 1965b] R. M. Solovay. The measure problem (abstract), *Notices of the American Mathematical Society* 12 (1965), 217.

[Solovay, 1969] R. M. Solovay. The cardinality of Σ_2^1 sets of reals, in: Bulloff, Jack J., Thomas C. Holyoke, and Samuel W. Hahn (editors), *Foundations of Mathematics*. Symposium papers commemorating the sixtieth birthday of Kurt Gödel. Springer, Berlin 1969, 58–73.

[Solovay, 1970] R. M. Solovay. A model of set theory in which every set of reals is Lebesgue measurable, *Annals of Mathematics* 92 (1970), 1–56.

[Specker, 1957] E. Specker. Zur Axiomatik der Mengenlehre (Fundierungs- und Auswahlaxiom), *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 3 (1957), 173–210.

[Steinitz, 1910] E. Steinitz. Algebraische Theorie der Körper, *Journal für die reine und angewandte Mathematik* 137 (1910), 167–309.

[Suslin, 1917] M. Y. Suslin. Sur une définition des ensembles mesurables *B* sans nombres transfinis, *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris* 164 (1917), 88–91.

[Suslin, 1920] M. Y. Suslin. Problème 3, *Fundamenta Mathematicae* 1 (1920), 223.

[Tarski, 1924] A. Tarski. Sur quelques théorèmes qui équivalent à l'axiome du choix, *Fundamenta Mathematicae* 5 (1924), 147–154. Reprinted in [1986] below, vol. 1, 41–48.

[Tarski, 1931] A. Tarski. Sur les ensembles définissables de nombres réels, *Fundamenta Mathematicae* 17 (1931), 210–239. Translated in Tarski [1983], 110–142.

[Tarski, 1933] A. Tarski. Pojęcie prawdy w językach nauk dedukcyjnych (The concept of truth in the languages of deductive sciences), *Prace Towarzystwa Naukowego Warszawskiego, Wydział III, Nauk Matematyczno-fizycznych (Travaux de la Société des Sciences et des Lettres de Varsovie, Classe III, Sciences Mathématiques et Physiques* #34 (1933). See also [1935] below.

[Tarski, 1935] A. Tarski. Der Wahrheitsbegriff in den formalisierten Sprachen (German translation of [1933] with a postscript), *Studia Philosophica* 1 (1935), 261–405. Reprinted in [1986] below, vol. 2, 51–198. Translated in [1983] below, 152–278.

[Tarski, 1951] A. Tarski. *A Decision Method for Elementary Algebra and Geometry* (prepared by J.C.C. McKinsey), University of California Press, Berkeley 1951. Second revised edition.

[Tarski, 1962] A. Tarski. Some problems and results relevant to the foundations of set theory, in: Nagel, Ernest, Patrick Suppes, and Alfred Tarski (editors). *Logic, Methodology and Philosophy of Science.* Proceedings of the 1960 [and first] International Congress (Stanford, California), Stanford University Press, Stanford 1962, 125–135. Reprinted in [1986] below, vol. 4, 115–125.

[Tarski, 1983] A. Tarski. *Logic, Semantics, Metamathematics. Papers from 1923 to 1938.* Translations by J.H. Woodger, second edition, Hackett, Indianapolis 1983.

[Tarski, 1986] A. Tarski. S. R. Givant and R. N. McKenzie, eds., *Collected Papers*, Basel, Birkhäuser 1986.

[Todorčević, 1984] S. Todorčević. Trees and linearly ordered sets, in: Kunen–Vaughan [1984], 235–293.

[Ulam, 1930] S. M. Ulam. Zur Masstheorie in der allgemeinen Mengenlehre, *Fundamenta Mathematicae* 16 (1930), 140–150. Reprinted in [1974] below, 9–19.

[Ulam, 1974] S. M. Ulam. W. A. Beyer, J. Mycielski, and G.-C. Rota, eds., *Sets, Numbers, and Universes.* Selected Works, MIT Press, Cambridge 1974.

[van Douwen, 1984] E. K. van Douwen. The integers and topology, in: Kunen–Vaughan [1984], 111-168.

[van Heijenoort, 1967] J. van Heijenoort, ed. *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931.* Harvard University Press, Cambridge 1967.

[Vitali, 1905] G. Vitali. Sul problema della misura dei gruppi di punti di una retta, Tip. Gamberini e Parmeggiani, Bologna 1905.

[von Neumann, 1923] J. von Neumann. Zur Einführung der transfiniten Zahlen, *Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae, sectio scientiarum mathematicarum* 1 (1923), 199–208. Reprinted in [1961] below, 24–33. Translated in van Heijenoort [1967], 346–354.

[von Neumann, 1925] J. von Neumann. Eine Axiomatisierung der Mengenlehre, *Journal für die reine und angewandte Mathematik* 154 (1925), 219–240. Berichtigung 155, 128. Reprinted in [1961] below, 34–56. Translated in van Heijenoort [1967], 393–413.

[von Neumann, 1928] J. von Neumann. Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre, *Mathematische Annalen* 99 (1928), 373–391. Reprinted in [1961] below, 320–338.

[von Neumann, 1928a] J. von Neumann. Die Axiomatisierung der Mengenlehre, *Mathematische Zeitschrift* 27 (1928), 669–752. Reprinted in [1961] below, 339–422.

[von Neumann, 1929] J. von Neumann. Über eine Widerspruchfreiheitsfrage in der axiomaticschen Mengenlehre, *Journal für die reine und angewandte Mathematik* 160 (1929), 227–241. Reprinted in [1961] below, 494–508.

[von Neumann, 1961] J. von Neumann. A. H. Taub, ed., *John von Neumann, Collected Works*, vol. 1, Pergamon Press, New York 1961.

[Vopěnka, 1962] P. Vopěnka. Construction of models of set theory by the method of ultraproducts (in Russian), *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 8 (1962), 293–304.

[Vopěnka, 1964] P. Vopěnka. The independence of the Continuum Hypothesis (in Russian), *Commentationes Mathematicae Universitatis Carolinae* 5 Supplement I (1964), 1-48. Translated in *American Mathematical Society Translations* 57 (1966), 85-112.

[Vopěnka, 1965] P. Vopěnka. Construction of a model for Gödel-Bernays set theory for which the class of natural numbers is a set of the model and a proper class in the theory, in: Addison-Henkin-Tarski [1965], 436–437.

[Vopěnka, 1967] P. Vopěnka. The general theory of ∇-models, *Commentationes Mathematicae Universitatis Carolinae* 8 (1967), 145-170.

[Wagon, 1985] S. Wagon. *The Banach-Tarski Paradox*, Encyclopedia of Mathematics and Its Applications, vol. 24, Cambridge University Press, Cambridge 1985. Paperback edition 1993.

[Wang, 1974] H. Wang. *From Mathematics to Philosophy,* Humanities Press, New York 1974.

[Wang, 1974a] H. Wang. The concept of set, in Wang [1974], 181–223. Reprinted in Benacerraf-Putnam [1983], 530–570.

[Whitehead and Russell, 1910] A. N. Whitehead and B. Russell. *Principia Mathematica*, vol. 1, Cambridge University Press, Cambridge 1910.

[Whitehead and Russell, 1912] A. N. Whitehead and B. Russell. *Principia Mathematica*, vol. 2, Cambridge University Press, Cambridge 1912.

[Whitehead and Russell, 1913] A. N. Whitehead and B. Russell. *Principia Mathematica*, vol. 3, Cambridge University Press, Cambridge 1913.

[Wiener, 1914] N. Wiener. A simplification of the logic of relations, *Proceedings of the Cambridge Philosophical Society* 17 (1914), 387–390. Reprinted in van Heijenoort [1967], 224–227.

[Wittgenstein, 1956] L. Wittgenstein. G. H. von Wright, R. Rhees, and G. E. M. Anscombe, eds., *Remarks on the Foundations of Mathematics,* Basil Blackwell, Oxford 1956. Second printing 1967.

[Zermelo, 1904] E. Zermelo. Beweis, dass jede Menge wohlgeordnet werden kann (Aus einem an Herrn Hilbert gerichteten Briefe), *Mathematische Annalen* 59 (1904), 514–516. Translated in van Heijenoort [1967], 139–141.

[Zermelo, 1908] E. Zermelo. Neuer Beweis f̈ur die M̈oglichkeit einer Wohlordnung, *Mathematische Annalen* 65 (1908), 107–128. Translated in van Heijenoort [1967], 183–198.

[Zermelo, 1913] E. Zermelo. Über eine Anwendung der Mengenlehre auf die Theorie des Schachspiels, in Hobson, Ernest W., and A.E.H. Love (editors), *Proceedings of the Fifth International Congress of Mathematicians*, Cambridge 1912, vol. 2, Cambridge University Press, Cambridge, 1913, 501–504.

[Zermelo, 1929] E. Zermelo. Über den Begriff der Definitheit in der Axiomatik, *Fundamenta Mathematicae* 14 (1929), 339–344.

[Zermelo, 1930] E. Zermelo. Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen ̈uber die Grundlagen der Mengenlehre, *Fundamenta Mathematicae* 16 (1930), 29–47. Translated with commentary by Michael Hallett in Ewald [1996], 1208–1233.

[Zermelo, 1931] E. Zermelo. Über Stufen der Quantifikation und die Logik des Unendlichen, *Jahresbericht der Deutschen Mathematiker-Vereinigung* 41 (1931), 85–88.

[Zermelo, 1935] E. Zermelo. Grundlagen einer allgemeinen Theorie der mathematicschen Satzsysteme, *Fundamenta Mathematicae* 25 (1935), 136–146.

[Zorn, 1935] M. Zorn. A remark on method in transfinite algebra, *Bulletin of the American Mathematical Society* 41 (1935), 667–670.