THE BULLETIN OF SYMBOLIC LOGIC
Volume 10, Number 4, Dec. 2004

ZERMELO AND SET THEORY

AKIHIRO KANAMORI[1]

Dedicated to the memory of Frau Gertrud Zermelo

Ernst Friedrich Ferdinand Zermelo (1871–1953) transformed the set theory of Cantor and Dedekind in the first decade of the 20th century by incorporating the Axiom of Choice and providing a simple and workable axiomatization setting out generative set-existence principles. Zermelo thereby tempered the ontological thrust of early set theory, initiated the delineation of what is to be regarded as set-theoretic, drawing out the combinatorial aspects from the logical, and established the basic conceptual framework for the development of modern set theory. Two decades later Zermelo promoted a distinctive cumulative hierarchy view of models of set theory and championed the use of infinitary logic, anticipating broad modern developments. In this paper Zermelo's published mathematical work in set theory is described and analyzed in its historical context, with the hindsight afforded by the awareness of what has endured in the subsequent development of set theory. Elaborating formulations and results are provided, and special emphasis is placed on the to and fro surrounding the Schröder-Bernstein Theorem and the correspondence and comparative approaches of Zermelo and Gödel. Much can be and has been written about philosophical and biographical issues and about the reception of the Axiom of Choice, and we will refer and defer to others, staying the course through the decidedly mathematical themes and details.

§1. Beginnings. Zermelo, born at the time Cantor was making his first incursions into the transfinite, would make the first large moves in set theory

Received March 8, 2004; accepted June 2, 2004.

[1]This is a much expanded version of an invited address, on the occasion of the 50th anniversary of the death of Zermelo, at the 12th International Congress of Logic, Methodology and Philosophy of Science held August 7–13, 2003 at Oviedo, Spain, and the author gratefully thanks the organizers for the invitation. The author also gratefully acknowledges the generous support of the Dibner Institute for the History of Science and Technology, the hospitality of the Department of Pure Mathematics and Mathematical Statistics at Cambridge University, and helpful and valuable comments by Heinz-Dieter Ebbinghaus, Juliet Floyd, Volker Peckhaus, Gregory Taylor, and the referee.

1079-8986/04/1004-0002/$7.70

after him. But those moves, conceptual and foundational, came at the crest of considerable youthful experience as a concrete mathematician working in applied mathematics and mathematical physics. Zermelo completed his *Dissertation* [1894] on the Weierstrassian calculus of variations at the University of Berlin. Turning to mathematical physics, Zermelo then became Max Planck's assistant at the Institute for Theoretical Physics at Berlin from 1894 to 1897. During this period Zermelo [1896] applied the Poincaré recurrence theorem for dynamical systems, giving its first really elegant proof, to raise a substantial objection to the elder and eminent Ludwig Boltzmann's longstanding contention that the kinetic theory of gases can explain irreversibility of macroscopic physical processes. Zermelo [1896a] thereupon engaged in a penetrating exchange with Boltzmann on the explanation of these processes.[2] Moving to Göttingen to work on his *Habilitation* on hydrodynamics, he completed it (Zermelo [1902]) in 1899 and thereupon began lecturing as a *Privatdozent* for Mathematics.

Zermelo would stay at Göttingen through the first decade of the new century, in what would become the most productive and influential period of his mathematical career.[3] He continued to publish and maintain an active interest in the calculus of variations and mathematical physics, writing with Hans Hahn an encyclopedia article Zermelo–Hahn [1904] on the first topic and translating into German J. Willard Gibbs's 1902 treatise *The Elementary Principles of Statistical Mechanics*.[4] But newly stimulated, Zermelo also turned to the main engagement of his mathematical career, set theory and its foundations. In the winter semester of 1900–1 he gave his first course of lectures on set theory, following Cantor's *Beiträge* [1895, 1897]. A byproduct was Zermelo's first set-theoretic paper, a note [1901] on cardinal arithmetic. In May 1903, he spoke on the work of Frege at the Göttingen Mathematical Society, comparing the concept of number in Cantor, Dedekind, and Frege.

In Zermelo's turn to set theory and its foundations the influence of David Hilbert was pivotal.[5] Hilbert at Göttingen was at the height of his powers.

[2]See Hollinger–Zenzen [1985] for a study of the dynamics and physical origins of irreversibility. Zermelo's result figured prominently in a debate between Boltzmann and Planck, tilting it in the latter's favor and establishing for Zermelo a reputation among the physicists.

[3]Peckhaus [1990: chap. 4] provides a biographical sketch and detailed account of Zermelo's years 1897–1910 at Göttingen.

[4]See Gibbs [1905]. Notably, the title of the treatise continues "developed with especial reference to the rational foundation of thermodynamics". Gibbs did not succeed in providing a rational foundation, but one can surmise that Zermelo was already predisposed to foundational studies.

[5]Many years later, Zermelo wrote in a report to the Emergency Society of German Science [Wissenschaft] (See Moore [1980]; the society supported poor scientists and had given Zermelo a fellowship, and the report was sent to the society with a letter dated December 3, 1930.): "Thirty years ago, when I was a *Privatdozent* at Göttingen, I came under the influence of D. Hilbert, to whom I am surely the most indebted for my mathematical development, and

Having completed his axiomatization of geometry *Grundlagen der Geometrie* [1899], Hilbert [1900a] extended his axiomatic method to the real numbers. In his famous list [1900] of problems for the 20th century the second problem was to establish the consistency of his axiomatization. The first problem was to establish the Continuum Hypothesis, and in connection with this he expressed [1900: 263] the desirability of "actually describing" a well-ordering of the real numbers.[6]

Zermelo's first substantial result in set theory was his independent discovery of Russell's Paradox. He then established [1904] the Well-Ordering Theorem, that every set can be well-ordered, provoking an open controversy about this initial explicit use of the Axiom of Choice. After providing a second proof [1908] of the Well-Ordering Theorem in response, Zermelo also provided the first full-fledged axiomatization [1908a] of set theory. In the process, he ushered in an abstract, generative view of sets that would come to dominate in the years to come.

Zermelo's independent discovery of the argument for Russell's Paradox is substantiated in a note dated 16 April 1902 found in the philosopher Edmund Husserl's *Nachlass*.[7] According to the note, Zermelo pointed out that any set M containing all of its subsets as members, i.e., $\mathcal{P}(M) \subseteq M$, is "inconsistent" by considering the subset $M_0 = \{x \in M \mid x \notin x\}$: If $M_0 \in M$, then of course $M_0 \in M_0$ *iff* $M_0 \notin M_0$. Schröder [1890: 245] had argued that Boole's "class 1" regarded as consisting of everything conceivable is inconsistent, and Husserl in a review [1891] had criticized Schröder's argument for not distinguishing between inclusion and membership. That inclusion may imply membership is the same concern that Bertrand Russell had to confront: For the Russell of the *The Principles of Mathematics* [1903] mathematics was to be articulated in an all-encompassing logic, a complex philosophical system based on universal categories. For his universal class U, it would have to be that $\mathcal{P}(U) \subseteq U$, leading to contradiction. Zermelo on the other hand did not push the argument in the direction of paradox as Russell had done, but was merely to regard it as an indication that the notion of set would have to be restricted.[8] Of course, much more significant than $\mathcal{P}(M) \nsubseteq M$ is that $\mathcal{P}(M)$ has higher cardinality than M. However, Cantor [1891] had actually introduced his diagonal argument to show that

I began to occupy myself with the foundational questions of mathematics, especially with the fundamental problems of Cantorian *set theory*, whose full significance I learned to appreciate only then, through the extremely fruitful collaboration of the Göttingen mathematicians."

[6]See Ewald [1996: 1096ff] for an updated translation of the introductory material and the first two problems from Hilbert's 1900 address. See Dreben–Kanamori [1997] for more on Hilbert and set theory.

[7]See Rang-Thomas [1981].

[8]In the earliest notes about axiomatization found in his *Nachlass*, probably written around 1905, Zermelo took the assertion $M \notin M$ to be an axiom, as well as the assertion that any "well-defined" set M has a subset not a member of M (see Moore [1982: 155]).

for any set X the collection of functions from X into a two-element set is of a strictly higher cardinality than that of X. The connection between sub*sets* and characteristic *functions* was not generally appreciated then, and Zermelo was just making the first moves in his abstract approach to sets.

A letter from Hilbert bears on issues of context, paradox and priority. Hilbert wrote to Frege on 7 November 1903 (Frege [1976: 79ff][1980: 51ff]):

> Many thanks for the second volume of your *Grundgesetze*, which I find very interesting. Your example at the end of the book (p. 253) [i.e., Russell's Paradox] was known to us here;* I found other even more convincing contradictions as long as 4–5 years ago; they led me to the conviction that traditional logic is inadequate and that the theory of concept-formation [Begriffsbildung] needs to be sharpened and refined. As I see it, the most important gap in the traditional structure of logic is the assumption made by all logicians and mathematicians up to now that a concept is already there if one can state of any object whether or not it falls under it. This does not seem adequate to me. What is decisive is the recognition that the axioms that define the concept are free from contradiction.
>
> *I believe Dr. Zermelo discovered it 3–4 years ago after I had communicated my examples to him.

In early 1897 Cantor had concluded with a *reductio* argument that the collection of all the alephs precluded the viability of sets as extensions of concepts, as otherwise there would be an aleph bigger than all alephs,[9] and this would lead by 1899 to a new engagement with "absolutely infinite or inconsistent multiplicities" of which more below. Cantor began corresponding with Hilbert on these matters,[10] and Hilbert devised his "even more convincing contradictions" around this time as mentioned in the above passage and Zermelo too became aware that "the theory of concept-formation [Begriffsbildung] needs to be sharpened or refined."

Peckhaus–Kahle [2002] describes "Hilbert's Paradox" taken from 1905 course notes at Göttingen, a paradox which can be cast in simple terms as follows: There is no set S satisfying (a) if $X \in S$, then its power set $\mathcal{P}(X) \in S$, and (b) if $T \subseteq S$, then its union $\bigcup T \in S$. Suppose that there were such an S. Then $\mathcal{P}(\bigcup S) \in S$ by (b) and then (a). But then, $\mathcal{P}(\bigcup S) \subseteq \bigcup S$, which is a contradiction! This is an elegant way to show that a conceptualization does not have an extension, more set-theoretic than Cantor's, with his alephs replaced by sets corresponding to beths, successive power-set cardinalities, and no well-orderings ostensibly involved. Hilbert actually started with the collection of natural numbers in his S and instead of (a) closed S off under the process of going from an X to X^X, the collection

[9] See Ferreirós [1999: 290ff] or Grattan-Guinness [2000: 117ff].

[10] See Ewald [1996: 923ff], especially for what seems to be the first letter from Cantor to Hilbert on these matters, dated 26 September 1897.

of functions: $X \to X$. Be that as it may, he emphasized that his paradox is "of purely mathematical nature".[11]

The logicians were later in arriving: Russell only absorbed Cantor's work in 1900; Russell's analysis of Cantor's diagonal proof led in the spring of 1901 to Russell's Paradox; and he famously communicated his paradox to Frege in June 1902, with devastating effect for the latter's *Grundgesetze*. Hilbert's letter was pointing out how they at Göttingen had assimilated such conundrums, and this may explain to some extent why Zermelo would publish a minor note [1901] on cardinal arithmetic, but not bother to publish the paradox argument. On the other hand, with increasing interest in the development of formal logic and the Frege–Russell logicist program Göttingen mathematicians and philosophers took renewed interest in the paradoxes.[12]

Who first devised Russell's Paradox, Zermelo or Russell? In his letter of 7 November 1903 to Frege, Hilbert wrote (as quoted above): "I believe Dr. Zermelo discovered [the paradox] 3–4 years ago". Husserl's note substantiating the discovery is dated 16 April 1902. Zermelo himself only wrote in a footnote to his [1908: 118]:[13] "I had, however, discovered this antinomy myself, independently of Russell, and had communicated it prior to 1903 [i.e., the year of publication of Russell's *The Principles of Mathematics*] to Professor Hilbert among others." In a letter of 10 April 1936 to Heinrich Scholz, Zermelo wrote:[14] "The set-theoretic antinomies were often discussed in the Hilbert circle around 1900 [um 1900 herum], and at that time I myself gave a precise formulation of the antinomy of the greatest cardinal which was later named after Russell (on the 'set of all sets not containing themselves')." With nothing more authoritative it remains unclear whether Zermelo actually found Russell's Paradox before Russell himself did in the spring of 1901.

The 1904 International Congress of Mathematicians, held in August at Heidelberg, was to be a generational turning point for the development of set theory. Hilbert [1905] spoke critically of the earlier work on the foundations of mathematics and stressed the importance of overcoming the paradoxes. Julius Kőnig delivered a lecture in which he provided a detailed argument to establish that $2^{\aleph_0}$ is not an aleph, i.e., that the continuum is not well-orderable. The argument combined the now familiar inequality

---

[11] A thin thread of connection runs from the operations in (the presented version of) Hilbert's Paradox through Zermelo's [1908a] generative axioms like Power Set and Union (cf. §3 below) and on to Zermelo's [1930] conditions (I) and (II) for normal domains (see §6 below). (a) and (b) for *sets* T are in fact the closure conditions for the cumulative hierarchy picture of the universe of sets with the Axiom of Foundation (cf. §§5, 6); it is just that one cannot take the union of S itself as then it would be the entire universe.

[12] See Peckhaus [2004].

[13] This is footnote 9 in van Heijenoort [1967: 191].

[14] See Peckhaus–Kahle [2002], quoting from the Scholz *Nachlass* at the University of Muenster.

$\aleph_\alpha < \aleph_\alpha^{\aleph_0}$ for α of cofinality ω with a result from Felix Bernstein's Göttingen dissertation [1901: 49] that alas does not universally hold.[15] Cantor was understandably upset with the prospect that the continuum would simply escape the number context that he had constructed for its analysis. Much of his life had been devoted to the Continuum Problem, but in his preoccupation with the second number class he had never entertained the basic distinction between regular and singular alephs.

Accounts differ on how the issue was resolved. According to Kowalewski [1950: 202], Zermelo found the gap in Kőnig's argument within a day of his lecture. However, the weight of evidence is for Felix Hausdorff having found the error.[16] Whatever the resolution, the torch had passed from Cantor to the next generation. Zermelo would be spurred to formulate his Well-Ordering Theorem and axiomatize set theory, and Hausdorff, to develop the higher transfinite in his study of ordertypes and cofinalities. And as with many incorrect proofs, there would be positive residues: Kőnig's inequality as soon generalized by Zermelo would become well-known as a basic restriction on cardinal exponentiation, 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.

§2. The first proof of the Well-Ordering Theorem. On 24 September 1904, a month and a half after the Heidelberg Congress, Zermelo sent to Hilbert his first proof of the Well-Ordering Theorem in a letter intended for publication in *Mathematische Annalen*, a letter so and promptly published (Zermelo [1904]). Zermelo, in his 33rd year and already estimable as a mathematical physicist, thereby effected a conceptual shift in mathematics.

Reversing Russell's progress from Cantor's correspondences to the identity map inclusion $\mathcal{P}(U) \subseteq U$, Zermelo considered functions $\gamma\colon\ \mathcal{P}(M) \to M$, specifically *choice functions*, those γ satisfying $\gamma(Y) \in Y$ for non-empty Y. This of course was the basic ingredient in Zermelo's [1904] formulation of what he soon called the Axiom of Choice. Russell the metaphysician had drawn elaborate philosophical distinctions and was forced by Cantor's diagonal argument into a dialectical confrontation with them. Zermelo the mathematician never quibbled over such distinctions and proceeded to resolve the problem of well-ordering sets mathematically.

In his mature presentation [1895] of his theory of cardinality Cantor had defined cardinal exponentiation in terms of the set of *all* functions from a set N into a set M. An arbitrary such function he termed a "covering [Belegung]" given by a "law", thus continuing his frequent use of "law" when

[15] 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 [1904] acknowledged the gap.

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

referring to functions. Arbitrary functions on arbitrary domains are now of course commonplace in mathematics, but several authors at the time referred specifically to the concept of covering; Philip Jourdain in his introduction to his English translation of Cantor's [1895, 1897] 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 . . . ." Zermelo [1904: 514] used the term "covering", but with his choice functions any residual sense of "law" was abandoned by him: " . . . we take an arbitrary covering γ and derive from it a definite well-ordering of the elements of M."

That part of Zermelo's proof which does not depend on the Axiom of Choice can be isolated in the following result, which establishes a basic correlation between functions $\gamma\colon \mathcal{P}(M) \to M$ and canonically defined well-orderings. The presentation is in Zermelo's terminology, but makes explicit the well-orderings involved as strict, i.e., irreflexive, relations. Significantly, Zermelo [1904] still had one foot in the Cantorian world where sets are structured, e.g., presented together with an implicit well-ordering.

THEOREM 1. *Suppose that* $\gamma\colon \mathcal{P}(M) \to M$. *Then there is a unique* $\langle W, \prec\rangle$ *such that* $W \subseteq M$ *and* $\prec$ *is a well-ordering of* W *satisfying*:
(a) *For every* $x \in W$, $\gamma(\{y \in W \mid y \prec x\}) = x$, *and*
(b) $\gamma(W) \in W$.

The picture here is that γ generates a well-ordering which according to (a) starts with

$$\begin{aligned} a_0 &= \gamma(\emptyset),\\ a_1 &= \gamma(\{a_0\}) = \gamma(\{\gamma(\emptyset)\}),\\ a_2 &= \gamma(\{a_0, a_1\}) = \gamma(\{\gamma(\emptyset), \gamma(\{\gamma(\emptyset)\})\}), \end{aligned}$$

and so continues as long as γ applied to the initial segment constructed thus far produces a new element. W is the result when according to (b) an old element is again named.[17]

PROOF OF THEOREM 1. Call $Y \subseteq M$ a *γ-set iff* there is a well-ordering R of Y such that for each $x \in Y$, $\gamma(\{y \in Y \mid y\ R\ x\}) = x$. The following are thus γ-sets (some of which may be the same):

$$\emptyset;\quad \{\gamma(\emptyset)\};\quad \{\gamma(\emptyset), \gamma(\{\gamma(\emptyset)\})\};\quad \{\gamma(\emptyset), \gamma(\{\gamma(\emptyset)\}), \gamma(\{\gamma(\emptyset), \gamma(\{\gamma(\emptyset)\})\})\}.$$

We shall establish:

($*$) If Y is a γ-set with a witnessing well-ordering R and Z is an γ-set with a witnessing well-ordering S, then $\langle Y, R\rangle$ is an initial segment of $\langle Z, S\rangle$, or vice versa.

[17] Note that if M is *transitive*, i.e., $M \subseteq \mathcal{P}(M)$, and γ is the identity on at least the elements in the above display, then we are generating the first several von Neumann ordinals (cf. §5 below). But of course, γ cannot be the identity on all of $\mathcal{P}(M)$.

Taking $Y = Z$ it will follow that any γ-set has a unique witnessing well-ordering.

For establishing $(*)$, we continue to follow Zermelo: By the comparability of well-orderings we can assume without loss of generality that there is an order-preserving injection $e\colon Y \to Z$ with range an S-initial segment of Z. It then suffices to show that e is in fact the identity map on Y: If not, let t be the R-least member of Y such that $e(t) \neq t$. It follows that $\{y \in Y \mid y\ R\ t\} = \{z \in Z \mid z\ S\ e(t)\}$. But then,

$$e(t) = \gamma(\{z \in Z \mid z\ S\ e(t)\}) = \gamma(\{y \in Y \mid y\ R\ t\}) = t,$$

a contradiction.

To conclude the proof, let W be the union of all the γ-sets. Then W is itself a γ-set by $(*)$ and so, with $\prec$ its witnessing well-ordering, satisfies (a). For (b), note that if $\gamma(W) \notin W$, then $W \cup \{\gamma(W)\}$ would be a γ-set, contradicting the definition of W. Finally, that (a) and (b) uniquely specify $\langle W, \prec\rangle$ also follows from $(*)$. $\dashv$

Zermelo of course focused on choice functions as given by the Axiom of Choice to well-order the entire set:

Corollary 2 (The Well-Ordering Theorem) (Zermelo [1904]). *If $\mathcal{P}(M)$ has a choice function, then M can be well-ordered.*

Proof. Suppose that $\varphi\colon \mathcal{P}(M) \to M$ is a choice function, i.e., $\varphi(X) \in X$ whenever X is non-empty, and define a function $\gamma\colon \mathcal{P}(M) \to M$ to "choose from complements" by: $\gamma(Y) = \varphi(M - Y)$. The resulting W of the theorem must then be M itself. $\dashv$

It is noteworthy that Theorem 1 leads to a new proof of Cantor's basic result that there is no bijection between $\mathcal{P}(M)$ and M, a proof that eschews diagonalization and provides a definable counterexample:

Corollary 3. *For any $\gamma\colon \mathcal{P}(M) \to M$, there are two distinct sets W and Y both definable from γ such that $\gamma(W) = \gamma(Y)$.*

Proof. Let $\langle W, \prec\rangle$ be as in Theorem 1, and let $Y = \{x \in W \mid x \prec \gamma(W)\}$. Then by (a) of Theorem 1, $\gamma(Y) = \gamma(W)$, yet $\gamma(W) \in W - Y$. $\dashv$

In the $\gamma\colon \mathcal{P}(M) \to M$ version of Cantor's diagonal argument, one would consider the definable set

$$A = \{\gamma(Z) \mid \gamma(Z) \notin Z\} \subseteq M.$$

If $\gamma(A) \notin A$, then we have the contradiction $\gamma(A) \in A$. If on the other hand $\gamma(A) \in A$, then $\gamma(A) = \gamma(B)$ for some B such that $\gamma(B) \notin B$. But then, $B \neq A$. However, no such B is provided with a *definition*.[18] Although the corollary is thus more informative, it should be pointed out that it

[18] This is also the main thrust of Boolos [1997], in which the argument for Theorem 1 is given *ab initio* and not connected with Zermelo [1904].

relies on significantly more set-theoretic apparatus than Cantor's diagonal argument.[19]

Theorem 1 has another notable consequence expressible in a setting where sets are not inherently well-ordered: Since the γ there need only operate on the *well-orderable* subsets of M, the $\mathcal{P}(M)$ in Corollary 3 can be replaced by the following set:

$$\{Z \subseteq M \mid Z \text{ is well-orderable}\}.$$

That this set, like $\mathcal{P}(M)$, is not bijective with M was first pointed out by Alfred Tarski [1939] through a less direct proof.

Zermelo himself did not make Theorem 1 explicit,[20] but that it has larger consequences is a testament to Zermelo's new approach. His main contribution with his Well-Ordering Theorem was the introduction of choice functions, leading to the postulation of the Axiom of Choice. But besides this Theorem 1 brings out Zermelo's delineation of the power set as a sufficient domain of definition for generating well-orderings. In fact, the Well-Ordering Theorem marks the beginning of the historical emergence of the power set as a distinctly set-theoretic, rather than presumptively logical, concept. With hindsight one can view the proof of Theorem 1 as a transfinite version of that of the Finite Recursion Theorem, the theorem for justifying definitions by recursion on the natural numbers. The Finite Recursion Theorem was established by Dedekind in his celebrated essay *Was sind und was sollen die Zahlen?* [1888: (126)] and also by Frege in his *Grundgesetze* [1893: thm. 263].[21] Whereas just the existence of infinite sets furnished the setting for their proofs, Zermelo used the power set, and his venture into the transfinite was avowedly impredicative: After specifying the collection of γ-sets its union is taken to specify a member of the collection. Poincaré [1906a: XIV] rejected Zermelo's proof because of this.[22] The main front of

---

[19] In particular, Theorem 1 features a prominent use of the Union Axiom, as well as uses of the Power Set and Separation Axioms to get the set of γ-sets (as well as to formalize well-orderings in set theory). Cantor's diagonal argument can be formalized in terms of class functions to avoid the Power Set Axiom, but it still requires the Separation Axiom. Finally, Zermelo's argument is impredicative, whereas Cantor's is not (cf. the discussion below).

[20] Tarski [1939: thm. 3] did have a version of Theorem 1; substantially the same version appeared in the expository work of Nicolas Bourbaki [1956: 43] (Chapter 3, §2, Lemma 3). Bourbaki's version is weighted in the direction of the application to the Well-Ordering Theorem: It supposes that for some $Z \subseteq \mathcal{P}(M)$, $\gamma\colon Z \to M$ with $\gamma(Y) \notin Y$ for every $Y \in Z$, and concludes that there is a $\langle W, \prec\rangle$ as in Theorem 1 except that its (b) is replaced by $W \notin Z$. From this version Bell [1995] developed a version in a many-sorted first-order logic and used it to recast Frege's work on the number concept.

[21] Dedekind termed his (126) "Theorem of definition by induction"; nowadays a distinction is made between *definitions* by recursion and *proofs* by induction. See Heck [1995] for an account of the theorem in Frege [1893].

[22] Poincaré [1906a: XIV] wrote at the end: "... although I am rather disposed to Zermelo's axiom, I reject his proof, which for an instant made me believe that aleph-one could indeed exist."

Poincaré's criticism of Zermelo and the logicists, a criticism that would have a decisive influence on Russell, was that they used impredicative definitions, definitions that contain a "vicious circle" in that they first specify a collection only to pick out a member.[23]

The proof of Theorem 1 is in fact essentially the argument for the Transfinite Recursion Theorem, the theorem that justifies definitions by recursion along well-orderings. This theorem was first properly articulated and established by von Neumann [1923, 1928] in his system of set theory. The difference is only that Zermelo's proof is defining the well-ordering itself. One can latterly view the situation as follows: The Axiom of Replacement is central to von Neumann's argument, and the axiom would eventually be adjoined to the axiomatization of set theory. Zermelo's proof stakes out what can be done up to the application of the axiom.[24] Seen in these various ways Zermelo's proof was seminal for modern set theory, especially when viewed against the backdrop of how well-orderability was being investigated at the time.

Cantor [1883: 550] had propounded the basic principle that every "well-defined" set can be well-ordered. However, he came to believe that this principle had to be established, and in a letter of 3 August 1899 to Dedekind gave a remarkable argument.[25] 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.[26] A multiplicity that can be thought of without contradiction as "being together" he called a "*consistent multiplicity* or a 'set' ['Menge']". Using what has now come to be known as the Burali-Forti Paradox, Cantor then pointed out that the class Ω of all ordinal numbers is an inconsistent multiplicity. He then 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. G. H. Hardy [1903] and Philip Jourdain [1904, 1905] also gave arguments involving the injection

[23] See Goldfarb [1988] for more about Poincaré and the logicists, and for a contrasting viewpoint, McLarty [1997].

[24] 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 a historical misrepresentation, but one that resonates with how acceptance of Zermelo's proof broke the ground for formal transfinite recursion.

[25] See Ewald [1996: 931ff]. Van Heijenoort [1967: 113ff] translated a letter as it appeared in Cantor [1932: 443ff]; however, as discovered by historians, that letter was a meshing of two letters into one by Zermelo as editor of Cantor [1932].

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

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] (see below).

Zermelo was presumably not privy to the 1899 Cantor–Dedekind correspondence when he established his Well-Ordering Theorem. Consonant with his observation on Schröder's inconsistent classes that no M can satisfy $\mathcal{P}(M) \subseteq M$, Zermelo's advance was to preclude the appeal to inconsistent multiplicities by shifting the weight away from Cantor's well-orderings with their *successive* choices to the use of functions on power sets making *simultaneous* choices. Decades later Zermelo, when editing Cantor's collected works [1932] and coming to the 1899 Cantor-Dedekind correspondence, chided Cantor for his reliance on successive choices and the doubts raised by the possible intrusion of inconsistent multiplicities. Zermelo noted that "it is precisely doubts of this kind that impelled the editor [Zermelo] a few years later [in 1904] to base his own proof of the well-ordering theorem purely upon the axiom of choice without using inconsistent multiplicities."[27] En passant, we note that in a foreshadowing, Dedekind [1888: (159)] got a denumerable subset of an infinite set through the idea of making successive choices, and he astutely set up what amounts to simultaneous choices for a well-defined definition by recursion.[28]

Many years later Gödel expressed a remarkable point of view that amounts to an anachronistic and ironic inversion. Gödel wrote in a letter of 8 November 1957 to Stanisław Ulam (Gödel [2003: 295]):

> I believe that his [von Neumann's] necessary and sufficient condition which a property must satisfy, in order to define a set, is of great interest, because it clarifies the relationship of axiomatic set theory to the paradoxes. That this condition really gets at the essence of things is seen from the fact that it implies the axiom of choice, which formerly stood quite apart from other existential principles. The inferences, bordering on the paradoxes, which are made possible by this way of looking at things, seem to me, not only very elegant, but also very interesting from the logical point of view. Moreover I believe that only by going farther in *this* direction, i.e., in the direction opposite to constructivism, will the basic problems of abstract set theory be solved.

[27]See Cantor [1932: 451] or van Heijenoort [1967: 117]. Zermelo [1908: 120] had already become aware through Jourdain [1905] of the Cantorian approach to well-ordering via inconsistent multiplicities and had dismissed Jourdain's rendition as "a mere word game".

[28]Zermelo [1909: 190, n. 5] later pointed out the implicit use of the Axiom of Choice here; see also the beginning of the next section for Zermelo [1909] and finiteness. The Axiom of Choice loomed large in Dedekind's last section, where the master faltered in his otherwise gap-less development by leaving the single gap that could and should have been filled: Dedekind [1888: (171)] amounts to the Axiom of Choice for finite families of sets, and instead of "selecting" elements "at pleasure" he should have established this by finite induction on the cardinality of the finite family. Zermelo [1908: 112] in a footnote (footnote 3 of van Heijenoort [1967: 187]) commenting on work of Peano pointed out the need for induction to establish the Axiom of Choice for finite families of sets.

By "his necessary and sufficient condition" Gödel was presumably referring to Axiom IV2 of von Neumann [1925: 225], to the effect that a class is a set exactly when there is no surjection from that class onto the universe V of sets. In words quoted by Ulam [1958: 13n], Gödel focused on the maximum character:

> The great interest which this axiom [von Neumann's Axiom IV2] has lies in the fact that it is a maximum principle, somewhat similar to Hilbert's axiom of completeness in geometry. For, roughly speaking, it says that any set which does not, in a certain well-defined way, imply an inconsistency exists. Its being a maximum principle also explains the fact that this axiom implies the axiom of choice.

The class Ω of all ordinal numbers is a proper class, so Axiom IV2 implies that there is a surjection of Ω onto V, and hence, as easily seen, a well-ordering of V. However, this is exactly the tack that Cantor was taking in the 1899 correspondence, arguing that an absolutely infinite or inconsistent multiplicity like V must get well-ordered in ordertype Ω. Whereas Zermelo had avoided such arguments by formulating and appealing to his Axiom of Choice, Gödel is regarding having a well-ordering of V as a *justification* of the Axiom of Choice! Gödel's writing in the first passage "The inferences, bordering on the paradoxes, ... " is an opaque reflection of historical issues, but now these inferences are "very interesting from the logical point of view."

Zermelo [1904: 516] went on to note without much ado that his Well-Ordering Theorem implies that every infinite cardinal number is an aleph and satisfies $\mathfrak{m}^2 = \mathfrak{m}$, and that the theorem secured Cardinal Comparability, i.e., for any two cardinal numbers $\mathfrak{m}$ and $\mathfrak{n}$, either $\mathfrak{m} \leq \mathfrak{n}$ or $\mathfrak{n} \leq \mathfrak{m}$, — so that the main issues raised by Cantor's *Beiträge* [1895] were at once resolved. Zermelo maintained that the Axiom of Choice 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 in itself. The axiom is consistent with Cantor's unitary view of the finite and transfinite, in that it posits for infinite sets an unproblematic feature of finite sets. On the other hand, the theorem did shift, in Zermelo's later emphasis, the weight from Cantor's well-orderings with their residually temporal aspect of numbering through *successive* choices to the use of a function on a power set for making *simultaneous* choices. Cantor's work had served to exacerbate a growing stress among mathematicians, who were already exercised by 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.[29] This can be viewed as a turning point for mathematics,

[29]See Moore [1982: chap. 2].

with the subsequent tilting toward the acceptance of the Axiom of Choice symptomatic of a conceptual shift in mathematics.

§3. The second proof and axiomatization. 1907, Zermelo's 36th year, was a high point for him in that he both was awarded, at Göttingen through Hilbert's help, the first lectureship in the new subject of mathematical logic ever awarded at a German university[30] and wrote two major articles, [1908], dated 14 July 1907 and [1908a], dated 30 July 1907.[31] Zermelo [1908] was a direct response to his critics; he not only provided a second proof of the Well-Ordering Theorem but also a detailed and wide-ranging reply that amounted to a spirited defense of his premises and argument as well as his first extended articulation of his expansive view about mathematics. Zermelo [1908a] provided the first full-fledged axiomatization of set theory; axiomatization was assuming a general methodological role in mathematics, but beyond that Zermelo through a simple and workable axiomatization set forth an abstract, generative view of sets. Before getting to the 1908 publications, however, we tuck in Zermelo [1909], which was actually dated earlier, May 1907, since it is a pivoting point for several larger themes.[32]

The thrust of Zermelo [1909],[33] also reported in [1909a],[34] was to recast the principle of mathematical induction in terms of finite sets and thus newly engage Dedekind and Poincaré. Zermelo [1909] began with number-free formulations of finite set. Dedekind [1888] in a context provided by the existence of an infinite set had established his principle of "complete induction" for the natural numbers, mathematical induction as we would now say.[35] Eschewing infinite sets, Zermelo proceeded to establish versions of the principle for finite sets, writing at the end that he had argued within his coming axiomatization of set theory. This exemplified Zermelo's emerging

[30]See Peckhaus [1992].

[31]Moore [1982: 158] has argued that these two publications, with the proximity of their dates, should be regarded as a unitary pair.

[32]According to Moore [1982: 163, n. 9] Zermelo had corresponded in 1907 with Poincaré about publishing his article in the *Revue de métaphysique et de morale*, and in a letter of 19 June 1907 Poincaré wrote that the article was too mathematical for the *Revue* and that he had proposed it to Mittag-Leffler for publication in *Acta Mathematica*. This is where the article appeared fully two years later. Zermelo [1908a: 262] referred to a paper fitting the description of [1909] as "in preparation", but this may only be acknowledging the publication delay.

[33]See Parsons 1987] for more on this paper in the context of developing arithmetic in set theory.

[34]Zermelo [1909a] was largely a report of his [1909] work presented at the 1908 International Congress of Mathematicians held at Rome.

[35]Strictly speaking, Dedekind first stated his theorem of complete induction for his "chains" in a general context ((59)). Only afterward did he appeal to the existence of an infinite set to introduce the natural numbers (by abstraction, so that they are a "free creation of the mind") and then stated the theorem of complete induction for the natural numbers ((80)).

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.[36] Zermelo [1909] wrote at the beginning: "... for me, every theorem stated about finite numbers is nothing other than a theorem about *finite sets*." (Zermelo's actual incorporation of the natural numbers into set theory would not be done in [1909] but in his axiomatization paper [1908a] in connection with his Axiom of Infinity: There is a set having the empty set 0 as a member and closed under the taking of singletons. Zermelo [1908a] wrote at the end of §1: "The set Z_0 contains the elements, 0, $\{0\}$, $\{\{0\}\}$, and so forth, and it may be called the *number sequence*, because its elements can take the place of numerals.[37]) Zermelo [1909] wrote in the conclusion:

> If these axioms, that I propose to enunciate completely in another article, are nothing more than purely logical principles, then the principle of [mathematical] induction is as well; if on the contrary they are intuitions of a special sort, then one can continue to regard the principle of [mathematical] induction as a result of intuition or as a "synthetic a priori judgment".

In Kantian terms Poincaré regarded the principle of mathematical induction as synthetic a priori, while Frege and the logicists would have it analytic. Zermelo was directing his remarks at Poincaré, but in a reply Poincaré [1909] merely pointed out the impredicativity at work. Later Poincaré [1909a] rejected Zermelo's [1908a] axiomatization altogether. Be that as it may, Zermelo would be vindicated to the extent that modern texts of set theory invariably show how his axioms subsume the natural numbers and mathematical induction.

To elaborate on the finiteness, the immediate sense of a set being finite is having a bijection with $\{0, \ldots, n-1\}$ for some natural number n. Zermelo provided two equivalent, non-numeric formulations, each exhibiting a heritage from Dedekind. The second formulation, thematically connected with Dedekind's [1888: §§7, 8] analysis of greatest elements for finite sets of natural numbers, was elegant: A set X is finite exactly when there is a well-ordering $\prec$ of X such that every non-empty subset of X not only has a $\prec$-least element but also a $\prec$-greatest element, i.e., the converse of $\prec$ is also a well-ordering.[38] In his incisive analysis Zermelo correlated his definitions of finite set with Dedekind's original [1888: (64)] definition, nowadays enshrined as: A set X is *Dedekind-finite* exactly when there is no bijection between X and a proper subset of X; otherwise, X is *Dedekind-infinite*. *Assuming* that the

[36]See Hallett [1984: 244ff] and Taylor [1993] for more on Zermelo's reductionism.

[37]Kurt Grelling, a student and follower of Zermelo and later of "heterological" fame, would base his *Dissertation* [1910] at Göttingen on Zermelo [1909]. Grelling advocated the founding of arithmetic on Zermelo's axiomatization of set theory. See Peckhaus [1994] for a biography and bibliography of Grelling.

[38]This definition was arrived at independently by Stäckel [1907].

natural numbers (in some set-theoretic rendition) do form a set, a set is Dedekind-infinite exactly when it has a denumerable subset. The simple assertion that a finite set is Dedekind-finite is equivalent to what is now called the Pigeonhole Principle (Dirichlet's Schubfachprinzip). Zermelo [1909: §6] established the converse, proceeding through his non-numeric definitions and explicitly employing the Axiom of Choice.[39] Tarski [1924] later provided not only another elegant definition of finite set[40] but also a survey of the various definitions of finite set, explicitly working in Zermelo's [1908a] axiomatization. Vindicating Zermelo's appeal to the Axiom of Choice, the earliest results in the 1960's on the independence of the Axiom of Choice via forcing established the consistency of having infinite, Dedekind-finite sets.

We now focus now on the major 1908 publications. Although the larger objections raised against Zermelo's [1904] proof had to do with the use of arbitrary functions on arbitrary subsets of a set, there were also specific objections raised about possible inconsistencies having to do with the class Ω of all ordinal numbers. After all, Zermelo had defined a well-ordering in a new way, and what prevented this well-ordering from being identifiable with Ω? 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]. This new tack is thematically related to Zermelo's proof of the Schröder–Bernstein Theorem, which he had already sent to Poincaré in January 1906 and which had a similar incentive (cf. §4 below). 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:

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.

[39]Dedekind [1888: (159), (160)] had first established that Dedekind-finite implies finite, making an implicit use of the Axiom of Choice as Zermelo [1909: 190, n. 5] pointed out. Zermelo was perhaps exercised by Poincaré's [1906a: X] critical quotation of a remark from Russell [1906: 49]: "But, so far as I know, we cannot prove that the number of classes contained in a [Dedekind-]finite class is always [Dedekind-]finite, or that every [Dedekind-]finite number is an inductive number [i.e., is in the closure of $\{0\}$ under the successor operation]." Zermelo already in a letter of 27 June 1905 to Hilbert wrote that the Axiom of Choice is needed to prove that Dedekind-finite implies finite; see Moore [2002a: 50].

[40]A set X is finite exactly when every non-empty subset of $\mathcal{P}(X)$ has a $\subseteq$-minimal element. As pointed out by Gregory Taylor (private communication), Zermelo in his last publication [1935] referred to Tarski's definition; Zermelo was dealing with well-founded relations, and Tarski's definition amounts to: X is finite exactly when the proper inclusion relation on $\mathcal{P}(X)$ is well-founded.

This new proof is less comprehensible without the first [1904] proof, but with the intersection approach no question could arise, presumably, about intrusions by classes deemed too large like Ω. While the first proof featured what we would now call a transfinite recursive construction of a well-ordering $\prec$, Zermelo in effect reconstituted the well-ordering with final segments $R(a) = \{b \mid a \preceq b\}$ under reverse inclusion, a natural ordering for sets.[41] Zermelo thus moved further from Cantor's structured sets by applying Hessenberg's reduction of well-orderings to set-theoretic relations. We today would regard the first [1904] proof as more economical: While that proof required only separation of the set of γ-sets from the power set of M and taking the union of γ-sets, the new proof required separation of the set of Θ-chains from the power set of the power set of M. In anticipation of his axiomatization Zermelo himself pointed out [1908: §1] the set existence principles that he was applying in his new proof: the Separation and Power Set Axioms, and of course, the Axiom of Choice. Although Zermelo neglected to mention the Union Axiom, so prominent in the first proof, it too was needed in the second proof.[42]

Notably, Zermelo's two proofs of his Well-Ordering Theorem correspond to two distinguishable conceptions of natural number.[43] The first proof corresponds to "building up", starting from zero and recursively applying the successor function, and the second proof corresponds to "paring down", getting at the closure of zero under the successor function by taking an intersection of all closed supersets. It is commonplace in modern mathematics that the closure of a set under a mapping can be defined as either the union of an increasing, recursively defined sequence of sets or the intersection of all closed supersets. However, the first instances occurred in the foundational investigation of the natural numbers and explicitly for a new situation in the two Zermelo proofs. Dedekind [1888] had provided the general theory of closures through paring down. In his terminology, given a set [System] S, a mapping [Abbildung] $\phi\colon S \to S$, and a subset A of S, *the* chain [Kette] of A is the closure of A under ϕ, i.e., $\bigcap\{K \subseteq S \mid A \cup \phi``K \subseteq K\}$. Dedekind had proceeded to define the natural numbers as the closure of $\{1\}$ under

[41]Zermelo [1908: §1] stressed how his new formulation of well-ordering "rests exclusively upon the elementary notions of set theory, whereas experience shows that, with the usual presentation, the uninformed are only too prone to look for some mystical meaning behind Cantor's relation $a \prec b$, which is suddenly introduced." Zermelo went on to make explicit a characterization which can be economically rendered as follows: A set M is well-orderable *iff* there is an $R\colon M \to \mathcal{P}(M)$ such that for any non-empty $P \subseteq M$, there is a unique $x \in P$ such that $P \subseteq R(x)$.

[42]Zermelo [1908] formulated the Axiom of Choice as positing for any set M consisting of non-empty, pairwise disjoint sets the existence of a set that meets each member of M in exactly one element. However, to apply the axiom to choose elements from arbitrary subsets of a given set, one needs first to establish a technical theorem that correlates such subsets with pairwise disjoint sets. That theorem (Zermelo [1908a: para. 28]) requires the Union Axiom.

[43]This is elaborated in George–Velleman [1998].

the successor operation. Frege in his *Begriffsschrift* [1879] and *Grundlagen* [1884] had taken essentially the same approach. Zermelo's Θ-chains are evidently a direct extension of Dedekind's chains into the transfinite. Just as Dedekind's approach to the natural numbers required second-order logic, as we would now say, and was impredicative, in that the set of natural numbers was itself one of the chains to be intersected, likewise Zermelo's approach required one more power set (as has been noted) and was analogously impredicative. On the other hand, while "building up" is not impredicative for the natural numbers, the first [1904] proof is avowedly impredicative, perhaps an inevitable concession to the transfinite.

Zermelo started his axiomatization paper [1908a] as follows:

> Set theory is that branch of mathematics whose task it is to investigate mathematically the fundamental notions "number", "order", and "function", taking them in their pristine, simple form, and to develop thereby the logical foundations of all of arithmetic and analysis; thus it constitutes an indispensable component of the science of mathematics.

Zermelo went to describe how the "antinomies" had precluded the viability of sets as extensions of concepts, so that Cantor's definition of set must be restricted, something that had not been successfully done. He continued:

> Under these circumstances there is at this point nothing left for us to do but to proceed in the opposite direction, and starting from set theory as it is historically given, to seek out the principles required for establishing the foundations of this mathematical discipline. In solving the problem we must, on the one hand, restrict these principles sufficiently to exclude all contradictions, and, on the other, take them sufficiently wide to retain all that is valuable in this theory.
>
> Now in the present paper I intend to show how the entire theory created by Cantor and Dedekind can be reduced to a few definitions and seven principles, or axioms, which appear to be mutually independent.

The first lines above are consonant with a pragmatic view, advocated in Zermelo's previous paper, of how mathematical axioms are to be justified [1908: 112]: "by analyzing the modes of inference that in the course of history have come to be recognized as valid and by pointing out that the principles are intuitively evident and necessary for science." Zermelo's approach would be successful in being sufficiently restrictive "to exclude all contradictions" and sufficiently wide "to retain all that is valuable," but he would moreover transform the set theory of Cantor and Dedekind by making explicit *new* set existence principles and promoting a generative point of view. Zermelo had begun working out an axiomatization as early as 1905, incorporating aspects of his own discovery of the Russell Paradox and addressing issues raised by his [1904] proof.[44] The mature presentation is a precipitation of seven axioms, and these do not just reflect "set theory as it is historically

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

given", but explicitly buttress his proof(s) of the Well-Ordering Theorem. Abstract set theory was thus launched from its mooring.[45]

Zermelo's axiomatization was methodologically in the spirit of Hilbert's *Grundlagen der Geometrie* [1899], positing a *domain* of objects with a *fundamental relation*, membership. An object is to be a *set* if it is either empty or has an object as a member; thus, incipiently *urelements* (now also called individuals or atoms), objects distinct from the empty set yet having no member and capable of belonging to sets, were allowed. Beyond this oddity from the modern perspective, Zermelo advanced the now-familiar view, analogous to Hilbert's for geometric objects and relations, of sets being solely structured by membership and generated by simple operations. As for Hilbert, consistency was important for Zermelo, but moreover his principles would have a *generative* function. He had already written in [1908: 124]: " ... if in set theory we confine ourselves to a number of established principles such as those that constitute the basis of our proof [of the Well-Ordering Theorem] — principles that enable us to form initial sets and to derive new sets from given ones — then all such contradictions [like Russell's] can be avoided."

Zermelo's seven set axioms constitute the now familiar theory Z, Zermelo set theory: Extensionality, Elementary Sets ($\emptyset$, $\{a\}$, $\{a, b\}$), Separation, Power Set, Union, Choice, and Infinity. With Extensionality Zermelo of course espoused the extensional viewpoint. 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. A generous view is that at the time there was no generally accepted system of logic, and so Zermelo was pragmatically indicating the path to be followed.[46] Another is that definiteness is not to be given by language, formal or otherwise, but rather by direct appeal to membership and generative set-forming operations as part of Zermelo's reductionism of mathematical concepts.[47] With Infinity and Power Set Zermelo provided for sufficiently rich settings for set-theoretic constructions. Tempering the logicians' extravagant and problematic "all"

[45]Concerning "abstract", Fraenkel in his text *Abstract Set Theory* [1953] distinguished between *abstract sets* (the nature of whose elements are not of concern) and *sets of points* (typically numbers). In the early years "general set theory" was also used with connotations similar to "abstract set theory". Cantor's study of sets of points evolved into point-set topology and descriptive set theory. Cantor and Dedekind certainly entertained abstract sets, but it was Zermelo who set the theory going on its way to modern set theory. The latter-day Skolem [1962] was still entitled *Abstract Set Theory*.

[46]See Ferreirós [1999: 323]. This accords with what Zermelo wrote when he returned to the notion of definite property in Zermelo [1929: 340]; see §6 below.

[47]See Taylor [1993: 546ff].

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 argumentation also brought him in proximity of Replacement (as was pointed out in connection with Theorem 1), the next axiom to be adjoined in the subsequent development of set theory.[48]

Zermelo's first result in his axiomatic theory was just the result of his Husserl note (§1), that every set M has (through Separation) a subset $\{x \in M \mid x \notin x\}$ not a member of M, with the consequence that there is no universal set. In modern texts of set theory this is also the first substantial result presented, but they usually take the opposite tack, showing that there is no universal set by *reductio* to Russell's Paradox.[49]

Although Hilbert's axiomatization of geometry may have served as a model for Zermelo's axiomatization of set theory and Dedekind's [1888] essay as a precursor, there are crucial differences having to do with subject matter and proof. By the second edition of his *Grundlagen* [1903: 16] Hilbert had incorporated his Completeness [Vollständigkeit] Axiom which made his axioms categorical, i.e., having a unique model up to isomorphism, in this case Euclidean space, the maximal possibility. Dedekind's presentation [1888] of the natural numbers had also been categorical, and so both in intent and outcome Dedekind and Hilbert had been engaged in the *analysis* of fixed subject matter. Zermelo's axioms were by no means categorical, and in fact he would in later life advocate an open-ended view with a hierarchy of models (see §6 below). This brings up the larger issue of the role of proof for articulating sets.

By the time of Dedekind [1888] proof had become basic to mathematics, and indeed his work did a great deal to enshrine proof as the vehicle for algebraic abstraction and generalization.[50] Like algebraic constructs sets were new to mathematics and would become incorporated by setting down the rules for their proofs. Just as calculations are part of the sense of numbers, so proofs would become part of the sense of sets, as their "calculations".

[48]It is notable that Cantor, in that letter of 3 August 1899 to Dedekind, advocated closure for his notion of ("consistent") set under the taking of unions and loose forms of Separation and Replacement (see Cantor [1932: 451] or van Heijenoort [1967: 117]), and, in an earlier letter of 20 September 1899 to Hilbert, advocated closure for his notion of ("completed") set under the taking of power set and loose forms of Separation and Replacement (see Moore [2002a: 44ff]).

[49]Zermelo applied his first result *positively* to generate specific sets disjoint from given sets for his recasting of Cantor's theory of cardinality in terms of immediately and mediately equivalent sets (see below).

[50]The first sentence of the preface to Dedekind [1888] is: "In science nothing capable of proof ought to be accepted without proof."

Just as Euclid's axioms for geometry had set out the permissible geometric constructions, the axioms of set theory would set out the 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 with nor bolstered by substantial antecedents. There was no fixed, intended subject matter. Like strangers in a strange land stalwarts developed a familiarity with sets guided step by step by the 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.

In future years the basic scaffolding provided by Zermelo's axioms, with their schematic simplicity and open-endedness, would win out as a working foundation for mathematics over the unwieldy and unsightly type theory of Russell, a fearful symmetry imposed by an artful dodger. Set theory would provide the underpinnings of mathematics, and Zermelo's axioms would resonate with emerging mathematical practice. To take one example, Maurice Fréchet's thesis [1906], of similar vintage to Zermelo's axiomatization and seminal to functional analysis, proposed the study of function spaces, formalizable with Power Set and Separation.

Zermelo's analysis 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 segregation of property considerations to Separation. Based on generative and prescriptive axioms, set theory would become more combinatorial, less logical. That Zermelo did not develop a nuanced view of logic, as particularly evidenced by his blanket reliance on definite properties, played to advantage, if anything, in emphasizing the set-theoretic. Even though Zermelo himself regarded the Axiom of Choice as a "logical" principle ([1904: 141][1930: 31]), its very isolation and its subsequent investigation by the Polish School[51] established it as distinctly set-theoretic. It was this sort of increasingly mathematical analysis that shifted the focus of mathematical logic away from the Frege–Russell logicist program.

Zermelo's reduction to seven axioms sufficient "for the entire theory created by Cantor and Dedekind" went hand in hand with his reductionist approach which, as Zermelo had written at the outset, would investigate number, order, and function "taking them in their pristine, simple form." Zermelo's formulation of the Axiom of Choice, an axiom which would seem to be naturally cast in terms of functions, exemplified this: *For any set M consisting of non-empty, pairwise disjoint sets there is a set that meets each member of M in exactly one element.* Zermelo's reductionism had already exhibited itself in his [1909] (as we have discussed) and in his second [1908]

[51] Especially Sierpiński, Tarski, and Kuratowski.

proof of the Well-Ordering Theorem, with the desired well-ordering given in terms of reverse inclusion. However, we would recognize today that the reduction was not complete, since Zermelo did not objectify the well-ordering *itself* as a set. For that, one needs to incorporate the ordered pair and relations into set theory.[52] Also, Zermelo wrote pointedly in [1908: 120]: "... In Cantor's theory 'order types' and 'cardinal numbers' are nothing but convenient *means of expression* [*Ausdrucksmittel*] for the comparison of sets with respect to the similarity or equivalence of their parts ... " This is more of a pronouncement than a recounting, since Cantor had a substantial commitment to his ordinal and cardinal numbers *qua* numbers. In any case, Zermelo proceeds in his axiomatization paper [1908a] to develop the "theory of [cardinal] equivalence" from his axioms:

First, Zermelo specified that for any *disjoint* sets M and N, M is *immediately equivalent* to N, exactly when there is a $\Phi \subseteq \mathcal{P}(M \cup N)$ such that every member of $M \cup N$ occurs in exactly one (unordered) pair $\{m, n\} \in \Phi$. Thus Φ serves as a set-theoretic reduction of a bijection between M and N. Next, appealing to there being no universal set Zermelo proved that for any sets M and N there is a set M' disjoint from $M \cup N$ such that M' is immediately equivalent to M. With this in hand, Zermelo then specified that for any sets M and N, M is *mediately equivalent* to N exactly when there is third set R disjoint from both yet immediately equivalent to each. Using this notion Zermelo proceeded to develop the theory of cardinal equivalence within his axiomatic framework.

Although Zermelo's treatment would nowadays be viewed as *ad hoc*, it both exemplifies his reductionism and how we operate in his axiomatic theory today. We too make unabashed use of set-theoretic operations, and we too appeal freely to the Power Set axiom to provide a setting, even when we are only interested in the subsets which are merely unordered pairs. And we too develop the theory of cardinality as far as possible in terms of sets, before getting to the matter of cardinal number objects. Indeed, Zermelo's development might be viewed as an elegant, minimalist approach to cardinality.[53] Of course, with the incorporation of the ordered pair, relations, and functions into set theory — and through their reconstrual, the real numbers — the reduction of mathematics to sets would become complete.

Zermelo's development culminated with his proof of what is at times called the Zermelo-Kőnig Inequality. In terms of cardinal numbers this theorem states: If $\mathfrak{m}_t < \mathfrak{n}_t$ for every t in a set T, then $\Sigma_t \mathfrak{m}_t < \Pi_t \mathfrak{n}_t$. Of course, Zermelo formulated and established his inequality entirely in terms of sets. The theorem subsumes Cantor's theorem $\mathfrak{m} < 2^{\mathfrak{m}}$ and the theorem of Kőnig

[52]This was done by Hausdorff [1914], the other great developer of Cantor's set theory in the next generation.

[53]Skolem as late as in 1957 (see his [1962]) presented the theory of cardinality without formalizing functions, taking Zermelo's approach instead.

[1904] from the Heidelberg congress, which was for the special case of T being countable. The result recalls the beginning and comes full circle, for as an outcrop of the cottage industry investigating choice principles the Zermelo-Kőnig Inequality was eventually shown to be equivalent to the Axiom of Choice.[54]

§4. The Schröder–Bernstein Theorem. Zermelo provided a proof of the Schröder-Bernstein Theorem in [1908a], and therein lies a tale involving Poincaré, Russell, Peano, and ultimately Dedekind. Zermelo's proof stands in relation to an earlier proof of Bernstein's as Zermelo's second [1908] proof of the Well-Ordering Theorem does to his first [1904] proof. Moreover, the motivations for the second proofs were to address similar issues of forestalling concerns about the role of number. Finally, Zermelo came to his proof of the Schröder-Bernstein Theorem before his second proof [1908] of the Well-Ordering Theorem, and so he could be seen to be predisposed to the [1908] approach and its reductionism.

In the theory of cardinality the Schröder-Bernstein Theorem is the basic assertion that for cardinal numbers $\mathfrak{m}$ and $\mathfrak{n}$, $\mathfrak{m} \leq \mathfrak{n}$ and $\mathfrak{n} \leq \mathfrak{m}$ implies $\mathfrak{m} = \mathfrak{n}$. The issue first arose in Cantor's development of cardinal numbers; the early history is interlaced with the question of whether every set can be well-ordered; and the theorem is of course immediate assuming the Axiom of Choice because of Zermelo's Well-Ordering Theorem. The Schröder–Bernstein Theorem commands a separate significance only with a proof involving no well-orderability assumptions, and the first correct such proof to appear in print was due to Bernstein and appeared in Borel [1898: 104–6].[55]

Zermelo's proof was for the synoptic formulation in terms of sets and mappings which was how Cantor himself had first raised the issue in 1882:[56]

$$\text{If } M' \subseteq M_1 \subseteq M \text{ and there is a bijection } \phi\colon M \to M', \\ \text{then there is a bijection: } M \to M_1.$$

The following is Zermelo's proof in brief (and just to affirm notation, $\subset$ denotes proper inclusion and $h``X$ denotes the image of set X under function h):

Set $Q = M_1 - M'$, and for $A \subseteq M$, define $f(A) = Q \cup \phi``A$. f is monotonic: if $A \subseteq B \subseteq M$, then $f(A) \subseteq f(B)$. Set $T = \{A \subseteq M \mid f(A) \subseteq A\}$.[57] Noting that T is not empty since $M \in T$, let $A_0 = \bigcap T$.

[54] Rubin–Rubin [1963: 75ff].

[55] Schröder [1898] claimed a proof, but it was flawed (cf. Korselt [1911]), and the theorem is sometimes called the Cantor-Bernstein theorem.

[56] See Cantor's 5 November 1882 letter to Dedekind in Ewald [1996: 874ff].

[57] Zermelo carefully pointed out that T is a set, through a typical appeal in his system to the Power Set and Separation Axioms.

Then $A_0 \in T$.[58] Moreover, $f(A_0) \subset A_0$ would imply by the monotonicity of f that $f(f(A_0)) \subseteq f(A_0) \subset A_0$, contradicting the definition of A_0. Consequently, we must have $A_0 = f(A_0) = Q \cup \phi``A_0$. It follows that $M_1 = A_0 \cup (M' - \phi``A_0)$ is a disjoint union, and of course so is $M' = \phi``A_0 \cup (M' - \phi``A_0)$. Hence, the function: $M_1 \to M'$ which is ϕ on A_0 and identity elsewhere is a bijection, and this suffices for the result.

Zermelo himself did not define the function f explicitly, but he did define T and $A_0 = \bigcap T$, and his argument turned on A_0 being a "fixed point" of f, i.e., $f(A_0) = A_0$.[59]

Bernstein's earlier proof, like those most often given today, depends on defining a countable sequence of sets by recursion. Poincaré's "petitio principii" criticism of the logicists was that "logical" developments of the natural numbers and their arithmetic inevitably presuppose the natural numbers and mathematical induction, and in connection with this Poincaré [1906: XXVI] pointed out the circularity of developing the theory of cardinality with the Schröder-Bernstein Theorem based on Bernstein's proof, and therefore on the natural numbers. This point had *mathematical* weight, and according to a footnote in his [1908a: 272–3] Zermelo had sent Poincaré his new proof of Schröder-Bernstein in January 1906,[60] and it appeared in Poincaré [1906a: XIV]. In the footnote Zermelo emphasized how his proof avoids numbers and induction altogether, and noted that Peano [1906] published a proof that was "quite similar".[61] Russell on first reading Zermelo [1908a] expressed delight with his proof of Schröder-Bernstein but went on to criticize his axiomatization of set theory.[62] In the first volume of Whitehead and Russell's *Principia Mathematica* [1910–13] there was no formal use of

[58] This step makes the proof impredicative, like Zermelo's argument used for Theorem 1.

[59] Kanamori [1997] pursues the fixed point idea as part of a unifying mathematical theme, one that continues next after Zermelo to Kuratowski [1922].

[60] Zermelo first gave his new proof in a letter of 28 June 1905 to Hilbert; see Moore [2002a: 50].

[61] Zermelo's footnote is footnote 11 of van Heijenoort [1967: 209], and Peano [1906] is dated 31 March 1906. The reference to Peano [1906] is part of a contretemps involving Poincaré: Poincaré [1906a: XIV] in publishing Zermelo's proof proceeded to make it part of his criticism of Zermelo's work based on the use of impredicative notions, the main front of Poincaré's critique of the logicists and Zermelo. Zermelo in a footnote to his [1908: 118] (footnote 8 of van Heijenoort [1967: 191]) expressed annoyance that Peano when referring to Poincaré [1906a] only mentioned Peano [1906], not Zermelo, in connection with the new proof of Schröder-Bernstein but went on the argue against Zermelo's espousal of the Axiom of Choice.

[62] Russell in a letter of 23 March 1908 to Zermelo (Zermelo *Nachlass* C 129/99) wrote: ". . . especially remarkable to me was your simple and elegant new proof of the Schröder-Bernstein Theorem." Russell's earlier letter of 15 March 1908 to Jourdain (Grattan-Guinness [1977: 109]) began: "I have only read Zermelo's article once as yet, and not carefully, except his new proof of Schröder-Bernstein, which delighted me." Russell then criticized the Separation Axiom as being "so vague as to be useless." For Russell, the paradoxes cannot be avoided in this way but had to be solved through his theory of types.

the class of natural numbers, and indeed the Axiom of Infinity was avoided; while this would not satisfy Poincaré, the theory of cardinality was developed using Zermelo's proof.[63]

It is now time to look at the mathematics more closely. There is really only one proof idea for establishing the Schröder-Bernstein Theorem. Suppose the hypothesis, that $M' \subseteq M_1 \subseteq M$ and there is a bijection $\phi\colon M \to M'$. Define sets

$$P_0 = M - M_1,\ Q_0 = M_1 - M',\ \text{and } P_{n+1} = \phi``P_n \text{ and } Q_{n+1} = \phi``Q_n$$

by recursion. The P_n's and Q_n's can be pictured as concentric rings nested inward: P_0 is the outer ring; Q_0 the next; then P_1; then Q_1; and so forth. Let

$$\overline{P} = \bigcup_n P_n \text{ and } \overline{Q} = \bigcup_n Q_n.$$

Then $\overline{P}$ is the closure of P_0 under ϕ, and $\overline{Q}$ is the closure of Q_0 under ϕ.

Zermelo's Q and A_0 are here Q_0 and $\overline{Q}$ respectively, and his proof results in the function: $M_1 \to M'$ which is ϕ on $\overline{Q}$ and identity elsewhere. That is, the rings Q_n are sent inward to Q_{n+1} successively and everything else in M_1 is fixed.

Bernstein (see Borel [1898: 104–6]) began with the hypothesis

$$A_1 \subseteq A,\ B_1 \subseteq B,\ A \text{ is bijective with } B_1, \text{ and } B \text{ is bijective with } A_1,$$

with the goal being to provide a bijection between A and B. However, we can convert to the previous scheme by taking A to be M; A_1 to be M_1; and through the bijective correspondence of B with A_1, take $M' \subseteq A_1 = M_1$ to be corresponding to B_1; so that there is a bijection $\phi\colon M \to M'$. Bernstein's proof then amounts to defining the P_n's and Q_n's as before by recursion but now defining a bijection: $M \to M_1$ which is ϕ on $\overline{P}$ and identity elsewhere. That is, the rings P_n are sent inward to P_{n+1} successively and everything else in M is fixed. Instead of Zermelo's $M_1 \to M'$ map Bernstein in effect developed the other possibility, the $M \to M_1$ map.[64]

The Zermelo and Bernstein proofs were compared early on, the modern and also historically focal point being one discussed in §3, that the closure of a set under a mapping can be defined either as the union of an increasing, recursively defined sequence of sets or the intersection of all closed supersets. Although the explicit stratifications $\overline{P} = \bigcup_n P_n$ and $\overline{Q} = \bigcup_n Q_n$ are helpful to motivate and survey the proofs, as seen from Zermelo's rendition one does not need these stratifications nor indeed to presuppose the natural numbers and recursion. The situation foreshadows, since Zermelo had come to his Schröder-Bernstein proof by at least January 1906, that of Zermelo's second

[63] The Schröder-Bernstein Theorem in *Principia* is $*73 \cdot 88$.

[64] Beginning with the same hypothesis as did Bernstein, Schröder [1898] had developed nested sequences of sets, e.g., what in the above terminology would be $M \supseteq \phi``M \supseteq \phi``(\phi``M) \ldots$, and took "limits", but did not in the end define a bijection: $A \to B$.

proof of the Well-Ordering Theorem. It is a testament to Zermelo's unity of vision.

In Whitehead and Russell's *Principia Mathematica*, after Zermelo's proof was given in $*73$, the Zermelo and Bernstein approaches were pictured and compared in $*94$ and the latter was carried out separately in $*95$. However, in the orderly development the issue was not the presupposition of the natural numbers; rather surprising to the modern eye, the prolonged and gratuitous labor in the logic of relations contended with the different starting hypotheses, i.e., the $M' \subseteq M_1 \subseteq M$ etc. hypothesis vs. the $A_1 \subseteq A$ and $B_1 \subseteq B$ etc. hypothesis.

The correlation of the different hypotheses had presented no difficulties in the earlier Peano [1906: 338], where a simple account of Bernstein's argument was provided in the $M' \subseteq M_1 \subseteq M$ formulation. Moreover, Peano provided a number-free proof for the existence of Bernstein's $M \to M_1$ bijection, exactly analogous to Zermelo's for the existence of his $M_1 \to M'$ bijection, also in light of Poincaré's critique.

Finally, Hausdorff in his classic *Grundzüge der Mengenlehre* [1914: 48ff] was quite clear in his comparison of the Bernstein and Zermelo proofs. Working toward the $M_1 \to M'$ bijection, he first provided a proof à la Bernstein and then presented Zermelo's proof. Without fuss, Hausdorff pointed out that both proofs amount to the same bijection sending each Q_n to Q_{n+1}. Modern mathematics would go this way, foregoing the logical preoccupations and difficulties of *Principia Mathematica* and carrying out increasingly set-theoretic constructions.

The role of Dedekind brackets this to and fro, both at the beginning and the end. Zermelo in that footnote [1908a: 272–3] also pointed out that his proof "rests solely upon Dedekind's chain theory [1888: IV]," and this has more substance than Zermelo realized at the time. Dedekind in fact had a proof of Schröder-Bernstein already in 1887, but the proof only appeared in 1932, both in a manuscript dated 11 July 1887 appearing in Dedekind's collected works [1932: 447–9] and in an enclosure in a letter of 29 August 1899 from Dedekind to Cantor appearing in Cantor's collected works [1932: 449].[65] As editor for the latter, Zermelo noted in a footnote that Dedekind's proof is "not essentially different" from that appearing in Zermelo [1908a], and expressed puzzlement that neither Cantor nor Dedekind published the proof (see Cantor [1932: 451]).

Dedekind's proof rested on a decomposition theorem that appeared in his [1888].[66] In Dedekind's terminology as described already before, given a set [System] S, a mapping [Abbildung] $\phi\colon S \to S$, and a subset A of S, *the* chain [Kette] of A is the closure of A under ϕ, i.e., $\bigcap\{K \subseteq S \mid A \cup \phi``K \subseteq K\}$. The decomposition theorem was the last theorem in section IV, "Mapping

[65] For a careful translation of the letter together with the enclosure, see Ewald [1996: 937–9].

[66] This is emphasized by Ferreirós [1999: 240].

of a System into Itself": Given a map ϕ, suppose that $\phi``K \subseteq L \subseteq K$. If $U = K - L$, U_0 is the chain of U and $V = K - U_0$, then:

$$K = U_0 \cup V \text{ and } L = \phi``U_0 \cup V.$$

Dedekind wrote afterward: "The proof of this theorem, of which (as of the two preceding) we shall make no use, may be left for the reader." Dedekind was remarkably terse here.[67] But the unions are clearly disjoint, and so if ϕ is injective, there is evidently a bijection between K and L which is ϕ on U_0 and identity elsewhere. This is exactly how Dedekind proved the Schröder-Bernstein Theorem in that 1887 manuscript. Thus, the Dedekind and Zermelo proofs use the same closure argument, though in Zermelo's terms instead of his $M_1 \to M'$ map Dedekind gets the $M \to M_1$ map.

An Eternal Return: After Cantor posed the problem in 1882 Dedekind solved it using his chains in 1887. This unknown, Bernstein in 1898 found a version of the argument depending on numbers and recursion; Zermelo in 1906, having absorbed Dedekind's chain theory gave a number-free proof; and Whitehead and Russell later elaborated on the matter with unforgiving detail in the logic of relations. In 1932, it became known that Dedekind himself had the simplest proof all along, one reflected in Zermelo's.

§5. Ramifications. Zermelo left Göttingen in 1910 to take up a position as ordinary Professor at the University of Zurich, but he had to vacate that position in 1916. Notably, Paul Bernays completed his *Habilitation* on analytic number theory at Zurich in 1912.[68] From 1921 Zermelo lived in Freiburg im Breisgau, and in 1926 he became an Honorarprofessor at the University of Freiburg. At least of the 1910's Zermelo later wrote that he could not work on the foundations of set theory owing to a "lengthy illness and isolation in a foreign country",[69] and he only published in the subject again starting in 1929. In this fallow period he did publish a couple of papers in the calculus of variations and an article [1914] on integral domains of complex numbers that relied on the Axiom of Choice.[70] Mostly notably, with his predilection for applications Zermelo published two articles on chess, [1913] on the possibility of a winning strategy and [1928] on the ranking of players in tournaments. Before getting to Zermelo's late work we tuck in a discussion of [1913] because of its connection with the modern developments in set theory and then go on to synopsize the relevant developments in set theory to set the stage for Zermelo's re-entry into the fray.

[67] Ferreirós [1999: 240] speculates as to why.

[68] See Specker [1979: 382], a short memoir of Bernays. Bernays entered university at Göttingen in 1909. It was probably through various interactions there and because of Zermelo's move that Bernays went to Zurich.

[69] From Zermelo's report to the Emergency Society of German Science [Wissenschaft] (See Moore [1980]), a report sent to the society with a letter dated December 3, 1930.

[70] Zermelo [1914] used the axiom to get a "Hamel basis" for the complex numbers over the rationals. This work was taken up by Emmy Noether [1916]; see Moore [1982: 173ff].

Zermelo's *An application of set theory to the theory of chess* [1913] is widely regarded as having established the first theorem of game theory. Zermelo first noted that though he focused on chess his considerations apply to "all similar games of reason". The main assumptions were that there are two competing players who alternately play moves and that there are only finitely many possible positions; however, no stopping rules were assumed (unlike how chess is actually played) so that there could be infinitely long plays. After having construed the problem of how to evaluate a chess position as a mathematical problem, Zermelo wrote [1913: 501]: "The method used for solving this problem in what follows is drawn from 'set theory' and 'logical calculus' and shows the fertility of these branches of mathematics in a case where nearly all aggregates considered are *finite*." This recalls the [1909] reductionist emphasis on finite sets. It is a measure of how far set-theoretic thinking has become embedded in mathematics that today we would regard the "application of set theory" in Zermelo's paper as merely set-theoretic formulation in set-theoretic notation. However, it must be remembered that such formulations for the mathematization of problems were still quite novel at the time.[71]

Zermelo [1913] first discussed the concept of being in a "winning position" in chess and advanced a version of what has since been referred to in game theory as Zermelo's Theorem: In chess, either White can force a win, or else Black can force a win, or else both players can force a draw. Turning to his main focus he then argued that if from a chess position q a player can force a win at all, then there is a natural number $t(q)$ such that he can force a win in at most $t(q)$ moves no matter how his opponent plays. As later clarified by Dénes Kőnig [1927], embedded in Zermelo's arguments was the concept of a *winning strategy*, a function for one of the players from positions to moves such that if he plays according to this function he will always win, and *determinacy*, the assertion that one of the players has a winning strategy. Kőnig [1927] pointed to an inadequacy in Zermelo's argument for $t(q)$, and in an appendix presented a new, simple proof of the result verbally communicated to him by Zermelo that we would now recognize as showing, in the parlance of game theory, that (finite length) *zero-sum two-person games of perfect information are strictly determined.* Soon afterwards Lázló Kalmár [1928] generalized this work by allowing infinitely many positions, and gave the first clear formulation and proof of Zermelo's Theorem.[72] Zermelo [1913] may not have initiated game theory, but on the other hand his work predates the pioneering work of Borel [1921] and von Neumann [1928b] on the now familiar minimax strategy.

[71] Even in the much later classic von Neumann–Morgenstern [1944: §§8–10] of game theory, special attention was paid to the benefits of defining games in set-theoretic terms.

[72] Schwalbe–Walker [2001] analyzes this early work by Zermelo, Kőnig, and Kalmár on games and provides an English translation of Zermelo [1913]; the analysis does not however connect Zermelo's argument in the appendix of Kőnig [1927] to determinacy.

The investigation of the determinacy of *infinite* length games is perhaps the most distinctive and intriguing development of modern set theory, and the subject was to expand across the breadth of set theory from combinatorics and forcing to large cardinals and inner model theory.[73] The beginning of this investigation was the basic result of Gale–Stewart [1953] that "open" games are determined, a result to which the subject would return and return, and the argument is essentially the same as Zermelo's.

Enduring his prolonged period of illness and inactivity Zermelo, perhaps perched on a magic mountain in the Black Forest, at least could have had the satisfaction of seeing that his articulations would take root and thrive. The Axiom of Choice became increasingly applied as an explicit construction principle throughout mathematics. Wide-ranging consequences were investigated 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. Zermelo's axiomatization had initially drawn ambivalent response among commentators, especially those exercised by the paradoxes, but the eventual success of the framework would be secured by its increasing mathematical use to structure and clarify arguments, and the underlying abstract, generative view of sets would become generally accepted by the mid-1930's.[74]

The work of Friedrich Hartogs [1915] is particularly notable for being an early confluence with historical import. Cardinal Comparability, the assertion that for any two cardinal numbers $\mathfrak{m}$ and $\mathfrak{n}$ either $\mathfrak{m} \leq \mathfrak{n}$ or $\mathfrak{n} \leq \mathfrak{m}$, had become a problem for Cantor by the time of his *Beiträge* [1895]. Hartogs showed, explicitly working in Zermelo's axiomatization *sans* the Axiom of Choice, that *Cardinal Comparability implies that every set can be well-ordered.* Thus, an evident consequence of the principle that every set can be well-ordered also implied the principle, and this first "reverse mathematics" result established equivalence among the well-ordering principle, Cardinal Comparability, and Axiom of Choice over a base theory. This was the first substantial use of Zermelo's axiomatization after his own in the Well-Ordering Theorem, and as with that theorem, the axiomatization served to ground the investigation of well-orderings in the post-Cantorian era.

In the 1920's fresh initiatives structured the loose Zermelian framework with new features and corresponding developments in the axiomatics. The three most important figures here were Abraham Fraenkel, John von Neumann, and Thoralf Skolem. Zermelo would be influenced by the first two and exercised by the third, particularly in their various attempts at coming to grips with what definite properties are to be for the Separation Axiom. Much has been written with various thematic emphases about these developments,[75] which are briefly synopsized below.

[73]See Kanamori [2003: chap. 6].

[74]See Moore [1982] for the history.

[75]See Moore [1982: 260ff], Hallett [1984: part 2], and Ferreirós [1999: XI].

Fraenkel suggested substantial innovations for Zermelo's axiomatization, promoting an algebraic approach that also led to the first independence results for the axioms. Starting from some urelements and initial sets and closing off under set-theoretic operations, Fraenkel [1922a] constructed a model in which the Axiom of Choice fails for a countable set consisting of pairs.[76] The analogy with Hilbert's *Grundlagen* [1899] was extended, since he too had provided models for establishing axiom independence. These models were typically algebraic closures, and this was to have a thematic reverberation back in set theory, where there was no fixed subject matter: Fraenkel's efforts brought to the fore the attitude latent in Zermelo [1908a] that the domain of set theory is to be an algebraic closure according to the axioms.[77] Fraenkel broached several innovations for Zermelo's axiomatization, innovations that would eventually be adopted but mainly through the more incisive analyses of others. The first was to exclude urelements, notwithstanding his model-building work, and ill-founded sets; the second was to clarify what definite properties are to be for the Separation Axiom; and the third was to close also under functional replacement as provided by the Axiom of Replacement.

Fraenkel [1921][1922: 234ff] raised the issue of categoricity for Zermelo's axiomatization and to get at it proposed his Axiom of Restriction [Beschränktheit]: There are no sets other than are necessary as per the axioms. While Hilbert had achieved categoricity for his axioms by incorporating a maximality condition with his Completeness Axiom, Fraenkel would enforce minimality by a now standard, although still clumsy sounding, restriction clause for closures, having in mind taking the intersection of all possible domains for sets. Fraenkel [1922: 234] specifically mentioned the superfluousness of urelements and ill-founded sets. These would eventually be dispensed with in the mainstream development of set theory owing to considerations of elegance, parsimony, and canonicity, but not through such an *ad hoc* device as Fraenkel's vague, meta-theoretic axiom.

Fraenkel's model building led him to realize [1922a: 286] that the Separation Axiom had to be clarified, and so he formulated set-theoretic propositional functions whose purpose was to render Separation via their algebraic

[76]Fraenkel [1922a] 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, keeping in mind the restriction of Extensionality to sets.

Much later Andrzej Mostowski [1939] forged a method with post-Gödelian sensibilities, bringing out the importance of groups of permutations leaving various urelements fixed, and the result models as well as later versions are now known as the *Fraenkel-Mostowski* models.

[77]Taylor [1993: 544ff] emphasizes this closure aspect of Zermelo [1908a] while observing that the Axioms of Choice and Separation are anomalous.

closure. Fraenkel [1922] also pointed out that at least in one way the algebraic closure approach had to be enriched: He observed the inadequacy of Zermelo's axioms for establishing that $E = \{Z_0, \mathcal{P}(Z_0), \mathcal{P}(\mathcal{P}(Z_0)), \dots\}$ is a set, where $Z_0 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots\}$ is Zermelo's infinite set from his Axiom of Infinity, and proposed the Axiom of Replacement idea to remedy this defect. Fraenkel [1922a: 286][1925: 254][1926: 132ff] proposed several recursive definitions for the propositional functions to figure in Separation and Replacement. Unfortunately, Replacement with these functions was later shown to be inadequate even to secure the set E in question and moreover can already be derived from Zermelo's original axioms.[78] Actually, Zermelo's axioms turn out to be very weak for handling recursive definitions: The union of E, with membership restricted to it, is a model of Zermelo set theory, yet this model does not contain $\{\emptyset, \mathcal{P}(\emptyset), \mathcal{P}(\mathcal{P}(\emptyset)), \dots\}$ nor its union, the countable set consisting of the hereditarily finite sets. Hence, Zermelo set theory cannot establish the existence of some simple countable sets consisting of finite sets and could be viewed as remarkably lacking in closure under finite recursion. Replacement simply serves in part to rectify the situation by rounding out the available sets.[79] Such examples show how necessary the Replacement idea is for basic set-theoretic constructions, despite how Replacement has sometimes been portrayed as only affecting large-cardinality sets and regarded as less crucial than the Zermelo axioms. In any case, it was von Neumann's formal incorporation of a *method* into set theory, transfinite recursion, that necessitated the full exercise of Replacement.

Von Neumann effected a counter-reformation of sorts by incorporating Cantor's transfinite numbers and even his absolutely infinite or inconsistent multiplicities into a distinctively new axiomatization of set theory and moreover began the process of focusing the subject matter of set theory on the cumulative hierarchy.

[78] See von Neumann [1928: 375ff]. Skolem [1923] independently pointed out the inadequacy of Zermelo's axioms for procuring the set E above and proposed Replacement in the framework of first-order logic. It is this proposal that would eventually be adopted, but in any case Skolem in a publication [1962: 13] appearing a year before his death wrote of Replacement, simply and without nuance: "Fraenkel introduced a further axiom which is more powerful with respect to the proof of the existence of large transfinite cardinals."

[79] See Mathias [2001] for more on the weakness of Zermelo set theory in this direction; Mathias provides a general method for constructing "slim" transitive proper class models of the theory and gets such a model in which the hereditarily finite sets do not form a set. Zermelo's Axiom of Infinity stated that there is a set Z such that $\emptyset \in Z$ and whenever $x \in Z$, so also $\{x\} \in Z$. One could in retrospect take the view that the thrust of an axiom of infinity should be to legitimize *one* instance of finite recursion that results in an infinite set. To ensure the existence of the set of hereditarily finite sets one could posit: There is a set Z such that $\emptyset \in Z$ and whenever $x, y \in Z$, so also $x \cup \{y\} \in Z$. This would round out Zermelo's theory in the finite domain, but Mathias's results establish that many sets "isomorphic" to the set of hereditarily finite sets still cannot be shown to exist.

For Cantor the transfinite numbers had become central to his investigation of definable sets of reals and the Continuum Problem, and sets had emerged structured with well-orderings and only as the developing context dictated, with the "set of" operation never iterated more than three or four times. For Zermelo his second, reverse-inclusion-chain 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. However, Zermelo's reductionism with respect to numbers mainly concerned reducing mathematical arguments to set-theoretic arguments from axioms and would give way to von Neumann's systematic incorporation of the transfinite numbers as *bona fide* sets.

Von Neumann [1923, 1928], and before him Dimitry Mirimanoff [1917, 1917a] and Zermelo in unpublished 1915 work, isolated the now familiar concept of *ordinal*, with the basic idea of taking precedence in a well-ordering simply to be membership.[80] 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-isomorphic to exactly one ordinal ordered by 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 several principles of generation for ordinal numbers would be subsumed by the Zermelian framework. For this and already to define the arithmetic of ordinals von Neumann saw the need to formalize transfinite recursion. And Replacement was necessary even for the very formulation, let alone the proof. With the ordinals in place von Neumann [1928, 1928a] completed the restoration of the Cantorian transfinite by *defining* the cardinals as the initial ordinals, codifying a strategy that had in fact appeared in the 1899 Cantor-Dedekind correspondence.

That correspondence, it will be remembered, featured Cantor's absolutely infinite or inconsistent multiplicities in distinctive juxtaposition with sets, and von Neumann's own axiomatization [1925] of set theory would dramatically incorporate these multiplicities in the first systematic treatment of sets together with proper classes. In a rectification of Fraenkel's functional approach to Separation, von Neumann axiomatized a notion of function taken as primitive and proceeded to establish a context sufficient for the full exercise of Replacement and hence for the Cantorian theory of transfinite numbers as transmuted to ordinals. He then encapsulated the distinction between sets and proper classes in his pivotal Axiom IV2: A class is a set exactly when there is no surjection from that class onto the universe V of sets. This axiom was alluded to at the end of §2 in connection with the Axiom of Choice. More crucially, IV2 together with Replacement obviated the need to have Separation at all, as von Neumann's functions handily supplanted

[80]See Hallett [1984: §8.1].

Zermelo's definite properties. Generally speaking, IV2 transformed Cantor's negative concept of inconsistent multiplicity with their taint of paradox into the positive concept of having a surjection onto V and hence onto the class Ω of all ordinals. This is reminiscent of how the age-old negative concept of infinite was transformed into the positive concept of being Dedekind-infinite, i.e., having a countable subset.

For reasons of parsimony and elegance the basic Zermelian framework will be reaffirmed as sufficient in the coming decades, when formalized in first-order logic. However, from von Neumann would be inherited both a predisposition to entertain proper classes in the mathematical development of set theory, a predisposition that would have fruitful consequences particularly in the theory of large cardinals, and also the finite axiomatizability of definability, which would be antithetical to Zermelo's later infinitistic vision but would play a significant role in the development of set theory, e.g., in the development of Gödel's constructible universe (cf. §8 below).

With ordinals and Replacement, set theory continued its shift toward a theory of a more definite transfinite subject matter, a process fueled by the incorporation of well-foundedness. Mirimanoff [1917: 51ff] was the first to study the well-founded sets, and the cumulative hierarchy is distinctly anticipated in his work. 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,[81] and defined the cumulative hierarchy 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.

Getting finally to Skolem's main contribution, his prescient [1923] made the proposal of using for Zermelo's definite properties for the Separation Axiom those properties expressible in first-order logic. 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]

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

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 the Skolem 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.

As for the emergence of first-order logic, 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 text Hilbert–Ackermann [1928], which has a considerable overlap with lecture notes for a course first given in 1917–8 by Hilbert at Göttingen,[82] reads remarkably like a modern text. In marked contrast to the formidable works of Frege and Russell with their forbidding notation and all-inclusive approach, the text 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.

As Hilbert was lecturing on logic in 1917–8 his former student Hermann Weyl brought out a notable monograph, *Das Kontinuum* [1918]. Endeavoring to provide a predicative foundation for analysis Weyl started with the natural numbers and constructed what is essentially a version of that part of the ramified theory of types in which quantification is restricted to variables ranging over the natural numbers. Referring specifically to Zermelo's definite property Weyl [1918: 36] thought of his approach as providing a satisfactory rendition; earlier, Weyl [1910] had raised the issue of how to characterize definiteness. As for temporal priority Skolem [1923: 152] wrote that he had in fact communicated his result on the relativism of set-theoretic notions to Bernstein in the winter of 1915–6.

§6. Second-order ZF and its models. After his long hiatus and in his late fifties Zermelo returned to the fray in 1929 for what would be a brief, final burst of activity in the foundations of set theory and logic. Exercised by new finitistic trends in mathematical logic Zermelo advocated an expansive,

[82]See Moore [1988: 188ff] and Sieg [1999: B].

infinitistic viewpoint, in what would turn out to be a rearguard action against the new mathematics of finitary metamathematics.[83] Zermelo delivered a series of broad-ranging lectures on foundational issues in Warsaw in May and June.[84] In his first publication [1929] on the foundations of set theory in nearly two decades Zermelo, roused by criticism of the vagueness of his definite property for the Separation Axiom, provided an axiomatization for the property in second-order terms.

In reference to his informal appeal to definiteness in his axiomatization paper Zermelo [1929: 340] wrote: "A generally accepted 'mathematical logic' to which I could have appealed did not exist at that time, as it does not today when each foundational researcher has his own system of logic [seine eigene Logistik]." He then criticized Fraenkel's approach to Separation through his recursively defined propositional functions as "constructive" and presupposing in their formulation the natural numbers — this of course runs afoul of Zermelo's set-theoretic reductionism. Zermelo advocated taking instead an "axiomatic" approach, nodding to von Neumann [1925] but preferring a simpler presentation. Starting generally from a system R of fundamental relations, Zermelo [1929: 344] stipulated that the collection of definite properties relative to R is to contain the relations in R; is to be closed under the logical connectives and first- and second-order quantification; and is to have no proper subcollection enjoying these features. Thus, Zermelo axiomatically presented the definite properties as an explicit *closure*, and this has an almost affecting thematic resonance with his functional closure proof of the Schröder-Bernstein Theorem (cf. §4 above), his second, intersection proof of the Well-Ordering Theorem, and his axiomatization of set theory as a closing off of the domain of set theory. The incentive of the first had been the presupposition of the natural numbers in an earlier proof, and this too resonates with Zermelo's criticism of Fraenkel's approach. Zermelo [1929: 344] pointed out that he had not presupposed the natural numbers. In any case, in his first explicit engagement with syntax he supported the use of a *finitary* language in his explication of definiteness.

Skolem [1930] responded with alacrity and in the same journal to Zermelo [1929]. Zermelo may have avoided the natural numbers, but Skolem pointed out the unrestrained use of set theory itself in Zermelo's axiomatization of definite property. Skolem was sensitive to the economy of resources to be used in the metalanguage, and this extended more crucially to the underlying logic: Skolem referred back to his [1923] and its first-order logic proposal

[83] Van Dalen–Ebbinghaus [2000] and Taylor [2002] both suggest an early genesis for Zermelo's infinitistic views, pointing to a one-page typescript, *Thesen über das Unendliche in der Mathematik*, found in Zermelo's *Nachlass* which set out his conviction that mathematics is of an infinitary character and its logic must be based on an infinitary language. However, the assumption that this typescript is from 1921 is put into question by Ebbinghaus [2003], who argues that it is only from 1942.

[84] For more on these lectures see Moore [1980: §10.2] and van Dalen–Ebbinghaus [2000].

for Separation and proceeded to argue against the new vagueness introduced by Zermelo in allowing second-order quantification. Finally, he informed Zermelo of (what we now call) Skolem's Paradox. Zermelo had already been preparing his next paper [1930] for publication in *Fundamenta*, and he may only have become aware of Skolem [1923] after seeing proofs of Skolem [1930].[85] In any case, exercised by adversity Zermelo made several changes that accentuated his position, and as he had done two decades before, he published a paper that would mark a signal advance.

Zermelo in his remarkable *On Boundary Numbers and Set Domains* [1930] offered his final axiomatization of set theory as well as a striking, synthetic view of a procession of natural models that would have a modern resonance. Appearing only six articles after Skolem [1930] in *Fundamenta*, Zermelo [1930] is ostensibly a response, more informal and rough around the edges than his writings decades earlier, but its dramatically new picture of set theory reflects gained experience and suggests the germination of ideas over a prolonged period. The article is a *tour de force* which set out principles that would be adopted in the further development of set theory and focused attention on the cumulative hierarchy picture, dialectically enriched by initial segments serving as natural models.

Zermelo [1930: §1] first formulated his axiom system, and though the presentation is opaque largely because of a second-order lens, the thrust of ZFC is there. Indeed, Zermelo used the term "Zermelo–Fraenkel" and the acronym "ZF" to indicate the result of adding Replacement to the Zermelo [1908a] axioms.[86] Zermelo proceeded to focus his investigations on "ZF′", the result of deleting Infinity as not being part of "general" set theory,[87] assuming Choice as an implicit, underlying "general logical principle", and adjoining Foundation.

As described in §5 above, Foundation in modern set theory ranks the universe of sets into a 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. In a notable inversion, what has come to be regarded as the underlying *iterative conception* became a heuristic for motivating the axioms of set theory generally. It is nowadays 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

[85]See Ebbinghaus [2003: 201]. Skolem [1930] himself noted that Zermelo "does not seem to know my Helsinki talk [i.e., Skolem [1923]]."

[86]"Zermelo-Fraenkel" was first invoked by Von Neumann [1928: 374] for this purpose.

[87]This recalls the strategy of Zermelo [1909]. As discussed below, Infinity will hold in all of his models except those of one "characteristic", namely ω, and there is an interaction there with Foundation in connection with the first "development" theorem.

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.

Modern ZFC is recognizable in Zermelo's presentation *except* for ambiguity about the applicability of Separation and Replacement. What properties, as given by propositional functions or logical formulas, are to be in the purview of these two axioms? The vagueness of definite property for Separation had prompted several efforts at remedy including Zermelo [1929], and Zermelo [1930: 31] himself wrote that "when appropriately interpreted" Replacement implies Separation. The discussion below of Zermelo's models clarifies how these two axioms are to be taken in his second-order context. There is also an issue about Zermelo's specific formulation of Foundation, and this too will be addressed in due course.

Zermelo [1930: §2] provided formulations now basic to modern set theory, but as with the ZFC axioms the presentation is opaque, here because of Zermelo's insistence on having urelements and fixing one as *the* empty set. Zermelo formulated the von Neumann ordinals, but starting generally from any urelement u: $u, \{u\}, \{u, \{u\}\}, \ldots$. (For convenience, these will be referred to as the *u-ordinals*; the usual (von Neumann) ordinals are a special case, and we shall avail ourselves of them and their usual notation in what follows.) Notably, Zermelo in unpublished 1915 work may have been the first to sketch the rudiments of the von Neumann ordinals.[88] Nonetheless, it is evident from his presentation that for Zermelo Cantor's ordinal numbers retained a prior sense; Zermelo's reductionism interestingly did not extend to identifying the numbers with sets, and the various u-ordinals only "represent" the ordinal numbers. With urelements in play, Zermelo investigated the various *normal domains*, models of ZF′ when the membership relation is restricted to them. In Zermelo's words, a normal domain has a "width" given by its *basis* consisting of urelements, and a "height" given by its *characteristic*, the supremum of ordinal numbers represented in it.

Zermelo's crucial observation was that there are simple set-theoretic conditions on the ordinals that secure his ZF′, conditions that newly underscore how the Zermelian sets are to be an algebraic closure of his axioms. In an inspired move, Zermelo took the characteristic κ of a normal domain P to be again an ordinal number, thereby "resolving" the Burali-Forti Paradox by having κ outside of P but within set theory. Zermelo's simple conditions are:

(I) κ is a regular cardinal, i.e., if $\alpha < \kappa$ and $F\colon \alpha \to \kappa$, then $\bigcup F``\alpha < \kappa$, and

(II) κ is a strong limit cardinal, i.e., if $\beta < \kappa$, then $2^\beta < \kappa$.

[88] See Hallett [1984: 278ff].

Zermelo initially observed that these conditions are *necessary*; he argued in terms of his representing u-ordinals, but we can proceed directly with the usual (von Neumann) ordinals: To establish (I), suppose that $\alpha < \kappa$ and $F : \alpha \to \kappa$. Since $\alpha \in P$, by a crucial use of Replacement $F``\alpha$ is a *set* in $P$. $\bigcup F``\alpha$ is thus a set, in fact an ordinal, in P, and hence $\bigcup F``\alpha < \kappa$. To establish (II), suppose that $\beta < \kappa$. Then $\mathcal{P}(\beta)$ is a set in $P$ by the Power Set Axiom. If to the contrary $2^\beta \geq \kappa$, there would be a $G : \mathcal{P}(\beta) \to \kappa$ such that $G``\mathcal{P}(\beta) = \kappa$. But by another crucial application of Replacement $G``\mathcal{P}(\beta)$ must be a set in P, which is a contradiction.

The intended applicability of Separation and Replacement can be informed by these proofs. Zermelo [1930: 30] had stated Separation in terms of propositional functions and provided the following footnote:

> Here the propositional function $f(x)$ can be quite *arbitrary* [ganz *beliebig*], as can the replacement function in (E) [the Replacement Axiom]. Thus none of the consequences of restricting these functions to a particular class are relevant for the point of view taken here. I reserve for myself a detailed discussion of the issue of 'definiteness" following my last note in this Journal [Zermelo [1929]] and the critical "Remarks" thereon by T. Skolem [Skolem [1930]].

Despite Zermelo [1929], Zermelo now foregoes all restrictions and lets his properties for Separation be arbitrary. This footnote has been variously projected into the larger conceptual context of Zermelo's advocacy of an infinitistic viewpoint in the mid-1930's.[89] While such projections are germane and broadly significant, there is also an immediate *mathematical* necessity in the context of Zermelo [1930]:

The above arguments for the necessity of conditions (I) and (II) both require that Replacement be applied without restriction. In modern terms, *Replacement should be taken as a single, second-order axiom quantifying over all possibilities*, yielding what we now call Second-Order ZF. It does not even suffice for Replacement to be a schema of second-order axioms, and the reference in the above cited footnote to Zermelo [1929], which would still sanction Separation and Replacement taken as schemata, is misleading.[90] Zermelo's exposition is generally less meticulous than it was two decades before and especially haphazard on the role of Replacement: Arguing for the necessity of condition (I), Zermelo only pointed out that the union of a set of u-ordinals is again a u-ordinal, but does not discuss how he has a set. For the necessity of (II), he does not explicitly associate u-ordinals to the subsets of a set, and when finally he appeals to Replacement it is for a limit case made redundant by (I).

---

[89]See Moore [1982: 269], Taylor [1993: 550], Van Dalen–Ebbinghaus [2000: 152], and Ebbinghaus [2003: 201, 205].

[90]Tait [1998] makes this point, albeit for a less direct reason than the necessity of (I) and (II).

Zermelo [1930: §3] continued with three "development" theorems for his normal domains. The first stated that each normal domain P is indeed stratified according to rank because of Foundation: Let Q be the basis and κ the characteristic of P. With the caveat that for general X, $\mathcal{P}(X)$ is to denote the collection of all sub*sets* of X, recursively define the cumulative ranks:

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

and conclude that

$$P = V_\kappa^Q = \bigcup_{\alpha<\kappa} V_\alpha^Q.$$

Zermelo emphasized the partitioning into disjoint layers $V_{\alpha+1}^Q - V_\alpha^Q$ and pointed out that each such layer contains a u-ordinal, and so the hierarchy is strict. It becomes evident that, consistent with his expansive view of Replacement, Zermelo is entertaining all possible subsets in his normal domains, so that the $\mathcal{P}(V_\alpha^Q)$ above has an absolute significance independent of the domain P.[91] However, there is an ambiguity, or at least a relativism, about the status of Q, Zermelo's "totality" ["Gesamtheit"] of urelements: Is it itself to be a set? Zermelo did not posit that it be a set in the sense of P, and the open-ended articulation above of Zermelo's scheme accommodates this with the ranks V_α^Q not necessarily being sets in the sense of P.[92] In the special case of Q being finite, Zermelo's axioms do ensure that Q and all the V_α^Q's are sets.

The first development theorem raises an issue about Foundation and Infinity. Zermelo [1930: 31] actually formulated Foundation both as stipulating that there are no infinite descending $\in$-chains, "Or equally: every partial domain T contains at least one element t_0 none of whose elements are in T." As is now well-known, the former form implies the latter only in the presence of substantial axioms, including the Choice. Aside from this, the latter, a second-order form of Foundation, was needed in the proof of the first development theorem: It is well-known that usual (set) Foundation implies such a strong form assuming *Transitive Containment*, that every set is a subset of a transitive set.[93] Furthermore, (first-order) Replacement and

[91] This is clarified by the first isomorphism theorem in [1930: §4] mentioned below.

[92] It is plausible, even natural, for the totality of urelements not to be a set; cf. Barwise [1975: 11], in which KPU is Kripke-Platek set theory with urelements and KPU$^+$ is that theory augmented with the axiom that the urelements form a set.

Gregory Taylor (private communication) avers that Zermelo intended for the basis Q to be a set in the sense of P; Taylor takes Zermelo [1930] and his later [1935] as forming a common framework and cites [1935: 141] where in a hierarchy of propositions Zermelo starts with a set.

[93] Suppose that T is a non-empty class, say with $x \in T$. Let t be a transitive set such that $\{x\} \subseteq t$ and consider the set $t \cap T$, given by Separation. By (set) Foundation there is a $t_0 \in t \cap T$ such that $t_0 \cap t \cap T = \emptyset$. But then, since t is transitive, $t_0 \cap T = \emptyset$. This

Infinity do imply Transitive Containment.[94] However, Zermelo is not assuming Infinity. In fact, it is a latter-day observation that Second-Order ZF with only (set) Foundation but without Infinity does not suffice to establish Transitive Containment, and in fact has models whose membership relation is ill-founded.[95] Hence, Foundation as Zermelo formulated it in second-order terms is necessary for his cumulative hierarchy analysis in the absence of Infinity, i.e., in case of normal domains with characteristic ω.

The second development theorem addressed *unit* domains, those normal domains P with a single urelement, $Q = \{u\}$, and provided information about their ranks.[96] Zermelo had defined a Beth-type cardinal-valued function as follows: $\Psi(0) = 0$; $\Psi(\xi + 1) = 2^{\Psi(\xi)}$; and for limit ordinals α, $\Psi(\alpha) = \sup_{\xi<\alpha} \Psi(\xi)$. Zermelo now proved that each V_α^Q has cardinality $\Psi(\alpha)$ for infinite α.[97] Here he pointed out that V_α^Q is a set because of Replacement, in an argument that should have been used to establish the necessity of (I).

In remarks following the second theorem Zermelo concluded that unit domains satisfy von Neumann's axiom IV2 (cf. §§2, 6), that a class is a set exactly when there is no surjection from that class onto the entire universe. Zermelo thus established the consistency of IV2 relative to his axioms, proceeding in his second-order context with natural models. Zermelo noted, "For unit domains, though not for arbitrary normal domains, von Neumann's axiom holds .... Restricting set theory just to 'unit domains' would rob it for the most part of its applicability." Moreover, Zermelo [1930: 45] returned with emphasis to this point to criticize the putative restrictiveness of von Neumann's axiom. However, the following result shows that Zermelo's reservations have only to do with the nature and size of the totality of urelements. $|X|$ denotes the cardinality of a set X in the presence of the Axiom of Choice, i.e., the least ordinal bijective with X; that a cardinal κ satisfying (I) and (II) satisfies $\bigcup_{\alpha<\kappa} \kappa^\alpha = \kappa$ is a simple exercise in cardinal arithmetic; recall that in our articulation of Zermelo's context the basis Q of a normal domain P may or may not be a set, though it is not assumed to be a set in the sense of P.

PROPOSITION 4. *For any normal domain P with basis Q and characteristic κ, von Neumann's axiom holds in P iff Q is a set satisfying $|Q| \leq \kappa$.*

---

argument may have been first given by Gödel, in his letter of 20 July 1939 to Bernays (Gödel [2003: 121]).

[94]For any set x and $n \in \omega$, recursively define $x_0 = x$, and $x_{n+1} = \bigcup x_n$. Then $\bigcup_n x_n$ is a transitive set containing x. The use of Replacement and Infinity in this argument was noted by Gödel in the letter cited in the previous footnote.

[95]See Vopěnka–Hajek [1963] and Hauschild [1966].

[96]Zermelo had fixed a urelement u_0 to be the empty set, so this u is presumably not to be u_0.

[97]Zermelo proved for finite α that V_α^Q has cardinality $\Psi(\alpha + 1)$. His indexing of the cumulative ranks started at 1 instead of 0.

Proof. Suppose first that von Neumann's axiom holds in P. Since κ itself is a proper class of P, κ is surjective onto P and hence by Replacement P would be a set. But then, Q can be separated from P as the *set* of its members that are empty.

For the converse, suppose that Q is a set satisfying $|Q| \leq \kappa$. Every set in P, being well-orderable, is bijective with some u-ordinal in P,[98] and hence has cardinality less than κ. Hence, one can prove by induction that for every $\alpha \leq \kappa$, $|V_\alpha^Q| \leq \bigcup_{\alpha<\kappa} \kappa^\alpha = \kappa$. So presuming that urelements are exempted, von Neumann's axiom for P amounts to: For any $X \subseteq P$, $X \in P$ *iff* $|X| < \kappa$.

The forward direction was already noted. For the converse, if $|X| < \kappa$, then applying Replacement and (I) to the function $F\colon X \to \kappa$ given by $F(x) =$ the least ξ such that $x \in V_{\xi+1}^P$, it follows that there is an $\alpha < \kappa$ such that $X \subseteq V_\alpha^Q$, and so $X \in V_{\alpha+1}^Q \subseteq P$. ⊣

The third development theorem provided a more refined hierarchy for normal domains based on the Ψ function, to wit the $(\alpha + 1)$th cumulative level is now to be formed by adjoining only subsets of the αth level of cardinality at most $\Psi(\alpha)$, and with it established that conditions (I) and (II) are also *sufficient* for V_κ^Q to be a normal domain. Actually, this "canonical" hierarchy is, strictly speaking, not needed here or for the later results of the article, but it does clarify how Zermelo viewed his normal domains and their interactions.

Those cardinals κ satisfying conditions (I) and (II) were called *Grenzzahlen* [*boundary numbers*] by Zermelo,[99] and when $\kappa > \omega$ are now called the *(strongly) inaccessible* cardinals. These cardinals are basic in the theory of large cardinals, a mainstream of modern set theory devoted to the investigation of strong hypotheses and consistency strength, and in fact are the modest beginnings of a natural linear hierarchy of stronger and stronger postulations extending ZFC.[100] It is through Zermelo [1930] that inaccessible cardinals became structurally relevant for set theory as the delimiters of natural models. Just 13 articles before Zermelo's in *Fundamenta*, Sierpiński-Tarski [1930] had formulated the inaccessible cardinals arithmetically as those uncountable cardinals that are not the product of fewer cardinals each of smaller power and observed that inaccessible cardinals are regular limit cardinals, the first large cardinal concept, from Hausdorff [1908: 443]. Be

[98] This fundamental von Neumann result, a consequence of (first-order) Replacement, was pointed out by Zermelo [1930: §2(6)].

[99] "Boundary number" is how "Grenzzahlen" has been translated, e.g., by Hallett in his translation of Zermelo [1930] in Ewald [1996]. The more literal "limit number" has its connotative advantages as well: Zermelo [1930] referred to Kant's antinomies, and Kant in his *Prolegomena* had distinguished between *Grenzen* and *Schranken*, with the first having entities beyond.

[100] See Kanamori [2003].

that as it may, in the early model-theoretic investigations of set theory the inaccessible cardinals provided the natural models as envisioned by Zermelo. Years later Shepherdson [1952] provided more formal proofs of Zermelo's results in a first-order context with sets and classes but without urelements, taking account of the relativity of concepts and isolating Zermelo's models as the transitive and super-complete models.[101] Very recently Uzquiano [1999] investigated models of Second-Order Zermelo set theory (no Replacement but taking Separation as a single second-order axiom) and showed that V_δ for limit ordinals $\delta > \omega$ are by no means the only possibilities and that there is already considerable variation at level ω.

Zermelo [1930: §4] proceeded with three isomorphism theorems that established a second-order categoricity of sorts for his axioms in terms of the cardinal numbers of the bases and the characteristics. Note that it is still plausible for the bases not to be sets, since Zermelo was taking cardinal number in a prior Cantorian sense and not as von Neumann ordinals.[102] The first isomorphism theorem stated that two normal domains with the same characteristic and bases of the same cardinality are isomorphic, the isomorphisms generated by bijections between the bases. Unbridled second-order Replacement is crucial here as well, this time to establish that the cumulative ranks of the two domains are level-by-level extensionally identical, and this clarified the sense in which the $\mathcal{P}(V_\alpha^Q)$ must have an absolute significance. There is a historical resonance with Dedekind [1888] who had established the categoricity of *his* second-order axioms for arithmetic. Dedekind needed finite recursion to exhibit bijective mappings between models, and Zermelo needed *transfinite* recursion and thus Replacement. The second isomorphism theorem stated that two normal domains with different characteristics and bases of the same cardinality are such that one is isomorphic to a cumulative rank of the other. The third isomorphism theorem stated that two normal domains with the same characteristic are such that one is isomorphic to a subdomain of the other. Hence, as Zermelo emphasized, a normal domain is characterized up to isomorphism by its *type*, the pair $\langle \mathfrak{q}, \kappa \rangle$ where $\mathfrak{q}$ is the cardinal number of the basis, which can be arbitrary, and κ is the characteristic, which must be ω or inaccessible; and given two types $\langle \mathfrak{q}, \kappa \rangle$ and $\langle \mathfrak{q}', \kappa' \rangle$, isomorphic embeddability is a consequence of $\mathfrak{q} \leq \mathfrak{q}'$ and $\kappa \leq \kappa'$.

Zermelo [1930: §5] concluded with a brief discussion of existence, consistency, and categoricity. Speculating on the possibilities for characteristics, Zermelo pointed out that ω is a characteristic, as starting with any normal

[101] A class is *super-complete iff* any sub*set* is an element. Shepherdson [1952: 227] wrote: "[Equivalent results] were obtained by Zermelo although in an insufficiently rigorous manner. He appeared to take no account of the relativity of set-theoretical concepts pointed out by Skolem." Skolem relativity has seemingly become entrenched, but Zermelo of course was deliberately working in a second-order context and decidedly opposed "Skolemism"!

[102] However, Fraenkel–Bar Hillel [1958: 92] found Zermelo [1930] not "stringent" and in particular found the concept of the cardinality of the basis "objectionable".

domain P with basis Q one can consider the subdomain V_ω^Q. He actually viewed this in terms of the "canonical" hierarchy of his third development theorem, and this brings out how V_ω^Q for him consists of the hereditarily finite subsets of P and reinforces for infinite Q how Q is not to be a set in the sense of P. This also suggests why Zermelo deliberately eschewed Infinity, which thus establishes the relative consistency of ZF′. Zermelo proceeded by analogy to the least inaccessible cardinal via the ordinal type of the next normal domain, and pointed out how, like ω, such a cardinal cannot be proved to exist in ZF′. This kind of positing by analogy with Infinity is now typical in the theory of large cardinals and is resonant with Cantor's own seamless account of number across the finite and the transfinite. Since already it is seen that ω may exist in one model but not another, Zermelo [1930: 45] wrote: "Our axiom system is *non-categorical*, which in this case is not a disadvantage but rather an *advantage*, for on this very fact rests the enormous importance and unlimited applicability of set theory."

In a sweeping climax Zermelo put forward the general hypothesis that "every categorically determined domain can also be interpreted [aufgefasst] as a 'set', i.e., can appear as an element in a (suitably chosen) normal domain" and postulated "the existence of an unbounded sequence of boundary numbers [Grenzzahlen]" as a new axiom of "meta-set theory". The hypothesis calls into question the contention that the bases are not necessarily sets, but he had started his remarks by discussing unit domains, and by now he was painting in broad strokes. The postulation would bijectively correlate the ordinal numbers with the inaccessible cardinals and so provide for an endless procession of models.[103] The open-endedness of Zermelo's original [1908a] axiomatization had been structured by Replacement and Foundation, but after synthesizing the sense of progression inherent in the new cumulative hierarchy picture and the sense of completion in the inaccessible cardinals, Zermelo advanced a new open-endedness with an eternal return of models. This dynamic view of sets and set theory was a marked departure from Cantor's (and later, Gödel's) focus on a fixed universe of sets. Through means dramatically different and complementary to Cantor's absolute infinite Zermelo dissolved the traditional antinomies of set theory through a dialectical interplay between the global and the local.[104] Furthermore, not only did Zermelo subsume von Neumann's axiom IV2, that principled means of handling classes too large, but by having such classes be elements in a next normal domain and therefore coming under the purview of his generative axioms like Power Set, Zermelo dissolved further antinomies like Hilbert's

[103] Tarski [1938] also and later posited arbitrarily large inaccessible cardinals via his Axiom of Inaccessible Sets; he was led to this axiom by cardinality and closure considerations, and he formulated it in such a way that it implies the Axiom of Choice. In contrast to Zermelo's informal, second-order approach Tarski could be seen to be working in first-order ZF.

[104] Tait [1998] provides a sophisticated account of this aspect of Zermelo's conception of set theory and draws out large cardinal reflection principles.

Paradox (see §1) and the related incompatibility of $2^\kappa > \kappa$ with κ being the cardinal number of the universal class, the original "contradiction" that Russell came to in 1900 while studying Cantor's work.

Zermelo [1930: 47] concluded grandly and polemically:

> The "ultrafinite antinomies of set theory" that scientific reactionaries and anti-mathematicians refer to so assiduously and passionately in their campaign against set theory, these seeming "contradictions", are only due to the confusing of set theory itself, which is non-categorically determined by its axioms, with its particular representing models: What appears in one model as an "ultrafinite non- or super-set" is in the next higher one already a fully valid "set" with cardinal number and ordinal type, and is itself a foundation stone for the construction of a new domain. The unlimited series of Cantor's ordinal numbers is matched by just as unlimited a double series of essentially different set-theoretic models, in each of which the whole classical theory is expressed. The two diametrically opposing tendencies of the thinking mind [denkenden Geistes], the ideas of creative *progress* and of collective *completion*, ideas that also lie at the basis of the Kantian "antinomies", find their symbolic representation and their symbolic reconciliation in the transfinite number series based on the concept of well-ordering, which in its unrestricted progress reaches no true conclusion, only interim stopping points, namely those "boundary numbers" ["Grenzzahlen"] that separate the higher from the lower model types. And thus, the set-theoretic "antinomies" lead, if properly understood, not to a narrowing or mutilation, but rather to a presently unsurveyable unfolding and enriching of mathematical science.

As set theory would go, Foundation and the corresponding cumulative hierarchy picture would provide the setting for a developing high tradition that had its first milestone in Kurt Gödel's development [1938] of the constructible hierarchy. Zermelo [1930] would be peripheral to this development, presumably because of its second-order lens and its lack of rigorous detail and of attention to relativism. Indeed, it was little cited for decades except as a source for Foundation and ZFC set theory. However, with the assimilation of set-theoretic rigor, increasing confidence in consistency, and the emergence of ZFC as the canonical set theory, there has been of late new appreciation of the sweep of Zermelo [1930] especially because of renewed interest in second-order logic. Zermelo himself did not pursue axiomatic set theory after his [1930], but took and developed its schematic picture in a new direction.

§7. Incompleteness and infinitary logic. Zermelo's final efforts in mathematics were directed at developing an infinitary logic based on well-founded relations, this largely in response to what he perceived as a threat to set theory and mathematics presented by the "finitistic prejudice" of Skolem.[105]

[105] Although this was the main thrust, Zermelo in his return to mathematical activity also published a couple of papers in the calculus of variations and applied mathematics in 1930 and 1931 respectively and (as we have seen) edited Cantor's collected works [1932].

Loosely speaking, Zermelo proposed logics based on quantifiers regarded as infinitary conjunctions or disjunctions, reverting to an old view but now newly cast with unrestricted cardinalities, and proofs not as formal deductions from axioms, but as semantic determinations of truth or falsity of a proposition through transfinite induction based on its well-founded construction from elementary propositions. In this Zermelo seemed to have been little aware or took little interest in his old mentor Hilbert's development of metamathematics, or proof theory, and its avowedly finitistic approach and economy of means for establishing mathematical consistency.[106] Zermelo's efforts would be quickly overshadowed and largely forgotten in the wake of concurrent developments in first-order logic, initially isolated as an uninterpreted formalism by Hilbert, which would lead to the emergence of mathematical logic as a field of mathematics.

Kurt Gödel virtually completed the mathematization of logic by submerging metamathematical methods into mathematics. The main vehicle was of course the direct coding, "the arithmetization of syntax", in his celebrated Incompleteness Theorem [1931], which transformed Hilbert's consistency program and led to the undecidability of the Decision Problem from Hilbert–Ackermann [1928] and the development of recursion theory. 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 model theory (semantics) of first-order logic, and secured its key instrumental property with the Compactness Theorem.[107]

Tarski [1933, 1935] then completed the mathematization of logic by providing his definition of truth, exercising philosophers to a surprising extent ever since. 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 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 (or semantic) 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

[106]Zermelo did maintain, in a letter of 4 February 1932 to Richard Courant found in the *Nachlass*, that Hilbert was "his first and only teacher in science."

[107]Notably, Gödel's dissertation adviser was Hans Hahn, who almost three decades before had written with Zermelo that encyclopedia article Zermelo–Hahn [1904] on the calculus of variations.

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.

The 1931 annual September meeting of the *Deutsche Mathematiker-Vereinigung* held at Bad Elster, like the 1904 Heidelberg International Congress, would mark a generational transition. Both the 60-year-old Zermelo and the 25-year-old Gödel spoke on the same afternoon, of the 15th, Zermelo on his proposals for infinitary logic and Gödel on his recently published Incompleteness Theorem. This was certainly a pivotal clash of viewpoints, one that was to play itself out in a subsequent, albeit brief, exchange of letters.[108]

Zermelo wrote to Gödel ([Gödel 2003a: 420ff]) within a week of the meeting, on 21 September, enclosing a copy of Zermelo [1930] and raised an issue based on the introductory sketch in Gödel [1931]. Zermelo questioned the assertion that the predicate

$$\overline{\mathrm{Bew}}[R(n);n]$$

is in the "system" where famously $R(n)$ is the nth formula in some recursive enumeration of the formulas in one free variable; $[R(n);n]$ is a term denoting the sentence resulting when the free variable is replaced by the numeral of n; Bew (abbreviating Beweis [proof]) is the provability predicate; and the overline is negation. Of course, that the displayed predicate is expressible in the system is crucial to Gödel's argument. Not understanding, Zermelo proposed to omit "Bew" and write instead

$$\overline{[R(n);n]}$$

and proceeded to point out that if *this* were in the system, then one would have the now paradigmatic contradiction that with q such that $R(q)$ is $\overline{[R(n);n]}$, $[R(q);q]$ is both true and false. Zermelo then averred: "Just as in the Richard and Skolem paradoxes, the mistake rests on the (erroneous) assumption that every mathematically definable notion is expressible by a 'finite combination of signs' (according to a *fixed* system!) — what I call the 'finististic prejudice'."

Gödel replied on 12 October ([2003a: 423ff]) at length, patiently and respectfully, and urged a reading of his paper beyond the introductory sketch to see that his provability predicate is indeed expressible in the system. Gödel painstakingly pointed out that $\overline{[R(n);n]}$ has no meaning and that Zermelo

[108] See Moore [2002] for more on this "controversy" and Taussky–Todd [1987: 37ff] for an affecting account of this one encounter between Zermelo and Gödel.

must mean "$[R(n); n]$ is not *correct* [*richtig*]". This should be rendered

$$\overline{W}[R(n); n]$$

where $W(x)$ (W presumably for Wahrheit [truth]) is to mean "x is a formula that expresses a true assertion." *This* is a unary predicate, and if it were expressible in the system then one would have a genuine contradiction as Zermelo pointed out.

This to and fro is striking on two counts. First, that Zermelo shifted to the improper $\overline{[R(n); n]}$, leaving implicit that it is to convey a truth value, is focally symptomatic of his conflation, or at least willful assimilation, of the syntactic and semantic in his development of logic. Second, with his clear intent of predicating truth Zermelo had come to the argument for what is now known as Tarski's Undefinability of Truth.[109]

Concerning the first, it should be noted that the difference between truth and provability was by no means as clear as it became after Gödel's work. Although we now readily refer to Gödel's "true but unprovable" propositions, he himself in his [1931] emphasized their "formal undecidability" and when referring to their actual status never used the word "true" ["wahr"] but only "correct" ["richtig"]. This may have been due to caution in the face of the Hilbert school's emphasis on provability and consistency as opposed to outright truth[110] but also because of Gödel's participation in the Vienna Circle, which strongly opposed metaphysics. Only after a couple of years did Gödel, in lectures, start to use "true" to describe his propositions.

Concerning the undefinability of truth in a formal system, Gödel maintained to Hao Wang [1996: 82] that he had established it in the summer of 1930, this having been the first observation of his numerical coding of formulas. (Gödel first announced the existence of formally undecidable propositions on 6 September 1930 at a conference in Königsberg.[111]) The undefinability of truth is commonly attributed to Tarski [1933: §5, thm. 1], and Gödel's first public account of it is in his lecture [1934: §7]. Be that as it may, the undefinability of truth seems to have been *communicated* first in this exchange between Zermelo and Gödel. Gödel toward the end of his reply actually pointed out that the undefinability of truth leads to a quick proof of incompleteness: The class of provable formulas *is* definable and the class of correct [richtig] formulas is not, and so there must be a correct but unprovable formula.[112]

[109]See Murawski [1998] and Krajewski [2004] for the interactions of Tarski and Gödel on the undefinability of truth.

[110]Cf. Feferman [1984].

[111]See Gödel [1986: 196ff].

[112]As Gödel pointed out, this proof "furnishes no construction of the undecidable proposition and is not intuitionistically unobjectionable"; he of course constructed a specific undecidable proposition.

Toward the end of his reply Gödel addressed Zermelo's remark about the "finitistic prejudice":

> That one can *not* capture all of mathematics in one formal system already follows according to Cantor's diagonal procedure, but nevertheless it remains conceivable that one could at least formalize certain subsystems of mathematics completely (in the syntactic sense). My proof shows that that is also impossible if the subsystem contains at least the concepts of addition and multiplication of whole numbers. ... To be sure, the relatively undecidable propositions are always decidable in higher systems, to which I have also expressly alluded in my paper (cf. p. 191, footnote 48a); but even in those higher systems undecidable propositions of the same kind remain, and so on ad infinitum.

Footnote 48a of Gödel's [1931], *sans* contextualizing remarks, is as follows:

> ... 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 decidable 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.

This prescient note would be an early indication of a steady intellectual progress on Gödel's part that would take him from the Incompleteness Theorem through pivotal relative consistency results for set theory via the constructible universe to speculations about large cardinal axioms. Much has been made of this footnote;[113] that Gödel mentioned it indicates that he himself put much store in it. The footnote is understood as asserting that first-order satisfaction for a formal system becomes definable in a "higher type" over the system, although "truth" and satisfaction would only be systematically discussed in Tarski [1933]. Gödel concluded his reply by thanking Zermelo for his [1930] paper, mentioning that he had already read it soon after its appearance and suggesting that he could impart his thoughts about it to Zermelo. Alas, this would not happen; Gödel's remarks on [1930] would have been illuminating, particularly if they were to bear on connections between his footnote 48a and Zermelo's approach through his Grenzzahlen.

In what turned out to be the final letter (Gödel [2003a: 430ff]) between them, Zermelo wrote back to Gödel on 29 October that he could now infer what Gödel had meant about provable propositions under a "finitistic restriction". However, Zermelo wrote of having an uncountable collection of propositions only countably many of which are provable, so that there must be undecidable propositions. Moreover, Zermelo skewed Gödel's "higher

[113] See e.g., Kreisel [1980: 183, 195, 197], a memoir on Gödel, and Feferman [1987], where the view enunciated in the footnote is elevated to "Gödel's doctrine".

systems" as not distinguishable by having new propositions but merely new means of proof. Zermelo continued:

> ... what a "proof" really is is itself not in turn provable, but must in some form be *taken for granted, presupposed* [angenommen, vorausgesetzt]. And here it is just a question: What does one understand by a proof? Quite generally, one understands by this a system of propositions such that under the assumption of the premises, the validity of the assertion can be made *evident* [*einsichtig*]. And there remains only the question, what counts as "evident"? In any case, *not merely* — and that you yourself show precisely — the propositions of some finitistic schema, which in your case too can always be *extended*.

Zermelo was evidently not taking in Gödel's proof, but impatient, was proceeding to broach themes of his approach through infinitary logic, in which there are to be uncountably many propositions, what a proof *is* is to be taken for granted, and the only remaining question as to proof is what counts as evident [einsichtig].

What would unify both Zermelo's approach to set theory and to logic would be the reliance on well-founded relations. Zermelo had already written on 7 October 1931 to his colleague Reinhold Baer (Weingartner-Schmetterer [1987: 45ff]):

> I believe I have at last found in my "*Foundation Principle*" ["Fundierungsprinzip"] the right instrument for explaining whatever is in need of elucidation. But nobody understands it, just as nobody has yet reacted to my Fundamenta article [1930] — not even my good friends in Warsaw.[114]

Well-founded relations would become central to set theory, but Zermelo's "logical" proposals involving them would be forgotten.

Zermelo's published account [1932] of his Bad Elster talk (see also the related [1932a]) described his transfinite, well-founded proposition systems [Satzsysteme]. But [1932] was also a manifesto of sorts that railed against the "finitistic prejudice", affirming:

> ... the true subject matter of mathematics is *not*, as many would have it, "combinations of signs" but *conceptually-ideal relations* [*begrifflich-ideale Relationen*] between the elements of a conceptually determined *infinite manifold* [*Mannigfaltigkeit*]. And thus are our systems of signs only devices, forever *incomplete* and shifting from case to case, that help our *finite* understanding to at least approximate step by step mastery of the infinite, which we cannot *immediately* and *intuitively* "survey" or grasp.

[114]This is revealing of the lack of response to Zermelo [1930]. In this letter Zermelo carried on unguardedly and revealingly: "At Bad Elster I avoided any direct polemic against Gödel both in the lecture itself and afterward: one should not frighten off enterprising beginners." "... the gentlemen will *have* to declare their hand finally when I publicly assert that Gödel's much-admired 'proof' is nonsense [Unsinn]; ... " Referring to his letter of 21 September: "He [Gödel] still has *not* answered my letter. Apparently he has nothing else to grouse [meckern] about."

Zermelo went on to criticize the relativism of Gödel's approach, though ironically Gödel, at least according to his subsequent writings, would not have disagreed with what Zermelo wrote above. For Gödel the incompleteness of formal systems is a crucial *mathematical* phenomenon to be reckoned with but also to transcend, whereas Zermelo, not having appreciated that logic has been submerged into mathematics, insisted on an infinitary logic that directly reflected transfinite reasoning.

Several years would pass before Zermelo [1935] provided more details for his theory of propositions, in what would turn out to be his last mathematical publication.[115] Zermelo began, significantly, by introducing the general notion of well-founded relation and its crucial stratification property.[116] A binary irreflexive relation $\prec$ on an X is *well-founded* exactly when every non-empty $Y \subseteq X$ has an $\prec$-minimal element, and such a relation provides a stratification of X through a rank function $\rho\colon X \longrightarrow$ ordinals given by: $\rho(x) = \sup\{\rho(y) + 1 \mid y \prec x\}$.[117] The Axiom of Foundation is just the assertion that the membership relation on sets is itself well-founded, and the cumulative hierarchy is the resulting stratification.

Zermelo [1935] then introduced his expansive concept of proposition [Satzsystem], positing that the natural component relation be well-founded.[118] Viewing the situation syntactically through modern eyes and presuming a formal language, this would seem to be *de rigueur* for parsing formulas, but Zermelo was working directly with propositions, assertions. This led to the concept of a well-founded hierarchy of propositions, in evident analogy to the [1930] cumulative hierarchies of sets. The bases this time are to consist of collections Q of elementary propositions; an important case is to have a domain D of elements and a collection of (finitary) relations and to take Q as consisting of all propositions $Ra_1, \dots, a_n$ for R an n-ary relation and $a_1, \dots, a_n \in D$. The successive levels $P^Q_{\alpha+1}$ are then to consist of negations $\neg\mathcal{A}$ for $\mathcal{A} \in P^Q_\alpha$ and conjunctions $\wedge\mathcal{K}$ and disjunctions $\vee\mathcal{K}$ for $\mathcal{K} \subseteq P^Q_\alpha$, in full exercise of the power set operation and as a general way of rendering quantification. As with the cumulative hierarchies of sets well-foundedness served to establish the universality of the hierarchical schemes and so the generality of inductive arguments.

[115]This despite "(First Communication) [(Erste Mitteilung)]" in the title. Hitler effected the reincorporation of the Saar on 15 January 1935. Zermelo refused to give the Hitler salute and was soon debarred from teaching at the University of Freiburg.

[116]This however was not the first time that well-founded relations were formulated. Well-founded relations were crucial for the descriptive set theory analysis of analytic sets in Luzin–Sierpiński [1918, 1923], and well-founded relations on the natural numbers were explicitly formulated in the systematic presentation of analytic sets in Luzin [1927: 50].

[117]Assume to the contrary that some member of X eludes this ranking. By well-foundedness let $x \in X$ be $\prec$-minimal with this property. But $\rho(x)$ is then defined after all, a contradiction.

[118]Taylor [2002] provides a detailed reconstruction of Zermelo's theory, and our account here follows his notationally.

With every proposition hierarchically analyzed Zermelo's notion of proof is a simple one. Any truth assignment for a basis Q of elementary propositions recursively induces a truth assignment through the hierarchy of the P^Q_α's. A proof of an implication $\wedge\mathcal{K} \rightarrow \mathcal{A}$, where $\mathcal{A}$ and every member of $\mathcal{K}$ is in P^Q_α, is then just an affirmation that any truth assignment for Q that assigns every member of $\mathcal{K}$ true also renders $\mathcal{A}$ true. Thus, a proof for Zermelo is much like an affirmation of satisfiability in a structure (interpretation of a formal language) for Tarski, except that on the one hand there is no explicit engagement with quantifiers, but on the other, the recursive definition is transfinite. With this simple view of proof, what we would now regard as a conflation of syntax with semantics, every proposition is either provable or else refutable with a counterexample truth assignment, and this conforms with an earlier declaration in Zermelo [1932a] that every mathematical proposition is decidable.

Zermelo had written in his second letter to Gödel that what a proof is "must in some form be *taken for granted, presupposed* [angenommen, vorausgesetzt]". But Zermelo had also written that the only remaining question as to proof is "what counts as 'evident' ['einsichtig']". Zermelo [1935: 144ff] wrote toward the end:

> ... A "proof" will contain *infinitely many* intermediate propositions in most cases, and it is not yet clear to what extent and through what auxiliary means it can be made self-evident [einleuchtend] to our *finite* understanding. At root, *any* mathematical proof, e.g., proof by complete induction, is thoroughly infinitary, and nonetheless we are capable of grasping it. On the face of it, there seems to be no fixed bounds to intelligibility.

Actually, proof by mathematical induction is not today considered "thoroughly infinitary" but paradigmatic in both being finitary and establishing universal propositions about infinite domains. In a sense, Zermelo in his final efforts had to face a basic problem with his semantic notion of proof which was solved for first-order logic by Gödel by his Completeness Theorem: The logical consequences though analyzable may not be surveyable, but if they are shown to be provable consequences according to finitary rules of proof, then they become recursively enumerable.

§8. Toward categoricity. What does Zermelo's infinitary logic have to do with his set theory? With his correlation of his cumulative hierarchies of propositions with those of sets, it is evident that Zermelo had in mind a joint development, and in fact a reduction of his infinitary logic to his cumulative hierarchy of sets.[119] Ebbinghaus [2003] in his penetrating analysis

[119]This is brought out by Taylor [2002], who also emphasizes the involvement of the Grenzzahlen, inaccessible cardinals. Zermelo [1935: 141ff] had written: "Proceeding upward, a well-stratified system of propositions can be closed off at will, e.g., at a well-defined 'Grenzzahl' π (cf. Zermelo [1930]), and then possess all the attributes of a set."

of Zermelo's concept of definite property, using items from his *Nachlass*, describes two shifts in Zermelo's thinking prompted by Skolem [1930]: First, though Zermelo [1929] had espoused a formulation of definite property in a finitary, albeit second-order, language, he thereafter pursued an approach through infinitary languages. Second, in terms of infinitary logic Zermelo pursued the possibility that every set is "categorically" definable, definable in a categorical system of axioms, and that there is such system for set theory, this in contradistinction to the non-categorical ZF$'$ of Zermelo [1930] with its procession of natural models. The infinitary logic would presumably be as developed in [1935] in correlation with [1930], but the tension between having a static universe to be given by a categorical axiom system and the [1930] "free unfolding" of models would not be resolved. Ebbinghaus [2003] describes how Zermelo at the end of his work in set theory was close to the recursive procedure for defining Gödel's constructible universe L, also formulated by 1935.[120] To expand on this point and generally to compare and contrast, we describe Gödel's separate progress in set theory as brought out by his lectures:

In his lecture [1933], Gödel expanded on the theme of footnote 48a from his [1931]. He propounded the view that the axiomatic set theory of Zermelo, Fraenkel, and von Neumann is "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, instead of having separate types with sets of type $n+1$ consisting purely of sets of type n, sets can be *cumulative* in the sense that sets of type n can consist of sets of *all* lower types. If S_n is the collection of sets of type n, then: S_0 is the type of the "individuals", and recursively, $S_{n+1} = S_n \cup \{X \mid X \subseteq S_n\}$. Second, the process can be continued into the transfinite, starting with the cumulation $S_\omega = \bigcup_n S_n$, proceeding through successor stages as before, and taking unions at limit stages. Gödel [1933: 46] credited Hilbert for pointing out the possibility of continuing the formation of types beyond the finite types. As for how far this cumulative hierarchy of sets is to continue, the "first two or three types already suffice to define very large ordinals" ([1933: 47]) which can then serve to index the process, and so on. Gödel observed that although this process has no end, this "turns out to be a strong argument in favor of the theory of types" ([1933: 48]). 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" ([1933: 48]). Thus, although he never mentioned Zermelo [1930], Gödel was thus entertaining its cumulative hierarchies, but as motivated by the theory of types. On the other hand, while never getting to Zermelo's Grenzzahlen Gödel emphasized the definability, both of ordinals

[120] Gödel told von Neumann about L in 1935; see e.g., Dawson [1997: 122ff].

at low types for the indexing of higher types and of propositions unprovable at one type but becoming provable in the next higher type.

Modern set theory was launched by Gödel's formulation of the model L of "constructible" sets, a model of set theory that established the relative consistency of the Axiom of Choice and the (Generalized) Continuum Hypothesis. In his first announcement Gödel [1938: 556] described L 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 described in his [1933] 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 to get a model of set theory. The extent of the ordinals was highlighted in his monograph [1940], based on lectures in 1938, in which he formally generated L set by set using a sort of Gödel numbering in terms of ordinals. In his [1939], Gödel presented the constructible hierarchy L essentially as it is presented today:

$$M_0 = \{\emptyset\};\ M_\beta = \bigcup_{\alpha<\beta} M_\alpha \text{ for limit ordinals } \beta;\ \text{and } M_{\alpha+1} = M'_\alpha,$$

where M' is "the set of subsets of M defined by propositional functions $\phi(x)$ over M," these propositional functions having been precisely defined. Gödel's construction affirmed 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 construction, schematized above in terms of the definable def operation, via transfinite recursion. As for Foundation, underlying the construction was the well-foundedness of sets, and significantly, footnote 12 of [1939] revealed that Gödel viewed his axiom A, that every set is constructible (now written $V = L$ following Gödel [1940]), as deriving its contextual sense from the cumulative hierarchy of sets regarded as an extension of the simple theory of types: "In order to give A 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)."

Gödel in his lecture [1939a] motivated L by referring explicitly to Russell's ramified theory of types. Gödel first described what amounts to the orders of that theory for the simple situation when the members of a countable collection of real numbers are taken as the "individuals" and new real numbers are successively defined via quantification over previously defined real numbers, and emphasized that the process can be continued into the transfinite. He then observed Gödel [1939a: 135] that this procedure can be applied to sets of real numbers, and the like, as "individuals", and moreover, that one can "intermix" the procedure for the real numbers with the procedure for sets of real numbers "by using in the definition of a real number quantifiers that refer to sets of real numbers, and similarly in still more complicated ways."

Gödel called a *constructible* set "the most general [object] that can at all be obtained in this way, where the quantifiers may refer not only to sets of real numbers, but also to sets of sets of real numbers and so on, *ad transfinitum*, and where the indices of iteration ... can also be arbitrary transfinite ordinal numbers". Gödel considered that although this definition of constructible set might seem at first to be "unbearably complicated", "the *greatest generality yields,* as it so often does, at the same time the *greatest simplicity*" ([1939a: 137]). Gödel was picturing Russell's ramified theory of types by first disassociating the types from the orders, with the orders here given through definability and the types represented by real numbers, sets of real numbers, and so forth. Gödel's intermixing then amounted to a recapturing of the complexity of Russell's ramification, the extension of the hierarchy into the transfinite allowing for a new simplicity.

Gödel went on to describe the universe of set theory, "the objects of which set theory speaks", as falling into "a transfinite sequence of Russellian [simple] types" ([1939a: 137]), the cumulative hierarchy of sets that he had described in [1933] and referred to in that footnote 12 of [1939]. He then formulated the constructible sets as an analogous hierarchy, the hierarchy of [1939], in effect introducing the Russellian ramifying orders through definability. In a comment bringing out the intermixing of types and orders, Gödel pointed out that "there are sets *of lower type* that *can* only *be defined* with the help of *quantifiers for sets of higher type*" ([1939a: 141]).

In his monograph [1940] Gödel had provided a formal presentation of L using an axiomatization of set theory with an antecedent in von Neumann [1925]. Gödel's formalization not only recalled von Neumann's [1925: II] analysis of "subsystems", but also shed light on von Neumann's main concern: the categoricity of his axiomatization. Recall (see §5 above) that Fraenkel [1922] had sought to close off the Zermelian generative axioms through an "axiom of restriction"; it was to pursue this that von Neumann had investigated subsystems for his axiomatization, but he concluded that there was probably no way to *formally* achieve Fraenkel's idea of a minimizing, and hence categorical, axiomatization. Gödel's axiom A, that every set is constructible, can be viewed as formally achieving this sense of categoricity, since, as he essentially showed in [1940], in axiomatic set theory L is a definable class, containing all the ordinals, that together with the membership relation restricted to it is a model of set theory, and L is a submodel of every other such class.

What are the points of contact, the similarities in approach and attitude, between Zermelo and Gödel? Both Zermelo and Gödel embraced direct transfinite reasoning, for Gödel actually viewed L as an outright construction in metamathematics.[121] Gödel latterly wrote in a letter of 7 March 1968 to Wang [1974: 8ff]: "there was a special obstacle which *really* made

[121] See Wang [1974: 8ff].

it *practically impossible* for constructivists to discover my consistency [of the Continuum Hypothesis] proof. It is the fact that the ramified hierarchy, which had been invented *expressly for constructive purposes,* had to be used in an *entirely nonconstructive way.*" This nonconstructive way was to prolong the ramified hierarchy using arbitrary ordinal numbers for the indexing, and for this the extent of the ordinal numbers as sets, the von Neumann ordinals, had to be sustained by Replacement. It is known from the Zermelo–Gödel correspondence that Gödel had read Zermelo [1930] soon after its appearance, and Replacement plays a central role there as we have seen.[122] In his first published sketch of L, Gödel [1939] stated his relative consistency results in terms of the axioms of Zermelo [1908] with definiteness as formalized in first-order logic, and although Zermelo [1930] is never cited, his Grenzzahlen make a cameo appearance when Gödel in his main statement of formal consistency asserted that M_Ω, where Ω is "the first inaccessible number", is a model of Zermelo's axioms together with Replacement.

Zermelo was stalking a categorical axiom system, and Gödel achieved one with $V = L$. However, that categoricity was achieved through minimization. In later years Gödel advocated a search for new axioms in a maximal direction antithetical to $V = L$, particularly axioms of large cardinal character. As described in §7, both Zermelo and Gödel urged venturing into higher systems to overcome inadequacies, and with Zermelo's ready incorporation of a proper class of inaccessible cardinals in his [1930] he would not have been unsympathetic to maximal approaches. Be that as it may, Zermelo in his efforts with his infinitary logic never even succeeded in carrying out a mathematical construction of any categorical domain of sets.

What was missing in Zermelo's approach? There are two salient gaps, one following from the other: Zermelo did not appreciate the mathematical importance of uninterpreted formalism as Hilbert did, particularly the operational efficacy of first-order logic, and so Zermelo could not take in the submergence of metamathematical methods into mathematics in the work of Gödel.[123]

[122]Kreisel in his memoir [1980: 190ff] on Gödel stressed the "non-elementary" character of the axioms of Zermelo [1930] and its cumulative hierarchy picture, attributed comparative significance to it with respect to Gödel's "thin" hierarchy of constructible sets, and suggested how Replacement had a growing importance in Gödel's thinking. The comparison between the Zermelo and Gödel hierarchies mostly serves a didactic purpose; Gödel, as described above, saw himself as extending Russell's ramified theory of types.

[123]In this Zermelo was certainly not alone. Russell wrote in a letter of 1 April 1963 to Leon Henkin: "It is fifty years since I worked seriously at mathematical logic and almost the only work that I have read since that date is Gödel's. I realized, of course, that Gödel's work is of fundamental importance, but I was puzzled by it. It made me glad that I was no longer working at mathematical logic. If a given set of axioms leads to a contradiction, it is clear that at least one of the axioms must be false. Does this apply to school-boys' arithmetic, and, if so, can we believe anything that we were taught in youth? Are we to think that $2 + 2$ is not 4, but 4.001? Obviously, this is not what is intended. ...

From early on Zermelo was exercised by Richard's Paradox, and how it retained its force for him is symptomatic. Richard's Paradox,[124] it will be recalled, is a definability version of Cantor's diagonal argument: Enumerate the finite strings of the 26 letters, retain only those that define real numbers, and let $u_1, u_2, u_3, \ldots$ be the enumeration of the corresponding real numbers, an enumeration therefore containing all real numbers definable by finitely many words. Now consider the real number with integral part 0 and nth digit in the decimal expansion the sequent of nth digit in the decimal expansion of u_n, where for $i < 9$ the sequent of i is $i + 1$ and for $i = 9$ the sequent is 0. This real number is thus definable, yet cannot be any u_n.

Just after introducing his Separation Axiom Zermelo [1908a: 264] aptly described how separating subsets from sets already given excludes contradictory notions like the set of all sets, but also how with definiteness "all criteria such as 'definable by means of a finite number of words', hence the 'Richard Paradox' ... vanish." Thus, definiteness was also to restrict conceptual resources in the formation of sets.[125] The collection $E = \{u_1, u_2, u_3, \ldots\}$ cannot be separated out from the real numbers, as definability in a natural language is not a definite property.[126]

Gödel [1931] wrote in his introductory sketch about his main line of argument: "The analogy of this argument with the Richard antinomy leaps to the eye." As described in the discussion of the Zermelo–Gödel correspondence, Zermelo misconstrued Gödel's argument and then averred: "Just as in the Richard and Skolem paradoxes, the mistake rests on the (erroneous) assumption that every mathematically definable notion is expressible by a 'finite combination of signs' (according to a *fixed* system!) ..."[127] In particular, Zermelo no longer took the thrust of Richard's Paradox to be the mathematical inaccessibility of the collection of definable real numbers, but having acceded to "combinations of signs" in "systems", uninterpreted formalisms, Zermelo saw the issue now to be the inadequacy of any one such finitary system for expressing every "mathematically definable notion." The

---

You note that we [Whitehead and Russell] were indifferent to attempts to prove that our axioms could not lead to contradictions. In this, Gödel showed that we had been mistaken. But I thought that it must be impossible to prove that any given set of axioms does not lead to a contradiction, and, for that reason, I had paid little attention to Hilbert's work."

[124] See Richard [1905]; for a discussion of Richard's Paradox contemporary with Zermelo's late work see Church [1934].

[125] Taylor [1993: 547ff] emphasizes this aspect of Separation and puts it on an equal footing with "limitation of size".

[126] Richard [1905] himself diagnosed the problem to be the taking of E as "totally defined" and this diagnosis was endorsed by Poincaré [1906a: VII] presumably because of the impredicativity.

[127] In that unguarded letter of 7 October 1931 to Reinhold Baer, Zermelo wrote (Weingartner-Schmetterer [1987: 45ff]): "The question of the antinomy of Richard and the Skolem doctrine *must* at last be discussed *seriously*, seeing that frivolous dilettantism is again at work to discredit the whole area of research ...."

connection with Cantor's diagonal argument in the role of defining a new real number is now more immediate, and Zermelo sees this as supporting his accommodation of uncountably many propositions via infinitary languages.

However Zermelo weighted Richard's Paradox he remained gripped by it and skeptical about definability cast as a mathematical concept. Tarski and Gödel wove first-order definability into mathematics, the former already in his [1931] formulating the definable sets of reals, and the latter proceeding from footnote 48a of his [1931] to the definition of L. In his monograph [1940] Gödel took pains to define definability, as we might now say, by proceeding in a weak second-order system that in effect provided a finite axiomatization of Separation. This formal presentation left no doubt about the construction and results, but obscured the metamathematical ideas. In later years, L would come to be presented more directly, as it was by Gödel in his [1939], in terms of first-order definability formalized in ZFC. Gödel [1940a: 176] wrote:

> One may at first doubt that this assertion [A, that every set is constructible] has a meaning at all, because A is apparently a metamathematical statement since it involves the manifestly metamathematical term "definable" or "constructible". But now it has been shown in the last few years how metamathematical statements can be translated into mathematics, and this applies also to the notion of constructibility and the proposition A, so that its consistency with the axioms of mathematics is a meaningful assertion.

We have now converged to the heart of the matter. *Zermelo aimed at categorical definability through direct engagement with transfinite reasoning. Gödel took the ordinal numbers as given and used them to index a hierarchy based on first-order definability.* He was able to carry out a *mathematical* construction by building on the Hilbertian focus on first-order logic and mathematization of its metamathematics and on the Russellian type hierarchy as ramified by orders through definability.

What, finally, about Zermelo's definite property for Separation? Through his [1929], with its second-order axiomatization of definite property, Zermelo bolstered the latent sense of his [1908a] axioms as specifying the domain of sets through algebraic closure. From his [1930] on, faced with "Skolemism" Zermelo took an expansive approach to definiteness that soon became based on an infinitary logic, but unlike much of his previous work this approach had no accompanying local closure feature. With definability made mathematical a new closure became possible: Gödel formulated L in [1940] explicitly as a closure of the ordinals based on a few "fundamental operations" and in [1939] more structurally as levels of a cumulative hierarchy with successive levels given by definable closure. Moreover, in a twist of fate the Skolem Paradox argument, that bugbear of Zermelo's, participated in a new rectification of Separation through a closure argument in Gödel's context:

Referring to his result that every constructible real appears by the ω_1th level M_{ω_1}, a result that has as an immediate consequence that the Continuum Hypothesis holds in L, Gödel [1940a: 178] wrote:

> You will see that this theorem is actually nothing else but an axiom of reducibility for transfinite orders, for it says that an arbitrary propositional function with integers as arguments is always formally equivalent with a propositional function of an order $< \omega_1$. So since an axiom of reducibility holds for constructible sets it is not surprising that the axioms of set theory hold for the constructible sets, because the axiom of reducibility or its equivalents, e.g., Zermelo's Aussonderungsaxiom [Separation Axiom], is really the only essential axiom of set theory.

The argument for the Continuum Hypothesis holding in L has as a significant feature a Skolem hull or closure argument, somewhat obscured in the formal presentation of Gödel [1940] but common fare nowadays in model theory.[128] The passage shows Gödel to be holding a remarkably synthetic, unitary view, equating as he does Russell's ill-fated Axiom of Reducibility in his type theory with Zermelo's Separation Axiom in set theory and regarding the latter, in the end, as the "the only essential axiom of set theory."

Lives, especially creative, struggling ones, can often be summed up with ironic juxtapositions. Zermelo did indeed transform the set theory of Cantor and Dedekind through the axiomatic incorporation of generative operations and new principles, and in the process drew out principles as specifically set-theoretic out of the presumptively logical. But in establishing an operational basis for set theory and promoting a predisposition for the set-theoretic analysis of a wide range of conceptualizations Zermelo provided the preconditions for the incisive work of Gödel and Tarski, the mathematization of metamathematics and the formalization of satisfiability. The latter's "definition of truth" is but a simple recursive definition in set-theoretic terms, a mapping of truth into set theory. And now is the irony:

Zermelo himself did not take the linguistic turn, in that he did not develop an uninterpreted formalism. Having separated out set theory from logic he did not appreciate, perhaps could not take in, the reincorporation of logic back into set theory. The important advances in the subsequent development of set theory, starting with Gödel's L, would be based on the definability of definability, of satisfiability in set-theoretic structures. Zermelo was first to surpass Cantor, but Zermelo himself was surpassed by Gödel.

REFERENCES

[1975] Jon Barwise, ***Admissible sets and structures***, Springer-Verlag, Berlin, 1975.

[128] Gödel in his publications does not credit Skolem, but Kreisel [1980: 199] wrote: "As he [Gödel] mentioned in conversation, the idea of the sort of argument involved, occurred to him when he learnt the Skolem argument as a student."

[1995] John L. Bell, *Type reducing correspondences and well-orderings: Frege's and Zermelo's constructions re-examined*, ***The Journal of Symbolic Logic***, vol. 60 (1995), pp. 209–221.

[1901] Felix Bernstein, *Untersuchungen aus der Mengenlehre*, 1901, inaugural dissertation, Göttingen, printed in Halle, reprinted in ***Mathematische Annalen***, vol. 61 (1905), pp. 111–155.

[1997] George Boolos, *Constructing Cantorian counterexamples*, ***The Journal of Philosophical Logic***, vol. 26 (1997), pp. 237–239, reprinted in [1998] below, pp. 339–341.

[1998] ———, ***Logic, logic, and logic*** (Richard Jeffrey, editor), Harvard University Press, Cambridge, 1998.

[1898] Emile Borel, ***Leçons sur la théorie des fonctions***, Gauthier-Villars, Paris, 1898.

[1921] ———, *La théorie du jeu et les équations intégrales á noyau symétrique*, ***Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris***, vol. 173 (1921), pp. 1304–1308, reprinted in [1972] below, pp. 901–904, translated in ***Econometrica***, vol. 21 (1953), pp. 97–100.

[1972] ———, ***Oeuvres de Emile Borel***, Editions de C.N.R.S., Paris, 1972.

[1956] Nicolas Bourbaki, ***Eléments de mathématique. I. Théorie des Ensembles. Chapter III: Ensembles ordonnés, cardinaux, nombres entiers***, Actualités Scientifiques et Industrielles #1243, Hermann, Paris, 1956.

[1891] Georg Cantor, *Über eine elementare Frage der Mannigfaltigkeitslehre*, ***Jahresbericht der deutschen Mathematiker-Vereinigung***, vol. 1 (1891), pp. 75–78, reprinted in [1932] below, pp. 278–280, translated in Ewald [1996], vol. 2, pp. 920–922.

[1895] ———, *Beiträge zur Begründung der transfiniten Mengenlehre. I*, ***Mathematische Annalen***, vol. 46 (1895), pp. 481–512, translated in [1915] below, reprinted in [1932] below, pp. 282–311.

[1897] ———, *Beiträge zur Begründung der transfiniten Mengenlehre. II*, ***Mathematische Annalen***, vol. 49 (1897), pp. 207–246, translated in [1915] below, reprinted in [1932] below, pp. 312–351.

[1915] ———, ***Contributions to the founding of the theory of transfinite numbers***, including translations of [1895] and [1897] above with introduction and notes by Philip E. B. Jourdain, Open Court, Chicago, 1915, reprinted by Dover, New York, 1965.

[1932] ———, *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, (Ernst Zermelo, editor), Julius Springer, Berlin, 1932, reprinted by Springer-Verlag, Berlin, 1980.

[1934] Alonzo Church, *The Richard paradox*, ***American Mathematical Monthly***, vol. 41 (1934), pp. 356–361.

[1997] John W. Dawson, Jr., ***Logical dilemmas. The life and work of Kurt Gödel***, A.K. Peters, Cambridge, 1997.

[1888] Richard Dedekind, ***Was sind und was sollen die Zahlen?***, F. Vieweg, Braunschweig, 1888, sixth, 1930 edition reprinted in [1932] below, pp. 335–390, second, 1893 edition translated in [1963] below, pp. 29–115, third edition translated with commentary in Ewald [1996], vol. 2, pp. 787–833.

[1932] ———, *Gesammelte mathematische Werke*, (Robert Fricke, Emmy Noether, and Öystein Ore, editors), vol. 3, F. Vieweg, Braunschweig, 1932, reprinted by Chelsea, New York, 1969.

[1963] ———, ***Essays on the theory of numbers***, translations by Wooster W. Beman, Dover, New York, 1963, (reprint of original edition, Open Court, Chicago, 1901).

[1997] Burton Dreben and Akihiro Kanamori, *Hilbert and set theory*, ***Synthèse***, vol. 110 (1997), pp. 77–125.

[2003] Heinz-Dieter Ebbinghaus, *Zermelo: Definiteness and the universe of definable sets*, ***History and Philosophy of Logic***, vol. 24 (2003), pp. 197–219.

[1996] William Ewald (editor), ***From Kant to Hilbert: A source book in the foundations of mathematics***, Clarendon Press, Oxford, 1996, in two volumes.

[1984] Solomon Feferman, *Kurt Gödel: Conviction and caution*, ***Philosophia Naturalis***, vol. 21 (1984), pp. 546–562, reprinted in [1998] below, pp. 150–164.

[1987] ———, *Infinity in mathematics: Is Cantor necessary?*, ***L'infinito nella Scienza*** (G. Toraldo di Francia, editor), Instituto della Enciclopedia Italiana, Roma, 1987, pp. 151–209, reprinted in [1998] below, particularly pp. 229–248.

[1998] ———, ***In light of logic***, Oxford University Press, New York, 1998.

[1999] José Ferreirós, ***Labyrinth of thought: A history of set theory and its role in modern mathematics***, Birkhauser Verlag, Basel, 1999.

[1921] Abraham Fraenkel, *Über die Zermelosche Begründung der Mengenlehre*, ***Jahresbericht der deutschen Mathematiker-Vereinigung***, vol. 30 (1921), pp. 97–98.

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

[1922a] ———, *Über den Begriff "definit" und die Unabhängigkeit des Auswahlaxioms*, ***Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse***, 1922, pp. 253–257, translated in van Heijenoort [1967], pp. 284–289.

[1925] ———, *Untersuchungen über die Grundlagen der Mengenlehre*, ***Mathematische Zeitschrift***, vol. 22 (1925), pp. 250–273.

[1926] ———, *Axiomatische Theorie der geordneten Mengen (Untersuchungen über die Grundlagen der Mengenlehre II)*, ***Journal für die reine und angewandte Mathematik***, vol. 155 (1926), pp. 129–158.

[1953] ———, ***Abstract set theory***, North Holland, Amsterdam, 1953.

[1958] Abraham Fraenkel and Yehoshua Bar-Hillel, ***Foundations of set theory***, North Holland, Amsterdam, 1958.

[1906] Maurice Fréchet, *Sur quelques points du calcul fonctionnel*, ***Rendiconti Circolo Matematico di Palermo***, vol. 22 (1906), pp. 1–74.

[1879] Gottlob Frege, ***Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens***, Nebert, Halle, 1879, reprinted by Olms, Hildesheim 1964, translated in van Heijenoort [1967], pp. 1–82.

[1884] ———, ***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.

[1893] ———, ***Grundgesetze der Arithmetik, Begriffsschriftlich abgeleitet***, vol. 1, Hermann Pohle, Jena, 1893, reprinted by Olms, Hildesheim 1962.

[1903] ———, ***Grundgesetze der arithmetik, Begriffsschriftlich abgeleitet***, vol. 2, Hermann Pohle, Jena, 1903, reprinted by Olms, Hildesheim 1962.

[1976] ———, ***Wissenschaftlicher Briefwechsel*** (Gottfried Gabriel et alii, editors), Felix Meiner Verlag, Hamburg, 1976.

[1980] ———, ***Philosophical and mathematical correspondence*** (Gottfried Gabriel et alii, editors), The University of Chicago Press, Chicago, 1980, abridged from [1976] above by Brian McGuinnes and translated by Hans Kaal.

[1953] David Gale and Frank M. Stewart, *Infinite games with perfect information*, ***Contributions to the theory of games*** (Harold W. Kuhn and Alan W. Tucker, editors), vol. 2, Annals of Mathematics Studies, no. 28, Princeton University Press, Princeton, 1953, pp. 245–266.

[1998] Alexander George and Daniel J. Velleman, *Two conceptions of natural number*, ***Truth in mathematics*** (H. G. Dales and G. Oliveri, editors), Clarendon Press, Oxford, 1998, pp. 311–327.

[1905] J. Willard Gibbs, ***Elementare Grundlagen der statistischen Mechanik***, Barth, Leipzig, 1905, Ernst Zermelo's German translation of Gibbs's ***Elementary Principles of Statistical Mechanics***, Charles Scribner, New York, 1902.

[1930] Kurt Gödel, *Die vollständigkeit der Axiome des logischen Funktionenkalküls*, ***Monatshefte für Mathematik und Physik***, vol. 37 (1930), pp. 349–360, reprinted and translated in [1986] below, pp. 102–123.

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

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

[1934] ———, *On undecidable propositions of formal mathematical systems*, (mimeographed lecture notes, taken by Stephen C. Kleene and J. Barkley Rosser), 1934, reprinted with revisions in [1986] below, pp. 346–371.

[1938] ———, *The consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis*, ***Proceedings of the National Academy of Sciences U.S.A.***, vol. 24 (1938), pp. 556–557, reprinted in [1990] below, pp. 26–27.

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

[1939a] ———, *Vortrag Göttingen*, 1939, text and translation in [1995] below, pp. 126–155, and the page references are to these.

[1940] ———, ***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, pp. 33–101.

[1940a] ———, ***Lecture [on the] consistency [of the] continuum hypothesis***, Brown University, 1940, in [1995] below, pp. 175–185, and the page references are to these.

[1986] ———, ***Collected works*** (Solomon Feferman, editor), vol. 1, Oxford University Press, New York, 1986.

[1990] ———, ***Collected works*** (Solomon Feferman, editor), vol. 2, Oxford University Press, New York, 1990.

[1995] ———, ***Collected works*** (Solomon Feferman, editor), vol. 3, Oxford University Press, New York, 1995.

[2003] ———, ***Collected works*** (Solomon Feferman and John W. Dawson, Jr., editors), vol. 4, Oxford University Press, Oxford, 2003.

[2003a] ———, ***Collected works*** (Solomon Feferman and John W. Dawson, Jr., editors), vol. 5, Oxford University Press, Oxford, 2003.

[1988] Warren Goldfarb, *Poincaré against the logicists*, ***History and philosophy of modern mathematics*** (William Aspray and Philip Kitcher, editors), Minnesota Studies in the Philosophy of Science, vol. 11, University of Minnesota Press, Minneapolis, 1988, pp. 61–81.

[1977] Ivor Grattan-Guinness, ***Dear Russell – Dear Jourdain***, 1977, Duckworth & Co., London, and Columbia University Press, New York.

[2000] ———, ***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.

[1910] Kurt Grelling, *Die Axiome der Arithmetik mit besonderer Berücksichtigung der Beziehungen zur Mengenlehre*, ***Ph.D. thesis***, Göttingen, 1910.

[1984] Michael Hallett, ***Cantorian set theory and limitation of size***, Logic Guides #10, Clarendon Press, Oxford, 1984.

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

[1979] ———, ***Collected papers of G. H. Hardy*** (I. W. Busbridge and R. A. Rankin, editors), Clarendon, Oxford, 1979.

[1915] FRIEDRICH HARTOGS, *Über das Problem der Wohlordnung*, ***Mathematische Annalen***, vol. 76 (1915), pp. 436–443.

[1966] K. HAUSCHILD, *Bemerkungen, das Fundierungsaxiom betreffend*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 12 (1966), pp. 51–56.

[1904] FELIX HAUSDORFF, *Der Potenzbegriff in der Mengenlehre*, ***Jahresbericht der deutschen Mathematiker-Vereinigung***, vol. 13 (1904), pp. 569–571.

[1908] ———, *Grundzüge einer Theorie der geordneten Mengen*, ***Mathematische Annalen***, vol. 65 (1908), pp. 435–505.

[1914] ———, ***Grundzüge der Mengenlehre***, de Gruyter, Leipzig, 1914, reprinted by Chelsea, New York 1965.

[1995] RICHARD G. HECK, JR., *Definition by induction in Frege's Grundgesetze der Arithmetik*, ***Frege's philosophy of mathematics*** (William Demopoulos, editor), Harvard University Press, Cambridge, 1995, pp. 295–333.

[1906] GERHARD HESSENBERG, ***Grundbegriffe der Mengenlehre***, Vandenhoeck & Ruprecht, Göttingen, 1906, reprinted from ***Abhandlungen der Fries'schen Schule, Neue Folge 1*** (1906), pp. 479–706.

[1899] DAVID HILBERT, ***Grundlagen der Geometrie. Festschrift zur Feier der Enthüllung des Gauss–Weber Denkmals in Göttingen***, Teubner, Leipzig, 1899, reprinted as ***Grundlagen der Geometrie***, Leipzig, Teubner, 1899.

[1900] ———, *Mathematische probleme, Vortrag, gehalten auf dem internationalem Mathematiker-Kongress zu Paris*, ***Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen***, (1900), pp. 253–297, reprinted in ***Archiv für Mathematik und Physik*** vol. 3 (1901), no. 1, pp. 44–63 and pp. 213–237, translated in ***Bulletin of the American Mathematical Society***, vol. 8 (1902), pp. 437–479, that translation reprinted in Felix E. Browder (editor), *Mathematical Developments Arising from Hilbert's Problems*, ***Proceedings of Symposia in Pure Mathematics***, vol. 28, American Mathematical Society, Providence, 1976, pp. 1–34.

[1900a] ———, *Über den Zahlbegriff*, ***Jahresbericht der deutschen Mathematiker-Vereinigung***, vol. 8 (1900), pp. 180–184, translated in Ewald [1996], vol. 2, pp. 1092–1095.

[1903] ———, ***Grundlagen der Geometrie. Festschrift zur Feier der Enthüllung des Gauss–Weber Denkmals in Göttingen***, 1903, second edition of [1899] above.

[1905] ———, *Über die Grundlagen der Logik und der Arithmetik*, ***Verhandlungen des dritten internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904*** (Leipzig) (A. Krazer, editor), B. G. Teubner, 1905, pp. 174–184.

[1928] DAVID HILBERT and WILHELM ACKERMANN, ***Grundzüge der theoretischen Logik***, Julius Springer, Berlin, 1928, second edition, 1938, third edition, 1949. Second edition was translated by Lewis M. Hammond, George G. Leckie, and F. Steinhardt as ***Principles of Mathematical Logic***, New York, Chelsea 1950.

[1985] HENRY B. HOLLINGER and MICHAEL J. ZENZEN, ***The nature of irreversibility***, D. Reidel, Dordrecht, 1985.

[1891] EDMUND HUSSERL, *Besprechung von E. Schröder, Vorlesungen über die Algebra der Logik (exakte Logik), vol. I, Leipzig, 1890*, ***Göttingische Gelehrte Anzeigen***, (1891), pp. 243–278, reprinted in Edmund Husserl, ***Aufsätze und Rezensionen (1890–1910), Husserliana Vol. XXII***, Nijhoff, The Hague 1978, pp. 3–43, translated by Dallas Willard in ***Edmund Husserl, Early Writings in the Philosophy of Logic and Mathematics. Collected Works V***, Kluwer, Dordrecht 1994, pp. 52–91.

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

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

[1905] ———, *On a proof that every aggregate can be well-ordered*, ***Mathematische Annalen***, vol. 60 (1905), pp. 465–470.

[1928] Lázló Kalmár, *Zur Theorie der abstrakten Spiele*, ***Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae, Sectio Scientiarum Mathematicarum***, vol. 4 (1928), pp. 65–85, (from 1946: *Acta Scientiarum Mathematicarum, Szeged*).

[1996] Akihiro Kanamori, *The mathematical development of set theory from Cantor to Cohen*, this Bulletin, vol. 2 (1996), pp. 1–71.

[1997] ———, *The mathematical import of Zermelo's Well-Ordering Theorem*, this Bulletin, vol. 3 (1997), pp. 281–311.

[2003] ———, ***The higher infinite***, second ed., Springer-Verlag, Heidelberg, 2003.

[1927] Dénes Kőnig, *Über eine Schlussweise aus dem Endlichen ins Unendliche*, ***Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae (Szeged), sectio scientiarum mathematicarum***, vol. 3 (1927), pp. 121–130.

[1905] Julius (Gyula) Kőnig, *Zum Kontinuum-Problem*, ***Verhandlungen des dritten internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904*** (A. Krazer, editor), B. G. Teubner, Leipzig, 1905, pp. 144–147, reprinted in ***Mathematische Annalen***, vol. 60 (1905), pp. 177–180 with Berichtigung, *ibid.*, p. 462.

[1911] Alwin Korselt, *Über einen Beweis des äquivalenzsatzes*, ***Mathematische Annalen***, vol. 70 (1911), pp. 294–296.

[1950] Gerhard Kowalewski, ***Bestand und Wandel. Meine Lebenserinnerungen. Zugleich ein Beitrag zur nueuren Geschichte der Mathematik***, R. Oldenbourg, München, 1950.

[2004] Stanisław Krajewski, *Gödel on Tarski*, ***Annals of Pure and Applied Logic***, vol. 127 (2004), pp. 303–323.

[1980] Georg Kreisel, *Kurt Gödel, 28 April 1906–14 January 1978*, ***Memoirs of the Fellows of the Royal Society***, vol. 26 (1980), pp. 149–224, Corrections vol. 27 (1981), p. 697 and vol. 28 (1982), p. 718.

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

[1931] Kazimierz Kuratowski and Alfred Tarski, *Les opérations logiques et les ensembles projectifs*, ***Fundamenta Mathematicae***, vol. 17 (1931), pp. 240–248, reprinted in Tarski [1986] vol. 1, pp. 551–559, and translated in Tarski [1983], pp. 143–151.

[1915] Leopold Löwenheim, *Über Möglichkeiten im Relativkalkül*, ***Mathematische Annalen***, vol. 76 (1915), pp. 447–470, translated in van Heijenoort [1967], pp. 228–251.

[1927] Nikolai Luzin, *Sur les ensembles analytiques*, ***Fundamenta Mathematicae***, vol. 10 (1927), pp. 1–95.

[1918] Nikolai Luzin and Wacław Sierpiński, *Sur quelques propriétés des ensembles (A)*, ***Bulletin de l'Académie des Sciences Cracovie, Classe des Sciences Mathématiques, Série A***, (1918), pp. 35–48.

[1923] ———, *Sur un ensemble non mesurable B*, ***Journal de Mathématiques Pures et Appliquées***, vol. 2 (1923), no. 9, pp. 53–72.

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

[1997] Colin McLarty, *Poincaré: Mathematics & Logic & Intuition*, ***Philosophia Mathematica***, vol. 5 (1997), no. 3, pp. 97–115.

[1917] Dimitry Mirimanoff, *Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles*, ***L'enseignment mathématique***, vol. 19 (1917), pp. 37–52.

[1917a] ———, *Remarques sur la théorie des ensembles et les antinomies Cantoriennes. I*, ***L'enseignment mathématique***, vol. 19 (1917), pp. 209–217.

[1980] GREGORY H. MOORE, *Beyond first-order logic: The historical interplay between mathematical logic and axiomatic set theory*, ***History and Philosophy of Logic***, vol. 1 (1980), pp. 95–137.

[1982] ———, ***Zermelo's Axiom of Choice. Its origins, development, and influence***, Springer-Verlag, New York, 1982.

[1988] ———, *The emergence of first-order logic*, ***History and philosophy of modern mathematics*** (William Aspray and Philip Kitcher, editors), Minnesota Studies in the Philosophy of Science, vol. 11, University of Minnesota Press, Minneapolis, 1988, pp. 95–135.

[2002] ———, *Die Kontroverse zwischen Gödel und Zermelo*, ***Kurt Gödel, Wahrheit und Beweisbarkeit*** (Bernd Buldt et alii, editors), vol. 2, öbv-a-hpt, Wien, 2002, pp. 55–64.

[2002a] ———, *Hilbert on the infinite: The role of set theory in the evolution of Hilbert's thought*, ***Historia Mathematica***, vol. 29 (2002), pp. 40–64.

[1939] ANDRZEJ MOSTOWSKI, *Über die Unabhängigkeit des Wohlordnungssatzes vom Ordnungsprinzip*, ***Fundamenta Mathematicae***, vol. 32 (1939), pp. 201–252, translated in [1979] below, vol. 1, pp. 290–338.

[1979] ———, ***Foundational studies. Selected works*** (Kazimierz Kuratowski et alii, editors), North-Holland, Amsterdam, 1979.

[1998] ROMAN MURAWSKI, *Undefinability of truth. The problem of priority: Tarski vs Gödel*, ***History and Philosophy of Logic***, vol. 19 (1998), pp. 153–160.

[1916] EMMY NOETHER, *Die allgemeinsten Bereiche aus ganzen transzendenten Zahlen*, ***Mathematische Annalen***, vol. 77 (1916), pp. 103–128.

[1987] CHARLES PARSONS, *Developing arithmetic in set theory without infinity: Some historical remarks*, ***History and Philosophy of Logic***, vol. 8 (1987), pp. 201–213.

[1906] GIUSEPPE PEANO, *Super theorema de Cantor-Bernstein*, ***Rendiconti del Circolo Matematico di Palermo***, vol. 21 (1906), pp. 360–366, reprinted in [1957–9] below, vol. 1, pp. 337–344.

[1957–9] ———, ***Opera scelte*** (U. Cassina, editor), Edizioni Cremonese, Rome, 1957–9, in three volumes.

[1990] VOLKER PECKHAUS, ***Hilbertprogramm und Kritische Philosophie. Das Göttinger Modell interdisziplinärer Zusammenarbeit zwischen Mathematik und Philosophie***, Vandenhoeck & Ruprecht, Göttingen, 1990.

[1992] ———, *Hilbert, Zermelo und die Institutionalisierung der mathematischen Logik in Deutschland*, ***Berichte zur Wissenschaftsgeschichte***, vol. 15 (1992), pp. 27–38.

[1994] ———, *Von Nelson zu Reichenbach. Kurt Grelling in Göttingen und Berlin, Bibliographie der Werke Kurt Grellings*, ***Hans Reichenbach und die Berliner Gruppe*** (Lutz Danneberg et alii, editors), Vieweg & Sohn, Braunschweig, 1994, pp. 53–86.

[2004] ———, *Paradoxes in Göttingen*, ***One hundred years of Russell's paradox. Papers from the 2001 Munich Russell conference*** (Godehard Link, editor), de Gruyter, Berlin, 2004.

[2002] VOLKER PECKHAUS and REINHARD KAHLE, *Hilbert's paradox*, ***Historia Mathematica***, vol. 29 (2002), pp. 157–175.

[1906] HENRI POINCARÉ, *Les mathématiques de la logique*, ***Revue de métaphysique et de morale***, vol. 14 (1906), pp. 17–34, translated in Ewald [1996], vol. 2, pp. 1038–1052.

[1906a] ———, *Les mathématiques et la logique*, ***Revue de métaphysique et de morale***, vol. 14 (1906), pp. 294–317, translated in Ewald [1996], vol. 2, pp. 1052–1071.

[1909] ———, *Réflexions sure les deux notes précédentes*, ***Acta Mathematica***, vol. 32 (1909), pp. 195–200.

[1909a] ———, *La logique de l'infini*, ***Revue de métaphysique et de morale***, vol. 17 (1909), pp. 461–482.

[2002] WALTER PURKERT, *Grundzüge der Mengenlehre – Historische Einführung*, ***Felix Hausdorff,*** *Gesammelte Werke*, vol. 2, 2002, pp. 1–89.

[1981] BERNHARD RANG and WOLFGANG THOMAS, *Zermelo's discovery of the "Russell paradox"*, ***Historia Mathematica***, vol. 8 (1981), pp. 15–22.

[1905] JULES RICHARD, *Les principes des mathématiques et le problème des ensembles*, ***Revue générale des sciences pures et appliquées***, vol. 16 (1905), p. 541, translated in van Heijenoort [1967], pp. 142–144.

[1963] HERMAN RUBIN and JEAN RUBIN, ***Equivalents of the Axiom of Choice***, North-Holland, Amsterdam, 1963.

[1903] BERTRAND A. W. RUSSELL, ***The principles of mathematics***, Cambridge University Press, Cambridge, 1903, later editions, George Allen & Unwin, London.

[1906] ———, *On some difficulties in the theory of transfinite numbers and order types*, ***Proceedings of the London Mathematical Society***, vol. 4 (1906), no. 2, pp. 29–53.

[1890] ERNST SCHRÖDER, ***Vorlesungen über die Algebra der Logik (exakte Logik)***, vol. 1, B. G. Teubner, Leipzig, 1890, reprinted in [1966] below.

[1898] ———, *Über zwei Definitionen der Endlichkeit und G. Cantor'sche Sätze*, ***Deutsche Akademie der Naturforscher, Nova Acta Academiae Caesareae Leopodino-Carolinae Germanicae Naturae Curiosorum***, vol. 71 (1898), pp. 303–362.

[1966] ———, ***Vorlesungen über die Algebra der Logik***, Chelsea, New York, 1966, three volumes.

[2001] ULRICH SCHWALBE and PAUL WALKER, *Zermelo and the early history of Game Theory*, ***Games and Economic Behavior***, vol. 34 (2001), pp. 123–137.

[1952] JOHN C. SHEPHERDSON, *Inner models for set theory – Part II*, ***The Journal of Symbolic Logic***, vol. 17 (1952), pp. 225–237.

[1999] WILFRIED SIEG, *Hilbert's programs: 1917–1922*, this BULLETIN, vol. 5 (1999), pp. 1–44.

[1930] WACŁAW SIERPIŃSKI and ALFRED TARSKI, *Sur une propriété caractéristique des nombres inaccessibles*, ***Fundamenta Mathematicae***, vol. 15 (1930), pp. 292–300, reprinted in Tarski [1986], vol. 1, pp. 289–297.

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

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

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

[1962] ———, ***Abstract set theory***, Notre Dame Mathematical Lecture Notes #8, Notre Dame, 1962.

[1970] ———, ***Selected works in logic*** (Jens E. Fenstad, editor), Universitetsforlaget, Oslo, 1970.

[1979] ERNST SPECKER, *Paul Bernays*, ***Logic Colloquium '78*** (Maurice Boffa, Dirk van Dalen, and Kenneth McAloon, editors), North-Holland, Amsterdam, 1979.

[1907] PAUL STÄCKEL, *Zu H. Webers Elementarer Mengenlehre*, ***Jahresbericht der deutschen Mathematiker-Vereinigung***, vol. 16 (1907), pp. 425–428.

[1998] WILLIAM W. TAIT, *Zermelo's conception of set theory and reflection principles*, ***The philosophy of mathematics today*** (Matthias Schirn, editor), Oxford University Press, Oxford, 1998, pp. 469–483.

[1924] ALFRED TARSKI, *Sur les ensembles finis*, ***Fundamenta Mathematicae***, vol. 6 (1924), pp. 45–95, reprinted in [1986] below, vol. 1, pp. 67–117.

[1931] ———, *Sur les ensembles définissables de nombres réels*, ***Fundamenta Mathematicae***, vol. 17 (1931), pp. 210–239, translated in [1983] below, pp. 110–142, and reprinted in [1986] below, vol. 1, pp. 517–548.

[1933] ———, *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.

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

[1938] ———, *Über unerreichbare Kardinalzahlen*, ***Fundamenta Mathematicae***, vol. 30 (1938), pp. 68–89, reprinted in [1986] below, vol. 2, pp. 359–380.

[1939] ———, *On well-ordered subsets of any set*, ***Fundamenta Mathematicae***, vol. 32 (1939), pp. 176–183, reprinted in [1986] below, vol. 2, pp. 551–558.

[1983] ———, ***Logic, semantics, metamathematics. Papers from 1923 to 1938***, second ed., Hackett, Indianapolis, 1983, translations by J. H. Woodger.

[1986] ———, ***Collected papers*** (Steven R. Givant and Ralph N. McKenzie, editors), Birkhäuser, Basel, 1986.

[1987] OLGA TAUSSKY-TODD, *Remembrances of Kurt Gödel*, ***Gödel remembered*** (Paul Weingartner and Leopold Schmetterer, editors), Bibliopolis, Naples, 1987.

[1993] R. GREGORY TAYLOR, *Zermelo, reductionism, and the philosophy of mathematics*, ***Notre Dame Journal of Formal Logic***, vol. 34 (1993), pp. 539–563.

[2002] ———, *Zermelo's Cantorian theory of systems of infinitely long propositions*, this BULLETIN, vol. 8 (2002), pp. 478–513.

[1958] STANISŁAW M. ULAM, *John von Neumann, 1903–1957*, ***Bulletin of the American Mathematical Society***, vol. 64 (1958), pp. 1–49.

[1999] GABRIEL UZQUIANO, *Models of second-order Zermelo set theory*, this BULLETIN, vol. 5 (1999), pp. 289–302.

[2000] DIRK VAN DALEN and HEINZ-DIETER EBBINGHAUS, *Zermelo and the Skolem paradox*, this BULLETIN, vol. 6 (2000), pp. 145–161.

[1967] Jean van Heijenoort (editor), ***From Frege to Gödel: A source book in mathematical logic, 1879–1931***, Harvard University Press, Cambridge, 1967, reprinted 2002.

[1923] JOHN VON NEUMANN, *Zur Einführung der transfiniten Zahlen*, ***Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae (Szeged), sectio scientiarum mathematicarum***, vol. 1 (1923), pp. 199–208, reprinted in [1961] below, vol. 1, pp. 24–33, translated in van Heijenoort [1967], pp. 346–354.

[1925] ———, *Eine Axiomatisierung der Mengenlehre*, ***Journal für die reine und angewandte Mathematik***, vol. 154 (1925), pp. 219–240, Berichtigung, ibd. vol. 155, p. 128, reprinted in [1961] below, vol. 1, pp. 34–56, translated in van Heijenoort [1967], pp. 393–413.

[1928] ———, *Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre*, ***Mathematische Annalen***, vol. 99 (1928), pp. 373–391, reprinted in [1961] below, vol. 1, pp. 320–338.

[1928a] ———, *Die Axiomatisierung der Mengenlehre*, ***Mathematische Zeitschrift***, vol. 27 (1928), pp. 669–752, reprinted in [1961] below, vol. 1, pp. 339–422.

[1928b] ———, *Zur Theorie der Gesellschaftsspiele*, ***Mathematisches Annalen***, vol. 100 (1928), pp. 295–320, reprinted in [1961] below, vol. 6, pp. 1–26.

[1929] ———, *Über eine Widerspruchsfreiheitsfrage in der axiomatischen Mengenlehre*, ***Journal für die reine und angewandte Mathematik***, vol. 160 (1929), pp. 227–241, reprinted in [1961] below, vol. 1, pp. 494–503.

[1961] ———, ***John von Neumann. Collected works*** (Abraham H. Taub, editor), Pergamon Press, New York, 1961.

[1944] John von Neumann and Oskar Morgenstern, ***Theory of games and economic behavior***, Princeton University Press, Princeton, 1944.

[1963] Petr Vopěnka and Petr Hajek, *Über die Gültigkeit des Fundierungsaxioms in speziellen Systemen der Mengentheorie*, ***Zeitschrift für mathematische Logic und Grundlagen der Mathematik***, vol. 9 (1963), pp. 235–241.

[1974] Hao Wang, ***From mathematics to philosophy***, Humanities Press, New York, 1974.

[1996] ———, ***A logical journey. From Gödel to philosophy***, The MIT Press, Cambridge, 1996.

[1910] Hermann Weyl, *Über die Definitionen der mathematischen Grundbegriffe*, ***Mathematisch-naturwissenschaftliche Blätter***, vol. 7 (1910), pp. 93–95 and 109–113.

[1918] ———, ***Das Kontinuum***, Verlag von Veit, Leipzig, 1918.

[1910–13] Alfred N. Whitehead and Bertrand A. W. Russell, ***Principia mathematica***, Cambridge University Press, Cambridge, 1910–13, in three volumes.

[1894] Ernst Zermelo, *Untersuchungen zur Variationsrechnung*, ***Ph.D. thesis***, University of Berlin, 1894, 97 pp.

[1896] ———, *Über einen Satz der Dynamik und die mechanische Wärmetheorie*, ***Annalen der Physik und Chemie, Neue Folge***, vol. 57 (1896), pp. 485–494.

[1896a] ———, *Über mechanische Erklärungen irreversibler Vorgänge. Eine Antwort auf Hrn. Boltzmanns "Entgegnung"*, ***Annalen der Physik und Chemie, Neue Folge***, vol. 59 (1896), pp. 793–801.

[1901] ———, *Über die Addition transfiniter Kardinalzahlen*, ***Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen***, (1901), pp. 34–38.

[1902] ———, *Hydrodynamische Untersuchungen über die Wirbelbewegungen in einer Kugelfläche. Erste Mitteilung*, ***Zeitschrift für Mathematik und Physik***, vol. 47 (1902), pp. 201–237.

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

[1908] ———, *Neuer Beweis für die Möglichkeit einer Wohlordnung*, ***Mathematische Annalen***, vol. 65 (1908), pp. 107–128, translated in van Heijenoort [1967], pp. 183–198.

[1908a] ———, *Untersuchungen über die Grundlagen der Mengenlehre I*, ***Mathematische Annalen***, vol. 65 (1908), pp. 261–281, translated in van Heijenoort [1967], pp. 199–215.

[1909] ———, *Sur les ensembles finis et le principe de l'induction complète*, ***Acta Mathematica***, vol. 32 (1909), pp. 185–193.

[1909a] ———, *Über die Grundlagen der Arithmetik*, ***Atti del IV Congresso Internazionale dei Matematici, Roma 1908*** (G. Castelnuovo, editor), vol. 2, Accademia dei Lincei, Roma, 1909, pp. 8–11.

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

[1914] ———, *Über ganze transzendente Zahlen*, ***Mathematische Annalen***, vol. 75 (1914), pp. 434–442, with Nachtrag, *ibid*, p. 442.

[1928] ———, *Die Berechnung der Turnier-Ergebnisse als ein Maximumproblem der Wahrscheinlichkeitsrechnung*, ***Mathematische Zeitschrift***, vol. 29 (1928), pp. 436–460.

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

[1930] ———, *Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre*, ***Fundamenta Mathematicae***, vol. 16 (1930), pp. 29–47, translated in Ewald [1996], vol 2, pp. 1208–1233.

[1932] ———, *Über Stufen der Quantifikation und die Logik des Unendlichen*, ***Jahresbericht der deutschen Mathematiker-Vereinigung (Angelegenheiten)***, vol. 41 (1932), pp. 85–88.

[1932a] ———, *Über mathematische Systeme und die Logik des Unendlichen*, ***Forschungen und Fortschritte***, vol. 8 (1932), pp. 6–7.

[1935] ———, *Grundlagen einer allgemeinen Theorie der mathematischen Satzsysteme (Erste Mitteilung)*, ***Fundamenta Mathematicae***, vol. 25 (1935), pp. 136–146.

[1904] Ernst Zermelo and Hans Hahn, *Weiterentwicklung der Variationsrechnung in den letzten Jahren*, ***Encyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen***, vol. 2, part 1, Teubner, Leipzig, 1904, pp. 626–641.

DEPARTMENT OF MATHEMATICS
BOSTON UNIVERSITY
BOSTON, MA 02215
E-mail: aki@math.bu.edu