THE BULLETIN OF SYMBOLIC LOGIC
Volume 18, Number 1, March 2012

IN PRAISE OF REPLACEMENT

AKIHIRO KANAMORI

Abstract. This article serves to present a large mathematical perspective and historical basis for the Axiom of Replacement as well as to affirm its importance as a central axiom of modern set theory.

The standard ZFC axioms for set theory provide an operative foundation for mathematics in the sense that mathematical concepts and arguments can be reduced to set-theoretic ones, based on sets doing the work of mathematical objects. As is widely acknowledged, Replacement together with the Axiom of Infinity provides the rigorization for transfinite recursion and together with the Power Set Axiom provides for sets of large cardinality, e.g., through infinite iterations of the power set operation. And in a broad sense, in so far as the concept of function (mapping, functor) plays a central role in modern mathematical practice, Replacement plays a central role in set theory.

A significant motivation for the writing of this article is to confirm—to rectify, if need be perceived—the status of Replacement among the axioms of set theory. From time to time hesitation and even skepticism have been voiced about the importance and even the need for Replacement.[1] The issues, animated by how, historically, Replacement emerged as an axiom later than the others, have largely to do with ontological commitment, about what sets

Received December 17, 2010.

Most of this article was written whilst the author was a senior fellow at the Lichtenberg-Kolleg of the Georg-August-Universität Göttingen for 2009–2010. He expresses his gratitude for the productive conditions and support of the Kolleg. The article served as the text for an invited address at the Workshop on Set Theory and the Philosophy of Mathematics held at the University of Pennsylvania, October 2010, and the author would like to express his gratitude to the organizer Scott Weinstein. For information or inspiration on some points, many thanks to Juliet Floyd, John Burgess, Heinz-Dieter Ebbinghaus, Colin McLarty, Adrian Mathias, Michael Rathjen, Erich Reck, Wilfried Sieg and the referee.

[1]This is brought out below with respect to the early contributors to the development of set theory and in §6 in connection with writings of George Boolos. The text Potter [2004] pursues an axiomatic development of set theory which avoids Replacement as being contextually unmotivated (cf. its pp. 296ff). The recent webpage discussion of Replacement at `http://cs.nyu.edu/pipermail/fom/2007-August/` drew out some skeptical attitudes.

1079-8986/12/1801-0002/$5.50

(there) are or should be. Replacement has been regarded as necessary only to provide for large cardinality sets and, furthermore, unmotivated by an ultimately ontological "iterative conception" of set.

The answer here, as well as the general affirmation, is that Replacement is a subtle axiom, with the subtlety lying in both what it says and what it does. Like the Axiom of Choice, Replacement has to do with functional correlation, but the latter turns on formalization, on definability. And the definability crucially leads to a full range of possibilities for recursive definition. Together with the Axiom of Foundation, Replacement is thus a crucial pillar of modern set theory, which, as a matter of method, is an investigation through and of recursion and well-foundedness. From a modern vantage point, Replacement also carries the weight of the heuristic of reflection, that any property ascribable to the set-theoretic universe should already be ascribable to a curtailed initial segment, as an underlying and motivating principle of set theory.

But stepping back even from this, Replacement can be seen as a crucial bulwark of *indifference to identification*, in set theory and in modern mathematics generally. To describe a prominent example, several definitions of the real numbers as generated from the rational numbers have been put forward—in terms of the geometric continuum, Dedekind cuts, and Cauchy sequences—yet in mathematical practice there is indifference to actual identification with any particular objectification as one proceeds to work with *the* real numbers. In set theory, one opts for a particular representation for an ordered pair, for natural numbers, and so forth. What Replacement does is to allow for articulations that these representations are not necessary choices and to mediate generally among possible choices. Replacement is a corrective for the other axioms, which posit specific sets and subsets, by allowing for a fluid extensionalism. The deepest subtlety here is also on the surface, that through functional correlation one can shift between tokens (instances, representatives) and types (extensions, classes), and thereby shift the ground itself for what the types are.[2]

In what follows, these themes, here briefly broached, are widely explored in connection with the historical emergence of set theory and Replacement. This account can in fact be viewed as one of the early development of set theory as seen through the prism of Replacement, drawing out how the Replacement motif gauges the progress from 19th Century mathematics to modern set theory. There is recursion emerging as method; early incentives for trans-species coordination, expressive of indifference to identification; provision for sets of relatively large cardinality; the crucial role in transfinite

[2]Willard Van Orman Quine with his theses of indeterminacy of reference and ontological relativity illuminated indifference to identification, and discussed in particular the ordered pair in his *Word and Object* [1960], §53. For recent articulations of indifference to identification see Richard Pettigrew [2008] and John Burgess [2009], from which the phrase was drawn.

recursion; and the formalization as principle of the heuristic of reflection. §1 casts Replacement as underlying modern mathematics through the motif $\{t_i \mid i \in I\}$, the roots of this involvement going back particularly to the infinitary initiatives of Richard Dedekind. §2 sets out the early history of Replacement in the work of the pioneers of set theory from an adumbration in correspondence of Cantor himself through to the formulations of Fraenkel and Skolem. §3 sets out the middle history of Replacement, largely the work of von Neumann on transfinite recursion and the emergence of the cumulative hierarchy picture. §4 sets out the later history of Replacement, the main focus being on Gödel's work with the constructible universe. §5 discusses the various forms of Replacement, bringing out its many emanations and applications. Finally, §6 sets out skeptical reactions about Replacement and recent involvements of the axiom and affirms on general grounds the importance of Replacement as a central axiom of modern set theory.

§1. Replacement underlying mathematics. The Axiom of Replacement is expressible in first-order set theory as the following schema: For any formula φ with free variables among $a, x, y, z_1, \ldots, z_n$ but not including b,

$$\forall z_1 \forall z_2 \ldots \forall z_n \forall a (\forall x \exists! y \varphi \longrightarrow \exists b \forall y (\exists x (x \in a \wedge \varphi) \longleftrightarrow y \in b)), \quad \text{(Rep)}$$

where $\exists!$ abbreviates the formalizable "there exists exactly one". Seen through unpracticed eyes this formalization suggests complications and fuss; in a historically resonant sense as we shall see, the first-order account conveys a fundamental principle about functional substitution as an inherent feature of modern mathematics. Informally, and as historically presented, Replacement asserts that for any definable class function F ($F(x) = y \longleftrightarrow \varphi$ in the above) and for any set a, the image $F``a = \{F(x) \mid x \in a\}$ is a set. Thus, Replacement is proximate whenever one posits $\{t_i \mid i \in I\}$ according to some specifiable correlation of t_i's to i's in an index set I. The many facets of $\{t_i \mid i \in I\}$ only emerged in mathematics, however, in its 19th Century transformation from a structural analysis of what there is to a complex edifice of conceptual constructions, and it is worthwhile to recapitulate this development. In what follows, "Replacement motif" is used to refer to $\{t_i \mid i \in I\}$ as a broader conceptualization than "Replacement" for versions of the later axiom.

As with the Axiom of Choice, before Replacement was articulated implicit assumptions had been made that now can be seen as dependent on the axiom. The seminal work of Richard Dedekind is particularly relevant, and in what follows we highlight its involvement with Replacement. In the development of set theory as a subject, the Replacement motif first occurred in Cantor's work in connection with cardinality (cf. §2) with the working assumption that $i \mapsto t_i$ was one-to-one, which is more faithful to "replacement". The injectivity would be importantly relaxed in important situations, however, and this is inherent in the consideration of more and more functions in mathematics. To be noted is that in discussing the Replacement motif in the early stages one is focusing on an emergent theme, and retroactive formalization will only require, naturally enough, a small part of the full modern schema.

Surely the first substantive appearance of $\{t_i \mid i \in I\}$ was in analysis, with the t_i's themselves functions and I the natural numbers. Initially, the functions were typically real functions given symbolically as polynomials, and the assignment $i \mapsto t_i$ would not itself have been considered a function. When infinite sequences of infinite series were considered, e.g., for nowhere differentiable functions as limits of continuous functions, new stress was put on the traditional approach to the infinite regarded only in potentiality. The Replacement motif became substantively involved when the infinite in actuality became incorporated into mathematics. This occurred with the emergence of the now basic construction of a set I structured by an equivalence relation E, the consideration of the corresponding equivalence classes $[i]_E$ for $i \in I$ leading to the totality $\{[i]_E \mid i \in I\}$. One sees here that, as the function $i \mapsto [i]_E$ becomes something to be reckoned with, it is quite general and, importantly, not one-to-one. This early move toward modern algebra went hand in hand with the development of set-theoretic conceptualizations, particularly in the work of Dedekind.

Already in the early [1857], Dedekind worked in modular arithmetic with the actually infinite residue classes themselves as unitary objects. His context was in fact $Z[x]$, the polynomials in x with integer coefficients, and with this he was the first to entertain a totality consisting of infinitely many infinite equivalence classes. The roots of these polynomials, first investigated by Dedekind, are of course what came to be called the algebraic numbers. In a telling passage discussing $Z[x]$ modulo a prime number p, Dedekind wrote (cf. Dedekind [1930–1932, vol. 1, pp. 46–47]):

> The preceding theorems correspond exactly to those of number divisibility, in that the whole system of infinitely many functions of a variable congruent to each other modulo p behaves here as a single concrete number in number theory, as each function of that system substitutes completely for any other; such a function is a representative of the whole class; each class possesses its definite degree, its divisors, etc., and all those traits correspond in the same manner to each particular member of the class. The system of infinitely many incongruent classes—infinitely many, since the degree may grow indefinitely—corresponds to the series of whole numbers in number theory.

One can arguably date the entry of the actual infinite into mathematics here, in the sense of infinite totalities serving as unitary objects within an infinite mathematical system,[3] as well as the beginnings of mathematical indifference to identification for equivalence classes and particular representatives.

In a fragment "Aus den Gruppen-Studien (1855–1858)" Dedekind [1930–1932, vol. 3, pp. 439–445], working notationally with finite groups, gave in effect the Homomorphism Theorem, that given a homomorphism of a group G onto a group H with a corresponding kernel K, there is an isomorphism between the quotient group G/K and H. Allowing for the historical happenstance of working in the finite, Dedekind here was first to *start* with a function and develop corresponding equivalence classes. The Replacement motif $\{t_i \mid i \in I\}$ comes into focus, soon to be put in a new light.

In Dedekind's *Nachlass* can be found sketches, conjectured to be from 1872,[4] of the now-familiar genetic generation of the integers as equivalence classes of pairs of natural numbers, a pair representing their difference, and of the rational numbers as equivalence classes of pairs of integers, a pair representing their ratio. With this approach to achieve the status of definitions in mathematics, one sees here the further beginnings of mathematical indifference to identification. Not only are there to be correlations with antecedent notions of integers and rationals, but also correlations between construals of integers and their reconstruals as rationals. Today one works indifferently with integers and rationals as algebraic systems.

Notably, equivalence classes would *not* be initially involved in the genetic generation of the reals from the rationals. Dedekind's [1872] cuts themselves represented the reals, and Cantor [1872] considered that he had defined the reals as fundamental, i.e., Cauchy, sequences of rationals but did not actually work with equivalence classes of such sequences.

Also in 1872 Dedekind worked his way down to the bedrock of the natural numbers, and the eventual result was his celebrated essay *Was sind und was sollen die Zahlen?* [1888]. Dedekind here took sets [Systeme] and mappings [Abbildungen] as basic notions; worked with unions and intersections of sets and compositions and inversions of mappings; and developed a theory of chains [Ketten], sets with one-to-one mappings into themselves.

[3]Carl Friedrich Gauss, in his *Disquisitiones Arithmeticae* and later, worked in modular arithmetic directly with residues and never entertained equivalence classes as unitary objects.

Bernard Bolzano in his *Paradoxien des Unendlichen* [1851] and elsewhere (see his mathematical works in Russ [2004]) had earlier advocated the actual infinite, but his advocacy went only so far as to discuss comparative mappings between infinite collections and moreover was not embedded in mathematical practice.

In the prefaces to both his celebrated essays [1872] and [1888] Dedekind recorded the autumn of 1858 as when he first came to the Dedekind cuts. Thus he had devised his best-known construction of an infinite system consisting of infinite totalities serving as unitary objects just a year after his analysis of $Z[x]$ modulo a prime.

[4]See Sieg and Schlimm [2005, pp. 134–138].

Dedekind famously defined (paragraph 64) an infinite set to be a set having a one-to-one mapping into a proper subset. With this he had come to a positive formulation of the actual infinite, one which is evidently the logical negation of Dirichlet's Pigeonhole Principle.[5] Dedekind then (in)famously "proved" (66) the existence of a Dedekind-infinite set by invoking "my own realm of thoughts". Working toward the natural numbers, Dedekind defined (71) a "simply infinite system" to be a set N for which there is a one-to-one mapping ϕ into itself and a $1 \in N$ not in the range such that N is the closure of $\{1\}$ under ϕ. From a Dedekind-infinite set he got to a simply infinite system, and abstracting from particulars but proceeding with the same notation, took (73) *the* natural numbers to be the members of N, with base element 1 and successor operation $n' = \phi(n)$. This was the paradigmatic move between token and type; Dedekind soon gave expression to this indifference to identification by affirming recursive definition.

Recursive definition is crucial for modern mathematics, the basic case being the formulation of $\{t_i \mid i \in I\}$ with I the natural numbers and t_{i+1} uniformly given in terms of t_i—with the shifting of the focus to the function $i \mapsto t_i$ itself. One has here the Replacement motif cast with iteration of procedure as a generalization of counting itself. Dedekind proceeded to establish the Recursion Theorem (126): Given a set Ω, a distinguished element $\omega \in \Omega$, and a mapping $\theta \colon \Omega \to \Omega$, there is one and only one mapping $\psi \colon N \to \Omega$ satisfying $\psi(1) = \omega$ and $\psi(n') = \theta(\psi(n))$ for every n.

Dedekind was the first to point out the need to posit the existence of a completed mapping in this way,[6] and with the theorem, he soon provided the now-familiar recursive definition of addition as a mapping, the recursive definition of multiplication in terms of addition, and the recursive definition of exponentiation in terms of multiplication. However, it is instructive to point out that Dedekind's argument at this foundational level, as seen through modern eyes, can be considered to be subtly circular:

To establish (126), Dedekind appealed to the preparatory (125) according to which one can recursively define finite mappings ψ_n on $\{1, \ldots, n\}$ unique in satisfying: $\psi_n(1) = \omega$ and $\psi_n(m') = \theta(\psi_n(m))$ for $m < n$. What is

[5]One cannot presume that Dedekind consciously negated the Pigeonhole Principle to get his definition of infinite set, but there are interactions and confluences: The Pigeonhole Principle seems to have been first applied in mathematics by Dirichlet in papers of 1842 (cf. Dirichlet [1889/97, pp. 579, 636]), one on Pell's equation and another in which the principle is applied to prove a crucial approximation lemma for his well-known Unit Theorem describing the group of units of an algebraic number field. The Pigeonhole Principle occurred in Dirichlet's *Vorlesungen über Zahlentheorie* [1863], edited and published by Dedekind. The occurrence is in the second, 1871 edition, in a short Supplement VIII by Dedekind on Pell's equation, and it was in the famous Supplement X that Dedekind laid out his theory of ideals in algebraic number theory, working directly with infinite totalities.

[6]Gottlob Frege in his 1893 *Grundgesetze* also established the Recursion Theorem; see Heck [1995] for an account.

the status of $\{\psi_n \mid n \in N\}$ at this point? Dedekind did consider sets [Systeme] of mappings, e.g., in (131), but to have construed the sequence $\langle\psi_n \mid n \in N\rangle$ as an objectified mapping at (125) would have been circular since it is (126) itself which posits recursively presented mappings. Dedekind then considered the desired mapping for (126) to be given by $\psi(n) = \psi_n(n)$ as a direct definition. He had deftly devolved first to the approximations ψ_n's, but without the ψ_n's collectively comprehended, how is one running through the indexing of the ψ_n? Seen through modern eyes sensitized to the ways of paradox, this "diagonal" stipulation would not be taken to be a contextually internal definition. The thrust of the Recursion Theorem (126) is to be able to pass from a given mapping to another mapping stipulated by recursion; the proof of (126) itself has this form, but with the givenness of $\langle\psi_n \mid n \in N\rangle$ not quite in hand.[7] However basic Dedekind was taking the notion of mapping, with (126) he had made an existence assertion about mappings, and logically speaking, an existence assertion cannot be arrived at without generative existence principles at work. And if one were to make such principles explicit for (126) as Dedekind had done elsewhere for sets [Systeme], one would be drawing in some version of the Replacement motif $\{\psi_n \mid n \in N\}$.

Looking briefly ahead to compare and contrast, how one articulates recursion would become a crucial theme in the formalization of both arithmetic and set theory. Thoralf Skolem [1923a], developed elementary arithmetic by taking the "recursive mode of thought" as basic and proceeding with a system of equation rules. *Primitive Recursive Arithmetic* is a subsequent formalization of Skolem's approach, in which function symbols are simply introduced for each (primitive) recursion and their generating rules in the manner of Dedekind's (126) are given as axioms. Like Dedekind, Skolem took function as a basic concept, but he made explicit the recursions. Paul Bernays [1941, pp. 11–12], in an exposition of an axiomatic set theory with classes, articulated Dedekind's (126) argument prior to introducing an axiom of infinity by formalizing functions as classes of ordered pairs and applying the "class theorem" or the predicative comprehension schema, which asserts that extensions of formulas without class quantifiers are classes. With this, $\{\psi_n \mid n \in N\}$ is a class and so also the resulting ψ. In ZFC set theory, functions are formalized as sets of ordered pairs; N can be taken to be the natural numbers as given by the Axiom of Infinity; one can appeal to Replacement to get $\{\psi_n \mid n \in N\}$ as a set; and Union establishes $\psi = \bigcup\{\psi_n \mid n \in N\}$ to be a set.[8]

---

[7]Something of this circularity surfaces in Landau's text *Grundlagen der Analysis* [1930], cf. its preface.

[8]This importantly is the argument when the Recursion Theorem is situated in set theory with Ω not necessarily a set. When Ω is a set, one can appeal to Power Set and Separation, getting $\{\psi_n \mid n \in N\}$ as a subset of $P(N \times \Omega)$. This is often done in expository accounts,

Returning to Dedekind [1888], the first use of the Recursion Theorem (126) was actually to establish (132) that *all simply infinite systems are isomorphic*. In (126) one takes Ω to be an arbitrary simply infinite system and, appealing to "complete induction" in both N and Ω, shows that the resulting θ is an isomorphism. The result is nowadays touted as establishing the second-order categoricity of the Dedekind–Peano axioms, although strictly speaking Dedekind was not proceeding in an axiomatic context. Replacement can again be seen as involved in providing for indifference to identification, to establish both the efficacy of a chosen representative token of a type and a crucial pliability whereby one can entertain various tokens as efficacious.

Many years later, Ernst Zermelo in his final axiomatization paper [1930b] also established a second-order categoricity result, one for his natural models of set theory. With transfinite recursion he pieced together their cumulative hierarchies, and Replacement—incorporated into his later axiomatization—is crucial for the process. Zermelo in his progress toward his first axiomatization [1908b] had studied *Was sind und was sollen die Zahlen?* carefully and had subsequently pointed out what turned out to be Dedekind's only other logical gap in his essay. Dedekind had established (160) that, as we would now say, Dedekind-finite sets are finite, but Zermelo [1909, p. 190, n. 5] pointed out that this requires a use of the Axiom of Choice, to choose from a set of sets of mappings. Also, Dedekind's proof relied on the Recursion Theorem (126). If formalization through reduction to sets is to be worked out, then Dedekind's work needs the Axiom of Choice, and in a methodologically similar way, Replacement, namely to procure certain sequences of mappings. These two principles are thus implicated, for the first time, in the formalization of the infinite in actuality, and they would become part and parcel of its accommodation in set theory.

Although Replacement would seem to have a crucial role underpinning mathematics through $\{t_i \mid i \in I\}$, textbook expositions of ZFC set theory most often introduce the axiom schema rather late in the development. This presumably has to do with several factors: As formally stated in first-order logic, the axiom schema is syntactically complicated, and its paraphrase in terms of classes adds a layer of conceptualization over sets; Replacement, together with the Axiom of Foundation, was adjoined to Zermelo's initial 1908 axioms significantly later; and the set-theoretic reduction of the ordered pair, and thereby function, induced ready appeals to the Power Set Axiom in preference to Replacement. Instead, one can proceed as in the text Bernays [1958] and introduce Replacement early, emphasizing the motif $\{t_i \mid i \in I\}$

---

but the direct appeal to Power Set is not cardinally parsimonious. Alternately, one can invoke the Dedekind chain theory by defining a function $\phi \subseteq (N \times \Omega) \times (N \times \Omega)$ by $\phi(\langle n, x\rangle) = \langle n', \psi(x)\rangle$ and taking the closure of $\langle 1, \omega\rangle$ under ϕ; Power Set and Separation are used here to get the closure.

as basic. The text sets out a formal apparatus with class terms $\{x \mid \varphi(x)\}$ and the conversion $\varphi(a) \longleftrightarrow a \in \{x \mid \varphi(x)\}$.

But even in an elementary text, one can develop axiomatic set theory to exhibit Replacement as reflective of the use of $\{t_i \mid i \in I\}$ in mathematical practice: After an informal look at first-order logic and class terms, introduce the initial axioms of Extensionality, Separation, Existence, Pairing, and Union and develop the basics through the Kuratowski ordered pair to relations and functions. With the work on Separation as a first-order schema having prepared the way, now introduce Replacement and emphasize the paraphrases in terms of class functions and $\{t_i \mid i \in I\}$. Put Replacement forthwith to use to establish, e.g., that given an equivalence relation the equivalence classes together form a set, and that given two sets their Cartesian product exists.[9] Proceed to introduce the Axiom of Infinity and the natural numbers. With Replacement, establish Dedekind's Recursion Theorem and with it the arithmetic of the natural numbers as well as Transitive Containment, i.e., that every set is a subset of a transitive set. With precedents set, proceed to the (von Neumann) ordinals, and with Replacement in full play establish the basic theory through to the Transfinite Recursion Theorem. As Cantor and Dedekind had come to see, ordinality should be prior to cardinality, and so only now introduce cardinality and the Power Set Axiom to give the concept heft. One thus sees that the substantive issues about the existence of large cardinality sets arise with Power Set, not Replacement. Finally, there is the Axiom of Foundation and the picture of the cumulative hierarchy of sets; with all sets now to appear in this recursively defined hierarchy, Replacement and Foundation work together to establish results by recursion for all sets.

§2. Early history of Replacement. The history of the emergence of Replacement as an integral axiom of modern ZFC set theory, like that of the Axiom of Choice, has to do most importantly with the emergence of set-theoretic methods and their formalization. With Choice, it was well-ordering and through it maximalization as with Zorn's Lemma, and with Replacement it was transfinite recursion and through it closure of the set-theoretic universe under processes involving large cardinality. With the thrust of Replacement involving functional correlation, its history, like that of Choice, has to do initially with the liberalization of the concept of function and the expansion of the concept of set. The motif $\{t_i \mid i \in I\}$ underlying mathematics, as described in the previous section, became rigorized as a set-existence principle, one that became central to set theory.

[9]The expected argument works with any adequate definition of the ordered pair, i.e., any for which one can uniquely recover the first and second coordinates. Conversely, Mathias [∞, sect. C] has observed that if as a schema Cartesian products exist for every adequate definition of ordered pair, then Replacement follows.

Replacement first occurred in an early form in terms of one-to-one correspondence in late work of Georg Cantor himself. The motif $\{t_i \mid i \in I\}$ had in any case been congenial to Cantor both in his investigation of cardinality and because of his generation principles for limit and higher cardinality ordinal numbers.[10] Having developed the transfinite landscape for over two decades, Cantor in correspondence with Hilbert and Dedekind newly formulated a crucial juncture for set existence.[11]

In a letter to Hilbert of 26 September 1897, Cantor stressed that "the totality of all alephs cannot be conceived as a determinate, well-defined *finished set* [*fertige Menge*]" since a cardinal number for it cannot be entertained without contradiction. He repeated this assertion in a letter to Hilbert of 2 October 1897 where he wrote further that a set can be thought of as *finished* if "it is possible ... to *think of all its elements as existing together*, ... " A year later, in a letter to Hilbert of 10 October 1898 elaborating his concept of finished set, Cantor concluded as a "theorem" the assertion that *if two multiplicities are cardinally equivalent* (*i.e., in one-to-one correspondence*) *and one is a finished set, then so is the other*. This is the first substantive anticipation of Replacement in set theory. In that letter Cantor made the observation that the totality of all subsets of the natural numbers is cardinally equivalent to the totality of all functions from the natural numbers into $\{0, 1\}$, a few years before Russell would make this connection in his theory of relations. With this observation applied to the "linear continuum" and a prior "proposition" that "[t]he multiplicity of all the subsets of a finished set M is a finished set", Cantor then derived that the "linear continuum" is a finished set. Thus, Cantor could be seen as adumbrating indifference to identification for the roughest criterion, one-to-one correspondence, and with it, expressing a now-standard plasticity for the real numbers.

In a well-known letter to Dedekind of 3 August 1899, Cantor, after confronting the Burali-Forti Paradox,[12] emphasized his distinction between consistent multiplicities—sets—and inconsistent, absolute multiplicities—like the totality of all ordinal numbers. Cantor in fact argued that if a definite multiplicity does not have an aleph as its cardinal number, then there would be a "projection" of the totality of all ordinal numbers into the multiplicity. Cantor wrote: "Two equivalent multiplicities either are both 'sets' or are both inconsistent." Thus his anticipation of Replacement now served less as a set existence principle and more as an articulation of dichotomy based on the possibilities for well-ordering. In any case, Replacement had been

[10] Bernays in his book [1958] introduced Replacement early for "general set theory", which was to correspond to Cantor's context for generating the transfinite numbers.

[11] The correspondence appears in Meschkowski and Nilson [1991], and some letters are translated in Ewald [1996], including those cited below except that of 10 October 1898.

[12] Cantor came to the "paradox" of the largest ordinal number before the appearance of Burali-Forti [1897]. Whether paradoxical or not, it had to be confronted and analyzed. See Moore and Garciadiego [1981] for the history of the Burali-Forti paradox.

anticipated in a pre-formal setting with one-to-one correlation taken as a basic notion.[13]

The early set-theoretic work of Zermelo, formative of course for the development of set theory, has bearing through thematic points of contact.[14] Zermelo [1904] famously made explicit an initial appeal to the Axiom of Choice and established the Well-Ordering Theorem, that every set can be well-ordered. The proof of the theorem presented a new mode of argument, one that can be viewed as an anticipation of the proof of the later Transfinite Recursion Theorem, the theorem that validates definitions by recursion indexed by well-orderings. This theorem was first properly articulated and established by von Neumann (see below) using Replacement. The difference is only that Zermelo's proof does not require functional correlation, as enabled by Replacement, as one is defining the well-ordering itself.

Zermelo's proof generated controversy about the acceptability of the Axiom of Choice, and largely in response, Zermelo subsequently published both a second proof [1908a] of the Well-Ordering Theorem and the first full-fledged axiomatization [1908b] of set theory. Zermelo's axioms were few and had a remarkable simplicity, except for the one "logical" axiom having to do with properties, the Axiom of Separation. As it provided for the crucial separating out of members of a given set according to a "definite" property, what these properties were came to be seen as in need of clarification as mathematical logic was further developed. As is well-known, the stronger property-based Axiom of Replacement played no role in Zermelo's axiomatization. Zermelo was proceeding pragmatically and parsimoniously to establish set theory as a discipline axiomatically given in the Hilbertian style and to put his Well-Ordering Theorem on a sound footing with a modicum of set-existence principles. Zermelo's second proof [1908a] of the Well-Ordering Theorem indeed coordinated with the axioms, particularly as concerns well-ordering. Taking an approach first used by Gerhard Hessenberg [1906, pp. 674ff], Zermelo cast a well-ordering as the set of its final segments under the reverse-inclusion ordering, thereby situating the concept in his axiomatic framework.

The work of Friedrich Hartogs [1915] on well-orderings and Cardinal Comparability is a conspicuous juncture having interaction with Replacement as motif but still historically and mathematically prior to the use of

[13]The little known A. E. Harward, in an article [1905] about cardinal arithmetic and well-orderings, also came to these issues and ideas entertained by Cantor. Though Harward took a distinction between his "unlimited classes" and "aggregates" as provisional, he emphasized how the class of all ordinal numbers is unlimited and like Cantor, anticipated Replacement as following from the meaning of the terms (p. 440): "Any class of which the individuals can be correlated one to one with the elements of an aggregate is itself an aggregate." Although Harward would remain obscure, he probably has the distinction of being the first to have anticipated Replacement in print. See Moore [1976] for more about Harward.

[14]See Kanamori [2004] for more on Zermelo.

Replacement as axiom. The thrust of [1915] is that for any set x there is a well-ordered set y not injectible into x. Cardinality Comparability—that for any two sets, one is injectible into the other or vice versa—then implies that x is injectible into y and hence is well-orderable. Thus, Cardinal Comparability implies the Axiom of Choice, a first "reverse mathematics" result establishing the equivalence among the Axiom of Choice, the Well-Ordering Theorem, and Cardinal Comparability.

In what follows, we give the spine of Hartogs' proof to better discuss how he proceeded with the details: Given a set x, let W consist of the well-orderings of subsets of x. Let E be the equivalence relation defined on W by $r\ E\ s$ exactly when r is order-isomorphic to s. Then $<$ on E-equivalence classes is well-defined by: $[r]_E < [s]_E$ exactly when s has an initial segment in $[r]_E$. Moreover, $<$ is a well-ordering of $\{[r]_E \mid r \in W\}$. Finally, this set cannot be injectible into x, else the range of the injection would have a well-ordering belonging to W and $<$ would be order-isomorphic to an initial segment of itself, which is a contradiction.

This argument, amounting to a positive subsumption of the Burali-Forti paradox, would have to be properly implementable without the Axiom of Choice in order to effect the implication from Cardinal Comparability to Choice. To affirm this, Hartogs emphasized implementation in the Zermelo [1908b] system without Choice. This early use of Zermelo's axiomatization is resonant with Zermelo's own, initial use to buttress his proof of the Well-Ordering Theorem.

Pursuing the above sketch, the first implementation issue is how to render a well-ordering and thus to have the set W. Norbert Wiener [1914] and Felix Hausdorff [1914, pp. 32ff, 70ff] rendered the ordered pair, and with this a theory of relations and functions can be developed in set theory. Hartogs was quite unlikely to be aware of this, but he had become aware of a known approach for rendering well-ordering with sets. In his appendix Hartogs acknowledged being informed by Hessenberg that well-orderings can be rendered in Zermelo's axiomatization through systems of final segments—which is what in fact Zermelo had done in his second proof [1908a] of the Well-Ordering Theorem. Thus, with Power Set and Separation, one can take W as a set, being a subset of $P(P(x))$ given by a definite property. Proceeding, once the equivalence relation E is formulated, $\{[r]_E \mid r \in W\}$, though an instance of the Replacement motif, can be shown to be a set by Separation and Power Set, separating from $P(W)$, and $<$ can be defined to complete the rendition in Zermelo's system without Choice.

But how is the equivalence relation E to be formulated? Here's the rub. Hartogs noted (p. 438): "The task of checking all notions and theorems ... is made somewhat difficult since the axiomatic presentation [Durcharbeitung] of set theory given by Zermelo does not yet extend to the theory of ordered and well-ordered sets." In particular, the move beyond Cantor's ordinal numbers as type to the concept of order-isomorphism between well-ordered sets, necessary to formulate E, had yet not been made in the axiomatic presentation. It is here, the situation elaborative of indifference to identification, that Replacement can serve. Indeed, in some modern accounts of Hartogs' theorem one simply uses Replacement to associate with each member of W the corresponding (von Neumann) ordinal and then takes the supremum to complete the argument. However, this would be how one would proceed on the other side of this historical cusp for Replacement, which is not actually necessary for the proof. Without an ordered pair and a theory of relations, Zermelo [1908b] did develop a "theory of equivalence" sufficient for rendering the Cantorian theory of cardinality; this theory was based on getting disjoint copies of two sets and in the union of the copies using the unordered pair to render one-to-one correspondence. Such a theory can be developed for rendering the Cantorian theory of ordinality with order-isomorphisms between Hessenberg-style well-ordered sets, thereby effecting a formalization of Hartogs' theorem and thus establishing that Cardinal Comparability implies Choice. Hartogs did not see his argument through to the end, but the techniques were available to do so at the time, techniques that did not require Replacement.

The set-theoretic work of Dimitry Mirimanoff [1917a], [1917b] sits interestingly intermediate between that of Cantor and of von Neumann. As with Hartogs, Mirimanoff was publishing at a time when Zermelo's axiomatization was not widely called upon and the reduction of ordered pair and function to sets was only beginning. His work goes in and out of the modern theory, but in any case led to a *published* appearance of a form of Replacement, one like Cantor's based on equivalence.

In connection with Russell's Paradox, Mirimanoff [1917a] formulated the *ordinary sets* as those sets x for which every descending $\in$-chain $\dots x_2 \in x_1 \in x$ is finite. These mediated by the Axiom of Choice are, of course, what we now call the well-founded sets. In connection with the Burali-Forti Paradox, Mirimanoff formulated what can be seen to be the (von Neumann) ordinals.[15] He specifically started with an urelement[16] e, getting at what we here call the e-ordinals, and pictured the first three as:

$$(e); (e, (e)); (e, (e), (e, (e))).$$

He (pp. 45–46) motivated these by considering a well-ordering, replacing its members by initial segments, then replacing these by the set of their

[15]Zermelo was most probably the first chronologically to have formulated the concept of (von Neumann) ordinal, and this by 1915; the rudiments of the theory appear in his *Nachlass* (cf. Hallett [1984, pp. 277ff]) and indications are there of collaboration with Paul Bernays (cf. Ebbinghaus [2007, 3.4.3.]).

[16]Urelements, also called atoms or individuals, are objects distinct from the empty set yet having no member. Zermelo allowed for urelements in his "domain" for his [1908b] axiomatization. Mirimanoff used the term node [noyau].

initial segments, and so forth. Although he thus took every well-ordering to be order-isomorphic to an e-ordinal, the envisioned, repeated replacement process can only get to an e-ordinal for *finite* well-orderings. In [1917b], Mirimanoff more assuredly described how to correlate a well-ordering with an e-ordinal by progressively replacing the ordering relation by membership from the front. But even then, the pivotal von Neumann result that every well-ordering is order-isomorphic to a (von Neumann) ordinal can only be rigorously proved by a recursion requiring Replacement. On the other hand, Mirimanoff [1917a, p. 47] offered a characterization of e-ordinal as those ordinary sets with e that are transitive and linearly ordered by membership. Such a characterization for ordinals does not appear in von Neumann's work, and has been attributed much later to Raphael Robinson [1937] as a simplification.[17] In [1917b], Mirimanoff elegantly developed the theory of ordinals much as it is done today. In Mirimanoff's terms, a set [ensemble] can exist or not; the set of all "ordinal numbers of Cantor" does not exist because of the Burali-Forti Paradox, but no decision is made on whether non-ordinary [extraordinary] sets exist or not.

Mirimanoff [1917a] proceeded to his solution of the "fundamental problem" of when sets exist for the case of ordinary sets. To this end he stated three "postulates" corresponding to Union, Power Set, and Replacement. While the first two stated that the ordinary sets are closed under the taking of unions and power sets respectively, the latter stated that *if any set exists, then so does any (cardinally) equivalent ordinary set*. The first 'set' is not qualified, and how he initially applied this principle brings out the conceptual distance from how we now work with Replacement. With his formulation Mirimanoff concluded forthwith that for each ordinal number of Cantor the corresponding e-ordinal exists, and with this he established as a preliminary result that a set of e-ordinals exists exactly when their corresponding ordinal numbers are bounded. Mirimanoff's Replacement thus served as a *trans-species* bridge between "the ordinal numbers of Cantor" and their set representations as e-ordinals, numbers and sets being regarded as of different species. This is interestingly alien to the present sense of Replacement as an axiom within set theory proper but is also consonant with Replacement's role in shifting between tokens and types and Cantor's original use.

For the solution of the "fundamental problem" for ordinary sets in general, Mirimanoff briefly described what can be seen as the cumulative hierarchy indexed by the ordinal numbers, and showed that an ordinary set exists exactly when it has a rank in this hierarchy. In his argument his several uses of Replacement are resonant with how it now serves to affirm the recursive definition of the cumulative hierarchy; however, there is steady coordination,

[17] Bernays in a letter of 3 May 1931 to Gödel actually provided the first direct definition of ordinal, as a transitive set each of whose members is transitive. This letter will be discussed in some detail in §3.

trans-species, with the "ordinal numbers of Cantor". The extent and details of Mirimanoff's work were presumably not fully appreciated in the next, formative years for set theory, and one can only speculate as to why.[18]

Replacement as an axiom henceforth to be reckoned with for the axiomatization of set theory first emerged in 1921 correspondence between Zermelo and Abraham Fraenkel.[19] Fraenkel was the first to investigate Zermelo's 1908 axioms with respect to possible independences, and their first exchange had to do with this. In a letter of early May 1921, Fraenkel astutely raised what would be a pivotal issue: With $Z_0 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots\}$ being the infinite set given by Zermelo's axiom of infinity and $Z_1 = P(Z_0)$ the power set, $Z_2 = P(Z_1)$, and so forth, how can Zermelo's axioms establish that $\{Z_0, Z_1, Z_2, \dots\}$ is a set? Without the union of this set, the existence of sets of cardinality $\aleph_\omega$ would not be provable.

Zermelo wrote back forthwith on 9 May (cf. Ebbinghaus [2007, p.137]):

> Your remark concerning $Z^* = Z_0 + Z_1 + Z_2 + \cdots$ seems to be justified, and I missed this point when writing my [axiomatization paper [1908b]]. Indeed, a new axiom is necessary here, but which axiom? One could try to formulate it as follows: If the objects $A, B, C, \dots$ are assigned to the objects $a, b, c, \dots$ by a one-to-one relation, and the latter objects form a set, then $A, B, C, \dots$ are also elements of a set M. Then one only needed to assign $Z_0, Z_1, Z_2 \dots$ to the elements of the set Z_0, thus getting the set $\Theta = \{Z_0, Z_1, Z_2, \dots\}$ and $Z^* = \bigcup \Theta$. However, I do not like this solution. The abstract notion of assignment it employs seems to be not "definite" enough. Precisely this was the reason for trying to replace it by my "theory of equivalence" [from [1908b]]. As you see, this difficulty is still unsolved. Anyway, I appreciate your having brought it to my attention.

Thus, almost immediately after being confronted with Fraenkel's example Zermelo had formulated Replacement as Cantor had earlier in correspondence, in terms of cardinal equivalence. Zermelo's expressed skepticism about the "abstract notion of assignment" not being " 'definite' enough" has to do with his definite property for the Separation Axiom, perhaps because of the recursive aspect of $\{Z_0, Z_1, Z_2, \dots\}$. The "abstract notion of assignment" has a resonance in Fraenkel's letter of 19 May, in which he ostensibly argued that the new axiom implies the Axiom of Choice, by replacing each set of a system of non-empty, pairwise disjoint sets by a member. This

[18] There seems to have been only two citations of his work, in Fraenkel [1922, p. 233] and Neumann [1925, p. 230], and these only in connection with the extraordinary sets. Mirimanoff had published in the French Swiss journal, *L'Enseignement Mathématique*, which may not have been readily seen in Germany in the difficult period after the war.

[19] See Ebbinghaus [2007, 3.5] for this correspondence and about Replacement; what follows draws substantially from this source.

first unpracticed reaction brings out the conceptual difficulties at the time about what should go for functional correlation.

Soon afterward, in a paper [1922] completed 10 July 1921, Fraenkel described the shortcoming he had found in Zermelo's axioms and introduced: "Axiom of Replacement. If M is a set and each element of M is replaced by 'a thing in the domain', then M turns into a set again." By a "thing in the domain" Fraenkel was referring to either a set or urelement in Zermelo's domain for his axioms. This is the first occasion when injectivity for "replacing" is no longer assumed, but on the other hand what the process of "replacing" is is left undeveloped. On 22 September, at the 1921 meeting of the Deutsche Mathematiker-Vereinigung at Jena, Fraenkel [1921] announced his results, and it was reported that Zermelo at the meeting accepted the axiom but voiced reservations about its scope.

On 6 July 1922, at the Fifth Congress of Scandinavian Mathematicians at Helsinki, Skolem [1923b] delivered an address on the axiomatization of set theory. Skolem actually sought to devalue axiomatic set theory, considering his most important result, as he remarked at the end, to be that set-theoretic notions are relative. However, his analysis not only featured the now well-known "Skolem's paradox" but was remarkably prescient and far-ranging about how to proceed in set theory. Skolem argued that Zermelo's system should be formalized in first-order logic, so that in particular his Axiom of Separation becomes a schema with first-order formulas rendering Zermelo's "definite" property. Skolem (in 4.) also pointed out the same deficiency that Fraenkel had, that $\{Z_0, P(Z_0), P(P(Z_0)), \dots\}$ cannot be proved to be a set, but moreover provided a semantic argument, which in modern terms amounts to the argument that the rank $V_{\omega+\omega}$ is a model of Zermelo's axioms without the set in question. Skolem then wrote (cf. Heijenoort [1967, p. 297]): "In order to remove this deficiency of the axiom system, we could introduce the following axiom: *Let U be a definite proposition that holds for certain pairs (a, b) in the domain B; assume, further, that for every a there exists at most one b such that U is true. Then, as a ranges over the elements of a set M_a, b ranges over all elements of a set M_b.*" Taken in Skolem's context as describing a first-order schema of propositions about the domain of all sets, this is the first substantively accurate statement of the modern axiom schema of Replacement.[20] In a succeeding footnote, Skolem sketched how the new axiom actually establishes that $\{Z_0, P(Z_0), P(P(Z_0)), \dots\}$ is a set, going through a proof for this instance of (Dedekind's) Recursion Theorem. One sees here the first clear account of Replacement and its role in a recursion.

[20]Actually, there is a slight variance. Modern Replacement would have "exactly one b" instead of Skolem's "at most one b". Skolem's stronger formulation can differ from Replacement for restricted versions of the schemas. For this, see Mathias [2001b, sect. 9, peroration]; its 9.32 also translates the Fraenkel–Skolem observation into an algebraic one nearer to the interests of many mathematicians.

Reviewing Skolem's paper, Fraenkel[21] simply stated that Skolem's considerations about Replacement correspond to his own, though evidently, Fraenkel's initial formulation conveyed only a loose idea of replacing. Later Fraenkel [1925, p. 271] did attempt a more formal account specifying the applicable replacing procedures but did not pursue this, opining that Replacement was too strong an axiom for "general set theory". Fraenkel's reservation now seems quite remarkable, and it is also to be noted that Skolem only wrote that "we could introduce" Replacement (in the above quotation). Whatever is the case, Fraenkel's name is now, of course, firmly entwined with Replacement, and this presumably has to do with the thrust of his publications on the axiomatics of set theory as well as the acknowledgments of those engaged with Replacement, von Neumann and Zermelo. The first reference to "Fraenkel's Axiom" appears in Neumann [1923, p. 347] and the first to "Zermelo–Fraenkel" as the axioms of Zermelo augmented with Replacement appears in von Neumann [1928a, p. 374], soon to be followed by Zermelo [1930b, p. 30]. It would be the work of von Neumann which would firmly establish Replacement as an important and needed axiom and orient set theory toward the first-order formulation first advocated by Skolem.

§3. Middle history of Replacement. Von Neumann effected a counter-reformation of sorts for ordinal numbers. The ordinal numbers had been focal to Cantor as separate entities from sets but peripheral to Zermelo with his emphasis on set-theoretic reductionism; von Neumann reconstrued then as *bona fide* sets, the (von Neumann) ordinals. In connection, von Neumann formalized transfinite recursion, this as he worked out a new axiomatization of set theory. Von Neumann's axiomatization was the first, via his I-objects and II-objects, to allow proper classes, as we would now say, together with sets; the paradoxes were systemically avoided by having only sets be members of classes. In all this work Replacement was a crucial feature from the first.

To briefly set the stage for transfinite recursion, emergent from Cantor's work with transfinite numbers are transfinite induction and transfinite recursion, the first a mode of proof and the second a mode of definition. Transfinite induction on a well-ordering is essentially just a contraposition of the main, least element property of well-ordering. Transfinite recursion, on the other hand, depends on having sufficient resources. In Whitehead and Russell's *Principia Mathematica*, volume 3 [1913], transfinite induction and transfinite recursion were articulated to the authors' specific purposes, the second workable in the theory of types context. In Hausdorff's *Grundzuge der Mengenlehre* [1914], V§5, transfinite induction and transfinite recursion were just presented as working principles. It would be the axiomatic formalization of transfinite recursion in set theory that draws in Replacement.

[21] *Jahrbuch über die Fortschritte der Mathematik* vol. 49 (1922), pp. 138–139.

Now to progressively describe von Neumann's work, he in [1923] proceeded informally, in part to bring out the lack of dependence on any particular axiomatic system, to set out the concept of ordinal, working the basic idea of taking precedence in a well-ordering just to be membership.[22] A Cantorian ordinal number had often been construed as the order type of the set of its predecessors, and von Neumann just forestalled the abstraction to order type. With this in hand, he set out the fundamental result that every well-ordering is order-isomorphic to an ordinal. In [1928b], von Neumann duly formalized the proof, making use of Replacement, and one sees here the first and basic use of Replacement to articulate canonical token representing type. With Replacement he moreover established the fundamental Transfinite Recursion Theorem, that given any class function F in two variables, there is a unique class function G on the ordinals such that $G(\alpha) = F(\alpha, G{\restriction}\alpha)$. As with Dedekind's Recursion Theorem (cf. §1) one pieces together initial segments; one proves by transfinite induction that for each ordinal β, there is exactly one function g_β with domain β and satisfying for all $\alpha < \beta$ that $g_\beta(\alpha) = F(\alpha, g_\beta{\restriction}\alpha)$. The shift from Dedekind to von Neumann is the shift from the finite to the transfinite, and Replacement is just what is needed to establish the inductive existence of g_β for limit ordinals $\beta > 0$, in the same way that for Dedekind's Recursion Theorem the existence of the desired function on the natural numbers is secured. The bootstrapping would have been subtle then but is straightforward now, and again it is the infinite in actuality that draws in the Replacement motif. Von Neumann moreover applied transfinite recursion to define (proper) classes, proceeding forthwith as Dedekind had done for the natural numbers to the definition of the ordinal arithmetical operations. The key method of modern set theory is transfinite recursion, and von Neumann thus established the intrinsic necessity of Replacement.[23]

[22]Von Neumann became aware of Zermelo's anticipation of the theory of ordinals; see Hallett [1984, p. 280] and Ebbinghaus [2007, p. 134].

[23]There is a local version of transfinite recursion provable without Replacement: If β is an ordinal, A is a set, and $f\colon P(A) \to A$, then there is a unique function $g\colon \beta \to A$ satisfying $g(\alpha) = f(g``\alpha)$. With a formalization of function as a set of ordered pairs one can apply Separation to the power set $P(\beta \times A)$ to get the set of approximating functions and then take the union. This approach is indeed analogous to Zermelo's first, [1904] proof of the Well-Ordering Theorem. With this local transfinite recursion, the attempt to define ordinal addition $\gamma + \alpha$ to a fixed γ and for α up to a specified β encounters the difficulty that $\gamma + \alpha$ may surpass β.

Notably, Potter [2004, p.183] presents this local version of transfinite recursion in his setting without Replacement. But his proof slides into an implicit appeal to Replacement even though the Power Set, Separation, Union approach was available to him. Moreover, with his theory of ordinal numbers based on equivalence classes of well-orderings he subsequently defines ordinal addition, not recursively because of the above mentioned difficulty, but directly by putting well-orderings in series. But then, he gives (p. 194) the recursion equations as if they were immediate consequences. However, the very formulation of the limit case

Many years later Fraenkel retrospectively acknowledged the importance of Replacement as brought out by von Neumann. Writing about his [1922], Fraenkel [1967, pp. 149ff] wrote: "I did not immediately recognize or take advantage of the full significance of my new axiom. Rather, this was done by von Neumann … " Fraenkel subsequently wrote (p. 169): "The importance of this axiom was now shown in a wholly unexpected way: von Neumann based the theory of transfinite numbers on my axiom in a way which showed that it is indispensable for this purpose."

Von Neumann's axiomatization of set theory, completed by 1923 for his eventual Budapest thesis and aired in [1925] and in full detail in [1928a], has several interactions with Replacement. He actually axiomatized the notion of function rather than the notion of set, noting [1925] (§2) that "every axiomatization uses the notion of function" and citing Replacement. Considering a rigorization of Zermelo's concept of "definite" property to be one of the accomplishments of his axiomatization, von Neumann proceeded to present his axioms for generating classes sufficient so that all definite properties can be correlated with classes according to his Reduction Theorem (§3), essentially the predicative comprehension schema, asserting that extensions of formulas without class quantifiers are classes. This has as a consequence a rigorization of Replacement as put forth by Fraenkel, which von Neumann took (§2, last footnote) in the class form: If F is a (class) function and a is a set, then the image $F``a$ is a set.

Replacement however is not actually an axiom of von Neumann's axiomatization. Rather, the taking of function as primitive was in part to the purpose of formulating his focal axiom IV 2, stated here in terms of sets and classes with V the class of all sets: *A class A is not (represented) by a set exactly when there is a surjection of A onto V.* Von Neumann thus transformed the negative concept of proper class, which had appeared in various guises, e.g., Cantor's inconsistent multiplicities, into the positive concept of having a surjection onto V. In fact, von Neumann had formalized a dichotomy based on possibilities for well-ordering that had been broached by Cantor in his letter of 3 August 1899 to Dedekind (cf. §2). IV 2 is an existence principle that plays the role of regularizing proper classes, much as the Axiom of Choice does for sets, by extending the Cantorian canopy of functional correspondence. IV 2 appropriately implies both Replacement and Choice in class forms, the latter asserting the existence of a global choice function on V.[24]

requires justification, ordinarily given with Replacement, but here possible with Power Set, Separation, and Union. It is hard to sublimate Replacement.

[24]For Replacement, if F is a class function whose range is not a set, then there is a surjection G of that range onto V; but then F composed with G is a surjection of the domain of F onto V, and consequently the domain is not a set. For Choice, since the class of all ordinals cannot be a set by the Burali-Forti argument, there is a surjection of the class onto V, and an inversion according to least preimages induces a well-ordering of V itself, and so there is a global choice function.

Both in his development of the ordinals and with his axiomatization, von Neumann rigorized and provided for the full extent of Cantor's later vision of set theory as set out in correspondence with Hilbert and Dedekind at the turn of the century, and Replacement was a basic underpinning for this.

In his last paper on axiomatization von Neumann [1929] provided an incisive analysis of his system by providing as a model the class of well-founded sets, an analysis that showed his axiom IV 2 to have a concrete plausibility. In modern terms, he in effect defined by transfinite recursion the *cumulative hierarchy* of well-founded sets through their stratification into cumulative "ranks" V_α, where

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

Mirimanoff [1917a, pp. 51ff] had been first to study the well-founded sets, and the cumulative hierarchy is distinctly anticipated in his work (cf. §2). In the axiomatic tradition Fraenkel [1922], Skolem [1923b], and von Neumann [1925] had considered the salutary effects of restricting the universe of sets to the well-founded sets. Von Neumann [1929] formulated in his functional terms the Axiom of Foundation, asserting that every non-empty class has a $\in$-minimal element, and observed that it is equivalent to the assertion that the cumulative hierarchy *is* the universe, $V = \bigcup_\alpha V_\alpha$. Moreover, he observed that in the presence of Foundation, Replacement and Choice (in class forms) are together actually *equivalent* to IV 2.[25,26] This result has the notable thematic effect of localizing the thrust of IV 2 to asserting the existence of just *one* class, a choice function on the universe. To conclude, von Neumann established that his axioms hold in the class of well-founded sets, thereby establishing the relative consistency of Foundation. The result is cited now as the first relative consistency result via "inner models" and about Foundation, but for von Neumann a major incentive was to affirm the plausibility of his axiom IV 2.

As for the underlying logic, von Neumann's work is evidently formalizable in a two-sorted first-order logic with variables for sets and variables for classes. However, as with his Reduction Theorem ([1925] §3), i.e., the predicative comprehension schema, quantification over classes is delimited and the full potency of second-order logic for sets is never invoked. As subsequently emended by Bernays and Gödel, this first-order aspect promoted the move toward Skolem's [1923b] suggestion of basing set theory on first-order

[25]To get IV 2, for any class A, A is the union of the layers $A \cap (V_{\alpha+1} - V_\alpha)$ by Foundation. If A is not represented by a set, then these layers are nonempty for arbitrarily large α by Replacement. But each such layer has a well-ordering by Choice, and these well-orderings can be put together to well-order all of A, again by Choice. Hence, there is a one-to-one correspondence between A and the class of all ordinals.

[26]Levy [1968] latterly showed by a clever argument that the Union Axiom for sets also follows from von Neumann's IV 2, so that IV 2 is equivalent to Replacement and Choice (in class forms) and Union (for sets) in the presence of the other axioms.

logic, and in particular, on regarding Replacement as a first-order schema of axioms.

Zermelo in his remarkable [1930b] offered his final axiomatization of set theory as well as a striking, synthetic view of a procession of natural models. This axiomatization incorporated for the first time both Replacement and Foundation (Zermelo's term), and thus the now-standard ZFC axiomatization is recognizable. However, Zermelo worked in a full second-order context, proceeding in effect with proper classes but not distinguishing a concept of class. Zermelo was actually presenting a dramatically new view of set theory as applicable through processions of models, each having a "basis" of urelements for some specific application and a cumulative hierarchy built on the basis, up to a height given by the "characteristic". He established these "*Grenzzahlen*" to be either the (von Neumann) ω or a strongly inaccessible cardinal and moreover established a second-order categoricity of sorts for his models as determined up to isomorphism by the cardinal numbers of the basis and of the characteristic. Significantly, Zermelo pointed out (p. 38) that those models starting with one urelement satisfy von Neumann's axiom IV 2. However, in Zermelo's approach IV 2 would not always hold, this depending on the cardinality of the basis of urelements.[27]

Whatever the relationship to von Neumann's IV 2 and work on Foundation, Zermelo's adoption of both Replacement and Foundation promoted the modern mathematical approach to set theory. In modern, first-order set theory, Replacement and Foundation focus the notion of set, with the first making possible the means of transfinite recursion, and the second making possible the application of those means to get results about *all* sets, they now appearing in the cumulative hierarchy.

Zermelo himself, however, was haphazard on the methodological role of Replacement. First, like Mirimanoff [1917a] Zermelo worked with ordinals starting from an urelement; did not make the von Neumann identification of the Cantorian ordinal numbers with ordinals; and just stated (p. 33) the von Neumann result that every well-ordering is isomorphic to an ordinal without mention of the role of Replacement. Then in his characterization (p. 34) of the characteristics of his models as the inaccessible cardinals, Zermelo did not explicitly associate his ordinals with the subsets of a set, and when finally he appealed explicitly to Replacement it is to a limit case made redundant by a previous assertion. For Zermelo, the importance of Replacement resided in its role in cofinality, as attested to by his mention of Hausdorff's work (pp. 33,34). Getting to his categoricity results, Zermelo established (p. 41) his first isomorphism result, that two models with the same characteristic and bases of the same cardinality are isomorphic, by extending a one-to-one correspondence between the bases through the two cumulative hierarchies. Unbridled second-order Replacement is crucial here, this time to establish

[27]See Kanamori [2004, p. 525] and generally for Zermelo's work in set theory.

that the cumulative ranks of the two models are level-by-level extensionally correlative, and this from one perspective requires a *trans-species* correlation across models. However, Zermelo did not mention Replacement at all here, nor indeed does he ever point out how Replacement is implicated in definitions by transfinite recursion. As mentioned above (cf. §1) there is a historical resonance with Dedekind [1888] who had established the categoricity of *his* second-order axioms for arithmetic. A final involvement of Replacement is in Zermelo's postulation of an "unlimited sequence of Grenzzahlen" and how one model "can also be conceived of as a 'set' " in a further model (p. 46). This entails a "reflection principle" which can be seen in terms of later developments (cf. §4) as amounting to a strong form of Replacement.[28]

In a note "On the set-theoretic model" [1930a] found in his *Nachlass*, Zermelo provided an analysis of set theory and a motivation for its axioms based on an "iterative conception".[29] He first set out his "set-theoretic model", and it is the cumulative hierarchy based on a totality of urelements. With this as a schematic picture he proceeded to motivate and justify his [1930b] axioms. This note reveals Zermelo to be actually the first who would retroactively motivate the axioms of set theory in terms of an iterative conception of sets as built up through stages of construction. For Replacement, Zermelo simply argued that, in modern terms, if a set's elements are replaced by members of a V_β, then the result would appear in $V_{\beta+1}$. Of this evidently circular argument, one of course bypassing any modern concern about definability, Zermelo wrote: "Of course, the assumption that the replacing elements belong to a *segment* of the development [i.e., a V_β], while being essential here, constitutes no real restriction." With this he was presumably importing his [1930b] picture of always having Grenzzahlen beyond, but circularity is still there at this further remove.

As set theory would develop, Replacement, Foundation, and the cumulative hierarchy picture would provide the setting for a developing high tradition that had its first milestone in Kurt Gödel's development of the constructible universe L. Zermelo [1930b] would be peripheral to this development, presumably because of its second-order lens and lack of rigorous detail. Gödel's work newly confirmed Replacement as a central axiom of set theory, and it was featured in a formal presentation of this work for which Gödel adapted an axiomatization of set theory due to Bernays, which itself was a transmutation of von Neumann's axiomatization.

§4. Later history of Replacement. Gödel's advances in set theory can be seen as part of a steady intellectual development from his fundamental

[28]The reflection aspect of Zermelo [1930b] was emphasized by William Tait [1998].

[29]For more on this note and its significance, see the author's introductory note appearing in Zermelo [2010].

work on completeness and incompleteness. In the well-known, prescient footnote 48a to his celebrated incompleteness paper [1931], Gödel had pointed out that the formation of ever higher types over the Russellian theory of types can be continued into the transfinite and that his undecidable propositions become decided if the type ω is added. Matters in a footnote, perhaps an afterthought then, took up fully one-third of a summary (cf. Gödel [1932, pp. 234ff]) dated 22 January 1931. Gödel pointed out that the enlargement of Z, first-order Peano arithmetic, with higher-order variables and corresponding comprehension axioms leads not only to the decidability of his undecidable propositions but also to new, undecidable propositions, all expressible in Z. He continued:

> In case we adopt a type-free construction of mathematics, as is done in the axiom system of set theory, axioms of cardinality (that is, axioms postulating the existence of sets of ever higher cardinality) take the place of type extensions, and it follows that certain arithmetic propositions that are undecidable in Z become decidable by axioms of cardinality, for example, by the axiom that there exist sets whose cardinality is greater than every α_n, where $\alpha_0 = \aleph_0$, $\alpha_{n+1} = 2^{\alpha_n}$.

This is Gödel's first remark on set theory of substance, and significantly, his example of an "axiom of cardinality" is evidently closely connected to the existence of the set that both Fraenkel [1922] and Skolem [1923b] had pointed to as the one to be secured by adding Replacement to Zermelo's 1908 axiomatization.

In an incisive lecture [1933o] Gödel expanded on his theme of higher types. He propounded the view that the axiomatic set theory "as presented by Zermelo, Fraenkel and von Neumann ... is nothing else but a natural generalization of the [simple] theory of types, or rather, it is 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. That is, with S_n to consist of the sets of type n newly construed, S_0 consists 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 is seen here as promoting the cumulative hierarchy picture, which had been advocated for set theory by Zermelo [1930b], [1930a], as an extension of the simple theory of types.

As for how far this cumulative hierarchy of sets is to continue, Gödel [1933o, p. 47] wrote:

> The first two or three [transfinite] types already suffice to define very large ordinals. So you can begin by setting up axioms for these

> first types, for which purpose no ordinal whatsoever is needed, then define a transfinite ordinal α in terms of these first few types and by means of it state the axioms for the system, including all classes of type less than α. (Call it S_α.) To the system S_α you can apply the same process again, i.e., take an ordinal β greater than α which can be defined in terms of the system S_α and by means of it state the axioms for S_β including all types less than β, and so on.

Gödel thus envisioned an "autonomous progression", in later terminology, with large ordinals definable in low types leading to higher types.

Gödel, according to Hao Wang [1981, p. 128] reporting on conversations with Gödel in 1976, had already been working on the continuum problem for some time and had devised what he considered to be a transfinite extension of Russell's *ramified* theory of types. He in effect had started working up the constructible hierarchy, where in modern terms, for any set x, $\mathrm{def}(x)$ is the collection of subsets of x definable over $\langle x, \in\rangle$ via a first-order formula allowing parameters from x, and:

$$L_0 = \emptyset;\ L_{\alpha+1} = \mathrm{def}(L_\alpha);\ \text{and } L_\delta = \textstyle\bigcup_{\alpha<\delta} L_\alpha \text{ for limit ordinals } \delta.$$

There is evidence however that Gödel could not get far pursuing his envisioned autonomous progression. Wang [1981, p. 129] reported how in Gödel's efforts to go up the constructible hierarchy he "spoke of experimenting with more and more complex constructions [of ordinals for indexing] for some extended period somewhere between 1930 and 1935." More pointedly, Georg Kreisel in his memoir [1980] of Gödel wrote (p. 193): "As early as 1931, Gödel alluded to some reservations [about Replacement]." Kreisel continued (p. 196, with $C_{\omega+\omega}$ his notation for $V_{\omega+\omega}$): "In keeping with his reservations, mentioned on p. 193, Gödel first tried to do without the replacement property, and to describe the constructible hierarchy L_α only for $\alpha <$ card $C_{\omega+\omega}$; in particular, without using von Neumann's canonical well-ordering [i.e., ordinals]. Instead, well-orderings had to be defined (painfully) in $C_{\omega+\omega} \ldots$"

The full embrace of Replacement led to a dramatic development. Set theory reached a new plateau with Gödel's formulation of the class $L = \bigcup_\alpha L_\alpha$ of *constructible* sets with which he established the relative consistency of the Axiom of Choice in mid-1935 and of the Continuum Hypothesis (CH) in mid-1937.[30] Gödel had continued the indexing of his hierarchy through *all* the ordinals as given beforehand to get a class model L of set theory and thereby to achieve *relative* consistency results, by showing that L satisfies Choice and CH. His early idea of using large ordinals defined in low types in a bootstrapping process would not suffice. Von Neumann's ordinals would

[30]See Dawson [1997, pp. 108,122]; in one of Gödel's *Arbeitshefte* there is an indication that he established the relative consistency of CH in the night of 14–15 June 1937.

be the spine for a thin hierarchy of sets, and this would be the key to both the Choice and CH results.

In his monograph [1940], based on 1938 lectures, Gödel provided a specific, formal presentation of L which made evident the methodological importance of Replacement. Gödel relied on an axiomatization of a class-set theory which, save for an economy of presentation, was the one detailed to him in a letter from Bernays of 3 May 1931.[31] Bernays' axiomatization itself was a transmutation of von Neumann's axiomatization which differed from it in two respects. First, while von Neumann had taken function as a primitive notion with his I-objects and II-objects being functions, Bernays reverted to collections, sets and classes. More importantly, Bernays like Zermelo [1930b] adopted Foundation, and with von Neumann's [1929] analysis of his characteristic axiom IV 2 as being equivalent under Foundation to the conjunction of Replacement and Choice in class forms, Bernays adopted the latter two instead. Bernays's approach thus had the effect of bringing von Neumann's axiomatization more in line with Zermelo [1930b]. But unlike in Zermelo's approach, there were no urelements, and as Bernays wrote in that May 1931 letter, "A *complete formalization*, and in fact in a *first-order* [*ersten Stufe*] framework, can be carried out without difficulty." To summarize, the features of Bernays' axiomatization that would commend its further use and influence were that it recast von Neumann's work to present a viable theory starting with sets and classes as primitive notions, and it incorporated Replacement and Foundation, as did Zermelo's later axiomatization, but in a first-order context and without the relativism of having urelements. Gödel had thus become aware of Replacement early on; his full espousal of the axiom in a first-order context affirmed its importance.

Gödel in his monograph carried out a careful development of "abstract" set theory through the ordinals and cardinals with features that have now become common fare. Gödel then used eight binary operations, producing new classes from old, to generate L set by set via transfinite recursion. This veritable "Gödel numbering" with ordinals bypassed the formalization of the $\mathrm{def}(x)$ operation and made evident certain aspects of L. Since there is a direct, definable well-ordering of L, choice functions abound in L, and Choice holds there.

Gödel's proof that L satisfies CH consisted of two separate parts, both depending on Replacement and both to become paradigmatic for inner model theory, that large part of modern set theory with beginnings in Gödel's work. Gödel established the implication $V = L$ implies CH and, in order to apply this implication within L, the absoluteness $L^L = L$—that the construction of L in L again gives L—to establish CH within L.

[31] See Gödel [2003, pp. 105ff]. Gödel in [1940] routinely acknowledged Bernays by citing his later, published account [1937], but it is clear from their correspondence that Gödel had assimilated Bernays' axiomatization through the letter from 1931.

The absoluteness of L depends inherently, albeit subtly, on Replacement. Replacement, affirming the interplay of token and type, bolsters the (von Neumann) ordinals as tokens for all well-orderings. Being an ordinal is absolute for transitive sets, i.e., being an ordinal in the sense of a set closed under membership is really to be an ordinal. With this, one can construct L, e.g., through Gödel's eight binary operations, in the sense of L and again get L. A basic reason why Gödel's early efforts with general well-orderings could not have succeeded is that one needs this absoluteness; one cannot just work with well-orderings or even equivalence classes of these. The absoluteness of ordinals as canonical for well-orderings is a pre-condition for the substantive arguments in L and now, in all inner model theory.

As for Gödel's argument for $V = L$ implies CH, it rests on what he termed an "axiom of reducibility"—which properly can be thought of as the fact that for any α the constructible subsets of L_α all belong to some L_β—and what is now known as a Skolem hull argument—by which that β can be taken to have the same cardinality as α when α is infinite. Replacement is crucial for the "axiom of reducibility". In referring to Russell's ill-fated axiom in his ramified theory of types, Gödel took his version to be a rectification. In his first announcement [1938] he wrote:

> [The] 'constructible' sets are defined to be those sets which can be obtained by Russell's ramified hierarchy of types, if extended to include transfinite orders. The extension to transfinite orders has the consequence that the model satisfies the impredicative axioms of set theory, because an axiom of reducibility can be proved for sufficiently high orders.

In his analysis of Russell's mathematical logic Gödel [1944, p. 147] again wrote about how with L he had proved an axiom of reducibility, emphasizing: " ... all impredicativities are reduced to one special kind, namely the existence of certain large ordinal numbers (or well-ordered sets) and the validity of recursive reasoning for them." Decades later Gödel wrote in a letter of 7 March 1968 to Wang [1974b, pp. 8–9]:

> ... there was a special obstacle which *really* made 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 ordinals for the indexing, and for this their extent had to be sustained by Replacement.

In late conversations, Gödel justified Replacement as follows (Wang [1996, p. 259]; see also Wang [1974a, p. 186]):

> 8.2.15. From the very idea of the iterative concept of set it follows that, if an ordinal number a has been obtained, the operation P of power set iterated a times from any set y leads to a set $P^a(y)$. But, for the same reason, it would seem to follow that, if instead of P, one takes some larger jump in the hierarchy of types, for example, the transition Q from x to the set obtained from x by iterating as many times as the smallest ordinal [not] of the well-orderings of x, $Q^a(y)$ likewise is a set. Now, to assume this for any conceivable jump operation—even for those that are defined by reference to the universe of all sets or by use of the choice operation—is equivalent to the axiom of replacement.

Gödel recalled here the lesson learned about the need for impredicative reference to the universe of all sets. Technically speaking, all that is needed to get Replacement is Gödel's $P^a(y)$'s together with the assertion that the class of all ordinals is regular with respect to (definable) class functions.[32]

With the "axiom of reducibility" Gödel had broached a new aspect of Replacement, that of reflection. Recalling Cantor's Absolute but cast in terms of the cumulative hierarchy, the heuristic of reflection surrounds the idea that the universe $V = \bigcup_\alpha V_\alpha$ cannot be characterized uniquely by a formula so that any particular property ascribable to it must already be ascribable to some rank V_α. In oral, necessarily brief remarks [1946] at a conference, Gödel voiced what would become a prominent way to motivate and formulate "strong axioms of infinity", now called large cardinal axioms, by reflection: "Any proof for a set-theoretic theorem in the next higher system above set theory", i.e., if the satisfaction relation for V itself were available, "is replaceable by a proof from such an axiom of infinity."

Gödel's L stood as a high watermark for set theory for quite a span of years, and during this period the Bernays–Gödel (BG) class-set theory maintained an expository sway. There was then a shift toward the more parsimonious ZFC set theory, especially after BG and ZFC were shown around 1950 to have the same provable consequences for sets. Concomitantly, Replacement became widely seen, as first envisioned by Skolem [1923b], as the first-order schema that it is taken to be today. Forthwith, new model-theoretic initiatives led to the formalization of reflection properties that put Replacement in a new light.

With the basic concepts and methods of model theory being developed by Tarski and his students at Berkeley, Richard Montague [1961] in his 1957 Berkeley dissertation had studied reflection properties in set theory and had shown that the axiom schema of Replacement is not finitely axiomatizable over the other axioms in a strong sense. Levy [1960a], [1960b] then exploited the model-theoretic methods to establish a broader significance for reflection

[32]That is, there is no such function with domain an ordinal and cofinal in the class of all ordinals.

principles. Sufficient for reflection is ZF, ZFC minus the Axiom of Choice. The ZF *Reflection Principle*, drawn from Montague [1961, p. 99] and Levy [1960a, p. 234], asserts that for any (first-order) formula $\varphi(v_1, \ldots, v_n)$ in the free variables as displayed and any ordinal β, there is a limit ordinal $\alpha > \beta$ such that for any $x_1, \ldots, x_n \in V_\alpha$,

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

where as usual φ^M denotes the relativization of the formula φ to M. The idea is to carry out a Skolem closure argument with the collection of subformulas of φ. Montague showed that this schema holds in ZF, and Levy showed that it is actually equivalent to the axiom schema of Replacement together with the Axiom of Infinity in the presence of the other axioms of ZF. Through this work the ZF Reflection Principle has become well-known as making explicit how reflection is intrinsic to the ZF system and as a new face for Replacement. Levy [1960a] cast the ZF Reflection Principle as motivation for stronger reflection principles, with the first in his hierarchy being the ZF principle but with "limit ordinal α" replaced by "inaccessible cardinal α". This motivated the strongly Mahlo cardinals, and, notably, was implicit in Zermelo [1930b] (cf. §3).

Reflection, and thus Replacement, was seen in a further new light through an axiomatic set theory proposed by Wilhelm Ackerman [1956]. His theory A is a first-order theory that can be cast as follows: There is one binary relation $\in$ for membership and one constant V; the objects of the theory are to be referred to as classes, and members of V as sets. The axioms of A are the universal closures of:

(1) Extensionality: $\forall z(z \in x \longleftrightarrow z \in y) \longrightarrow x = y$.
(2) Comprehension: For each formula ψ not involving t,

$$\exists t \forall z(z \in t \longleftrightarrow z \in V \wedge \psi).$$

(3) Heredity: $x \in V \wedge (t \in x \vee t \subseteq x) \longrightarrow t \in V$.
(4) Ackermann's Schema: For each formula ψ in free variables $x_1, \ldots, x_n, z$ and having no occurrence of V,

$$x_1, \ldots, x_n \in V \wedge \forall z(\psi \longrightarrow z \in V) \longrightarrow \exists t \in V \forall z(z \in t \longleftrightarrow \psi).$$

This last, a comprehension schema for sets, is characteristic of Ackermann's system. It forestalls Russell's Paradox, and its motivation was to allow set formation through properties independent of the whole extension of the set concept and thus to be considered sufficiently definite and delimited.

Ackermann [1956] himself argued that every axiom of ZF, when relativized to V, can be proved in A. However, Levy [1959] found a mistake in Ackermann's proof of Replacement, and whether the schema can be derived from Ackermann's Schema remained an issue. Toward a closer correlation with ZF, Levy came to the idea of working with A*: A together with the Axiom of Foundation relativized to V. As for ZF, Foundation focuses the sets with a stratification into a cumulative hierarchy. Levy [1959] showed that, leaving aside the question of Replacement, A^* establishes substantial reflection principles. On the other hand, he also showed through a sustained axiomatic analysis that for a sentence σ of set theory (so without V): *If σ relativized to V is provable in* A^*, *then σ is provable in* ZF. The thrust of this work was to show that Ackermann's Schema can be assimilated into ZF—a somewhat surprising result—and that ZF and A^* have about the same theorems for sets.

Levy and Robert Vaught in their [1961] later observed by an inner model argument that, as for ZF and Foundation, if A is consistent, then so is A^*. They then went on to confirm that the addition of Foundation to A was substantive; they showed that Ackermann's Schema is equivalent to a reflection principle in the presence of the other A^* axioms, and that A^* establishes the existence of $\{V\}$ and the power classes $P(V)$, $P(P(V))$, and so forth.

Years later, returning to the original issue about Replacement, William Reinhardt [1970] in his 1967 Berkeley dissertation built on Levy–Vaught [1961] to establish for A^* what Ackermann could not establish for A: *Every axiom of* ZF, *when relativized to V, can be proved in* A^*. Thus, A^* and ZF do have exactly the same theorems for sets. Replacement is the central axiom schema to establish in order to get ZF, and once done, Ackermann's set theory exhibited it in a new light.

§5. **Forms of Replacement.** Replacement is central to ZFC, but it also has a complicated form, and as befits the situation it has been analyzed in a variety of ways. In this section we discuss various forms of Replacement, starting from simple modulations and proceeding to more substantive ones. To repeat, we take Replacement to be the following schema: For any formula φ with free variables among $a, x, y, z_1, \dots, z_n$ but not including b,

$$\forall z_1 \forall z_2 \dots \forall z_n \forall a (\forall x \exists ! y \varphi \longrightarrow \exists b \forall y (\exists x (x \in a \wedge \varphi) \longleftrightarrow y \in b)), \quad \text{(Rep)}$$

where $\exists !$ abbreviates the formalizable "there exists exactly one". The Replacement schema easily implies the Separation schema, but in modern presentations of ZFC both are kept by convention, especially as in comparative investigations of subsystems of ZFC one wants to retain versions of Separation.

An initial observation with historical relevance is that Replacement is equivalent to the form requiring the functional correlation to be one-to-one. If F is a class function, then G defined via the ordered pair by $G(x) = \langle x, F(x) \rangle$ is one-to-one. So, for any set a, $G``a = \{\langle x, F(x) \rangle \mid x \in a\}$ is a set by "one-to-one Replacement". But then, one can as usual appeal to Union and Separation to get to the range of this set, i.e., $\{F(x) \mid x \in a\}$.[33]

[33] I do not know whether, without Separation, Replacement is derivable from "one-to-one Replacement".

A syntactic complication of Replacement is that in its formal statement one has to be attentive to passive parameters. Levy [1974] considered the elimination of parameters. Let R_0 (his notation) be the parameter-free version of Replacement, in the sense that Rep is altered so that the only free variables occurring in φ are x, y and the universal quantification $\forall z_1 \forall z_2 \ldots \forall z_n$ is deleted. Let P_2 be the restriction of the Power Set Axiom asserting that for any set x the totality of two-element subsets if x is a set. Levy established that in the presence of Extensionality, Pairing, Union, and P_2, R_0 implies Replacement. Hence, for a quick presentation of ZFC one can rightly give only the parameter-free R_0. Earlier, Harvey Friedman in his 1967 Ph.D. thesis (cf. [1971b]) had also shown that a restricted, "more explicit" group of axioms including a parameter-free version of Replacement entails ZFC, but indirectly in terms of consistency by going through the construction of the constructible universe L.

Levy called $\mathrm{R}^{\leftarrow}$ another variant of Rep where the "$\longleftrightarrow$" is relaxed to "$\longrightarrow$". That is, for any set a one concludes only that there is a set b that subsumes the image $F``a$ under the class function. Of course, with Separation, $\mathrm{R}^{\leftarrow}$ is equivalent to full Replacement. ZFC is sometimes presented with $\mathrm{R}^{\leftarrow}$ instead of Replacement, giving Separation an independent status. Levy showed that Extensionality, Empty Set, Pairing, Union, Power Set, a parameter-free version of Separation, and $\mathrm{R}^{\leftarrow}$ together do *not* imply Replacement. On the other hand, recently Mathias [2007] showed that Extensionality, Empty Set, Pairing, Union, Power Set, Δ_0-Separation, Foundation, Transitive Containment (that every set is a member of a transitive set), and $\mathrm{R}^{\leftarrow}$ *do* imply Replacement. Here, Δ_0-Separation is the Separation schema restricted to Δ_0, or bounded, formulas in the Levy hierarchy, i.e., those formulas of set theory that can be rendered with quantifiers only of form $\forall v \in w$ and $\exists v \in w$.

A more substantive variant of Rep is Collection, where both the "$\longleftrightarrow$" is relaxed to $\longrightarrow$" and the "$\exists ! y$" is relaxed to "$\exists y$". That is, one finally gives up the historically given functionality of φ so that for any x there is some witness y and concludes for any set a that there is a set b containing witnesses for every $x \in a$. Collection emerged in the late 1960s with Replacement having become less explicitly a class function principle and more explicitly a schema of formulas, suggestive of the relaxation of the $\forall x \exists ! y$ to $\forall x \exists y$. In any case, Collection is equivalent over the other axioms to Replacement: To establish Collection from Replacement, suppose that $\forall x \exists y \varphi$ and a is a set. For each $x \in a$, let α_x be the least α such that φ holds for a $y \in V_{\alpha+1}$. With Replacement we can consider the ordinal $\beta = \sup\{\alpha + 1 \mid x \in a\}$, and V_β can serve as the set b to confirm Collection. Power Set is necessary here, in the sense that Andrzej Zarach [1996] showed with a forcing argument that over a base theory without Power Set, Replacement does not imply Collection.

Replacement and Collection are distinct in intuitionistic set theory IZF and in constructive set theory CZF, theories based on intuitionistic logic in which the appeal to least ranks V_α ("Scott's trick") as above is not available. Nicolas Goodman [1985] showed using Kripke models that if one adjoins a class parameter (that is, a new predicate symbol) then intuitionistically one cannot show that Replacement implies Collection. Inspired by this, Harvey Friedman and Andrej Scedrov [1985] showed that intuitionistic set theory as formulated with Replacement, IZF_R, does not prove Collection, which is used in IZF. It would remain open whether IZF_R and IZF have the same proof-theoretic strength. Michael Rathjen [2005] showed that for the weaker CZF, Replacement and a strong form of Collection do have the same proof-theoretic strength.

Refinements of the various schemas have to do with restricting them according to complexity of formula, as with Δ_0-Separation. For the formulations, we briefly recall the standard Levy hierarchy of set-theoretic formulas: $\Sigma_0 = \Pi_0 = \Delta_0$ are the bounded formulas, where again the quantifiers are only of form $\forall v \in w$ and $\exists v \in w$. Then a formula is Σ_{n+1} if it is of the form $\exists v \varphi$ where φ is Π_n, and Π_{n+1} if it is of the form $\forall v \varphi$ where φ is Σ_n. It is to be pointed out that the very efficacy of concept classification according to this hierarchy depends on Replacement, or more directly, Collection. Collection amounts to having

$$\forall v \in w \exists v_0 \varphi \longleftrightarrow \exists v_1 \forall v \in w \exists v_0 \in v_1 \varphi,$$

and this shows how the bounded quantifier $\forall v \in w$ can be inductively absorbed in the complexity analysis. As expected, Σ_n-Replacement refers to the Replacement schema restricted to the Σ_n formulas, and so forth. Π_n-Collection implies Σ_{n+1}-Collection, for if $\exists w \varphi$ is Σ_{n+1} where φ is Π_n, then in $\forall x \exists y \exists w \varphi$ for Collection one can pair y and w. Also, the simple argument getting from Replacement to Separation shows that Σ_{n+1}-Replacement implies Σ_n-Separation.

The prominent set theory with restricted schemas is Kripke–Platek (KP) set theory: Extensionality, Empty Set, Pairing, Union, Δ_0-Separation, and Δ_0-Collection.[34] KP can carry the weight of substantive recursive procedures, and in particular the construction of Gödel's L. On the other hand, Σ_1-Separation is needed over KP to establish the fundamental von Neumann result that every well-ordering of a set is order-isomorphic to an ordinal. Latterly, Mathias [2001b] showed that over KP + Power Set + Choice, Σ_1-Separation is equivalent to every well-ordering of a set being order-isomorphic to an ordinal.[35] Finally, over KP + Power Set, Σ_2-Replacement suffices to define recursively the cumulative hierarchy, as being a V_α is Π_1.

[34]Mathias [2001b] includes Π_1-Foundation in KP.

[35]See Mathias [2001b, 3.18]; this paper has much more on restricted schemas of Replacement and variants. See Mathias [2006] for such schemas in weaker set theories.

The work on restricted set theories and schemas have certainly brought out the logical dependencies of the classical development surrounding Replacement. In particular, they have brought out the crucial aspects of Replacement as functioning separately from Power Set and so from large cardinality sets. At the same time, much of the work has focused attention on Replacement as the principle of set theory to be reckoned with, especially in its providing the basis for well-founded recursion, the importance of method coming to the fore.

§6. Replacement vindicated. In this last, more freewheeling section, we address issues that have invited skepticism, or at least hesitation, about Replacement based largely on grounds of ontological commitment and we provide an affirmation on general grounds for Replacement's central role in set theory. The large mathematical and historical perspective and basis for Replacement provided in the previous sections should already serve to address issues having to do with thematic importance and historical emergence, and so it is that we here turn to the further issues.

As mentioned in §3, in the *Nachlass* note [1930a] Zermelo first motivated the axioms of set theory in terms of the cumulative hierarchy picture. From the late 1960s, what has come to be regarded as *the iterative conception*, conceiving sets as built up through stages of construction, has become well-known as a heuristic for motivating the axioms of set theory generally.[36] This has opened the door to a metaphysical appropriation in the following sense: It is as if there is some notion of set that is "there", in terms of which the axioms must find justification as being true or false. But set theory has no particular obligations to mirror some "prior" notion of set, especially one like the ultimately ontological iterative conception, arrived at after the fact.

When Replacement has been justified according to the iterative conception, the reasoning has in fact been circular as it was in Zermelo [1930a], with some feature of the cumulative hierarchy picture newly adduced solely for this purpose. George Boolos [1971] argued that neither Replacement nor Choice, in their providing for a store of sets, is evident from the iterative conception. Subsequently, he [1989] observed that Replacement and Choice do follow from a certain "limitation of size" conception of set that he called FN; this is evident since FN amounts to an espousal of von Neumann's axiom IV 2 (cf. §3). With the two, the iterative and limitation of size concepts, Boolos ends: "Perhaps one may conclude that there are at least two thoughts 'behind' set theory."

Another thought is that set theory is what it is, a historically given, autonomous field of mathematics proceeding with its own self-fueling methods

[36] Joseph Shoenfield [1967, pp. 238ff] and [1977], Hao Wang [1974a], Dana Scott [1974], and George Boolos [1971], [1989] motivate the axioms of set theory in terms of iterative conceptions.

and procedures. We continue the dialectic with Boolos as the specificities of his writings invite discussion as he intended and provide an opportunity to raise important, general considerations.

In his late "Must we believe in set theory?" [2000], Boolos quite remarkably argued against the existence of the least fixed point $\kappa = \aleph_\kappa$ of the aleph function. According to Boolos,

> The burden of proof should be, I think, on one who would adopt a theory so removed from experience and the requirements of the rest of science (including the rest of mathematics) as to claim that there are κ objects.

Boolos was of course aware that the existence of κ is established with Replacement, and is inclined to jettison Replacement as not following from the "natural" iterative conception.

First, "belief" is a vague notion, and there is little one can profitably discuss about it as a concept, especially in mathematics and as concerns ontology.

Second, there are varieties of "experience". Though Boolos anticipates the charge that asking whether κ exists from an external vantage point is to fall into metaphysical error, this counterpoint is nonetheless quite relevant. From an external vantage point κ could seem a magic mountain to the foot soldier encumbered with dated provisions and determined to climb at some previously ordained pace. For Henri Poincaré $\aleph_1$ was inaccessible, and to one who takes successive counting to be the overwhelming primal act, 2^{100} is analogously inaccessible. But for Cantor, the new transfinite landscape was set out with generating principles for which the climb to κ would have been congenial. And Hausdorff subsumed κ in his schematic investigation of the uncountable transfinite.

κ is approachable in set theory via $\kappa_0 = \aleph_0$, $\kappa_{n+1} = \aleph_{\kappa_n}$, and $\kappa = \sup \kappa_n$. Like 2^{100}, κ is to be understood in terms of the operations and procedures that went into its formulation. What there is to grasp of κ is its recursive definition, legitimized by Replacement, together with the Cantorian notion of the alephs and what further can be proved from these about κ. One works *schematically*, in so far as one can work, the mathematical experience focused by guidelines for proceeding, these amenable to axiomatic presentation.

Once the Cantorian theory of the transfinite is taken in, one sees that $\kappa = \aleph_\kappa$ could be less than $2^{\aleph_0}$, the cardinality of the continuum, the subject of investigation of classical analysis. As mathematics was transmuting to a complex edifice of conceptual constructions, Hilbert [1900] wrote in connection with his axiomatization of the reals:

> Under the conception described above [the axiomatic method], the doubts which have been raised against the existence of the totality of all real numbers (and against the existence of infinite sets generally) lose all justification, for by the set of real numbers

> we do not have to imagine, say, the totality of all possible laws according to which the elements of a fundamental sequence can proceed, but rather—as just described—a system of things, whose mutual relations are given by the *finite and closed* system of axioms [for complete ordered fields], and about which new statements are valid only if one can derive them from the axioms by means of a finite number of logical inferences.

Third, and the most sweeping, of dialectical points vs. Boolos, which we expand to be about Replacement generally, concerns the use and "requirements" for mathematics. For set theory itself, the previous sections have provided ample evidence for the importance of Replacement in various directions. To press the point, in ZF without Replacement neither the (von Neumann) ordinal $\omega + \omega$ nor the rank V_ω nor Zermelo's particular infinite set $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \dots\}$ can be shown to exist. Keeping in mind the now-standard formulation of the Axiom of Infinity as the existence of the (von Neumann) ordinal ω, one simple model of ZF minus Replacement having none of these sets is $\langle W, \in\rangle$, where $W_0 = \omega$, $W_{n+1} = P(W_n)$, and $W = \bigcup\{W_n \mid n \in \omega\}$. This is a drastic failure vis-à-vis all of functional substitution, recursion, and indifference to identification.[37] One can of course just posit the existence of any particular V_α and allow for recursions up to some fixed length β, but then analogous issues arise at the level of $V_{\alpha+\beta}$.

Proceeding in the cumulative hierarchy, the ranks $V_{\omega+1}$ and $V_{\omega+2}$ correspond via cardinality to reals and real functions, and this has advanced the presumption *contra* Replacement that classical mathematics, at least, can be accommodated at these low levels of the cumulative hierarchy. However, the roughest indifference to identification, that according to cardinal equivalence, has to be acknowledged and then affirmed by considerable coding and dexterity. With ordered pairs and functions, one starts to move from just ontological considerations to analysis and method, and these require in the set-theoretic context several iterations of the 'set of' operation, which in terms of the cumulative hierarchy amounts to a climb of several ranks.

Moving away from the iterative conception for such analyses of set-theoretic representation, it is cardinally more parsimonious and theoretically more coherent to consider for a regular uncountable cardinal κ the sets hereditarily of cardinality less than κ, i.e.,

$$H_\kappa = \{x \mid |\mathrm{tc}(x)| < \kappa\},$$

[37] See Mathias [2001a] for constructions of models of ZF minus Replacement exhibiting such weaknesses, but containing all the ordinals. Model 13 of Mathias [2006, section 7] shows easily that rank cannot be defined in ZF minus Replacement. For an extreme failure of Replacement, see Mathias [2010] for a model of Bourbaki's 1949 version of set theory with ordered pair treated as a primitive notion, in which some unordered pair fails to exist.

where $\mathrm{tc}(x)$ is the transitive closure of x, informally $\mathrm{tc}(x) = x \cup \bigcup(x) \cup \bigcup\bigcup(x) \ldots$. The existence of the H_κ cannot be established in ZF without Replacement.[38]

The H_κ model ZFC$^-$, the ZFC axioms minus Power Set, whereas V_α for limit $\alpha > \omega$ model the ZFC axioms except Replacement. Set theorists from the 1970s on have regularly and widely appealed to the H_κ as arbitrarily large models of ZFC$^-$ approximating V, especially as Replacement has become so intrinsic to all of set theory.[39] A substantial point is that one can carry out forcing over H_κ as a model of ZFC$^-$, getting forcing extensions of ZFC$^-$; without Replacement it becomes problematic to control the proliferation of forcing terms.

It is the rendering of mathematical proof rather than of mathematical objects which is substantive, and metamathematical investigations have established the necessity of substantial resources. In an example that has become widely cited for calibration with Replacement, H. Friedman [1971a] showed in 1968 that to establish Borel Determinacy would require the use of the cumulative hierarchy up to V_{ω_1}, with a level-by-level analysis revealing that determinacy at each new level of the ω_1-level Borel hierarchy would require one more iteration of the power set operation. Martin [1975] in 1974 duly established Borel Determinacy, later [1985] providing a purely inductive proof that made more evident the correlation of the Borel and the cumulative hierarchies. Borel Determinacy is thus an incisive example of the methodological involvement of Replacement.

Years later H. Friedman [1981] considerably expanded his [1971a] work to establish new independences for propositions about Borel sets, some requiring the strength of large cardinals. In that part of the work closest to [1971a], Friedman established Borel "diagonalization" and "selection" theorems that follow from Borel Determinacy and have a notable simplicity, yet still require the use of the cumulative hierarchy up to V_{ω_1}. With I the unit interval of reals, one such proposition is: *Every symmetric binary Borel relation on I contains or is disjoint from a Borel function on I.*

Martin's Borel Determinacy result itself has through the years found wide-ranging applications. In the latest, Jan Reimann and Theodore Slaman [2010] established the Martin-Löf randomness of non-computable reals with respect to continuous probability measures, with a necessary use of infinitely many iterations of the power set operation. The original Fraenkel–Skolem

[38] Replacement is needed to proceed generally from regular uncountable cardinals κ to their H_κ, but there is the issue of having regular uncountable cardinals at all without Replacement. One can readily observe though that, e.g., in the model $\langle W, \in \rangle$ above of ZFC minus Replacement, there are certainly uncountable well-orderings but not the set H_{ω_1} consisting of the hereditarily countable sets.

[39] See Foreman and Kanamori [2010]; All three of its volumes exhibit extensive use of the H_κ.

motivation for Replacement in the first place was thus called upon for getting randomness for reals.

We next consider set theory not so much in its foundational role but as of a piece with modern mathematics. Since set theory emerged as a sophisticated field of mathematics in the early 1960s with the creation of forcing, both the conceptual space it provides and the methods that it engages have become more and more integrated into the broad fabric of modern mathematics. In this regard the Replacement motif $\{t_i \mid i \in I\}$, when unrestricted in colonizing new domains, implicates Replacement.

The conceptual space provided by the set-theoretic universe as buttressed by Replacement has been increasingly brought into play in modern mathematics to generate examples and counterexamples from analysis to topology. General topology particularly has become imbued with set-theoretic constructions involving large cardinalities, to the extent that set theory and topology have here a large intersection of methods and procedures. Two conspicuous examples are Mary Ellen Rudin's construction [1971] of a Dowker space of cardinality $\aleph_\omega^{\aleph_0}$ and Peter Nyikos's [1980] derivation of the normal Moore space conjecture from a measure extension "axiom", one proved consistent relative to a strongly compact cardinal by Kenneth Kunen. The first result refuted a conjecture posed in 1951, and the second confirmed a conjecture, relative to large cardinals, posed in 1962. Natural questions about general topological spaces had thus been posed early on, and substantive elucidations only emerged as set theory itself became common coin.

The growth of category theory from the middle of the 20th Century has brought to the fore the Replacement motif $\{t_i \mid i \in I\}$, and in so far as this development is to be formalized in axiomatic set theory, it evidently implicates Replacement as immanent.[40] Of course, if one aspires to a categorical foundation for mathematics then a reduction to set theory is not a primary concern. However, it is notable that the converse reduction of set theory to category theory seems to run afoul of the problem of how to handle the axiom schema of Replacement. Replacement is what mainly needs to be accommodated, and category theory seems unable to meet the challenge.

Categorical imperatives have become particularly topical of late because of the involvement of Grothendieck universes in Andrew Wiles's proof of the Shimura–Taniyama modularity for elliptic curves, establishing Fermat's Last Theorem. In terms of set-theoretic resources, the straightforward recasting of the proof in set theory initially set the bar for establishing the famous statement of classical number theory at ZFC plus having many inaccessible cardinals (cf. McLarty [2010]). Recently, Mark Kisin [2009] provided a conceptually simpler proof of the modularity, one avoiding the

[40]The (Dedekind) Recursion Theorem is directly implicated in the common construction for a functor F: $Set \to Set$ of the colimit $0 \to F0 \to FF0 \to \cdots$.

Grothendieck apparatus and straightforwardly renderable in ZFC. In a continuing reduction, McLarty [∞] has avowed (January 2011) that both the Wiles and Kisin proofs of Fermat's Last Theorem can be formalized in ZFC without Replacement and with only Δ_0-Separation.

The issue of "purity of method" emerges here in significant fashion. The classic example is the Prime Number Theorem having been first proved using sophisticated methods of complex analysis and then decades later having been given an elementary proof formalizable in Peano Arithmetic. One looks similarly for an elementary proof of Fermat's Last Theorem. It would be very interesting and quite significant to find an elementary proof, and one would think this to be a definite new conjecture of ongoing mathematics.

Whatever purity of method would outwardly dictate however, it is also very interesting and quite significant that there are proofs at all of Fermat's Last Theorem and that these can be rendered in ZFC or restricted systems. Replacement in particular is seen to play a conspicuous role both in allowing for the straightforward rendition of proofs and for providing the conceptual space for the comparative and complexity analysis of proofs. The refining or transformation of a proof to one depending on weaker resources is itself a proof process. If for Fermat's Last Theorem a proof is found evidently formalizable in Peano Arithmetic, it would be part of a historical process in which Replacement has played a substantive role. In any case, Replacement was seen to underpin a proof, and transcending specific statements, like the Fermat or the larger Kimura–Taniyana modularity for arithmetic algebraic geometry, proofs as arguments evolving to methods are free-standing and autonomous in modern mathematics.

Modern mathematics is, to my mind, a historically given, complex edifice based on *conceptual constructions*. With its richness, variety, and complexity any discussion of the nature of mathematics cannot but accede to the primacy of its history and practice. Mathematics is in a broad sense self-generating and self-authenticating, and alone competent to address issues of its correctness and authority.

What brings us mathematical knowledge? The carriers of mathematical knowledge are *proofs*, more generally arguments and constructions, as embedded in larger contexts. Mathematical knowledge does not consist of theorem statements and does not consist of more and more "epistemic access", somehow, to "abstract objects" and their workings. Moreover, mathematical knowledge extends not so much into the statements, but back into the means, methods and definitions of mathematics, sometimes even to its axioms. Statements gain or absorb their senses from the proofs made on their behalf. A statement may have several different proofs each investing the statement with a different sense, the sense reinforced by different refined versions of the statement and different corollaries from the proofs. The movement from one proof to another is itself a proof, and generally the analysis of proofs is an activity of proof and argument. In modern mathematics, proofs and arguments have often achieved an independent status beyond their initial contexts, with new spaces of mathematical objects formulated and investigated because certain proofs can be carried out. Proofs are not merely stratagems or strategies; they and thus their evolution are what carries forth modern mathematics.

What about existence in mathematics? Questions like "What is a number?" and "What is a set?"' are not mathematical questions, and any answers would have no operational significance in mathematics. *Within* mathematics, an existence assertion may be an axiom, a goal of a proof, or one of the junctures of a proof. Existence is then submerged into context, and it is at most a matter of whether one regards the manner in which an existence assertion is deduced or incorporated as coherent with the context. An existence axiom, like the Axiom of Infinity, simply sets the stage with a means for proof. An existence proof, especially one that depends on the law of the excluded middle and is not constructive, may call for a different proof, but it is still a proof of some kind.

As for the Axiom of Replacement, it is a central axiom of set theory, indeed the one to reckon with in verifications of the axioms. Replacement is the generative existence principle that legitimizes functional substitution, recursion, and through it, the means for proving results about *all* sets. Replacement is intrinsic to inner model theory, starting with the work of Gödel, as the theory depends on the absolute and canonical nature of the ordinals and the fine details of the many transfinite recursions that define inner models. Replacement is essential for forcing, in both the range of transfinite recursions necessary to build forcing extensions and the ability to control the proliferation of forcing terms. Through functional correlation Replacement gives expression to an expansive indifference to identification, allowing for various canonical tokens to stand for types and articulates the interaction and identification of type and token. In this bolstering of the concept of set through functional substitution, recursion, and indifference to identification, Replacement has become part of the sense of set in modern set theory.

REFERENCES

WILHELM ACKERMANN

[1956] *Zur Axiomatik der Mengenlehre*, ***Mathematische Annalen***, vol. 131, pp. 336–345.

PAUL BENACERRAF AND HILARY PUTNAM

[1983] ***Philosophy of mathematics. Selected readings***, second ed., Cambridge University Press, Cambridge.

PAUL BERNAYS

[1937] *A system of axiomatic set theory—Part I*, ***The Journal of Symbolic Logic***, vol. 2, pp. 65–77, reprinted in Müller [1976, pp. 1–13].

[1941] *A system of axiomatic set theory—Part II*, ***The Journal of Symbolic Logic***, vol. 6, pp. 1–17, reprinted in Müller [1976, pp. 14–30].

[1958] ***Axiomatic set theory***, Studies in Logic and the Foundations of Mathematics, North-Holland, Amsterdam, with a historical introduction by Abraham A. Fraenkel.

BERNARD BOLZANO

[1851] ***Paradoxien des Unendlichen***, C. H. Reclam, Leipig.

GEORGE S. BOOLOS

[1971] *The iterative concept of set*, ***The Journal of Philosophy***, vol. 68, pp. 215–231, reprinted in Boolos [1998], pp. 13–29.

[1989] *Iteration again*, ***Philosophical Topics***, vol. 17, pp. 5–21, reprinted in Boolos [1998], pp. 88–105.

[1998] ***Logic, logic, and logic***, Harvard University Press, Cambridge MA, edited by Richard Jeffery.

[2000] *Must we believe in set theory*?, ***Between logic and intuition: Essays in honor of Charles Parsons*** (Gila Sher and Richard Tieszen, editors), Cambridge University Press, Cambridge, preprinted in Boolos [1998, pp., 120–132], pp. 257–268.

CESARE BURALI-FORTI

[1897] *Una questione sui numeri transfiniti*, ***Rendiconti del circolo matematico de Palermo***, vol. 11, pp. 154–164.

JOHN P. BURGESS

[2009] *Putting structuralism in its place*, on website: `www.princeton.edu/~jburgess`.

GEORG CANTOR

[1872] *Über die Ausdehnung eines Satzes aus der Theorie der trigonometrischen Reihen*, ***Mathematische Annalen***, vol. 5, pp. 123–132.

JOHN W. DAWSON

[1997] ***Logical dilemmas: The life and work of Kurt Gödel***, A. K. Peters, Wellesley, MA.

RICHARD DEDEKIND

[1857] *Abriss einer Theorie der höheren Kongruenzen in bezug auf einen reellen Primzahl-Modulus*, ***Journal für die Reine und Angewandte Mathematik***, vol. 54, pp. 1–26, reprinted in Dedekind [1930–1932, vol. 1, pp. 40–67].

[1872] ***Stetigkeit und irrationale Zahlen***, F. Vieweg, Braunschweig, translated with commentary in Ewald [1996, pp. 765–779].

[1888] ***Was sind und was sollen die Zahlen?***, F. Vieweg, Braunschweig, third, 1911 edition translated with commentary in Ewald [1996, pp. 787–833].

[1930–1932] ***Gesammelte mathematische Werke***, F. Vieweg, Baunschweig, edited by Robert Fricke, Emmy Noether and Öystein Ore; reprinted by Chelsea Publishing Company, New York, 1969.

GUSTAV LEJEUNE DIRICHLET

[1863] ***Vorlesungen über Zahlentheorie***, F. Vieweg, Braunschweig, edited by Richard Dedekind; second edition 1871, third edition 1879.

[1889/97] ***G. Lejeune Dirichlet's Werke***, Reimer, Berlin.

HEINZ-DIETER EBBINGHAUS

[2007] ***Ernst Zermelo. An approach to his life and work***, Springer, Berlin.

William Ewald

[1996] ***From Kant to Hilbert: A source book in the foundations of mathematics***, Clarendon Press, Oxford.

Matthew Foreman and Akihiro Kanamori

[2010] ***Handbook of set theory***, Springer, Berlin.

Abraham A. Fraenkel

[1921] *Über die Zermelosche Begründung der Mengenlehre*, ***Jahresbericht der Deutschen Mathematiker-Vereinigung***, vol. 30, pp. 97–98.

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

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

[1967] ***Lebenskreise. Aus den Erinnerungen eines jüdischen Mathematikers***, Deutsche Verlags-Anstalt, Stuttgart.

Harvey M. Friedman

[1971a] *Higher set theory and mathematical practice*, ***Annals of Mathematical Logic***, vol. 2, pp. 325–357.

[1971b] *A more explicit set theory*, ***Axiomatic set theory*** (Dana S. Scott, editor), Proceedings of Symposia in Pure Mathematics, vol. 13(1), American Mathematical Society, Providence, RI, pp. 49–65.

[1981] *On the necessary use of abstract set theory*, ***Advances in Mathematics***, vol. 41, pp. 209–280.

Harvey M. Friedman and Andrej Scedrov

[1985] *The lack of definable witnesses and provably recursive functions in intuitionistic set theories*, ***Advances in Mathematics***, pp. 1–13.

Kurt Gödel

[1931] *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter System I*, ***Monatshefte für Mathematik und Physik***, vol. 38, pp. 173–198, reprinted and translated in Gödel [1986, pp. 144–195].

[1932] *Über Vollständigkeit und Widerspruchsfreiheit*, ***Ergebnisse eines mathematischen Kolloquiums***, vol. 3, pp. 12–13, reprinted and translated in Gödel [1986, pp. 24–27].

[1933o] *The present situation in the foundation of mathematics*, printed in Gödel [1995, pp. 45–53].

[1938] *The consistency of the axiom of choice and of the generalized continuum-hypothesis*, ***Proceedings of the National Academy of Sciences of the United States of America***, vol. 24, pp. 556–557, reprinted in Gödel [1990, pp. 26–27].

[1940] ***The consistency of the axiom of choice and of the generalized continuum hypothesis with the axioms of set theory***, Annals of Mathematics Studies, vol. 3, Princeton University Press, Princeton, reprinted in Gödel [1990, pp. 33–101].

[1944] *Russell's mathematical logic*, ***The philosophy of Bertrand Russell*** (Paul A. Schilpp, editor), Library of Living Philosophers, vol. 5, Northwestern University, Evanston, IL, reprinted in Gödel [1990, pp. 119–141].

[1946] *Remarks before the Princeton bicentennial conference on problems in mathematics*, reprinted in Gödel [1990, pp. 150–153].

[1986] ***Collected works: Publications 1929–1936***, vol. I, Clarendon Press, Oxford, edited by Solomon Feferman *et al.*

[1990] ***Collected works: Publications 1938–1974***, vol. II, Oxford University Press, New York, edited by Solomon Feferman *et al.*

[1995] ***Collected works: Unpublished essays and lectures***, vol. III, Oxford University Press, New York, edited by Solomon Feferman *et al.*

[2003] ***Collected works: Correspondence A–G***, vol. IV, Clarendon Press, Oxford, edited by Solomon Feferman *et al.*

NICOLAS D. GOODMAN

[1985] *Replacement and collection in intuitionistic set theory*, ***The Journal of Symbolic Logic***, vol. 50, pp. 344–348.

MICHAEL HALLETT

[1984] ***Cantorian set theory and the limitation of size***, Oxford Logic Guides, vol. 10, Clarendon Press, Oxford.

FRIEDRICH HARTOGS

[1915] *Über das Problem der Wohlordnung*, ***Mathematische Annalen***, vol. 76, pp. 438–443.

A. E. HARWARD

[1905] *On the transfinite numbers*, ***Philosophical Magazine***, vol. (6)10, pp. 439–460.

FELIX HAUSDORFF

[1914] ***Grundzüge der Mengenlehre***, de Gruyter, Leipzig.

RICHARD HECK JR.

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

GERHARD HESSENBERG

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

DAVID HILBERT

[1900] *Über den Zahlbegriff*, ***Jahresbericht der Deutschen Mathematiker-Vereinigung***, vol. 8, pp. 180–184, translated with commentary in Ewald Ewald [1996, pp. 1092–1095].

AKIHIRO KANAMORI

[2004] *Zermelo and set theory*, this BULLETIN, vol. 10, pp. 487–553.

MARK KISIN

[2009] *Moduli of finite flat group schemes, and modularity*, ***Annals of Mathematics***, pp. 1085–1180.

GEORG KREISEL

[1980] *Kurt Gödel*, ***Biographical Memoirs of Fellows of the Royal Society***, vol. 26, pp. 149–224, vol. 27 (1981), p. 697; vol. 28 (1982), p. 719.

EDMUND LANDAU

[1930] ***Grundlagen der Analysis***, Akademische Verlagsgesellschaft, Leipzig, translated as ***Foundations of analysis***, Chelsea Publishing Company, 1951.

AZRIEL LEVY

[1959] *On Ackermann's set theory*, ***The Journal of Symbolic Logic***, vol. 24, pp. 154–166.

[1960a] *Axiom schemata of strong infinity in axiomatic set theory*, ***Pacific Journal of Mathematics***, vol. 10, pp. 223–238, reprinted in *Mengenlehre*, Wissensschaftliche Buchgesellschaft Darmstadt, 1979, pp. 238–253.

[1960b] *Principles of reflection in axiomatic set theory*, ***Fundamenta Mathematicae***, vol. 49, pp. 1–10.

[1968] *On von Neumann's axiom system for set theory*, ***American Mathematical Monthly***, pp. 762–763.

[1974] *Parameters in the comprehension axiom schemas of set theory*, ***Proceedings of the Tarski symposium***, Proceedings of Symposia in Pure Mathematics, vol. 25, American Mathematical Society, Providence, RI, pp. 309–324.

AZRIEL LEVY AND ROBERT L. VAUGHT

[1961] *Principles of partial reflection in the set theories of Zermelo and Ackermann*, ***Pacific Journal of Mathematics***, vol. 11, pp. 1045–1062.

D. ANTHONY MARTIN

[1975] *Borel determinacy*, ***Annals of Mathematics***, vol. 102, pp. 363–371.

[1985] *A purely inductive proof of Borel determinacy*, ***Recursion theory*** (Anil Nerode and Richard A. Shore, editors), Proceedings of Symposia in Pure Mathematics, vol. 42, American Mathematical Society, Providence, RI, pp. 303–308.

ADRIAN R. D. MATHIAS

[2001a] *Slim models of Zermelo set theory*, ***The Journal of Symbolic Logic***, vol. 66, pp. 487–496.

[2001b] *The strength of Mac Lane set theory*, ***Annals of Pure and Applied Logic***, pp. 107–234.

[2006] *Weak systems of Gandy, Jensen and Devlin*, ***Set theory: Centre de Recerca Matemàtica, Barcelona 2003–2004*** (Joan Bagaria and Stevo Todorcevic, editors), Trends in Mathematics, Birkhäuser Verlag, Basel, pp. 149–224.

[2007] *A note on the schemes of replacement and collection*, ***Archive for Mathematical Logic***, vol. 46, pp. 43–50.

[2010] *Unordered pairs in the set theory of Bourbaki* 1949, ***Archive for Mathematical Logic***, vol. 94, pp. 1–10.

[∞] *Bourbaki and the scorning of logic*, to appear.

COLIN MCLARTY

[2010] *What does it take to prove Fermat's last theorem? Grothendieck and the logic of number theory*, this BULLETIN, pp. 359–377.

[∞] *Set theory for Grothendieck's number theory*, `www.cwru.edu/artsci/philo/Groth%20found26.pdf`.

HERBERT MESCHKOWSKI AND WINFRIED NILSON

[1991] ***Georg Cantor. Briefe***, Springer, Berlin.

DIMITRY MIRIMANOFF

[1917a] *Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles*, ***L'Enseignement Mathématique***, vol. 19, pp. 37–52.

[1917b] *Remarques sur la théorie des ensembles et les antinomies cantoriennes. I*, ***L'Enseignement Mathématique***, vol. 19, pp. 209–217.

Richard M. Montague
[1961] *Fraenkel's addition to the axioms of Zermelo*, ***Essays on the foundations of mathematics*** (Yehoshua Bar-Hillel, E. I. J. Poznanski, Michael O. Rabin, and Abraham Robinson, editors), Magnes Press, Jerusalem, dedicated to Professor A. A. Fraenkel on his 70th birthday, pp. 91–114.

Gregory H. Moore
[1976] *Ernst Zermelo, A. E. Harward, and the axiomatization of set theory*, ***Historia Mathematica***, vol. 3, pp. 206–209.

Gregory H. Moore and Alejandro R. Garciadiego
[1981] *Burali-Forti's paradox: a reappraisal of its origins*, ***Historia Mathematica***, vol. 8, pp. 319–350.

Gert H. Müller
[1976] ***Sets and classes. On the work of Paul Bernays***, Studies in Logic and the Foundations of Mathematics, vol. 84, North-Holland, Amsterdam.

Peter J. Nyikos
[1980] *A provisional solution to the normal Moore space problem*, ***Proceedings of the American Mathematical Society***, pp. 429–435.

Richard Pettigrew
[2008] *Platonism and Aristoteleanism in mathematics*, ***Philosophia Mathematica***, vol. 16, pp. 310–332.

Michael Potter
[2004] ***Set theory and its philosophy***, Oxford University Press, Oxford.

Willard V. O. Quine
[1960] ***Work and object***, MIT Press, Cambridge.

Michael Rathjen
[2005] *Replacement versus collection and related topics in constructive Zermelo–Fraenkel set theory*, ***Annals of Pure and Applied Logic***, vol. 136, pp. 156–174.

Jan Reimann and Theodore A. Slaman
[2010] *Effective randomness for continuous measures*, preprint.

William N. Reinhardt
[1970] *Ackermann's set theory equals ZF*, ***Annals of Mathematical Logic***, vol. 2, pp. 189–249.

Raphael M. Robinson
[1937] *The theory of classes, a modification of von Neumann's system*, ***The Journal of Symbolic Logic***, vol. 2, pp. 29–36.

Mary Ellen Rudin
[1971] *A normal space X for which $X \times I$ is not normal*, ***Fundamenta Mathematicae***, vol. 73, pp. 179–186.

Steve Russ
[2004] ***The mathematical works of Bernard Bolzano***, Oxford University Press, Oxford.

DANA S. SCOTT

[1974] *Axiomatizing set theory*, ***Axiomatic set theory*** (Thomas Jech, editor), Proceedings of Symposia in Pure Mathematics, vol. 13(2), American Mathematical Society, Providence, RI, pp. 207–214.

JOSEPH R. SHOENFIELD

[1967] ***Mathematical logic***, Addison-Wesley, Reading MA.

[1977] *Axioms of set theory*, ***Handbook of mathematical logic*** (K. Jon Barwise, editor), North-Holland, Amsterdam, pp. 321–344.

WILFRIED SIEG AND DIRK SCHLIMM

[2005] *Dedekind's analysis of number: systems and axioms*, ***Synthèse***, vol. 147, pp. 121–170.

THORALF SKOLEM

[1923a] *Begründung der elementaren Arithmetik durch die rekurrierende Denkweise ohne Anwendung scheinbarer Veränderlichen mit unendlichem Ausdehnungsberich*, ***Skrifter utgit av Videnskabsselskapets i Kristiania, I. Matematisk-naturvidenskabelig klasse***, vol. 6, pp. 1–38, translated in van Heijenoort [1967, pp. 302–333].

[1923b] *Einige Bermerkungen zur axiomatischen Begründung der Mengenlehre*, ***Wissenschaftliche Vorträge gehalten auf dem Fünften Kongress der Skandinavischen Mathematiker in Helsingfors vom 4. bis 7. Juli 1922***, Akademische Buchhandlung, Helsinki, translated in van Heijenoort [1967, pp. 290–301], pp. 217–232.

WILLIAM W. TAIT

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

JEAN VAN HEIJENOORT

[1967] ***From Frege to Gödel: A source book in mathematical logic, 1879–1931***, Harvard University Press, Cambridge MA, reprinted 2002.

JOHN VON NEUMANN

[1923] *Zur Einführung der transfiniten Zahlen*, ***Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae (Szeged), sectio scientiarum mathematicarum***, vol. 1, pp. 199–208, reprinted in von Neumann [1961, 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, pp. 219–240, Berichtigung ibid. *155*, 128; reprinted in von Neumann [1961, pp. 34–56]; translated in van Heijenoort [1967, pp. 393–413].

[1928a] *Die Axiomatisierung der Mengenlehre*, ***Mathematische Zeitschrift***, vol. 27, pp. 669–752, reprinted in von Neumann [1961, pp. 339–422].

[1928b] *Über die Definition durch transfinite Induktion und verwandte Fragen der allgemeinen Mengenlehre*, ***Mathematische Annalen***, vol. 99, pp. 373–391, reprinted in von Neumann [1961, pp. 320–338].

[1929] *Über eine Widerspruchfreiheitsfrage in der axiomatischen Mengenlehre*, ***Journal für die Reine und Angewandte Mathematik***, vol. 160, pp. 227–241, reprinted in von Neumann [1961, pp. 494–508].

[1961] ***Collected works***, vol. 1, Pergamon Press, New York, edited by Abraham H. Taub.

HAO WANG

[1974a] *The concept of set*, ***From mathematics to philosophy***, Humanities Press, New York, reprinted in Benacerraf and Putnam [1983, 530–570], pp. 181–223.

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

[1981] ***Popular lectures on mathematical logic***, Van Nostrand Reinhold, New York.

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

Alfred North Whitehead and Bertrand Russell

[1913] ***Principia mathematica***, vol. 3, Cambridge University Press, Cambridge.

Norbert Wiener

[1914] *A simplification of the logic of relations*, ***Proceedings of the Cambridge Philosophical Society***, pp. 387–390, reprinted in van Heijenoort [1967, pp. 224–227].

Andrzej M. Zarach

[1996] *Replacement $\nrightarrow$ collection*, ***Gödel '96. Logical foundations of Mathematics, Computer Science and Physics—Kurt Gödel's legacy*** (Petr Hájek, editor), Lecture Notes in Logic, vol. 6, Springer, Berlin, pp. 307–322.

Ernst Zermelo

[1904] *Beweis dass jede Menge wohlgeordnet werden kann*, ***Mathematische Annalen***, vol. 59, pp. 514–516, reprinted and translated in Zermelo [2010, pp. 114–119].

[1908a] *Neuer Beweis für die Moglichkeit einer Wohlordnung*, ***Mathematische Annalen***, vol. 65, pp. 107–128, reprinted and translated in Zermelo [2010, pp. 120–159].

[1908b] *Untersuchungen über die Grundlagen der Mengenlehre I*, ***Mathematische Annalen***, vol. 65, pp. 261–281, reprinted and translated in Zermelo [2010, pp. 188–229].

[1909] *Sur les ensembles finis et le principe de l'induction complète*, ***Acta Mathematica***, vol. 32, pp. 185–193, reprinted and translated in Zermelo [2010, pp. 236–253].

[1930a] *Über das mengentheoretische Model*, printed and translated in Zermelo [2010, pp. 446–453].

[1930b] *Über Grenzzahlen und Mengenbereiche*: *Neue Untersuchungen über die Grundlagen der Mengenlehre*, ***Fundamenta Mathematicae***, vol. 16, pp. 29–47, reprinted and translated in Zermelo [2010, pp. 400–431].

[2010] ***Collected works***, vol. 1, Springer, Berlin, edited by Heinz-Dieter Ebbinghaus and Akihiro Kanamori.

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