A Variant of Church's Set Theory with a Universal Set in which the Singleton Function is a Set

Flash Sheridan
101 Alma Street #704
Palo Alto, California 94301-1008 USA
http://pobox.com/~flash

3 January 2013, minor revisions 21 December 2013

Abstract

A Platonistic set theory with a universal set, CUSι, in the spirit of Alonzo Church's "Set Theory with a Universal Set," is presented; this theory uses a different sequence of restricted equivalence relations from Church's, such that the singleton function is a 2-equivalence class and hence a set, but (like Emerson Mitchell's set theory, and unlike Church's), it lacks unrestricted axioms of sum and product set. The theory has an axiom of unrestricted pairwise union, however, and unrestricted complements. An interpretation of the axioms in a set theory similar to Zermelo-Fraenkel set theory with global choice and urelements (which play the rôle of new sets) is presented, and the interpretations of the axioms proved, which proves their relative consistency.

The verifications of the basic axioms are performed in considerably greater generality than necessary for the main result, to answer a query of Thomas Forster and Richard Kaye. The existence of the singleton function partially rebuts a conjecture of Church about the unification of his set theory with Quine's New Foundations, but the natural extension of the theory leads to a variant of the Russell paradox.

An abridged version of this article will appear in Logique et Analyse. *This unabridged version is to be made available on the web site of the Centre National de Recherches de Logique, http://www.logic-center.be/Publications/Bibliotheque.*

Contents

List of Theorems

0. Introduction, Context, and Related Work

1 Philosophical Introduction and Motivation

> *Die Zeit ist nur ein psychologisches Erforderniss zum Zählen, hat aber mit dem Begriffe der Zahl nichts zu thun.*
>
> Time is only a psychological necessity for numbering, it has nothing to do with the concept of number.
>
> (Frege, *Die Grundlagen der Arithmetik* §40, tr. J.L. Austin)

> ταῦτα δὲ πάντα μέρη χρόνου, καὶ τό τ᾽ ἦν τό τ᾽ ἔσται χρόνου γεγονότα εἴδη, ἃ δὴ φέροντες λανθάνομεν ἐπὶ τὴν ἀίδιον οὐσίαν οὐκ ὀρθῶς. λέγομεν γὰρ δὴ ὡς ἦν ἔστιν τε καὶ ἔσται, τῇ δὲ τὸ ἔστιν μόνον κατὰ τὸν ἀληθῆ λόγον προσήκει, τὸ δὲ ἦν τό τ᾽ ἔσται περὶ τὴν ἐν χρόνῳ γένεσιν ἰοῦσαν πρέπει λέγεσθαι—κινήσεις γάρ ἐστον...
>
> And these are all portions of Time; even as "Was" and "Shall be" are generated forms of Time, although we apply them wrongly, without noticing, to Eternal Being. For we say that it "is" or "was" or "will be," whereas, in truth of speech, "is" alone is appropriate to It; "was" and "will be," on the other hand, are properly said of the Becoming which proceeds in Time, since [both of] these are motions...
>
> (Plato, *Timæus* 37E, tr. R.G. Bury (Loeb), corrected slightly)

This paper is part of an effort towards a whole-heartedly Platonistic[1] set theory which avoids the set-theoretic paradoxes,[2] but still contains such Fregean sets as the universal set and Frege-Russell cardinals. The standard method of avoiding the set-theoretic paradoxes, Zermelo-Fraenkel set theory, has been pragmatically successful, but suffers from what I have called elsewhere[3] half-hearted Platonism: It relies for its philosophical justification on a metaphor of constructing sets in time (or something like it),[4] which violates a crucial tenet of

[1] The relevant aspect of Platonism for the current discussion will simply be that mathematical objects are not temporal; concurrence with any of Plato's ideas about non-mathematical objects is not necessary. See further below.

[2] Particularly the Russell Paradox as it affected Frege's foundational program ([Frege 1903], afterword), but also the Burali-Forti Paradox of the set of all ordinals, and the Mirimanoff Paradox [Mirimanoff 1917b] of the set of all well-founded sets.

[3] [Sheridan 1982, 1989]; summarized in [Forster 1995] pp. 141-2.

[4] E.g., [Parsons 1977], [Gödel 1964] footnote 12, and [Almog 2008] pp. 550-1, 570-1. Even if this temporal metaphor is accepted, the theory would seem to violate it, by allowing sets to be constructed at an earlier level via quantification over sets constructed at a later level; but that is an internal matter for those who accept the metaphor. Unbeknownst to me, Church had noted, with perhaps a hint of scepticism, this impredicativity of ZFC in the notes for his Coble Memorial Lectures [Princeton University Church Archives, box 15, folder 10, typescript "Outline and Background Material, Arthur B. Coble Memorial Lectures"/"Sets of the Model Transfinitely Generated" page numbered 2, 39th page in folder, also mimeograph page 4]. (See below for the need for unwieldy

mathematical Platonism: that mathematical objects are independent of time.[5]

The main philosophical advance in Church's theory, I would claim (Church is silent on his philosophical motivation[6]), is the rejection of a general comprehension axiom schema; such axioms seem also to use the suspect metaphor of temporal construction of sets.[7] Instead Church's theory posits the existence of Fregean sets denied by ZF, such as the universal set and Frege-Russell cardinals, via atemporal operations such as symmetric difference and equivalence classes as sets. Church's (restricted) axioms of generalized Frege cardinals shed, I hope, some light on later work by neo-Fregeans,[8] and may help to rescue Frege's definition by abstraction of cardinal numbers. This is still, I believe, the most natural definition, and Frege's insistence that the definitions of numbers reflect their application remains the best available, albeit partial, explanation of the unreasonable effectiveness of mathematics.[9]

Church's main pragmatic advance is a double use of standard Zermelo-Fraenkel set theory with global choice, both within his theory as axioms restricted to well-founded sets, and metatheoretically as the basis for his (appar-

---

citations to the archives.) See also [Holmes 2001] for an argument that the theory justified by the iterative conception is actually Zermelo Set Theory with Σ_2 replacement. According to Professor Holmes, "this contain[s] an error, which Kanamori pointed out to me and which I know how to fix."

[5] Plato *Timæus*, 37E; [Frege 1884] §40.

[6] The following remark, in a paper on which Church was working at around the same time, might be indicative of Church's state of mind, but this is speculative: "To avoid impredicativity the essential restriction is that quantification over any domain (type) must not be allowed to add new members to the domain, as it is held that adding new members changes the meaning of quantification over the domain in such a way that a vicious circle results." "Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski," *The Journal of Symbolic Logic* volume 41, Number 4, Dec. 1976.

[Anderson 1998] p. 136 suggests that Church was "usually seen as a quite traditional Platonic Realist," at least in his mature period, with a caveat about Church eschewing the label because of its association with the thesis that only universals are real, which is not necessary for the present variety of Platonism. Church, at least in his early period, was not enough of a Platonist not to show some skepticism about the Axiom of Choice [Enderton 2008] p. 8, though in "Set Theory with a Universal Set" he uses an extremely strong form of it. (Church does, however, use Hailperin's finite axiomatization of Quine's New Foundations, rather than Quine's original comprehension schema, in his later theories; this is probably for technical reasons, though Hailperin's axioms might be seen as less Platonistically offensive.) Church was apparently working on a paper entitled "Frege on the Philosophy of Time" before he started work on his set theory [box 15, Folder 8, April 17, 1969]; I have not yet been able to obtain the manuscript from the Church Archives.

Church speculates [Church 1974a], pp. 298-9 about axiomatic possibilities for blaming the antinomies on *intermediate* sets, i.e., sets which are not *low* (i.e., equinumerous to a well-founded set), and whose complements are also not low; more specifically, on sets which are "balanced on the hazardous edge between low sets and intermediate sets." I am not aware of any progress on this approach. While it might apply to the Burali-Forti and Mirimanoff Paradoxes, it does not seem to apply to the Russell Class, which would contain, for instance, all normal singletons.

[7] See also [Sheridan 2005] and [Forster & Libert 2011] for an argument that a comprehension axiom amounts to a claim of implausible fixed points in simple set theoretic operations such as adjunction.

[8] See especially [Burgess 2005]. Much of the neo-Fregean program was beginning while I was doing my initial work—some of it at Oxford while I was writing the initial version of this paper—but I was largely unaware of its achievements until I resumed work on this paper more than a decade later.

[9] See [Heck 2013], p. 41 ff, p. 222 ff.

ently uncompleted) relative consistency proof. The inclusion of the axioms of ZFGC (as the well-founded subtheory, i.e., restricted to well-founded sets), while it makes the use of the theory as a foundation for mathematics easier for the working mathematician, does lay the theory open to the charge of philosophical hypocrisy. The coupling between the details of ZF, and Church's and my theories, is relatively weak, however; the ZF-like axioms are a clearly-distinguished subset. ([Sheridan 1985] shows that Church's theory can be viewed as a conservative extension of ZFGC, and a similar result seems clear for the current theory.) I have tried to keep my uses of the main one, Well-Founded Replacement, relatively isolated, though I have not succeeded as much as I had hoped.

Both of these features distinguish Church's theory from Quine's New Foundations [Quine 1937a], which in some other respects his theory resembles. Church ends his initial article with speculation that his theory might be unified with New Foundations. I found this philosophically objectionable, since NF has a comprehension axiom and lacks a clear philosophical motivation.[10] A later, unpublished, theory by Church attempts to converge with New Foundations, but he seems to have abandoned it; see the historical introduction below.

My main contribution is to distinguish theories of Church's sort further from New Foundations, by providing a variant in which the singleton function is a set. This is completely impossible in New Foundations.[11] I also provide a fully-worked out relative consistency proof. Church never published a full consistency proof, and the version in his archives, which refers to two earlier attempts, seems to have been abandoned as well. His notes written in 1989 suggest the need for "a new approach." My consistency proof also avoids the use of compactness needed by Church; I provide an interpretation for the full sequence of restricted equivalence relations, rather than an arbitrary finite subsequence.

How successful my endeavor was, however, is unclear: A natural extension of my theory is subject to a variant of the Russell Paradox, involving the set of all non-self-membered sets equinumerous to the universe,[12] and my equivalence classes (unlike Church's) are not closed under sum set and product set, though Mitchell's theory suffers a similar limitation. (My theory does, however, satisfy unrestricted pairwise union.)

The paradox, though it is not directly relevant to the consistency proof here, suggests that my equivalence classes ran afoul of what has since been called the Bad Company problem.[13] Church's equivalence classes are not obviously subject to the same difficulty, and are, as noted, closed under sum and product set. Thus it could be argued that the limitations of my endeavor are an argument for Church's conjecture about unification with New Foundations,[14] or possibly

[10] [Forster 1995] pp. 26-7; [Holmes 1998] p. 12; [Maddy 2011] p. 136, citing [Fraenkel, Bar-Hillel, and Lévy 1973] p. 164. Holmes' chapter eight does attempt to provide a philosophical motivation for stratification, but not from an atemporal perspective.

[11] [Holmes 1998], p. 110, 131.

[12] See the discussion following the 1-Isomorphism Lemma (19.4), below. Cp. also Holmes' proof of the non-set-hood of the membership relation, [Holmes 1998] p. 43.

[13] [Burgess 2005], pp. 164 ff.; [Boolos 1990], pp. 249–251; [Dummett 1991], pp. 188-9.

[14] Especially the cumulative hierarchical aspects of his 1975 and later unpublished archive notes,

Forster's "Naturam expellas furca" claim[15]

1.1 Criticism and Alternatives

Thomas Forster has criticized Church's type of construction as a Potemkin Village:[16] *"this technique is not a great deal of use for constructing models of a theory T unless T has an easy word problem.* Set theories with an easily solvable word problem are unlikely to be of interest."[17] I don't think anyone has disagreed with this criticism; both Church's unpublished follow-on work [Church Archives box 15, folder 11, untitled manuscript, page numbered 3, 3rd page in folder], and [Malitz 1976], written under Church's supervision, redefine equality recursively for the sake of more powerful axioms. Mitchell's concluding comment addresses this possibility as well [Mitchell 1976], pp. 30-1. My own contemplation of more powerful axioms extending my system led to a variant of the Russell Paradox, noted above and discussed below. Even before encountering this paradox, I had expressed scepticism that interesting extensions of these theories would allow models using the same simple techniques.

Oberschelp's theory [1973] also avoids my difficulties, which may indicate the wisdom of the limitations of his approach. I feel that Oberschelp's work deserves far more attention than it has received, but the presentation is difficult, and part of the consistency proof (p. 48) is a reference to another paper ([Oberschelp 1964a]) with a different formalism.

2 Historical Introduction

When the rough draft of this paper was substantially finished, I learned of papers in Church's archives at Princeton University on later set theories with a universal set. Those theories are out of scope for the technical sections of this paper, but the papers I have obtained from the archives, and its catalog, are the main source for this historical introduction. I have so far only received portions of the relevant papers from the archives, pending Princeton's procurement of scanning equipment. The page numbers on these papers are often missing or incorrect, so I have erred on the side of explicitness below in citing them.

On June 24th 1971, Alonzo Church presented a paper entitled "Set Theory with a Universal Set" to the Tarski Symposium, at the University of California at Berkeley.

---

e.g. Church Archives Box 15, folder 10, typescript "Outline and Background Material, Arthur B. Coble Memorial Lectures"/"Sets of the Model Transfinitely Generated" page numbered 2, 37th page.

[15][Forster 2006] p. 240, presumably alluding to Horace's *Epistles,* I. x. 24, about the necessity of the cumulative hierarchy for avoiding the paradoxes.

[16]The term "Potemkin Village," which is probably unfair to its namesake (a Governor-General showing villages in the Crimea to Catherine the Great), is used for constructions placed *only* where an observer will be looking for them. I believe the term was first used by me as a summary of Forster's criticism at his Stanford lecture on Church's theory, 11 April 2005.

[17][Forster 2001], p. 6.

In a fifty-page manuscript dated July 1971, labelled "Notes as to Set Theory with a Universal Set" (photocopy in archives box 47, Folder 10), Church states that "As even the amended model of April 1971, ... is not yet satisfactory, we make a new start using the outline of June 1971." This seems to be an eventually-abandoned attempt at another consistency proof for the full CUS; the mathematics does not seem final, and the photocopy, if not the manuscript, ends abruptly.

In 1974, a version of the paper was printed in the conference proceedings, with a "minor but essential modification" to the definition of "the equivalences characterizing the model" [Church 1974a], p. 307, footnote 11.

The published paper presents three main features:

1. A sequence of equivalence relations generalizing equinumerosity,
2. A set of axioms for a set theory with a universal set and some equivalence classes, including Frege-Russell cardinals, as sets,
3. A model (actually an interpretation [Shoenfield 1967], though Church does not use the term) of the axioms, restricted to a finite subsequence of the sequence of equivalence relations, with length a fixed arbitrary natural number m.

The paper presents no proofs; Church states (p. 307) that the "details of the verification... are straightforward but (if m > 0) laborious." I believe most readers have found this an understatement.

I am not aware of any accounts of Church's talk (though Emerson Mitchell was present, see below; oddly, no correspondence with Mitchell is listed in the index to the Church archives). Given the length of time necessary to understand the published paper, it seems unlikely that Church had time to present much more than the definitions of the sequence of equivalence relations, the axioms, and a high-level sketch of the interpretation.

Church later issued an undated two-page correction to footnote 4 of the paper, covering a tangential remark on standard von Neumann and Bernays set theories, which was not relevant to his new theory.

In 1973, Arnold Oberschelp published an article, apparently independently, with a technique similar to Church's, but using urelements rather than displaced sequences for the construction, and with a richer model which included the singleton function as a set [Oberschelp 1973]. Note that part of the consistency proof in both [Oberschelp 1973] (p. 48) and [Friedrichsdorf 1979] (p. 382) is merely a reference to [Oberschelp 1964a], which uses a significantly different formalism. Neither Church nor Oberschelp seems to have been aware of the other's work at the time, and the underlying similarity between the two techniques is not necessarily obvious. (Indeed, though I cited [Oberschelp 1973] in [Sheridan 1989], I did not realize its significance until [Sheridan 1990].) See the comments below on my definition of j-isomorphism, and the limitations in proving its absoluteness in the interpretation.

In 1974, according to Forster [2001], Urs Oswald independently rediscovered Church's method of permutation models, published in his ETH Zürich Ph.D. thesis of 1976, *Fragmente von "New Foundations" und Typentheorie.* Forster calls Oswald's discovery simultaneous with Church's, but Church's original paper was presented in 1971. Werner Depauli-Schimanovich also makes a claim for the priority of his 1971 doctoral thesis in his Arxiv web article [2008], which I have not evaluated.

In the fall of 1974, Church presented a lecture entitled "Notes on a Relative Consistency Proof of Axioms A – K of Church's Set Theory with a Universal Set" [Church 1974b]. Church mailed me a copy in 1984, with a handwritten notation (in a different handwriting, presumably Church's): "Student notes of 1974 lectures by Alonzo Church." The notes are eleven pages of quite dense mathematics, but only cover the case $m = 0$, i.e., omitting the equivalence class axioms L_j, whose verification Church implied was laborious but not straightforward. (I do not believe that this is an overstatement.) Church's comment on these notes, in his archive, is apparently "Probably not of much value - but possibly worth some reflection" [box 47, Folder 5]

In 1984 I requested that Church send me any further relevant work, but never received a response; my attempts to see him in Los Angeles in 1985 and 1987 failed due to his absence in the Bahamas. Until the publishing of the catalog of his archives on the web [http://arks.princeton.edu/ark:/88435/fx719m49m], I had assumed that he did no further work on his set theory, but the catalog lists some further lectures which sound relevant, which I have so far been unable to obtain from the archives. I cannot tell whether Church's disinclination to send me his later work represents a repudiation of it.

On September 23-25, 1975, at the University of Illinois at Urbana-Champaign, Church delivered the Arthur B. Coble Memorial Lectures, entitled "Set theory with a universal set."[18] Despite having the same title, this set theory was far more complicated than his 1974 paper; one of its main goals was to model Hailperin's finite axiomatization of Quine's New Foundations, though Church's notes (summarized below) indicates that he fails to model axiom P6. This topic deserves more investigation than it has received, but is outside the scope of this paper, and I have so far only been able to acquire a portion of the archives on this topic.

Princeton Archives Box 15, folder 10 consists of three versions of what I will call the Coble Theory. Two are hand-written, with the first apparently a corrected version of the second; the third is apparently a typewritten version of the corrected handwritten manuscript, with three hand-written corrections. The typewritten version also exists in a mimeographed copy in the possession of the late Professor Herbert Enderton of the University of California at Los

[18] http://www.math.uiuc.edu/Colloquia/coble_history.html, History section. Church's Princeton archives list the title of this and the lecture below as "Set Theory on a Universal Set"; this substitution of "on" for "with" seems to be a transcriber's mistake rather than a deliberate change—the University of Illinois web page lists the usual title, and no subsequent source ever seems to have used "on" rather than "with."

Angeles (which he graciously allowed me to scan), and apparently in a version at the University of California at Berkeley, which I have not seen. The UCLA copy lacks the handwritten corrections, but contains an extra typewritten sheet, specifying (1) two of the corrections made to the Princeton copy, and (2) a more sweeping correction: "the probability there must be substantive changes at various places, especially on the last two pages of either manuscript of the notes."

Princeton Archives Box 15, folder 11 contains hand-written notes for a somewhat different, incompletely-developed theory, which I will call the Folder 11 Theory. I will refer to the published 1974 theory (the main basis for my theory) as the 1974 Theory.

Some striking features of the Coble construction are:

1. "a transfinite array of relations, one for each ordinal m, that are left unspecified, the intention being that different set theories result by different choices of the invariance relations $\mathrm{inv}^{\mathrm{m}}$." [Church Archives, Box 15, folder 10, 5th page, numbered 5.] The independent construction in the middle part of the current paper could be similarly described, though the details are quite different.

2. Construction of a non-trivial identity relation for the model, which may avoid Forster's Potemkin Village criticism, discussed above. [Church Archives ibid.], [Forster 2001], p. 6.

3. The Coble set theory may have influenced the 1974 theory, e.g., the Axiom of Substitutivity in the earlier theory seems superfluous, but is necessary in the later theory, since the relation i (corresponding to identity) is constructed, and differs from the base equality. Keeping Substitutivity in the Basic Axioms allows Church to keep the axiom letters the same between the two theories.

4. The unnumbered last (19th) page of Professor Enderton's mimeographed copy of the Coble Lecture notes mentions "the probability there must be substantive changes at various places, especially on the last two pages of either manuscript of the notes." This correction is missing from the Princeton Archives copy.

5. "Given any set, there exist its complement and the set of complements of its members. The set of low sets, the set of intermediate sets, the set of high sets exist." ["Outline and Background Material, Arthur B. Coble Memorial Lectures"/"Sets of the Model Transfinitely Generated," section "Set Existence" Box 15, folder 10, page numbered 12, 47th page in folder]

Some features of the Folder 11 construction are:

1. Church calls this "the model of the Quine set theory which we seek to set up." [Church Archives box 15, folder 11, untitled manuscript, page numbered 2, 2nd page in folder.]

2. An explicit classification of constructed sets as low, high, low intermediate, high intermediate, and fully intermediate. [Church Archives box 15, folder 11, untitled manuscript, page numbered 3, 3rd page in folder.] This differs, at least in presentation, from the Coble Theory.

3. Church abruptly abandons a definition after clause 60, because of “the (later) discovery that the model obtained does not satisfy Hailperin’s P6.” He says that “an informal and partly heuristic account follows.” [Church Archives box 15, folder 11, untitled manuscript, page numbered 15, 15th page in folder.] Church’s statement of P6 is “(∃u) . x e u $\equiv_x$ (y) . <y, ι‘x> e v”. [Church Archives Box 15, folder 10, typescript “Outline and Background Material, Arthur B. Coble Memorial Lectures”/“Sets of the Model Transfinitely Generated,” section “Set Existence” unnumbered page, 48th page in folder] Hailperin’s original formulation [Hailperin 1944], p. 10, is “(α) (Eβ) (x) [x∈β ≡ (u)(<u, ιx>∈α)].” The following page begins with the observation that “The foregoing definition by recursion is a first draft.” Church goes through various stages of believing that he can or cannot prove P6 or Hailperin’s other axioms, and it is not clear what the final resolution was for this article:

   Two unnumbered pages note that the “proof of P6 requires modification of the above” [33rd page; different wording on the 34th page]. The unnumbered 36th page is headed “Analysis Directed Towards Proof of P6,” but the page numbered 40 states “it is not immediately clear that this amendment will successfully result in a model in which all nine of the Hailperin axioms hold...” The following page, numbered 42, is headed “Proof of the Main Lemma (for P6) from Lemmas 1-5,” but the manuscript ends abruptly after the next page, five lines into the proof of case 1b.

   Later items in the Church Archives catalog, which I have not yet been able to obtain, have titles mentioning P6 as well—I do not yet know if the later work overcomes this difficulty.

   Church seems to have continued working on this approach until at least “box 47, Folder 2: Notebook: Recursion clauses for inv^{m} as revised Sept. 1980.” I speculate that he had abandoned it by 1984, when he sent me [1974b] without mentioning the later theory. The undated addendum to the Coble lecture typed mimeograph, noting that substantive changes were probably needed, may have seemed at the time to be only a temporary setback.

4. Church notes “other divergences from the Quine set theory... the set of all sets *a* of pairs... evidently does not exist in the model” [Church Archives box 15, folder 11, untitled manuscript, p. 40, 40th page in folder], which demonstrates that the Folder 11 Theory is inferior in this respect to his 1974 theory as well as New Foundations.

In the Abo Akademi in Turku, Finland, on March 22, 1976, Church presented a shortened version of the Urbana, Illinois lectures [box 49, Folder 7].

In 1976, Emerson C. Mitchell was granted a Ph.D. from the University of Wisconsin at Madison for "A model of set theory with a universal set," which cites Church and builds on his technique to provide a set theory with unrestricted power set, but which lacks some of Church's other axioms [Mitchell 1976].[19] Mitchell was present at Church's 1971 lecture.[20] Mitchell's only bibliographic reference is to [Church 1974a]; the complexity of Mitchell's construction is reminiscent of Church's later unpublished work, but this is not proof that Church showed it to him. Cp. the resemblance between Church's unspecified sequence of equivalence relations in his later theories and my own.

In 1979 Ulf Friedrichsdorf published an article [Friedrichsdorf 1979] building on [Oberschelp 1973].

The Church archives list a number of notebooks on set theory from 1975 to 1983; some of these at least (e.g., box 15, Folder 10) seem to be on combining his set theory with New Foundations via Hailperin's finite axiomatization [Hailperin 1944].

In 1989 Church wrote about some of his notes, apparently the uncompleted consistency proof, "These notes are old [1971] but might be reconsidered for the sake of *some truth in it,* which might guide a new approach" [box 47, Folder 10]. This dissatisfied comment might be applied to all of the original work done in this area. For a survey of the field, the reader is referred to the articles and book by Forster in the bibliography, whose perspective is decidedly different.

3 Organization

The overall goal is to prove the equiconsistency of CUSι with ZF, by defining an interpretation of CUSι in a base theory equiconsistent with ZF, and proving the interpretation of each axiom from the base theory. This will establish the relative consistency of CUSι: Any proof of an inconsistency from the axioms of CUSι can be translated into a proof of an inconsistency in the base theory.

The central section of this paper (Part II, which was originally written separately, for Professor Church's cancelled ninetieth birthday festschrift) defines an interpretation of a partially-specified ill-founded set theory, and proves the interpretation of the Axiom of Extensionality. (The proofs of the other Basic Axioms—the first group of axioms, below, which are restricted versions of axioms of ZF—which constitute Part I, are simpler.) Unlike Church's and Mitchell's interpretations, but like Oberschelp's, my interpretation uses urelements in the rôle of the new, ill-founded, sets; this avoids having to rearrange the old sets to make room for the new. To answer a query of Forster and Kaye when a much earlier attempt at this result was presented as a doctoral thesis, and also to keep open the possibility of iterating this type of construction, or to do it with other

[19] Note that the spelling of Mitchell's first name on his thesis is an error.

[20] Abstract of [Mitchell 1976] in *The Journal of Symbolic Logic*, March 1977, Vol. 42, No. 1, p. 148.

set theories as a base theory, Part II was done in much greater generality than is needed for the main result, and with limited use of Choice and Foundation. (It is not clear at this point that either endeavor was worth the effort.) In particular, rather than using the specific sequence of restricted equivalence relations $\bumpeq^{j}$ defined in Part III, it was done with an arbitrary sequence $\sim^{j}$ satisfying certain properties.

After the interpretation is defined and the interpretation of the Axiom of Extensionality proven for the partially-defined interpretation defined in terms of $\sim^{j}$ in Part II, in Part III the required properties of $\sim^{j}$ are shown to hold for $\bumpeq^{j}$. This establishes the interpretation of Extensionality for the specific theory under investigation, CUSι, with the specific equivalence relations $\bumpeq^{j}$. The interpretations of the new axioms of interest can then finally be established, along with their consequences of interest, such as the existence of Frege-Russell cardinals and complements. In Part III both Global Choice and Foundation are assumed for the base theory, which reduces the generality but simplifies the derivations.

The verification of the rest of the Basic Axioms, which constitutes Part I, is also done in considerably more generality than necessary, but with a different and weaker set of requirements: primarily on the form of the definition of the new membership relation, plus a requirement that new sets be ill-founded, and some sanity requirements on well-foundedness in terms of the new membership relation.

One of the consequences of the requirements on the partially-specified equivalence relations, needed for Extensionality in Part II (which was written first), is, in effect, that the new sets are too large to be low. The verification of the other Basic Axioms (which were proven later) are simpler in Part I, and rely largely upon this fact; this may make verification of these axioms for other possible theories easier. The key to this result in the specific case in Part III is the Replacing at Level*j Construction, which for any non-degenerate $\bumpeq^{j}$ equivalence class, takes an arbitrary object and embeds it into the transitive closure of an object in the given equivalence class. Given that result, the verification that the results of Part I apply to the interpretation in Part III is far simpler than the application of Part II.

Note that I generally follow Church in referring to both $\sim^{j}$ and $\bumpeq^{j}$ informally as equivalence relations, though they are actually only provably equivalence relations on the well-founded sets. (Conveniently, this will not matter in the base theory, and hence in the consistency proof; it only affects discussion of results within the theory of interest.) Indeed, it is not even clear that symmetry or reflexivity hold for objects equivalent to the Universal Set, for even the first of Church's relations in his interpretation, nor for my $\bumpeq^{1}$ in mine. Church did not address this point in his published writings, though it is presumably the motivation for the restriction in his axiom of generalized Frege cardinals, and would have needed to be addressed in his full consistency proof. (The point does not arise in his surviving lecture notes for case m=0, and I was not able to find mention of it in his abandoned consistency proof in the archives.) It is possible that he was aware of subtleties which eluded me, since the obvious extension of my theory, asserting that these relations are unrestricted equivalence relations (plus some natural assumptions about the existence of mappings), runs into a variant of the Russell Paradox which his equivalence relations apparently avoid. It is even possible that he was aware of this when suggesting a unification of his theory with Quine's New Foundations, but there is no evidence of this.

The proofs in Parts I and II could be applied to Church's original theory; the generality in Part II was crafted to include Church's j-equivalence relations as well as my own. The Replacing at Level*j Construction in Part III would require substantial modification, but would be substantially easier for Church's relations. Such a proof would be significant, given Church's apparent abandonment of the consistency proof for his full system. It would not be complete, however, since Church's theory has unrestricted axioms of sum and product set, which my theory does not; they depend on the details of Church's equivalence relations, which do not seem to have been addressed in the abandoned consistency proof in the archives. The replacing result seems to be true for Church's theory, as noted in [Sheridan 1989], but Church does not seem to have addressed it in his surviving writings, though his presentation suggests that it may have been part of his motivation for the definition of his equivalence relations.

4 Discussion of the Axioms

The base theory in which results will be proven will be the Basic Axioms (below), plus a book-keeping axiom about urelements. Use of Choice and Foundation will be avoided in Parts I and II, except for some explicitly-mentioned uses of a consequence of Foundation near the end of Part II, but will be needed extensively in Part III. The Basic Axioms are equivalent to their usual counterparts in the presence of Foundation. The relative consistency of the form of global choice used, and of the book-keeping axiom, are unproblematic: global choice by the well-know result of Gödel, and the book-keeping axiom by a trivial use of the technique of [Church 1974a], or in the simpler form presented in [Forster 2001].

Note that since the base theory must allow urelements, Extensionality is restricted to non-empty objects. The theory of interest is largely neutral about the existence of urelements, though the Axiom of Generalized Frege Cardinals implies that the collection of all empty objects is a set; Well-Founded Replacement forbids this object to be the size of the Universe. Hence the interpretation presented below excludes urelements.

CUSι, the theory of interest, includes the Basic Axioms, but necessarily excludes Foundation. It also avoids dependence on Choice, though it is consistent to add it in the strong form used here, as it is incidentally true in the interpretation presented. (The global well-ordering used does not mention set membership, so it is unaffected by the reinterpretation of the membership relation.) The theory also includes the Restricted Axiom of Generalized Frege Cardinals, asserting that for the sequence of equivalence relations $\bumpeq^{j}$ (for $j \in \omega$), any well-founded set has a set of all sets to which it is $\bumpeq^{j}$, for each $j \in \omega$. The restriction to well-founded sets is for the purpose of the consistency proof; neither Church's technique nor Oberschelp's seems sufficient to provide unrestricted axioms of generalized Frege cardinals. Note that the equivalence classes themselves are not restricted to well-founded sets. For example, the set of all singletons, (which is the union of at most two $\eqsim^1$ equivalence classes) contains the singleton containing the Universal Set, and hence is ill-founded. For simplicity, $\eqsim^0$ is the universal relation, so the (unique) 0-equivalence class is the universal set.

Note that in the absence of Foundation, some of the restrictions on the Basic Axioms become significant. Sum Set, for instance, does not apply to the new sets, again because of the limitations in technique. Church's proof of his theory's consistency would have needed to demonstrate that his combinations of equivalence classes were closed under unrestricted sum set, but this seems to depend on the details of his equivalence relations, and does not apply to my modifications.

Part I
Language, Definitions, Basic Axioms, and ∈†-Interpretations

5 Language

The primitive symbols of the language, in addition to the usual first-order logical apparatus, are "=", "∈", "∅", "Υ", and (for use with a strong form of Global Choice) "$\mathscr{G}$". Two special symbols are needed for use in book-keeping axioms, "Υ" and "∅"; an explicit symbol for ∅ is needed for use in distinguishing it from urelements, and "Υ" will denote an assumed injection of the sets into the urelements.

Several symbols will be often used rather like primitive symbols, but are in fact defined terms: "$\not\equiv$" denotes exclusive disjunction, i.e., $P \not\equiv Q \equiv_{df} (P \vee Q)$ & $\neg$ (P & Q). $\in_1$, $\in_2$, and $\in_3$ will be the ill-founded set membership relations of interest in Parts I, II, and III respectively; $\in_1$ and $\in_2$ will be partially specified (in somewhat different ways) in Parts I and II, to show general results about broad classes of interpretations; $\in_3$ will be the specific membership relation used in Part III to show the relative consistency of CUSι. Since Church uses "∈" without a subscript to denote the new membership relation, rather than the old one (or for purposes of emphasis), I will often use "$\in_0$" as a synonym for "∈". Once I have defined "$\in_1$," a formula followed by a subscript "$_1$" will abbreviate that formula with "$\in_1$" substituted for all (including implicit) occurrences of "∈." Similarly for "$_2$" and "$_3$."

Limited use is made of class terms as a syntactic convenience without ontological commitment, as in [Quine 1969] and [Levy 1979] §3.1. Definition schemas are explicitly marked as such with "$\equiv_{dfs}$" and distinguished from single definitions marked with "$\equiv_{df}$". Following Church's modification of the Peano/Russell practice, dots and double dots are sometimes used informally as substitutes for brackets.

6 Definitions

These definitions are not intended to be surprising (where possible they are simply from [Levy 1979]), but the weakness of the base theory requires more elaboration than usual, and the proofs require tedious attention to primitive notation. The only surprising point is the definition of well-foundedness, not directly, but in terms of ill-foundedness and unending chains, which is important in the absence of Dependent Choices. The reader should feel free to skip these definitions (and the development of addition, below) on the assumption that the definitions do indeed mean what they are supposed to mean.

6.1 Logic

"$\Rightarrow$" and "$\Longleftrightarrow$" indicate implication and equivalence with least close possible binding. "$\equiv_{df}$" of course, has looser binding still. "$\equiv_{dfs}$" indicates a definition schema.

"$\not\equiv$" indicates exclusive disjunction, inequality for truth values. Informally, exclusive disjunction is associative: $(P \not\equiv Q) \not\equiv R \ldots$ iff an odd number of $P, Q, R \ldots$ are true iff $P \not\equiv (Q \not\equiv R \ldots)$.

"$\exists! x.\, \phi(x)$" abbreviates "$\exists x.\, \phi(x)\ \&\ :\forall x'.\, \phi(x') \rightarrow x' = x$", where ϕ is a predicate.

"$\iota x.\, \phi(x)$", if $\exists!\, x.\, \phi(x)$, denotes that x, and otherwise is undefined, where ϕ is a predicate with one free variable.

6.2 Sets and Membership

nonempty(x) $\equiv_{df} \exists y \in x$; $z \notin x \equiv_{df} \neg z \in x$.

set(x) $\equiv_{df}$ $x = \emptyset \vee$ nonempty(x); **urelement**(x) $\equiv_{df}$ $\neg$set(x).

SET[ϕ] $\equiv_{dfs}$ $\exists x.\, \text{set}(x)\ \&\ \forall z.\, z \in x \equiv \phi(z)$, where ϕ is a predicate with one free variable. (Note that the definition does not require that this x be unique, though extensionality would imply this.)

$\{x, y\} =_{df} \iota p.\, \forall w.\, w \in p \equiv (w = a \vee w = b)$; $\langle x, y\rangle =_{df} \{\{x\}, \{x, y\}\}$, i.e., the Kuratowski ordered pair.

$x \simeq y \equiv_{df} \forall z.\, z \in x \equiv z \in y$. (Read "x is **coextensive** with y.") (Thus $x \simeq_0 y \equiv \forall z.\, z \in_0 x \equiv z \in_0 y$, and [once I have defined "$\in_1$"] $x \simeq_1 y \equiv \forall z.\, z \in_1 x \equiv z \in_1 y$.)

Unique(y) $\equiv_{df} \forall x.\, x \simeq y \rightarrow x = y$; x and y are **disparate**$\equiv_{df} \neg\, x \simeq y$.

$x \subseteq y \equiv_{df}$ set(x) $\&\ :\forall z.\, z \in x \rightarrow z \in y$.

$x \subset y \equiv_{df} x \subseteq y\ \&\ \exists z \in y.\, z \notin x$.

6.3 Mapping

maps(f, a, b) $\equiv_{df}$ ∀p ∈ f ∃x ∈ a ∃y ∈ b. p = ⟨x, y⟩ &
 ∀x ∈ a ∃! y ∈ b ∃p ∈ f. p = ⟨x, y⟩ &
 ∀y ∈ b ∃x ∈ a ∃p ∈ f. p = ⟨x, y⟩.
 I.e., the function f maps a onto b, not necessarily one-to-one.
function(f) $\equiv_{df}$ ∃a, b. maps(f, a, b).
domain(f) $=_{df}$ ıa. set(a) & ∃b. maps(f, a, b).
range(f) $=_{df}$ ıb. set(b) & ∃a. maps(f, a, b).
 Note that, unlike $\text{maps}_{\text{formula}}$, below, if maps(f, a, b), then f's domain is a and range is b.
$\textbf{maps}_{\textbf{1-1}}$(f, a, b) $\equiv_{df}$ ∀p ∈ f ∃x ∈ a ∃y ∈ b. p = ⟨x, y⟩ &
 ∀x ∈ a ∃! y ∈ b ∃p ∈ f. p = ⟨x, y⟩ &
 ∀y ∈ b ∃! x ∈ a ∃p ∈ f. p = ⟨x, y⟩.
 Note that $\text{maps}_{1\text{-}1}$(f, a, b) implies maps(f, a, b).
a≈b $\equiv_{df}$ ∃f. $\text{maps}_{1\text{-}1}$(f, a, b). (Read "a is **equinumerous** to b.")
FUNCTION(ϕ, a) $\equiv_{dfs}$ ∀x ∈ a. ∃! y. ϕ(x, y).[21]
$\textbf{maps}_{\textbf{formula}}$(ϕ, a, b) $\equiv_{dfs}$ FUNCTION(ϕ, a) & ∀x ∈ a. ∃y ∈ b. ϕ(x, y) & ∀y ∈ b. ∃x ∈ a. ϕ(x, y).

6.4 Well-Foundedness

unending-chain(c) $\equiv_{df}$ nonempty(c) & ∀x ∈ c ∃y ∈ c. y ∈ x.
ill-founded(w) $\equiv_{df}$ ∃c. w ∈ c & unending-chain(c).
wf(w) $\equiv_{df}$ ¬ ill-founded(w). (Read "**well-founded**.")
low(x) $\equiv_{df}$ ∃f∃w. wf(w) & maps(f, w, x).

We will not need the notion of a low class, since the restricted Axiom of Replacement will imply that any such would correspond to a set. Informally, say that there are **many** P's if the class of P's does not correspond to a low set. Church [1974a], page 298 defines a set as low if it is equinumerous to a well-founded set; it is not hard to show the two definitions equivalent in the presence of a global well-ordering. Unbeknownst to me, Church's abandoned consistency proof (Box 47, Folder 10) has a predicate "retrogressive," which is similar to my unending chain: x ε r $\rightarrow_x$ (Ey) . y ε r . y ε x.

transitive(a) $\equiv_{df}$ ∀y∀z. z ∈ y & y ∈ a. → z ∈ a.

6.5 Infinity and Ordering

Dedekind-infinite(x) $\equiv_{df}$ ∃y. y ⊂ x & x ≈ y.
Dedekind-finite(x) $\equiv_{df}$ ¬Dedekind-infinite(x).
totally-linearly-orders(R, a) $\equiv_{dfs}$
 ∀x ∈ a. ¬xRx &
 ∀x ∈ a ∀y ∈ a ∀z ∈ a. xRy & yRz → xRz &
 ∀x ∈ a ∀y ∈ a. xRy ∨ x = y ∨ yRx.

[21] I will systematically confuse function symbols with relation symbols which I have proved functional.

well-orders(R, a) $\equiv_{dfs}$
totally-linearly-orders(R, a) &
$\forall z \subseteq a.\,$nonempty(z) $\rightarrow \exists m \in z\ \forall w \in z.\, \neg wRm$.
ordinal(x) $\equiv_{df}$ transitive(x) & wf(x) & well-orders($\in$, x) & set(x) & $\forall z \in x.\,$set(z).
$\omega =_{df} \iota w.\, \forall x.\, x \in w \equiv.$ Dedekind-finite(x) & ordinal(x).
If α and β are ordinals, define $\alpha<\beta$ iff$_{df}$ $a \in \beta$; $\alpha \leq \beta$ iff$_{df}$ $a \in \beta \vee \alpha = \beta$.

It may seem redundant to require that an ordinal be both well-founded and well-ordered by $\in$; but the obvious proof of my version of well-foundedness from well-ordering requires the existence of the given set's intersection with an unending chain, which in turn apparently requires well-foundedness.

I will use small Greek letters as variables for ordinals. As the ordinals less than some fixed ordinal μ perform the same function here as do the natural numbers less than or equal to some fixed natural number m in [Church 1974a], and Church uses j in such contexts, I will here use i, j, k, and n extensively as variables for ordinals less than or equal to μ.

6.6 Class Abstracts

"$\{x \mid \phi(x)\}$", or, for emphasis "$\{x \mid \phi(x)\}_0$" indicates $\iota s.\, set_0(s)$ & $\forall x.\, x \in_0 s \equiv \phi(x)$, if that exists, otherwise merely the virtual class (i.e., predicate) $\phi(x)$. Analogously "$\{x \mid \phi(x)\}_1$". The latter notion is of little interest if ϕ was defined in terms of "$\in_0$" rather than "$\in_1$". Note that a class abstract$_0$ is never an urelement$_0$.
"$\{x \in y \mid \phi(x)\}$" abbreviates "$\{x \mid x \in y$ & $\phi(x)\}$."
For $\tau(y)$ a term, $\{\tau(y) \mid \phi(y)\} =_{dfs} \{x \mid \exists y.\, \phi(y)$ & $x = \tau(y)\}$.

"Δ" normally means **symmetric difference**. For typographic convenience, "Δ" will be used in Part II for symmetric difference in the sense of $\in_2$; "δ" will mean symmetric difference in the sense of $\in_0$. I.e., $x\,\delta\, y =_{df} \{z \mid z \in_0 x \not\equiv z \in_0 y\}_0$, and $x\,\Delta\, y =_{df} \{z \mid z \in_2 x \not\equiv z \in_2 y\}_2$. (" $\boldsymbol{\partial}$ ", defined in a later section, will also be distinct.)
$\Sigma x =_{df} \{z \mid \exists y.\, z \in y$ & $y \in x\}$; $x \cup y =_{df} \{z \mid z \in x \vee z \in y\}$.
$\cap x =_{df} \{z \mid \forall y \in x.\, z \in y\}$; $x \cap y =_{df} \{z \mid z \in x$ & $z \in y\}$.
$a - b =_{df} \{x \in a \mid x \notin b\}$.
POW(a)$=_{df} \{x \mid x \subseteq a\}$.
Define for a term τ and ordinal α, $\bigcup_{\alpha \leq j < \mu} \tau(j) =_{dfs} \{x \mid \exists j\ \exists y.\, \alpha \leq j < \mu$ & $x \in y \in \tau(j)\}$. Similarly $\bigcup_{j \leq \mu} \tau(j) =_{dfs} \{x \mid \exists j\ \exists y.\, j \leq \mu$ & $x \in y \in \tau(j)\}$.
Define $f`x =_{df} \iota y.\, \langle x, y\rangle \in f$; $\phi^{\leftharpoonup} y =_{df} \iota x.\, \phi(x) = y$. (Read "$\phi$ **inverse** of y.")
$\phi``a =_{dfs} \{\phi(x) \mid x \in a\}$, for a a set; for a an urelement, $\phi``a =_{dfs} a$. (Read "the **image** of a under ϕ.") More explicitly, and partially following [Levy 1979], p. 27 for the non-empty case, $\phi``a =_{dfs} a$, if empty(a), else $\{y \mid \exists x \in a\,.\, y = \phi(x)\}$. Note that the obvious simpler definition in terms of class abstracts would have meant that the value of " for any formula and any urelement would have been the empty set, but it will be important for results about j-isomorphism that it instead be the urelement itself.
$\phi^{\Leftarrow}(h) =_{df} \{x \mid \phi(x) = h\}$.

Define $f|a =_{df} \{\langle x, y\rangle \in f \mid x \in a\}$. (Read "f **restricted** to a.")

7 The Axioms

7.1 The Basic Axioms

Extensionality: $\forall a \forall b.\ \text{nonempty}(a)\ \ \&\ \ \forall z:\ z \in a \equiv z \in b.\ \Rightarrow a = b$

Null Set: $\forall x.\ x \notin \emptyset$

Pair: $\forall x \forall y \exists p \forall w.\ w \in p \equiv (w = x \vee w = y)$

Well-Founded Sum Set: $\forall z.\ \text{wf}(z) \Rightarrow \exists u \forall x.\ x \in u \equiv .\ \exists y.\ x \in y\ \ \&\ \ y \in z$

Well-Founded Power Set: $\forall x.\ \text{wf}(x) \Rightarrow \exists p \forall z.\ z \in p \equiv z \subseteq x$

Infinity: $\exists w \forall x.\ x \in w \equiv .\ \text{Dedekind-finite}(x)\ \ \&\ \ \text{ordinal}(x)$

Well-Founded Replacement: a schema, one instance for each two-place predicate φ:
$\forall a.\ \text{wf}(a)\ \ \&\ \ \text{FUNCTION}(\varphi, a) \Rightarrow \exists b.\ \text{maps}_{\text{formula}}(\varphi, a, b)$

7.2 Global Choice

A global well-ordering, as in [Church 1974a]:

Axiom Schema of Global Well-Ordering:

$\forall x.\ \neg x\mathscr{G}x\ \ \&$
$\forall x \forall y.\ x\mathscr{G}y\ \ \&\ \ y\mathscr{G}z \rightarrow x\mathscr{G}z\ \ \&$
$\forall x.\ \varphi(x) \Rightarrow \exists y.\ \varphi(y)\ \ \&\ \ \forall z.\ \varphi(z) \rightarrow y\mathscr{G}z \vee y{=}z$

For convenience below, we will use a slight rearrangement of the global well-ordering, in which ∅ is the first element. I.e., define $x\mathscr{G}'y$ iff $(x{=}\emptyset\ \ \&\ \ y{\neq}\emptyset) \vee (x{\neq}\emptyset\ \ \&\ \ y{\neq}\emptyset\ \ \&\ \ x\mathscr{G}y)$. By abuse of notation, I will use $\mathscr{G}$ for $\mathscr{G}'$.

7.3 Foundation

Axiom of Foundation: $\forall x.\ \text{wf}(x)$

7.4 Base Theory

RZFU (the base theory) is the Basic Axioms plus the following axiom:

Urelement Bijection Axiom:

$\forall x.\ \text{set}(x) \rightarrow \exists! u.\ u = \Upsilon(x)\ \ \&$
$\forall x \forall u.\ u = \Upsilon(x) \Rightarrow \text{urelement}(u)\ \ \&\ \ \text{set}(x)\ \ \&$
$\forall x \forall y \forall u.\ u = \Upsilon(x)\ \ \&\ \ u = \Upsilon(y) \Rightarrow x = y\ \ \&$
$\forall u.\ \text{urelement}(u) \Rightarrow \exists x.\ u = \Upsilon(x)$

The last clause is not used in the proof of the Basic Axioms Theorem; it is only needed for the final construction, hence "Injection" rather than "Bijection" in the name of the axiom in the 1993 version of this paper.

This axiom will be used, via a Cantor-Schroeder-Bernstein-Dedekind construction, to show a bijection from the class of urelements to the class of indexes, defined below, which will be used to keep track of the new, ill-founded, sets.

7.5 CUSι

CUSι will consist of the Basic Axioms plus the following axioms, where $\backsimeq^j$ (j-isomorphism) will be defined below (III.18.2).

Restricted Axiom of Generalized Frege Cardinals:

$\forall j \in \omega\ \forall b.\ wf(b) \Rightarrow \exists F\ \forall x.\ x \in F \equiv b \backsimeq^j x$

Note that, while b is restricted to well-founded sets, x is not. Thus, given a reasonable amount of transitivity (which will be non-trivial), sets j-isomorphic to a well-founded set may also have generalized Frege cardinals.

Unrestricted Axiom of Symmetric Difference:

$\forall x \forall y \exists z \forall w.\ w \in z \equiv (w \in x \not\equiv w \in y)$

Note that, since the 0-cardinal of anything is the universal set, symmetric difference also gives us unrestricted complementation. Trivially this gives us unrestricted union of disjoint sets, and hence (the non-trivial case for) adjunction (i.e., the existence of $x \cup \{y\}$.) It does not seem to give us general pairwise union, however, so the following axiom is also necessary. Pairwise union together with complement will give pairwise intersection, of course, by the usual identity: $a \cap b = \sim(\sim a \cup \sim b)$.

Unrestricted Axiom of Pairwise Union:

$\forall x \forall y \exists z \forall w.\ w \in z \equiv (w \in x \vee w \in y)$

8 Elementary Lemmata

Uniqueness of Pairs

The set required by the Pair Axiom will be unique by Extensionality, since the required set is nonempty. (Note that the name is slightly inaccurate, since the case x=y implies the existence of singletons as well.) This uniqueness is not necessarily preserved in an arbitrary interpretation, though it will be for any interpretation of interest.

Well-Founded Pairwise Union

Observe that the Well-Founded Sum Set Axiom gives us pairwise union for well-founded sets in the Base Theory:

Lemma 8.1 (Pairwise Union for Well-Founded Sets)**.** $\forall x \forall y\ wf(x)\ \&\ wf(y) \rightarrow \exists z \forall w.\ w \in z \equiv (w \in x \vee w \in y)$

Proof. Let x and y be well-founded; by the Pair Axiom, {x, y} exists. It is well-founded; assume not: So there is an unending chain c containing {x, y}. Therefore either x or y is in c, and hence is ill-founded, contradiction.

So by Well-Founded Sum Set Axiom, $\exists u \forall r.\ r \in u \equiv .\ \exists s.\ r \in s\ \ \&\ \ s \in \{x, y\}$, which implies $\forall r.\ r \in u \equiv .\ r \in x \vee r \in y$. □

Thus it is unnecessary to add an axiom for well-founded pairwise union to the base theory. The unrestricted version for CUSι will depend on the details of the sequence of equivalence relations; it is not necessarily true for an arbitrary ∈†-interpretation, defined below.

9 ∈†-Interpretations and Proof of the Basic Axioms

I define a type of interpretation, an ∈†-interpretation, and show that any such interpretation over the base theory automatically satisfies the Basic Axioms except for Extensionality. The proofs are straightforward, since the axioms are restricted to well-founded sets, whose rôles do not change in the interpretation. The use of urelements avoids much of the tedium of [Church 1974b] and, to a lesser extent, [Mitchell 1976] and [Forster 2001]. It may also make similar interpretations of different ill-founded set theories more convenient, since it eliminates the initial need to verify the uninteresting old axioms and allows immediate attention to Extensionality and the new axioms.

An ∈†-interpretation will be a relation (called $\in_1$) defined in the form below, together with two ancillary two-place predicates Φ and ϒ', satisfying the additional requirements below. $\in_1$ will differ from the base membership relation only in that urelements (in the old sense) become ill-founded sets in the new sense. Φ and ϒ' are partly-specified but otherwise arbitrary in Part I. It may be easier, for now, for the reader to think of ϒ' as the injection ϒ of sets to urelements required by the Urelement Bijection Axiom, though in Part III a rearrangement will be necessary to avoid too many unused urelements. (In Part II, either ϒ or ϒ' would suffice, so for simplicity ϒ will be used.)

9.1 ∈†-Interpretations

Let $\in_1$, ϒ', and Φ be two-place formulæ defined in the language of the base theory. Let "$\in_0$" denote the usual membership relation; the "$_0$" will merely emphasize the distinction from the newly-defined membership relation, and avoids confusion with Church's notation, which adopts the opposite convention. (Similarly, subscripts 0 and 1 will be used to distinguish other formulæ defined in terms of the old and new membership relations. Where these formulæ already include subscripts, a comma will be used to separate the 0 or 1.) Abbreviate **unaltered**(x) $\equiv_{df}$ $\forall y.\ y \in_0 x \equiv y \in_1 x$, and define **altered**(x) $\equiv_{df}$ ¬unaltered(x).

(In Parts II and III, which treat $\in_2$ and $\in_3$ respectively, these terms will be redefined for convenience to suit the context.)

Definition Schema: $\in_1$, Υ', and Φ constitute an ∈†-interpretation iff_{dfs}

$\in_1$ Definition

$x \in_1 y \equiv$
(a) urelement(y) & $\exists$L. y = Υ'(L) & Φ(L, x)
$\vee$
(b) $x \in_0 y$.

Υ' Injection Requirement

$\forall x \forall u.\ u = \Upsilon'(x) \Rightarrow \text{urelement}_0(u)\ \&\ \text{set}_0(x)\ \&$
$\forall x \forall y \forall u.\ u = \Upsilon'(x)\ \&\ u = \Upsilon'(y) \Rightarrow x = y$

Ill-Foundedness Requirements

(1) $\forall x.\ \text{altered}(x) \rightarrow \text{ill-founded}_1(x)$
(2) $\forall x \forall y.\ \text{ill-founded}_1(x)\ \&\ x \subseteq_1 y \rightarrow \text{ill-founded}_1(y)$
(3) $\forall x \forall y.\ \text{ill-founded}_1(x)\ \&\ x \in_1 y \rightarrow \text{ill-founded}_1(y)$

Discussion I will prove, in the Base Theory, for an arbitrary ∈†-interpretation, the interpretation of each of the Basic Axioms except Extensionality. The current goal of this result is a relative consistency proof for the special case of an ∈†-interpretation which is my interpretation of CUSι in the Base Theory, but the result might also be useful for other ill-founded set theories.

The domain of the interpretation is the same as that of the ground model. At this level of generality, however, without Foundation in the base theory or Extensionality in the interpretation, the sets of the ground model need not be definable within the interpretation. This will be different for the relation $\in_3$ in Part III.

The altered objects are urelements_0, whose membership is decided by clause (a) of the definition above. Informally, the altered objects will sometimes be called the new sets, where "set" is used in the sense of the new membership relation, since they are urelements in the sense of the old.

Ill-Foundedness Requirements (2) and (3) are not as trivial as they seem, since we don't have pairwise union in general for the new sets. Unrestricted pairwise union will be true in the interpretation of CUSι, but is not necessarily true in general for ∈†-interpretations.

Informally, the most obvious ways to prove (2) and (3) fail. If we have an ill-founded_1 set w, and an unending-chain_1(c), and wish to show that a superset_1 s of w, (or a set x containing_1 w) is ill-founded_1, we could show the existence of $c' = c \cup \{s\}$ (respectively $c \cup \{x\}$.) There are two obvious approaches: First, we could try to show the existence of such a c' in the new theory, but this presumably would require an unrestricted axiom of pairwise union in the new theory. Second, we could try to show the existence of a suitable unending chain in the base theory; but the given c might not even be a set in the base theory.

As an alternative, we could try to find a low_1 subset_1 of c which contains_1 x and is still an unending-chain_1; Replacement in the base theory might then give us the required union of the new chain and {w}. The Axiom of Dependent Choices is the obvious candidate for constructing such a subchain. This was the motivation for the even weaker Low Chain Axiom in some of my previous work: ∀a∀c. unending-chain(c) & a ∈ c ⇒ ∃d. low(d) & unending-chain(d) & a ∈ d, which is a consequence of either Dependent Choices or Foundation. A still weaker alternative would be the Chain Adjunction Axiom: ∀a∀c. unending-chain(c) & a ∈ c & s ∈ a ⇒ ∃d. unending-chain(d) & s ∈ d. This is normally a consequence of the Low Chain Axiom (given low pairwise union in the interpretation), or of unrestricted pairwise union, or even merely unrestricted adjunction: take c ∪ {s} as d. At the current level of generality, demonstrating the interpretation of these axioms would be inconvenient at this stage of the proof, so I adopt the Ill-Foundedness Requirements instead.

9.2 Basic Axioms Theorem

Theorem 9.1 (Basic Axioms Theorem)**.** For an arbitrary ∈†-interpretation $\in_1$, the interpretations in terms of $\in_1$ of the Basic Axioms except Extensionality are provable from the Base Theory.

The proofs for each of the Basic Axioms except Extensionality will take the remainder of Part I, but I begin with a simple lemma. (Henceforward I will use heavily the convention noted above, about complex expressions using subscript zero or one to distinguish notions defined in terms of the new membership relation from those defined in terms of the old.)

Lemma 9.2 (Well-Foundedness Lemma)**.** $\text{wf}_1(x) \rightarrow \text{wf}_0(x)$.

Recall the definitions: ill-founded(w) $\equiv_{df}$ ∃c. w ∈ c & unending-chain(c), and unending-chain(c) $\equiv_{df}$ nonempty(c) & ∀x ∈ c ∃y ∈ c. y ∈ x.

Proof. First, note trivially that no unending-chain, nor any member of one, is empty. So no member_0 of an unending-chain_0 is an urelement_0 by Ill-Foundedness Requirement (1). Thus any unending-chain_0 is unaltered, as is its every member_0, hence also its every member_1; hence any unending-chain_0 is also an unending-chain_1.

Prove the contrapositive of the lemma, i.e., assume $\text{ill-founded}_0(w)$, and show $\text{ill-founded}_1(w)$. So ∃c. $w \in_0 c$ & $\text{unending-chain}_0(c)$. I.e., $\text{nonempty}_0(c)$ & $\forall x \in_0 c\ \exists y \in_0 c.\ y \in_0 x$. By the preceding, c is unaltered, as is its every member. Thus $\text{nonempty}_1(c)$ & $\forall x \in_1 c\ \exists y \in_1 c.\ y \in_1 x$, and $w \in_1 c$, as required. □

9.2.1 Proofs of the Interpretations of the Basic Axioms except Extensionality in an Arbitrary $\in\dagger$-Interpretation

Null Set: $\forall x.\ x \notin_1 \emptyset$

Proof. Ø is not an urelement (by fiat), so case (a) of the definition of $\in_1$ does not apply. Case (b) is never true because Ø is empty_0, so nothing is an element_1 of Ø. □

Pair: $\forall x \forall y \exists p \forall w.\ w \in_1 p \equiv (w = x \vee w = y)$

Proof. Choose x and y arbitrary objects in the base theory (note that the interpretation does not change equality or the universe.) By the uninterpreted axiom, $\exists p \forall w.\ w \in_0 p \equiv (w = x \vee w = y)$. This p is not an urelement, hence is unaltered. Thus $\forall w.\ w \in_1 p \equiv w \in_0 p \equiv (w = x \vee w = y)$, as required. □

Well-Founded Sum Set: $\forall z.\ \text{wf}_1(z) \Rightarrow \exists u \forall x.\ x \in_1 u \equiv .\ \exists y.\ x \in_1 y\ \ \&\ \ y \in_1 z$

Proof. Since z is well-founded_1, by Ill-Foundedness Requirement (1) it is unaltered. z is well-founded_0 by the Well-Foundedness Lemma. By the axiom in the base theory, $\exists u' \forall x.\ x \in_0 u' \equiv .\ \exists y.\ x \in_0 y\ \ \&\ \ y \in_0 z$. If this u' is an urelement_0, let u be Ø, and the result follows trivially; otherwise let u be u'. Thus u is unaltered. Show:

$\forall x.\ x \in_1 u \equiv .\ \exists y.\ x \in_1 y\ \ \&\ \ y \in_1 z$. Choose an arbitrary x.

Part 1: Assume $x \in_1 u$; show $\exists y.\ x \in_1 y\ \ \&\ \ y \in_1 z$.

Consider the y whose existence is required by the uninterpreted version of this axiom, with $x \in_0 y\ \ \&\ \ y \in_0 z$. But $y \in_0 z$ and z is unaltered, so $y \in_1 z$. And $x \in_0 y$, so y is not an urelement_0, hence unaltered. So $x \in_1 y$, as required.

Part 2: Assume $\exists y.\ x \in_1 y\ \ \&\ \ y \in_1 z$; show $x \in_1 u$.

Since u is unaltered, it will suffice to show $x \in_0 u$. z is unaltered, so $y \in_0 z$. Also z is well-founded_1, thus by Ill-Foundedness Requirement (3), so is y, and hence by (1) y is unaltered. Thus $x \in_0 y$, so by the use above of the axiom in the base theory, $x \in_0 u'$, which has the same members_0 as u, since u is unaltered, as required. □

Well-Founded Power Set: $\forall x.\ \text{wf}_1(x) \Rightarrow \exists p \forall z.\ z \in_1 p \equiv z \subseteq_1 x$

Proof. Since x is well-founded_1, it is unaltered. x is well-founded_0 by the Well-Foundedness Lemma. By the axiom in the base theory, $\exists p \forall z.\ z \in_0 p \equiv z \subseteq_0 x$. p is not an urelement_0 (it contains_0 at least Ø) so it is unaltered. Take an arbitrary z; show:

$z \in_1 p \equiv z \subseteq_1 x$

Part 1: Assume $z \in_1 p$; show $z \subseteq_1 x$.

I.e., choose arbitrary $y \in_1 z$; show $y \in_1 x$. p is unaltered, so $z \in_0 p$; thus $z \subseteq_0 x$. My definition of $\subseteq$ specifically excludes urelements, so z is not an urelement$_0$, hence is unaltered. Since $y \in_1 z$, $y \in_0 z$; but $z \subseteq_0 x$, so $y \in_0 x$. x is unaltered, so $y \in_1 x$, as required.

Part 2: Assume $z \subseteq_1 x$; show $z \in_1 p$.

p is unaltered, so it will suffice to show $z \in_0 p$, which is equivalent to $z \subseteq_0 x$ by choice of p. So take an arbitrary $y \in_0 z$; show $y \in_0 x$.

Subcase 2a: z is altered, hence ill-founded$_1$ by Ill-Foundedness Requirement (1). But $z \subseteq_1 x$, and x is well-founded$_1$, contradicting Ill-Foundedness Requirement (2).

Subcase 2b: z is unaltered. Since $y \in_0 z$, $y \in_1 z$. $z \subseteq_1 x$, so $y \in_1 x$. But x is unaltered, so $y \in_0 x$, as required. □

Infinity: $\exists w \forall x.\ x \in_1 w \equiv .$ Dedekind-finite$_1$(x) & ordinal$_1$(x)

The proof will require three results, below. Let ω denote the set required by the uninterpreted axiom; it will suffice to show that this set also has the properties required by the interpretation of the axiom. Since ω is non-empty$_0$ and hence unaltered, it will suffice to show that ∀x. Dedekind-finite$_1$(x) & ordinal$_1$(x) ≡ Dedekind-finite$_0$(x) & ordinal$_0$(x).

The interpretation of the axiom will follow from three results, the first of them trivial: the Set Lemma, the Ordinal Absoluteness Theorem, and the Dedekind Infinite Absoluteness Lemma.

Lemma 9.3 (Set Lemma)**.** ∀z. set$_0$(z) → set$_1$(z).

Proof. I.e., expanding definitions, show $z = \emptyset \vee$ nonempty$_1$(z). If $z=\emptyset$, we are done, since equality is unchanged. Otherwise, by the corresponding clause of the definition of set$_0$(z), nonempty$_0$(z). So z is unaltered, hence nonempty$_1$, hence a set$_1$. □

Theorem 9.4 (Ordinal Absoluteness Theorem)**.** ∀α. ordinal$_0$(α) ≡ ordinal$_1$(α).

The property of being an ordinal is absolute, i.e., is true of an object in the sense of $\in_1$ iff it is true of that object in the sense of $\in_0$. This will permit omitting subscripts 0 and 1 when saying that something is an ordinal.

Proof.

Part 1: $\text{ordinal}_0(\alpha) \rightarrow \text{ordinal}_1(\alpha)$. Assume $\text{ordinal}_0(\alpha)$.

Expanding the definition, show: $\text{transitive}_1(\alpha)$ & $\text{wf}_1(\alpha)$ & $\text{well-orders}_1(\in, \alpha)$ & $\text{set}_1(\alpha)$ & $\forall z \in_1 \alpha.\ \text{set}_1(z)$. The proof of the clauses will be in the following order: (i) $\text{set}_1(\alpha)$, (ii) $\forall z \in_1 \alpha.\ \text{set}_1(z)$, (iii) $\text{transitive}_1(\alpha)$, (iv) $\text{wf}_1(\alpha)$, (v) $\text{well-orders}_1(\in, \alpha)$.

Observation: Since $\text{set}_0(\alpha)$ & $\forall z \in_0 \alpha.\ \text{set}_0(z)$, α is unaltered, and so are all its members_0; hence all its members_1 are unaltered as well. (Informally, this implies that the entire transitive closure is unaltered, but the definition of transitive closure requires some extra machinery and will not be introduced until Part III.)

Conjunct i: Show $\text{set}_1(\alpha)$. Note that $\text{set}_0(\alpha)$, by the definition of $\text{ordinal}_0(\alpha)$, so by the Set Lemma we are done.

Conjunct ii: Show $\forall z \in_1 \alpha.\ \text{set}_1(z)$. Let $z \in_1 \alpha$. α is a set_0, hence unaltered; so $z \in_0 \alpha$, hence $\text{set}_0(z)$, by the corresponding clause of the definition of $\text{ordinal}_0(\alpha)$. So $\text{set}_1(z)$ by the Set Lemma.

Conjunct iii: Show $\text{transitive}_1(\alpha)$, i.e., $\forall y\, \forall z.\ z \in_1 y$ & $y \in_1 \alpha.\ \rightarrow z \in_1 \alpha$. Since α is an ordinal_0, it is a set_0 and hence unaltered, as is its every member_0. So, assuming $z \in_1 y$ & $y \in_1 \alpha$, we have that $y \in_0 \alpha$ and y is unaltered. Therefore $z \in_0 y$. Since α is transitive_0, $z \in_0 \alpha$, so $z \in_1 \alpha$ as required.

Conjunct iv: Show $\text{wf}_1(\alpha)$. Assume not; i.e., $\exists c.\ \alpha \in_1 c$ & $\forall z{\in_1}c\ \exists y{\in_1}c.\ y{\in_1}z$. If this c were unaltered, the following would be much simpler. Still, membership_1 in c is a formula in the language of the base theory, since $\in_1$ is defined in terms of it. So by the Axiom of Replacement (restricted to well-founded sets) in the base theory, we can show the existence of the subset_0 of α containing all members_1 of c. (This would be simpler with the Axiom of Separation, which is clearly a consequence of Replacement. With it we would simply define c', below, as the subset of α containing members_1 of c.) Consider the formula ϕ which maps z to z if $z{\in_1}c$, and otherwise to an arbitrary object. (For the sake of definiteness, define it as follows. By the assumption for the sake of a contradiction, since $\alpha \in_1 c$, $\exists y{\in_1}c.\ y{\in_1}\alpha$. Choose this y; since α is unaltered, y is also a member_0 of α.) Call the image of α under this mapping c'; since α is well-founded_0, it exists by the uninterpreted Axiom of Replacement. It is a subset_0 of a well-founded_0 set, hence well-founded_0. By the Axiom of Pair and the Well-Founded Pairwise Union theorem (both in the base theory) $c' \cup \{\alpha\}$ exists; call this d. Note that every member_0 of d is either α or a member_0 of α.

Claim: d is an unending chain_0 containing_0 α; so α is ill-founded_0, contradicting the hypothesis. Let z be an arbitrary member_0 of d; show that there is a w such that $w \in_0 d$ & $w \in_0 z$.

Case 1: $z=\alpha$. I.e., show there is a w such that $w \in_0 d$ & $w \in_0 \alpha$. By the assumption for the sake of a contradiction, $\exists y{\in_1}c.\ y{\in_1}\alpha$. α is unaltered, so $y{\in_0}\alpha$. By the definition of c', $y{\in_0}c'$, hence $y{\in_0}d$, as required.

Case 2: $z \neq \alpha$. Then $z \in_0 c'$, so $z \in_1 c$. c is an unending chain$_1$, so there is a w such that $w \in_1 c$ & $w \in_1 z$. z is a member$_0$ of α, hence unaltered, so $w \in_0 z$. Since α is transitive$_0$, $w \in_0 \alpha$. Thus by the definition of c', $w \in_0 c'$; hence $w \in_0 d$, as required.

Conjunct v: Show well-orders$_1(\in_1, \alpha)$, i.e.,

(a) totally-linearly-orders$_1(\in_1, \alpha)$ &

(b) $\forall z \subseteq_1 \alpha$. nonempty$_1(z) \rightarrow \exists m \in_1 z\ \forall w \in_1 z\ .\ w \notin_1 m$.

(a) totally-linearly-orders$_1(\in_1, \alpha)$: By inspection of the definition of totally-linearly-orders, this depends only on equality and membership$_1$ in α or in members$_1$ of α. Since α and all its members$_0$ or members$_1$ are unaltered, this subcase follows from totally-linearly-orders$_0(\in_0, \alpha)$.

(b) $\forall z \subseteq_1 \alpha$. nonempty$_1(z) \rightarrow \exists m \in_1 z\ \forall w \in_1 z\ .\ w \notin_1 m$: So let z be a non-empty$_1$ subset$_1$ of α. Since α is well-founded$_1$ by the proof of the preceding conjunct, by Ill-Foundedness Requirement (2), z is well-founded$_1$ as well, hence unaltered. Thus z is a nonempty$_0$ subset$_0$ of α, so by the corresponding clause of the definition of α's being an ordinal in the sense of $\in_0$, $\exists m \in_0 z\ \forall w \in_0 z\ .\ w \notin_0 m$. Since $m \in_0 z$, $m \in_0 \alpha$, and hence m is unaltered. Thus $m \in_1 z$ & $\forall w \in_1 z\ .\ w \notin_1 m$, as required.

Part 2: ordinal$_1(\alpha) \rightarrow$ ordinal$_0(\alpha)$. Assume ordinal$_1(\alpha)$.

Observation: Since wf$_1(\alpha)$, and thus (by Ill-Foundedness Requirement 3) $\forall z \in_1 \alpha$. wf$_1(z)$, α is unaltered, and so are all its members$_1$; hence all its members$_0$ are unaltered as well.

Conjunct i: Show set$_0(\alpha)$. α is a set$_1$, so $x = \emptyset$ or nonempty$_1(\alpha)$. Case 1: $\alpha = \emptyset$. Equality is unchanged, so set$_0(\alpha)$. Case 2: nonempty$_1(\alpha)$. α is unaltered, so nonempty$_0(\alpha)$.

Conjunct ii: Show $\forall z \in_0 \alpha$. set$_0(z)$. Let $z \in_0 \alpha$. α is unaltered, so $z \in_1 \alpha$, and hence z is a well-founded$_1$ set$_1$ and unaltered. So if z is nonempty$_1$, it is nonempty$_0$; otherwise it is ∅; in either case, z is a set$_0$.

Conjunct iii: Show transitive$_0(\alpha)$, i.e., $\forall y\ \forall z.\ z \in_0 y$ & $y \in_0 \alpha.\ \rightarrow z \in_0 \alpha$. Assume $z \in_0 y$ & $y \in_0 \alpha$; show $z \in_0 \alpha$. By the observation above, α and y are unaltered, so $z \in_1 y$ and $y \in_1 \alpha$. Therefore $z \in_1 \alpha$ and hence $z \in_0 \alpha$.

Conjunct iv: Show wf$_0(\alpha)$. α is well-founded$_1$, so this is immediate from the Well-Foundedness Lemma.

Conjunct v: Show well-orders$_0$($\in_0$, α), i.e.,
(a) totally-linearly-orders$_0$($\in_0$, α) &
(b) $\forall z \subseteq_0 \alpha$. nonempty$_0$(z) → $\exists m \in_0 z\ \forall w \in_0 z$. $w \notin_0 m$.

(a) totally-linearly-orders$_0$($\in_0$, α): As in the proof of the other direction, this follows by inspection, since α is unaltered, as are all its members$_0$ and members$_1$.

(b) $\forall z \subseteq_0 \alpha$. nonempty$_0$(z) → $\exists m \in_0 z\ \forall w \in_0 z$. $w \notin_0 m$: So let z be a non-empty$_0$ subset$_0$ of α. Since z is nonempty$_0$, it is unaltered, hence also a non-empty$_1$ subset$_1$ of α. Thus there is an m such that $m \in_1 z$ & $\forall w \in_1 z$. $w \notin_1 m$. As before, since $m \in_1 z$, $m \in_1 \alpha$, and hence m is unaltered. Thus $m \in_0 z$ & $\forall w \in_0 z$. $w \notin_0 m$, as required, which concludes the proof of the Ordinal Absoluteness Theorem. □

Lemma 9.5 (Dedekind Infinite Absoluteness Lemma)**.** For any ordinal α, Dedekind-infinite$_0$(α) ≡ Dedekind-infinite$_1$(α).

Proof.

Part 1: Assume α is an ordinal with Dedekind-infinite$_0$(α); show Dedekind-infinite$_1$(α).

I.e., assume $\exists y.\ y \subset_0 \alpha$ & ∃f. maps$_{1\text{-}1,0}$(f, y, α), and show $y \subset_1 \alpha$ & ∃g. maps$_{1\text{-}1,1}$(g, y, α). (Recall that ⊂ denotes being a proper subset, and that my definition of subset is restricted to sets.)

α and y are both sets$_0$, hence unaltered, and equality is unchanged by the interpretation; so $y \subset_1 \alpha$. α is a proper superset$_0$ of y, hence nonempty$_0$; so f is nonempty$_0$ and hence unaltered. Ordered pairs$_0$ are also nonempty$_0$ and unaltered; by inspection of the definition of maps$_{1\text{-}1}$, this ensures that maps$_{1\text{-}1}$(f, y, α) in the sense of $\in_1$ as well.

Part 2: Conversely, assume $\exists y.\ y \subset_1 \alpha$ & ∃f. maps$_{1\text{-}1,1}$(f, y, α), and show $y \subset_0 \alpha$ & maps$_{1\text{-}1,0}$(f, y, α).

The first conjunct is much as before, since y is well-founded$_1$, hence unaltered. For the second conjunct, first note that, trivially, if every member of a set is well-founded, it is well founded. Thus since α and y are well-founded$_1$, every pair$_1$ of elements$_1$ from either α or y is well-founded$_1$, and hence likewise for every ordered (Kuratowski) pair. Therefore by the first conjunct of the definition of maps$_{1\text{-}1}$, f is well-founded$_1$, hence unaltered. Therefore, as in the previous case, maps$_{1\text{-}1,0}$(f, y, α). □

Corollary 9.6. ∀x. Dedekind-finite$_1$(x) & ordinal$_1$(x) ≡ Dedekind-finite$_0$(x) & ordinal$_0$(x).

Thus ω is also the set of all Dedekind-finite ordinals in the sense of $\in_1$, as required, which completes the proof of the interpretation of the Axiom of Infinity.

Well-Founded Replacement: a schema, one instance for each two-place predicate φ: $\forall a.\ wf_1(a)\ \ \&\ \ FUNCTION_1(\varphi, a) \Rightarrow \exists b.\ maps_{formula,1}(\varphi, a, b)$.

Proof. Recalling definitions, $FUNCTION_1(\varphi, a) \equiv_{dfs} \forall x \in_1 a.\ \exists! y.\ \varphi(x, y)$, and $maps_{formula,1}(\varphi, a, b) \equiv_{dfs} \forall x \in_1 a.\ \exists! y \in_1 b.\ \varphi(x, y):\ \ \&\ \ : \forall y \in_1 b.\ \exists x \in_1 a.\ \varphi(x, y)$.

So assume a is well-founded$_1$, with φ an arbitrary formula which is functional over a, in the sense of $\in_1$. Let φ_0 be φ with all its occurrences of $\in_1$ expanded into its definition, so it is a formula in terms of $\in_0$; obviously $\forall x \forall y\ \varphi(x, y) \equiv \varphi_0(x, y)$. Since a is well-founded$_1$, it is unaltered and well-founded$_0$ by the Well-Foundedness Lemma. Thus, since equality is unchanged in the interpretation, $\forall x \in_0 a.\ \exists! y.\ \varphi_0(x, y)$, i.e., $FUNCTION_0(\varphi_0, a)$. So by the uninterpreted axiom, $\exists b.\ maps_{formula,0}(\varphi_0, a, b)$.

(Expanding the definition, this is $\forall x \in_0 a.\ \exists! y \in_0 b.\ \varphi_0(x, y):\ \ \&\ \ : \forall y \in_0 b.\ \exists x \in_0 a.\ \varphi_0(x, y)$.) If this b is an urelement, then Ø will also satisfy the requirements, so take Ø as b instead; thus in either case b is unaltered. Since a and b are unaltered, equality is unchanged, and $\varphi(x, y)$ is equivalent to its definitional expansion $\varphi_0(x, y)$, we have $\forall x \in_1 a.\ \exists! y \in_1 b.\ \varphi(x, y):\ \ \&\ \ : \forall y \in_1 b.\ \exists x \in_1 a.\ \varphi(x, y)$. I.e., $maps_{formula,1}(\varphi, a, b)$ as required. □

This establishes the Basic Axioms Theorem.

Part II
Extensionality and Arbitrary Restricted Equivalence Relations

In Part II, I introduce a somewhat different partially-specified membership relation, $\in_2$, and show that it satisfies Extensionality. This membership relation is defined in terms of an arbitrary series of relations satisfying requirements in the section "$\sim^j$ Requirements" (11) below.

The use of Choice is avoided in the base theory, as is Foundation, except, near the end of this Part, for explicitly-mentioned uses of the **Lowness$_0$ Assumption:** $\forall s.\, \mathrm{low}(s)$. I append a superscript "^{S}" to the names of theorems which assume this. In Part III I will show that a specific series of restricted equivalence relations satisfies the requirements in this part, which will establish that my interpretation satisfies Extensionality. The proof of Extensionality in this part is quite involved and special-purpose; readers whose interests are more general may wish to skip to Part III.

10 Elementary Lemmata

As in [Church 1974a], p. 303, I will omit the straightforward proofs that the axioms of Separation, Replacement, and Power Set remain true when their restrictions are loosened from well-founded sets to low sets, and also a similar argument which shows that the union of two low sets is a low set.

Theorem Schema 10.1 (Separation Restricted to Low Sets ($\mathrm{Sep}_{\mathrm{low}}$))**.** $\forall a.$ $\mathrm{low}(a) \Rightarrow \exists b.\, \mathrm{low}(b)\ \&\ \forall x.\, x \in b \equiv .\, x \in a\ \&\ \phi(x)$.

Corollary 10.2. A subclass of a low set is a low set.

Theorem Schema 10.3 (Replacement Restricted to Low Sets ($\mathrm{Rep}_{\mathrm{low}}$))**.** $\forall a.\, \mathrm{low}(a)\ \&\ \mathrm{FUNCTION}(\phi, a) \Rightarrow \exists b.\, \mathrm{low}(b)\ \&\ \mathrm{maps}_{\mathrm{formula}}(\phi, a, b)$.

Theorem 10.4 (Power Set Restricted to Low Sets ($\mathrm{Pow}_{\mathrm{low}}$))**.** $\forall x.$ $\mathrm{low}(x) \Rightarrow \exists p.\, \mathrm{low}(p)\ \&\ \forall z.\, z \in p \equiv\ z \subseteq x$.

Theorem 10.5 (Low Pairwise Union ($\mathrm{Union}_{\mathrm{low}}$))**.** $\forall a, b.$ $\mathrm{low}(a)\ \&\ \mathrm{low}(b) \Rightarrow \exists u.\, \mathrm{low}(u)\ \&\ \forall z.\, z \in u \equiv\ z \in a \vee z \in b$.

It is then easy to prove the following two results.

Lemma 10.6 (Low Symmetric Difference Lemma ($\mathrm{Diff}_{\mathrm{low}}$))**.** $\mathrm{low}(x)\ \&\ \mathrm{low}(y) \Rightarrow\ \mathrm{SET}[x\ \delta\ y]\ \&\ \mathrm{low}(x\ \delta\ y)$.

First note that, since $P \not\equiv Q \rightarrow P \vee Q$, $x\ \delta\ y = \{z \mid z \in x \not\equiv z \in y\}$ is a subclass of $x \cup y = \{z \mid z \in x \vee z \in y\}$. Since x and y are low, then $x \cup y$ is also low, by $\mathrm{Union}_{\mathrm{low}}$ (10.5). So $x\ \delta\ y$ exists and is low by $\mathrm{Sep}_{\mathrm{low}}$ (10.1).

Lemma 10.7. If a is low, then a − b exists and is low.

Likewise by $\mathrm{Sep}_{\mathrm{low}}$ (10.1).

10.1 Weak Arithmetic

To bypass a long uninteresting proof, I avoid induction on the natural numbers. This also keeps open the possibility of application to Quine's *New Foundations,* in which full induction fails even for the natural numbers even with the addition of the Axiom of Counting [Forster 1992], p. 30. Natural proofs that the ordinals are linearly ordered seem to require some form of induction [Forster 1992], p. 44. The behavior of ordinals in Oberschelp is even more obscure. In lieu of induction on ω, I will need only a few simple arithmetic facts. (With the stronger assumptions in Part III, arithmetic will become much easier.)

First note that, even with my unusual definitions, an ordinal is well-founded. By definition, of course, the ordinals less than some ordinal are linearly ordered.

10.1.1 Ordinal Addition

All we require ordinal addition for is the elementary properties below, primarily of oddness and evenness. I do not even need to show that every natural number is either odd or even but not both; I simply will use only such natural numbers. If we assumed definition by recursion on ordinals, the ordinary definition of addition would suffice. Without definition by recursion, we could still define "+0," "+1," and "+2" everywhere, and define addition on the finite ordinals via (Cantor) cardinal addition; see [Levy 1979], §III.3. To spare the reader's patience, and for greater applicability of my construction, I will instead omit the development of the definition, and present only the elementary properties of ordinal addition which I actually need.

Define $\mathbf{0} =_{\mathrm{df}} \varnothing$. $\mathbf{1} =_{\mathrm{df}} \{\varnothing\}$. $\mathbf{2} =_{\mathrm{df}} \{\varnothing, \{\varnothing\}\}$.

Define **odd**(a)$\mathrm{iff}_{\mathrm{df}}$ $\exists n, k \in \omega.\, n = k + k + 1$ & $n \approx a$; **even**(a)$\mathrm{iff}_{\mathrm{df}}$ $\exists n, k \in \omega.\, n = k + k$ & $n \approx a$. **Odd-or-even**(a)$\mathrm{iff}_{\mathrm{df}}$ odd(a) $\not\equiv$ even(a). The **parity** of x is odd (even) $\mathrm{iff}_{\mathrm{df}}$ x is odd (even). (N.b., these predicates may apply to sets, not just to natural numbers.)

10.2 Required Properties of +

The Required Properties of +, a two-place function on the ordinals, are:

(i) $\alpha+\mathbf{0}= \alpha; \alpha+\mathbf{1}= \alpha \cup \{\alpha\}$; $\alpha+\mathbf{2}= (\alpha + 1) + 1$.

(ii) $\forall x.\, \neg\mathrm{odd}(x) \vee \neg\mathrm{even}(x)$.

(iii) Parity Property: If odd-or-even(a) and odd-or-even(b) then {odd(a δ b) $\iff$ [odd(a) $\not\equiv$odd(b)]} and {even(a δ b) $\iff$ [odd(a) $\equiv$ odd(b)]}.

(iv) $\forall \alpha, \beta.$ ordinal(α) & ordinal(β) & $\alpha < \beta \Rightarrow \alpha + 1 \leq \beta$.

Thus if a and b are odd or even, then a δ b is odd iff a is odd and b is even or *vice versa;* a δ b is even iff a and b are both even or both odd.

$\alpha + 1$ will be a set by the Sum Set Axiom since ordinals are well-founded. A useful notion for informal exposition: $\alpha{-}1 =_{\mathrm{df}} \iota\beta.\, \beta + 1 = \alpha$.

Lemma 10.8. The empty set is even.

$0 = 0 + 0$, and the empty set is trivially an ordinal, Dedekind finite, and equinumerous to itself.

Lemma 10.9. Any singleton is odd.

Similarly, any singleton is equinumerous to $1 = 0 + 0 + 1$.

Lemma 10.10. odd-or-even(a) $\rightarrow$ low(a).

Since odd-or-even(a), $\exists n \in \omega.\, n \approx a$. Since n is an ordinal, it is well-founded. Since $n \approx a$, $\exists f.\, \mathrm{maps}_{1\text{-}1}(f, n, a)$, and hence a is low.

Lemma 10.11. odd-or-even(a) & odd-or-even(b) $\Rightarrow$ odd-or-even(a δ b).

Since $\{P \not\equiv Q\} \not\equiv \{P \equiv Q\}$ is a tautology, the symmetric difference of two odd or even sets is odd or even, by the Parity Property.

Lemma 10.12. There are no ordinals α, β such that $\alpha < \beta < \alpha + 1$.

Assume $\alpha \in \beta \in \alpha \cup \{\alpha\}$. Then $\beta \in \alpha$ or $\beta = \alpha$; in either case $\alpha \in \beta \in \alpha$. Thus $\{\alpha, \beta\}$ is an unending chain, and α and β are ill-founded, contradiction. The linear ordering of the ordinals would allow us to prove Required Properties of + (iv) from this lemma.

11 $\sim^{j}$ Requirements

Let μ be an ordinal.[22] Let **j-rep**(ξ) be a two-place function, $\xi \sim^{j} \zeta$ a three-place predicate, and **rank** a one-place function satisfying the following conditions:

(α) $\forall j, k \leq \mu\ \forall x, y.\, j \leq k\ \&\ x \sim^{k} y \Rightarrow x \sim^{j} y$,
(β) $\forall x, y.\, x \sim^{0} y$,
(γ) $\forall x, y.\, x \sim^{\mu} y \equiv x = y$,
(δ) $\forall j \leq \mu\ \forall b\ \exists r.\, r = \text{j-rep}(b)$,
(ϵ) For $0 \leq j \leq \mu, x \sim^{j} y$ iff $\text{j-rep}(x) = \text{j-rep}(y)$,
(ζ) $\forall h.\, \text{rank}(h) \leq \mu$, and $\forall g.\, \text{rank}(g) = j \Rightarrow \exists x.\, g = \text{j-rep}(x)$,
(η) $\text{rank}(\text{0-rep}(\varnothing)) = 0$, and $\neg\exists s\colon \text{low}(s)\ \&\ \forall d.\, d \in s \leftrightarrow \exists x.\, \text{1-rep}(x) = d$.

In prose, say "g is a **j-rep**" iff_{df} $\exists x.\, g = \text{j-rep}(x)$. A j-rep g is **rankable** iff_{df} it is in the domain of rank.

The main requirement on the given sequence of equivalence relations is (α), increasing strictness; $\sim^{0}$ and $\sim^{\mu}$ can be appended to any sequence satisfying it. Requirements (δ), (ϵ), and (ζ) are for the existence of representative functions, and can be satisfied for arbitrary equivalence relations in the presence of either Global Choice or Foundation. For the full proof of the consistency of CUSι, strengthened requirements are necessary; one addition is a generalization of (η), requiring that a well-behaved object whose j-rep is not j-*prolific* (see below) already have a $\sim^{j}$ equivalence class in the base theory. Requirement (η) is basically case $j = 0$ of this condition. Once I have defined "j-*prolific*" and "*daughter*", below, (η) will be a definitional expansion of "0-prolific(0-rep($\varnothing$))"; the sole purpose of this is the result below, Degeneracy/Diversity Property (i) (14.1).

I do not require the converse of the latter conjunct of (ζ); some j-reps may not be rankable. For example, with one simple possible definition of j-rep, the j-rep of a badly behaved object, for any $j \leq \mu$, might be the singleton containing that object, which therefore would have no obvious unique rank. (In Part III I will adopt a definition which avoids this difficulty.)

Requirements (δ) and (ϵ) imply that $\sim^{j}$ is an equivalence relation:

Lemma 11.1. $\forall x, y, z.$ (i) $x \sim^{j} x$ & (ii) $x \sim^{j} y \equiv y \sim^{j} x$ & (iii) $x \sim^{j} y\ \&\ y \sim^{j} z \rightarrow x \sim^{j} z$.

Proof. By (δ) j-rep is everywhere defined; by (ϵ) the preceding are equivalent to the following trivial propositions: (i) $\text{j-rep}(x) = \text{j-rep}(x)$ & (ii) $\text{j-rep}(x) = \text{j-rep}(y) \iff \text{j-rep}(y) = \text{j-rep}(x)$ & (iii) $\text{j-rep}(x) = \text{j-rep}(y)\ \&\ \text{j-rep}(y) = \text{j-rep}(z) \Rightarrow \text{j-rep}(x) = \text{j-rep}(z)$. □

11.1 Inverses

Lemma 11.2. $\forall a \forall x.\, \mu\text{-rep}(x) \in a \iff x \in \mu\text{-rep}^{\leftarrow}\text{“}a$.

[22] If μ has a predecessor, it corresponds to the arbitrary natural number m in [Church 1974a].

Proof. Since no two objects are $\sim^{\mu}$, μ-rep must be one-one, and $\forall x, y.\, x =$ μ-rep$^{\leftarrow}$y iff y = μ-rep(x). Thus μ-rep$^{\leftarrow}$"a = {x | ∃y. y ∈ a & x = μ-rep$^{\leftarrow}$y} = {x | ∃y. y ∈ a & y = μ-rep(x)} = {x | μ-rep(x) ∈ a}. □

By the definition of "$\Leftarrow$", k-rep$^{\Leftarrow}$h = {x | k-rep(x) = h}. In general, this will be a virtual class in the base theory; it is to be distinguished from the possible corresponding set in the interpretation.

11.2 Restricted Equivalence Relations

Note that in the context of Separation and Replacement restricted to low sets, relations which are equivalence relations in ZF may be equivalence relations only on some collection of well-behaved sets, for some useful definition of "well-behaved." E.g., ordinary equinumerosity, which is Church's equiv^1, need be an equivalence relation only on the low sets.

We could use such a restricted relation to define a full equivalence relation which corresponds on the well-behaved objects to the given equivalence relation, though I will not pursue this alternative. Assume $\sim^j$ is well-behaved on some sequence of classes A^j, for $0 < j < \mu$. Call the members of the A^j, **j-agreeable** objects. Define $x \sim^j_{new} y$ iff$_{df}$ $x, y \in A^j \;\&\; x \sim^j y \;.\vee\; x, y \notin A^j \;\&\; x = y$. This will imply that the j-equivalence class of an object which is not j-agreeable is a singleton.

This notion of agreeability corresponds to the restriction of the axioms L^j to well-founded sets in [Church 1974a] p. 305. I use a somewhat different notion, j-purity, below, in the j-Pure j-Isomorphism Absoluteness Theorem (III.20.10). A further later requirement on $\sim^j$ is the generalization of requirement (η) that, among such j-pure objects, those whose j-reps are not *j-prolific* (defined below) already have $\sim^j$ equivalence classes (generally singletons) in the base model. This corresponds to Church's dependent clause, "not empty at level j," [Church 1974a] p. 306.

11.3 Representative Functions

Say that functions ρ^j constitute **representative functions** for $\sim^j$ iff $\forall j, x, y.$ $\rho^j(x) = \rho^j(y) \iff x \sim^j y$. In Part II, I am simply assuming the existence of such representative functions, but will here make a few remarks about the general possibility of generating them. Given an arbitrary equivalence relation, defining a representative function seems to require either Choice or the trick in Scott [1955]; see Levy [1969]. Church uses a strong form of choice, a global well-ordering, and uses as the representative of a class the first member of that class. This makes the construction simpler and, given the degree of Platonism required for a Universal Set, does not seem terribly objectionable. But given New Foundations' inconsistency with Choice [Specker 1953], and the arguments in [Forster 1985], it seemed interesting to see whether the construction could be done only with weak choice. Scott's trick does not require the Axiom of Foundation, but does limit the construction to Generalized Frege Cardinals having some member with a cumulative hierarchy rank. (Church's construction in [1974a] has a similar limitation for different reasons.) No-one seems to have investigated techniques of generating such representatives using specific properties of the given equivalence relations.

12 Descent

Define **daughter**(h, g) iff_{df} $\exists j < \mu\ \exists x.\ j = \text{rank}(g)\ \&\ j\text{-rep}(x) = g\ \&\ j{+}1\text{-rep}(x) = h$. (Read "h is a daughter of g.") Informally, a daughter of g is a member of j + 1-rep"j-rep$^{\Leftarrow}$g, where j = rank(g). Note that h need not be rankable, but g must be. Trivially we have:

Lemma 12.1. If μ is a successor ordinal and $\exists x.\ \mu - 1\text{-rep}(x) = g\ \&\ \mu\text{-rep}(x) = h\ \&\ \text{rank}(g) = \mu - 1$, then daughter(h, g).

Since all objects have the same 0-rep (which has rank 0), any 1-rep is a daughter of the unique 0-rep.

Informally, j-ancestor(h) will be the only member of j-rep"k-rep$^{\Leftarrow}$h, where $k = \text{rank}(h)$ and $j < k < \mu$. Define, for $j < \text{rank}(h) < \mu$, **j-ancestor**(h) $=_{\text{df}}$ $\iota g.\ \exists a.\ j\text{-rep}(a) = g\ \ \&\ \ [\text{rank}(h)]\text{-rep}(a){=}h$. This g in fact exists and is unique. That is,

Lemma 12.2. Assume rank(h) = k, with $j < k < \mu$; then $\exists!\, g \exists a.\ j\text{-rep}(a) = g\ \ \&\ \ k\text{-rep}(a) = h$.

Proof. This proof is more complicated than it need be, in order to make it clear that there is no concealed use of Choice.

Claim 1: Such g exists; i.e., $\exists g\ \exists a.\ j\text{-rep}(a) = g\ \&\ k\text{-rep}(a) = h$. Since h has rank k, by (ζ) there is an a such that k-rep(a) = h. Let g = j-rep(a).

Claim 2: For any z, if k-rep(z) = h, then j-rep(z) = g. Assume k-rep(z) = h = k-rep(a). Thus by (ϵ) $z \sim^k a$. Since $j < k$, by (α) this implies $z \sim^j a$; so, again by (ϵ), j-rep(z) = j-rep(a) = g.

Therefore, for any g', if $\exists z.\ j\text{-rep}(z) = g'\ \&\ k\text{-rep}(z) = h$, then g' = g. Thus g is unique. ☐

Lemma 12.3. $\forall x.\ j < k < \mu\ \&\ \text{rank}(k\text{-rep}(x)) = k\ \Rightarrow\ j\text{-ancestor}(k\text{-rep}(x)) = j\text{-rep}(x)$.

Proof. Trivially, j-rep(x) satisfies the existence portion of the definition of j-ancestor, substituting x for a, j-rep(x) for g, k-rep(x) for h, and k for rank(h). By the preceding, j-rep(x) is the unique such. ☐

Define **j-prolific**(g) iff_{df} $\text{rank}(g) = j\ \&\ \neg\exists s\colon \text{low}(s)\ \&\ \forall d.\ d \in s\ \leftrightarrow\ \text{daughter}(d,g)$. Informally, something is j-prolific iff its rank is j and it has many daughters. The unique 0-rep is 0-prolific, since by (η) there are many 1-reps, all of which are daughters of the unique 0-rep.

13 μ+1-tuples and Sprigs

My $\mu + 1$-tuples are based on Church's $m + 2$-tuples, [1974a], p. 307, with the following differences. The primary difference is that the $\mu + 1$-tuples are not themselves the new sets; they are codes for urelements_0 (in the base theory) which are the new sets_2 (in the interpretation, once I have defined $\in_2$). Second, since μ need not be finite, and I need not worry about collisions, I can use the ordinary definition of sequence rather than Church's deliberately awkward definition. The third difference is that in a Church m+2-tuple, the last component is a flag indicating complementation; in mine, this is signalled by the 0-component, since (using 0-similarity) I assimilate complementation to symmetric difference with the Universe. The fourth difference is that Church's first component, a low set of exceptions, corresponds to my μ-component, a low set of μ-reps, since μ-similarity is equality.

A set of ordered pairs L is a **μ+1-tuple** iff_{df} $\exists r.\, \text{maps}(L, \mu + 1, r)$. Abbreviate L'j to L^j, and call it L's **j-component**. L will usually be denoted "$\big(L^0\, L^1\, \ldots\, L^n \ldots L^\mu\big)$"; I follow Church in using "$\big(\,\big)$" rather than "$\langle\,\rangle$" for this sort of tuple, and omit commas. A μ+1-tuple may have components which are urelements, but attention below will be restricted to μ+1-tuples whose components are all sets.

Specifying the components of a $\mu + 1$-tuple ensures that it exists and is low:

Lemma Schema 13.1. Given a term $\Lambda(\xi)$ defined for $\xi \leq \mu$, there is a unique low $\mu + 1$-tuple L such that $\forall j \leq \mu.\, L^j = \Lambda(j)$.

Proof. Let $\psi(j) = \langle j, \phi(j)\rangle$. By Rep_{low} (10.3), $\exists L.\, \text{maps}_{\text{formula}}(\psi, \mu{+}1, L)\ \&\ \text{low}(L)$; clearly L is as required, and is unique by the Axiom of Extensionality in the base theory. □

I will henceforth use this result without comment.

Informally, the intent is for new sets to be represented by urelements, tagged with a sequence of length μ+1, conventionally represented (L^0 ... L^j ... L^μ), L for short. The idea is that x is a member of a new set (old urelement, tagged by L) if there are an *odd* number of j's such that j-rep(x) is in L^j.

Thus the universal set will be the urelement with tag ({0-rep(∅)} ∅ ... ∅), since everything has the same 0-rep, and 1 is odd. The set of all pairs will, in Part III, be tagged by (∅ {1-rep(2)} ∅ ... ∅), and the singleton function by (∅ ∅ {2-rep(<∅,{∅}>)} ∅ ... ∅). The complement of ω will be tagged by ({0-rep(∅))} ∅ ... μ-rep"ω). Machinery will be developed below, first to formalize the notion of an odd number of j's, and then to restrict the new sets to those needed for the interpretation.

More formally, the *sprig of* a $\mu + 1$-tuple $\big(L^0\, L^1\, \ldots\, L^n \ldots L^\mu\big)$ *for* an object x will be a partially-defined sequence from 0 to μ, with its value for j, j-rep(x) if $\text{j-rep}(x) \in L^j$, and otherwise undefined. Define

$$\textbf{sprig}(\big(L^0\, L^1\, \ldots\, L^n \ldots L^\mu\big), x) =_{\text{df}} \{\langle j, \text{j-rep}(x)\rangle \mid j \leq \mu\ \&\ \text{j-rep}(x) \in L^j\}.$$

Lemma 13.2. $\forall L\ \forall b.\ \mu + 1\text{-tuple}(L) \Rightarrow \mathrm{SET}[\mathrm{sprig}(L, b)]\ \&\ \mathrm{low}(\mathrm{sprig}(L, b))$.

Proof. (Much as in the previous lemma.) Assume $\exists k \leq \mu.\ k\text{-rep}(b) \in L^k$. (If not, sprig(L, b) is simply $\varnothing$.) Define, for $j \in \mu, \phi(j) = \langle j, j\text{-rep}(b)\rangle$ if $j\text{-rep}(b) \in L^j$; otherwise $\phi(j) = \langle k, k\text{-rep}(b)\rangle$. By $\mathrm{Rep}_{\mathrm{low}}$ (10.3), $\exists s.\ \mathrm{maps}_{\mathrm{formula}}(\phi, \mu + 1, s)\ \&\ \mathrm{low}(s)$; clearly this s is sprig(L, b). □

A $\mu + 1$-tuple is *fathomable* iff its sprig for every object is odd or even: Define **fathomable**(L) $\equiv_{\mathrm{df}}$ $\mu + 1\text{-tuple}(L)\ \&\ \forall x.\ \text{odd-or-even}(\mathrm{sprig}(L, x))$.

13.1 Componentwise Symmetric Difference

If $(L^0\ L^1\ \ldots\ L^n \ldots\ L^\mu)$ and $(M^0\ M^1\ \ldots\ M^n \ldots\ M^\mu)$ are $\mu + 1$-tuples, define $(L^0\ L^1\ \ldots\ L^n \ldots\ L^\mu)\ \boldsymbol{\partial}\ (M^0\ M^1\ \ldots\ M^n \ldots\ M^\mu) =_{\mathrm{df}} (L^0\ \delta\ M^0\quad L^1\ \delta\ M^1\ \ldots\ L^n\ \delta\ M^n \ldots\quad L^\mu\ \delta\ M^\mu)$. Call this their **componentwise symmetric difference**. Note that this definition is ill-formed unless each $L^n\ \delta\ M^n$ exists.

Lemma 13.3. For any $\mu + 1$-tuples L and M, if each L^j and M^j is low, then $L\ \boldsymbol{\partial}\ M$ exists.

Proof. Since every L^j and M^j is low, each $L^j\ \delta\ M^j$ exists (and is low) by $\mathrm{Diff}_{\mathrm{low}}$ (10.6). Therefore $(L^0\ \delta\ M^0\quad L^1\ \delta\ M^1\ \ldots\quad L^n\ \delta\ M^n \ldots\quad L^\mu\ \delta\ M^\mu)$ exists by 13.1. □

Lemma 13.4 (Sprig Difference Lemma)**.** If L and M are $\mu + 1$-tuples and $L\ \boldsymbol{\partial}\ M$ exists, then $\forall x.\ \mathrm{sprig}(L, x)\ \delta\ \mathrm{sprig}(M, x) = \mathrm{sprig}(L\ \boldsymbol{\partial}\ M, x)$.

Proof. By 13.2, both sprig(L, x) and sprig(M, x) are low. Thus $\mathrm{sprig}(L, x)\ \delta\ \mathrm{sprig}(M, x)$ exists and is low, by $\mathrm{Diff}_{\mathrm{low}}$ (10.6). Choose j arbitrary less than or equal to μ, and set $c = \langle j, j\text{-rep}(x)\rangle$. Note that anything not of this form is not a member of sprig(L, x), sprig(M, x), $\mathrm{sprig}(L\ \boldsymbol{\partial}\ M, x)$, or $\mathrm{sprig}(L, x)\ \delta\ \mathrm{sprig}(M, x)$. So

$c \in \mathrm{sprig}(L, x)\ \delta\ \mathrm{sprig}(M, x) \iff$
$c \in \mathrm{sprig}(L, x) \not\equiv c \in \mathrm{sprig}(M, x) \iff$
$j\text{-rep}(x) \in L^j \not\equiv j\text{-rep}(x) \in M^j \iff$
$j\text{-rep}(x) \in L^j\ \delta\ M^j \iff$
$c \in \mathrm{sprig}(L\ \boldsymbol{\partial}\ M, x)$.

Thus $\forall c.\ c \in \mathrm{sprig}(L, x)\ \delta\ \mathrm{sprig}(M, x) \iff c \in \mathrm{sprig}(L\ \boldsymbol{\partial}\ M, x)$. So by Extensionality in the base theory (since both sprig and δ are defined in terms of class abstracts, which cannot be urelements), $\mathrm{sprig}(L, x)\ \delta\ \mathrm{sprig}(M, x) = \mathrm{sprig}(L\ \boldsymbol{\partial}\ M, x)$. □

Theorem 13.5 (Sprig Parity Theorem)**.** Assume L and M are fathomable $\mu + 1$-tuples and $L\ \boldsymbol{\partial}\ M$ exists; then $\mathrm{odd}(\mathrm{sprig}(L\ \boldsymbol{\partial}\ M, x)) \iff \mathrm{odd}(\mathrm{sprig}(L, x)) \not\equiv \mathrm{odd}(\mathrm{sprig}(M, x))$.

Thus $\mathrm{sprig}(L\ \boldsymbol{\partial}\ M, x)$ will be odd iff sprig(L, x) is odd and sprig(M, x) is even or *vice versa.*

Proof. L **∂** M exists by hypothesis. So by the Sprig Difference Lemma (13.4), sprig(L **∂** M, x) = sprig(L, x) δ sprig(M, x). Recall the first half of the Parity Property : odd-or-even(a) & odd-or-even(b) ⇒ (odd(aδb) ≡ [odd(a) ≢ odd(b)]). Since L and M are fathomable, sprig(L, x) and sprig(M, x) are both odd or even; thus sprig(L **∂** M, x) = [sprig(L, x) δ sprig(M, x)] is odd iff odd(sprig(L, x)) ≢ odd(sprig(M, x)). □

13.2 Diversity and Degeneracy

Church's construction ([1974a], p. 306, definition of "cardinal m-tuple") does not use any m + 2-tuple of the form (c, ∅, ∅, … ∅, 1), since its new membership would be the same as the old membership of the set c; this would lead to a violation of Extensionality. Mitchell, however, [Mitchell 1976], p. 7, type A_{000} does allow analogous tuples, eliminating instead the old sets; similarly [Forster 1992], page 119, clause 1. That approach would eliminate the following difficulties; the advantage of my approach is a substantial simplification of the proof of the interpretations of the Basic Axioms. My approach is still more complicated on this point than Church's: in his, the (unused) m + 2-tuple (c, ∅, ∅, … ∅, 1) would correspond to the set c; in mine, the urelement coded by the μ + 1-tuple (∅ … ∅ μ-rep"a) would behave analogously to the set a. I will call such tuples *degenerate:* A μ + 1-tuple is degenerate iff its only nonempty component is the last. Two μ + 1-tuples are diverse iff their j-components are unequal for some j less than μ. Formally:

Define **degenerate**(L) iff$_{df}$ μ + 1-tuple(L) & ∀j < μ. L^j = ∅.

Define **diverse**(L, M) $\equiv_{df}$ μ + 1-tuple(L) & μ + 1-tuple(M) & ∃j < μ. $L^j \neq M^j$.

Define **indiverse**(L, M) $\equiv_{df}$ ¬diverse(L, M).

Note that, for μ + 1-tuples L and M without urelements as components, diverse(L, M) ≢ degenerate(L **∂** M), and degenerate(L) ≢ diverse(L, (∅∅ … ∅)).

If I had followed Mitchell's alternative and allowed degenerate μ+1-tuples below, the following result would be unnecessary; I could simply use below the following instance of the Sprig Difference Lemma (13.4): sprig((L^0 L^1 … L^n … [L^μ δ μ-rep"a]), x) = sprig((L^0 L^1 … L^n … L^μ), x) δ sprig((∅ … ∅ μ-rep"a), x).

Lemma 13.6. Assume x is an object, a is a set, L is a μ + 1-tuple, and L^μ δ μ-rep"a exists.
If x ∈ a, then sprig((L^0 L^1 … L^n … [L^μ δ μ-rep"a]), x) =
sprig((L^0 L^1 … L^n … L^μ), x) δ {⟨μ, μ-rep(x)⟩}, and
if x ∉ a, then sprig((L^0 L^1 … L^n … [L^μ δ μ-rep"a]), x) =
sprig((L^0 L^1 … L^n … L^μ), x).

Proof. Note first that by the definition of "**∂**", (L^0 L^1 … L^n … [L^μ δ μ-rep"a]) = (L^0 L^1 … L^n … L^μ) **∂** (∅ … ∅μ-rep"a). So by the Sprig Difference Lemma (13.4), sprig((L^0 L^1 … L^n … [L^μ δ μ-rep"a]), x) = sprig((L^0 L^1 … L^n … L^μ), x) δ sprig((∅ … ∅ μ-rep"a), x). Trivially, sprig((∅ … ∅ μ-rep"a), x) = {⟨μ, μ-rep(x)⟩} iff μ-rep(x) ∈ μ-rep"a iff (since μ-rep is one-one) x ∈ a; it equals ∅ otherwise. Thus if x ∈ a, then sprig((L^0 L^1 … L^n … [L^μ δ μ-rep"a]), x) equals

$\mathrm{sprig}\big((L^0\, L^1\ \dots\ L^n \dots\ L^\mu), x\big)\ \ \delta\ \ \{\langle \mu, \mu\text{-rep}(x)\rangle\}$; otherwise it equals
$\mathrm{sprig}\big((L^0\, L^1\ \dots\ L^n \dots\ L^\mu), x\big)\ \ \delta\ \ \varnothing\ =\ \mathrm{sprig}\big((L^0\, L^1\ \dots\ L^n \dots\ L^\mu), x\big)$. □

14 Indices and Urelements

Define **INDEX**(L) $\equiv_{\mathrm{df}}$

a. $\mu + 1\text{-tuple}(L)\ \&\ \forall j \le \mu.\, \mathrm{set}(L^j)$,
b. $\mathrm{low}(\bigcup_{j\le\mu} L^j)$,
c. $\exists j < \mu\ \exists x.\, x \in L^j$,
d. $\forall j < \mu\ \forall a \in L^j.\, \mathrm{rank}(a) = j\ \&\ j\text{-prolific}(a)$,
e. $\forall a \in L^\mu\ \exists x.\, a = \mu\text{-rep}(x)$,
f. $\forall x.\, \text{odd-or-even}(\mathrm{sprig}(L, x))$.

Note the prohibition in clause (a) of urelements as components L^j. The formalism is neutral on whether urelements are members of these components, but their primary rôle will be through the membership of their μ-reps in L^μ.

Routine verification, below, shows that *INDEX* has three additional properties:

Proposition 14.1 (Degeneracy/Diversity Properties)**.**

g. INDEX(L) & INDEX(M) & diverse(L, M) $\Rightarrow$ INDEX(L **∂** M),
h. INDEX(L) & low(a) $\Rightarrow \mathrm{INDEX}\big((L^0\, L^1\ \dots\ L^n \dots\ [L^\mu\ \ \delta\ \ \mu\text{-rep}``a])\big)$,
i. $\mathrm{INDEX}\big((\{0\text{-rep}(\varnothing)\}\ \varnothing\ \dots\ \varnothing)\big)$.

For greater generality, we could take the above, not as a definition of INDEX, but as a minimum requirement, provided we add the Degeneracy/Diversity Properties as additional conditions on INDEX(L); in this Part the only use I make of the "⇐" part of the definition of INDEX is in the proof of these three properties. This will allow us to replace the condition "j-prolific(a)" in clause (d) by a stronger predicate, when applying the results of this section to the specific consistency proof for CUSι.

Lemma 14.2. INDEX(L) $\Rightarrow\ \forall j \le \mu.\ \mathrm{low}(L^j)$.

Proof. Observe that $L^j \subseteq \bigcup_{j\le\mu} L^j$; by (b) the latter is low, thus so is L^j by $\mathrm{Sep}_{\mathrm{low}}$ (10.1). □

Degeneracy/Diversity Property g. Assume INDEX(L) & INDEX(M) & diverse(L, M); show INDEX(L **∂** M).

Proof. By the preceding, each L^j and M^j is low. Thus by 13.3, L **∂** M is a $\mu + 1$-tuple. Let K = L **∂** M. Showing that K is an index requires verifying that it satisfies clauses a–f of the definition of *INDEX*. We have just proven the first conjunct of clause (a).
Show:

a.2: $\forall j \le \mu.\, \mathrm{set}(K^j).\, K^j = L^j\ \delta\ M^j$; each L^j and M^j is a low set, hence so is $L^j\ \delta\ M^j$ by $\mathrm{Diff}_{\mathrm{low}}$ (10.6).

b: $\mathrm{low}(\bigcup_{j\leq\mu} K^j)$. Note that, since $\bigcup_{j\leq\mu} L^j$ and $\bigcup_{j\leq\mu} M^j$ are low sets, so is $\bigcup_{j\leq\mu} L^j \cup \bigcup_{j\leq\mu} M^j$. But $\forall j \leq \mu.\, K^j = L^j \;\delta\; M^j \subseteq L^j \cup M^j$; thus $\bigcup_{j\leq\mu} K^j$ is a subset of $\bigcup_{j\leq\mu} L^j \cup \bigcup_{j\leq\mu} M^j$, and hence is low.

c: $\exists j < \mu\ \exists x.\, x \in K^j$. L and M are diverse by hypothesis, so $\exists j < \mu.\, L^j \neq M^j$; L^j and M^j are both low sets, hence by Extensionality, $\exists x. (x \in L^j \;\&\; x \notin M^j) \vee (x \notin L^j \;\&\; x \in M^j)$. Thus $x \in L^j \;\delta\; M^j = K^j$.

d: $\forall j < \mu\ \forall a \in K^j.\, \mathrm{rank}(a) = j \;\&\; j\text{-prolific}(a)$. If $j < \mu \;\&\; a \in K^j$, then $a \in L^j \not\equiv a \in M^j$, hence $a \in L^j \vee a \in M^j$. In either case, $\mathrm{rank}(a) = j \;\&\; j\text{-prolific}(a)$.

e: $\forall b \in K^\mu\ \exists x.\, b = \mu\text{-rep}(x)$. If $b \in K^\mu$, then $b \in L^\mu$ or $b \in M^\mu$. In either case, $\exists x.\, b = \mu\text{-rep}(x)$.

f: $\forall x.\, \text{odd-or-even}(\mathrm{sprig}(K, x))$. By the Sprig Difference Lemma (13.4), $\mathrm{sprig}(L, x) \;\delta\; \mathrm{sprig}(M, x) = \mathrm{sprig}(L \;\boldsymbol{\partial}\; M, x)$. By clause (f) of the definitions of "INDEX(L)" and "INDEX(M)," $\mathrm{sprig}(L, x)$ and $\mathrm{sprig}(M, x)$ are each odd or even; thus by 10.11, so is $\mathrm{sprig}(L \;\boldsymbol{\partial}\; M, x)$. □

Degeneracy/Diversity Property h. Assume INDEX(L) & low(a); show $\mathrm{INDEX}(\big(L^0\, L^1 \ldots L^n \ldots [L^\mu \;\delta\; \mu\text{-rep``}a]\big))$.

This would be an instance of g with $M = \big(\varnothing\, \varnothing \ldots \varnothing \ldots \mu\text{-rep``}a\big)$, except that clause (c) specifically excludes degenerate indices.

Proof. First, *Claim:* $L^\mu \;\delta\; \mu\text{-rep``}a$ is a low set. $L^\mu$ is low by 14.2. $\mu\text{-rep``}a$ is a low set by $\mathrm{Rep}_{\mathrm{low}}$ (10.3), since a is low by hypothesis. Therefore $L^\mu \delta\; \mu\text{-rep``}a$ is a low set by $\mathrm{Diff}_{\mathrm{low}}$ (10.6).

Set $M = \big(L^0\, L^1 \ldots L^n \ldots [L^\mu \;\delta\; \mu\text{-rep``}a]\big)$, and verify that M satisfies clauses a–f.

a.1: $\mu + 1\text{-tuple}(M)$. By 13.1.

a.2: $\forall j \leq \mu.\, \mathrm{set}(M^j)$. *Case 1:* $j < \mu$. For $j < \mu$, $M^j = L^j$, and L is an index. Thus $\mathrm{set}(M^j)$ by the corresponding clause of the definition of "INDEX(L)". *Case 2:* $j = \mu$. M^μ is $L^\mu \delta\; \mu\text{-rep``}a$, which is a set by the preceding claim.

b: $\mathrm{low}(\bigcup_{j\leq\mu} M^j)$. $\bigcup_{j\leq\mu} M^j$ is a subset of $\bigcup_{j\leq\mu} L^j \cup \mu\text{-rep``}a$, which is low by $\mathrm{Union}_{\mathrm{low}}$ (10.5), since $\bigcup_{j\leq\mu} L^j$ is low by the corresponding clause of the definition of "INDEX(L)".

c & d: True because $M^j = L^j$ for $j < \mu$.

e: $\forall b \in M^\mu\ \exists x.\, b = \mu\text{-rep}(x)$. Assume $b \in M^\mu = L^\mu \;\delta\; \mu\text{-rep``}a$. Thus $b \in L^\mu \not\equiv b \in \mu\text{-rep``}a$. If $b \in L^\mu$, then (since L is an index) by the corresponding clause of the definition of "INDEX(L)", $\exists x.\, b = \mu\text{-rep}(x)$, as required. If $b \in \mu\text{-rep``}a$, then $\exists x \in a.\, b = \mu\text{-rep}(x)$, by the definition of ``.

f: $\forall x.\, \text{odd-or-even}(\mathrm{sprig}(\big(L^0\, L^1 \ldots L^n \ldots [L^\mu \;\delta\; \mu\text{-rep``}a]\big), x))$.
Since L is an index, $\mathrm{sprig}(L, x)$ is odd or even; any singleton is odd. By 13.6, $\mathrm{sprig}(\big(L^0\, L^1 \ldots L^n \ldots [L^\mu \;\delta\; \mu\text{-rep``}a]\big), x)$ equals either
$\mathrm{sprig}(\big(L^0\, L^1 \ldots L^n \ldots L^\mu\big), x)$ or $\mathrm{sprig}(\big(L^0\, L^1 \ldots L^n \ldots L^\mu\big), x)\,\delta\,\{\langle\mu, \mu\text{-rep}(x)\rangle\}$.
Thus it is odd or even, in the latter case by 10.11. □

Degeneracy/Diversity Property i. $\mathrm{INDEX}(\big(\{0\text{-rep}(\varnothing)\}\; \varnothing \ldots \varnothing\big))$.

Proof. Set $\upsilon = \big(\{0\text{-rep}(\emptyset)\}\ \emptyset\ \ldots\ \emptyset\big)$.

a.1: $\mu + 1$-tuple(υ). By 13.1.

a.2: $\forall j \leq \mu.\, \text{set}(\upsilon^j)$. By the Axiom of Pair and the definition of set.

b: $\text{low}(\bigcup_{j\leq\mu} \upsilon^j)$. $\bigcup_{j\leq\mu} \upsilon^j = \{0\text{-rep}(\emptyset)\}$, which is equinumerous to $\{\emptyset\}$, hence low.

c: $\exists j < \mu\ \exists x.\, x \in \upsilon^j$. Set $j = 0$ and $x = 0\text{-rep}(\emptyset); 0 < \mu\ \&\ 0\text{-rep}(\emptyset) \in \{0\text{-rep}(\emptyset)\} = L^0$.

d: $\forall j < \mu\ \forall a \in \upsilon^j.\, \text{rank}(a) = j\ \&\ j\text{-prolific}(a)$. The only instance is $j = 0\ \&\ a = 0\text{-rep}(\emptyset)$. By the first conjunct of (η), $\text{rank}(0\text{-rep}(\emptyset)) = 0$. Expanding definitions,

$0\text{-prolific}(0\text{-rep}(\emptyset)) \iff$

$\text{rank}(0\text{-rep}(\emptyset)) = 0\ \&\ \neg\exists s\colon \text{low}(s)\ \&\ \forall d.\, d \in s \leftrightarrow \text{daughter}(d, 0\text{-rep}(\emptyset)) \iff$

$\text{rank}(0\text{-rep}(\emptyset)) = 0\ \&\ \neg\exists s\colon \text{low}(s)\ \&\ \forall d.\, d \in s \leftrightarrow [\exists x.\, 0 = \text{rank}(0\text{-rep}(\emptyset))\ \&\ 0\text{-rep}(x) = 0\text{-rep}(\emptyset)\ \&\ 1\text{-rep}(x) = d]$.

Since $\text{rank}(0\text{-rep}(\emptyset)) = 0$ and any two objects have the same 0-rep, this reduces to:

$\neg\exists s\colon \text{low}(s)\ \&\ \forall d.\, d \in s \leftrightarrow [\exists x.\, 1\text{-rep}(x) = d]$, which is the second conjunct of (η).

e: $\forall b \in \upsilon^\mu\ \exists x.\, b = \mu\text{-rep}(x)$. υ^μ is empty.

f: $\forall x.\, \text{odd-or-even}(\text{sprig}(\big(\{0\text{-rep}(\emptyset)\}\ \emptyset\ \ldots\ \emptyset\big), x))$. Since the 0-rep of anything is $0\text{-rep}(\emptyset)$, $\forall x.\, \text{sprig}(\upsilon, x) = \{\langle 0, 0\text{-rep}(\emptyset)\rangle\}$, which has one member and hence is odd by 10.9. □

14.1 Urelements and *

By the Urelement Bijection Axiom, we have a function $\Upsilon(x)$ injecting the sets into the urelements. (We could also use the partially-specified function Υ' from Part I, but for this Part that level of generality is not necessary.) I will define the function "*" as a restriction of Υ, and will abbreviate $^*(x)$ to $^*x$. Define $^*x =_{df} \Upsilon(x)$, provided INDEX(x); undefined otherwise. By this definition and the Urelement Bijection Axiom, we have the following:

Lemma 14.3 (* Properties Lemma)**.**

a. $\forall x.\ \text{INDEX}(x) \rightarrow \exists!\, u.\, u = {}^*x$,

b. $\forall x, u.\, u = {}^*x \Rightarrow \text{urelement}(u)\ \&\ \text{INDEX}(x)$,

c. $\forall x, y, u.\, u = {}^*x\ \&\ u = {}^*y \Rightarrow x = y$.

In prose, read "INDEX(L)" as "L is **an index**." Let the function *index* be the inverse of the function *. I.e., define **index**(u) $=_{df} \iota x.\, u = {}^*x$. Thus $\forall x.\, \text{INDEX}(x) \rightarrow \text{index}({}^*x) = x$. By (a), $\exists!\, u.\, u = {}^*x$, which establishes the existence part of the definition of index($^*x$); (c) establishes uniqueness. In prose call x **the index of** $^*x$, and call $^*x$, **x's urelement**.

15 $\in_2$

Define $x \in_2 y$ iff$_{df}$

(a) $\exists L.\, y = {}^*L$ & INDEX(L) & odd(sprig(L, x)) $\vee$
(b) $x \in_0 y$.

Note that since $y = {}^*L \Rightarrow$ urelement$_0$(y), the two clauses are mutually exclusive, and set$_0$(y) $\Rightarrow x \in_2 y \equiv x \in_0 y$. Redefine (analogously to the definition in Part I) **unaltered**(x) iff$_{df}$ $\forall z.\, z \in_0 x \equiv z \in_2 x$; **altered**(x) iff$_{df}$ $\neg$unaltered(x). Thus

Lemma 15.1. set$_0$(y) $\rightarrow$ unaltered(y), and altered(y) $\rightarrow \exists L.\, y = {}^*L$.

The two cases in the definition of $\in_2$ correspond to the six cases of Church's definition [1974a], page 306. Considerable simplification is achieved by the use of urelements (in place of Church's *i-analogue* function) and the definition of sprig, though at the cost of the Urelement Bijection Axiom and the non-primitive notations "*", "INDEX", "sprig", and "odd" in the definition.

Note that Church's use of Compactness ([1974a], p. 307) is here unnecessary, since this construction uses the full sequence of partially-defined restricted equivalence relations, rather than Church's initial segment of $\sim^j$'s, for $j \leq m$, with unspecified length m.

Observe that the definition of $\in_2$ immediately gives us many $\sim^j$-equivalence classes as sets$_2$ with $\sim^j_0$ in, regrettably, the sense of the old membership relation:

Observation 15.2 (Equivalence Class Observation)**.** Let a be an object, with $j < \mu$. If L is an index with $L^j = \{\text{j-rep}(a)\}$, and $L^k = \varnothing$ for $j \neq k$, then $\forall x.\, x \in_2 {}^*L \equiv x \sim^j_0 a$.

Proof. Since ${}^*L$ is an urelement$_0$, clause (b) of the definition of $\in_2$ is false. The first two conjuncts of clause (a) are true by hypothesis, so $x \in_2 {}^*L \equiv$ odd(sprig($(L^0\ L^1\ \ldots\ L^n \ldots\ L^\mu)$, x)). But $L^k = \varnothing$ for $k \neq j$, and $L^j = \{\text{j-rep}(a)\}$. Thus sprig(L, x) $= \{\langle j, \text{j-rep}(x)\rangle\}$ if j-rep(x) = j-rep(a), and $\varnothing$ otherwise. A singleton is odd and the empty set is not; thus $x \in_2 {}^*L \equiv$ j-rep(x) = j-rep(a). By $\sim^j$ Requirements (ϵ), $x \sim^j a$ iff j-rep(x) = j-rep(a), so $x \in_2 {}^*L \equiv x \sim^j a$. □

Note that what we actually want is this result with "$\sim^j$" replaced by its interpretation. Say that $\sim^j$ is **absolute** iff$_{dfs}$ $\forall x, y.\, x \sim^j_0 y \equiv x \sim^j_2 y$. Consideration of this requirement leads naturally to Oberschelp's comprehension schema; see [Sheridan 1990]. Trivially, though, the Equivalence Class Observation gives us:

Corollary 15.3. For any $j < \mu$ and any a, if $\sim^j$a is absolute and $(\varnothing \ldots \{\text{j-rep}(a)\} \ldots \varnothing)$ is an index, then $\forall x.\, x \in_2 {}^*L \equiv x \sim^j_2 a$.

I.e., ${}^*L$ is a's Frege j-cardinal in the sense of the new membership relation.

Note that this result does not require the Lowness$_0$ Assumption.

Lemma 15.4 (Universal Set Lemma)**.** $\forall y.\, y \in_2 {}^*(\{\text{0-rep}(\varnothing)\}\ \varnothing\ \ldots\ \varnothing)$.

Proof. Let $\upsilon = (\{\text{0-rep}(\varnothing)\}\ \varnothing\ \ldots\ \varnothing)$. Since (by Degeneracy/Diversity Property i) υ is an index, $\upsilon^0 = \{\text{0-rep}(\varnothing)\}$ and $\upsilon^j = \varnothing$ otherwise, then $\forall x.\, x \in_2 {}^*\upsilon \equiv x \sim^0 \varnothing$. But $\forall x, y.\, x \sim^0 y$, so $\forall x.\, x \in_2 {}^*\upsilon$. □

16 ∂ Theorem and Symmetric Difference$_2$

Theorem 16.1 (∂ Theorem)**.** If a, b, and a **∂** b are indexes, then
$\forall x. [(x \in_2 {}^*a) \not\equiv (x \in_2 {}^*b)] \iff x \in_2 {}^*(a\ \boldsymbol{\partial}\ b)$.

In other terms, ${}^*a\ \Delta\ {}^*b \simeq_2 {}^*(a\ \boldsymbol{\partial}\ b)$.

Proof. Since a, b, and a **∂** b are indexes, then a and b are fathomable μ+1-tuples, and ${}^*a$, ${}^*b$, and ${}^*(a\ \boldsymbol{\partial}\ b)$ exist. Thus substituting the first clause of the definition of "$\in_2$" three times (noting that clause (b) is false and the first two conjuncts of a are true), the theorem is equivalent to $[\mathrm{odd}(\mathrm{sprig}(a, x)) \not\equiv \mathrm{odd}(\mathrm{sprig}(b, x))] \iff \mathrm{odd}(\mathrm{sprig}([a\ \boldsymbol{\partial}\ b], x))$, which follows from the Sprig Parity Theorem (13.5). □

Lemma 16.2 (Degenerate Lemma)**.** Assume a is a low set and $(L^0\, L^1 \ldots L^\mu)$ is an index. Then $\forall x. x \in_2 {}^*(L^0\, L^1 \ldots [L^\mu\ \delta\ \mu\text{-rep``}a]) \iff . x \in_2 {}^*(L^0\, L^1 \ldots L^\mu) \not\equiv x \in_2 a$.

I.e., ${}^*L\ \ \Delta\ \ a\ \simeq_2\ {}^*(L^0\, L^1 \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a])$.

(If I had not excluded degenerate μ + 1-tuples from being indexes, this result would be simply ${}^*L\ \ \Delta\ \ a\ \simeq_2\ {}^*(L^0\, L^1 \ldots L^\mu)\ \boldsymbol{\partial}\ {}^*(\varnothing \ldots \varnothing\ \ \mu\text{-rep``}a)$.)

Proof. Since a is a set, it is unaltered, and hence $x \in_2 a \equiv x \in_0 a$. By Degeneracy/Diversity Property h, $(L^0\, L^1 \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a])$ is an index.

Case 1: $x \notin_0 a$. Then by the second case of 13.6, $\mathrm{sprig}((L^0\, L^1 \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a]), x) = \mathrm{sprig}((L^0\, L^1 \ldots L^\mu), x)$, so by the definition of $\in_2$, $x \in_2 {}^*(L^0\, L^1 \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a]) \iff . x \in_2 {}^*(L^0\, L^1 \ldots L^\mu)$. Since $x \notin_2 a$, the right side is equivalent to $x \in_2 {}^*(L^0\, L^1 \ldots L^\mu) \not\equiv x \in_2 a$, as required.

Case 2: $x \in_0 a$, so $x \in_2 a$; hence the required result is $x \in_2 {}^*(L^0\, L^1 \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a]) \not\equiv x \in_2 {}^*(L^0\, L^1 \ldots L^\mu)$. By the first case of 13.6, $\mathrm{sprig}((L^0\, L^1 \ldots [L^\mu\ \delta\ \mu\text{-rep``}a]), x) = \mathrm{sprig}((L^0\, L^1 \ldots L^\mu), x)\ \delta\ \{\langle \mu, \mu\text{-rep}(x)\rangle\}$. Since $\{\langle \mu, \mu\text{-rep}(x)\rangle\}$ is odd, by the Parity Property $\mathrm{sprig}((L^0\, L^1 \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a]), x)$ is odd iff $\mathrm{sprig}((L^0\, L^1 \ldots L^\mu), x)$ is not. This reduces to the required result by two applications of the definition of $\in_2$. □

Lemma 16.3 (Indiverse Lemma)**.** Assume $(M^0\, M^1 \ldots M^\mu)$ and $(L^0\, L^1 \ldots L^\mu)$ are indiverse indices. Then $\forall x. x \in_2 \mu\text{-rep}^{\leftarrow}\text{``}[L^\mu \delta M^\mu] \iff . x \in_2 {}^*(L^0\, L^1 \ldots L^\mu) \not\equiv x \in_2 {}^*(M^0\, M^1 \ldots M^\mu)$.

I.e., ${}^*L\ \Delta\ {}^*M\ \simeq_2\ \mu\text{-rep}^{\leftarrow}\text{``}[L^\mu\ \delta\ M^\mu]$.

(Thus, $\mu\text{-rep}^{\leftarrow}\text{``}[L^\mu\ \delta\ M^\mu]$ behaves as would the urelement corresponding to $(L\ \boldsymbol{\partial}\ M) = (\varnothing \ldots \varnothing\ [L^\mu\ \ \delta\ \ M^\mu])$, were degenerate μ + 1-tuples allowed in the domain of *.)

Proof. L and M are indexes, so $L^\mu\ \delta\ M^\mu$ is a low set. Thus by $\mathrm{Rep}_{\mathrm{low}}$ (10.3), $\mu\text{-rep}^{\leftarrow}\text{``}[L^\mu\ \delta\ M^\mu]$ is a $\mathrm{set}_0$, hence unaltered; thus by 11.2, $\mu\text{-rep}(x) \in_0 [L^\mu\ \delta\ M^\mu] \iff x \in_0 \mu\text{-rep}^{\leftarrow}\text{``}[L^\mu\ \delta\ M^\mu] \iff x \in_2 \mu\text{-rep}^{\leftarrow}\text{``}[L^\mu\ \delta\ M^\mu]$.

By the definition of $\in_2$, the Parity Property, and the Sprig Difference Lemma (13.4), $x \in_2 {}^*\big(L^0\ L^1\ \ldots L^\mu\big) \not\equiv x \in_2 {}^*\big(M^0\ M^1\ \ldots M^\mu\big) \iff$ odd(sprig(L, x) δ sprig(M, x)) $\iff$ odd(sprig(L $\boldsymbol{\partial}$ M, x)). Since L and M are indiverse, for $j < \mu$, $L^j = M^j$, so $L \,\boldsymbol{\partial}\, M = \big(\varnothing \ldots \varnothing\, [L^\mu\ \delta\ M^\mu]\big)$. Thus sprig(L $\boldsymbol{\partial}$ M, x) is $\{\langle\mu, \mu\text{-rep(x)}\rangle\}$ if μ-rep(x) $\in L^\mu\ \delta\ M^\mu$, and $\varnothing$ otherwise. Since a singleton is odd and the empty set is not, $x \in_2 {}^*\big(L^0\ L^1\ \ldots L^\mu\big) \not\equiv x \in_2 {}^*\big(M^0\ M^1\ \ldots M^\mu\big) \iff \mu$-rep(x) $\in_0 L^\mu\,\delta\, M^\mu \iff$ (by the above) $x \in_2 \mu\text{-rep}^{\leftarrow}$“$[L^\mu\ \delta\ M^\mu]$. □

Theorem 16.4 (Symmetric Difference$_2$ Theorem (s))**.** $\forall a \forall b \exists z \forall w.\, w \in_2 z \iff (w \in_2 a \not\equiv w \in_2 b)$.

I.e., for any a and b, $a \,\Delta\, b$ exists; the symmetric difference$_2$ of any two objects is a set$_2$. The Lowness$_0$ Assumption is required to show that the following cases are exhaustive.

Proof.
Case 1: a and b are low$_0$ and unaltered. By the former assumption and Diff$_{low}$ (10.6), a δ b is a set$_0$, hence unaltered. So $\forall w.\, w \in_0 a\ \delta\ b \iff (w \in_0 a \not\equiv w \in_0 b)$, and hence $\forall w.\, w \in_2 a\ \delta\ b \iff (w \in_2 a \not\equiv w \in_2 b)$.

Case 2: a and b are altered; then $\exists L, M.\ a = {}^*L\ \&\ b = {}^*M$. *Subcase 2.a:* Diverse(L, M). Then by Degeneracy/Diversity Property g, L $\boldsymbol{\partial}$ M is an index, so by the $\boldsymbol{\partial}$-Theorem (16.1), $a\ \Delta\ b \simeq_2 {}^*(L\ \boldsymbol{\partial}\ M)$. *Subcase 2.b:* Indiverse(L, M). Then by 16.3, ${}^*L\ \Delta\ {}^*M \simeq_2 \mu\text{-rep}^{\leftarrow}$“$[L^\mu\ \delta\ M^\mu]$.

Case 3: a is low$_0$ and unaltered, and b is altered; then $\exists L.\, b = {}^*L$. By 16.2, $a\ \Delta\ {}^*L$ will be ${}^*\big(L^0\, L^1\ \ldots [L^\mu\ \delta\ \mu\text{-rep}$“$a]\big)$.

Case 4: b is low$_0$ and unaltered, and a is altered. The proof is the same as for the preceding case, via renaming of variables. □

Since the symmetric difference of an object and the universal set is the object's complement, this gives us, under the Lowness$_0$ Assumption a complement for every object. I.e.,

Theorem 16.5 (Complements Theorem (s))**.** $\forall a \exists z \forall w.\, w \in_2 z \equiv\ w \notin_2 a$.

Proof. By the Universal Set Lemma (15.4), $\forall y.\, y \in_2 {}^*\big(\{0\text{-rep}(\varnothing)\}\ \varnothing\ \ldots\ \varnothing\big)$. By 16.4s, $\forall a \exists z \forall w.\, w \in_2 z \iff (w \in_2 a \not\equiv w \in_2 {}^*\big(\{0\text{-rep}(\varnothing)\}\ \varnothing\ \ldots\ \varnothing\big))$. But (P $\not\equiv$ The True) $\iff \neg$P, so $w \in_2 z \equiv\ \neg w \in_2 a$, as required. □

Theorem 16.6 ($\varnothing$-Theorem (s))**.** $\forall a, b.\,$ nonempty$_2$(a) & $a \neq b$ & $a \simeq_2 b \Rightarrow \exists c.$ INDEX(c) & empty$_2$(*c).

Proof.
Case 1: a and b are low$_0$ and unaltered. Thus since $\forall z.\, z \in_2 a \equiv\ z \in_2 b$, so $\forall z.\, z \in_0 a \equiv\ z \in_0 b$, and also nonempty$_0$(a). But this violates the Axiom of Extensionality in the base theory, contradiction.

Case 2: a and b are altered; then $\exists L, M.\, a = {}^*L\ \&\ b = {}^*M$. *Subcase 2.a:* Indiverse(L, M). Then by 16.3, $\varnothing \simeq_2 {}^*L\,\Delta\,{}^*M \simeq_2 \mu\text{-rep}^{\leftarrow}$“$[L^\mu\,\delta\,M^\mu]$. Since every member of L^μ or M^μ is a μ-rep, this implies that $L^\mu\ \delta\ M^\mu$ is empty$_0$. Since L^μ

and M^μ are $sets_0$, Extensionality in the base theory implies that $L^\mu = M^\mu$. Since L and M are indiverse, $L^j = M^j$ for $j < \mu$, and hence L = M, contradicting the assumption.

Subcase 2.b: Diverse(L, M). Then by Degeneracy/Diversity Property g, $L \,\boldsymbol{\partial}\, M$ is an index, so by the $\boldsymbol{\partial}$-Theorem (16.1), $\varnothing \simeq_2 a \,\Delta\, b \simeq_2 {}^*(L \,\boldsymbol{\partial}\, M)$, as required.

Case 3: a is low_0 and unaltered, and b is altered (or *vice versa,* via renaming of variables); then $\exists L.\,\text{INDEX}(L)\ \&\ b = {}^*L$. By Degeneracy/Diversity Property h, $\big(L^0\ L^1\ \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a]\big)$ is an index. By 16.2, $\varnothing \simeq_2 a \,\Delta\, {}^*L \simeq_2 {}^*\big(L^0\, L^1\ \ldots [L^\mu\ \ \delta\ \ \mu\text{-rep``}a]\big)$, as required. □

Thus we can reduce any violation of Extensionality in the interpretation to a counterexample involving the $empty_2$ set_2. This theorem was the goal of all but the end of this Part, and could have been done in much greater generality; so far the L^j's have served only to code collections of predicates, and other coding schemes would have served as well. The final section shows that no such violation of Extensionality is possible, and makes heavy use of the ordering of predicates imposed by the use of ordinals less than or equal to μ, and of the j-prolificity requirement in clause (d.2) of the definition of INDEX.

17 Coercion

For an index L, define **j-coercion**(L) $=_{df}$ $L^j \,\cup\,$ j-ancestor“$(\bigcup_{j+1\leq k<\mu} L^k)$ ∪ j-rep“μ-rep$^{\leftarrow}$“L^μ. (For $k < \mu$, every member of L^k has rank k, so j-ancestor is defined on L^k for k greater than j and less than μ.) Informally, j-coercion(L) = $L^j \cup$ j-ancestor“$L^{j+1} \ldots \cup$ j-ancestor“$L^n \ldots \cup$ j-ancestor“$L^{\mu-1} \cup$ j-rep“μ-rep$^{\leftarrow}$“L^μ. The point to this definition is to collect the j-reps of members of $L^j \ldots L^\mu$, which might be relevant to $membership_2$ in L, for use in the Ancestorless Lemma (17.3) and $Nonemptiness_2$ Lemma (17.4), below.

Lemma 17.1. If L is an index, then j-coercion(L) is a low_0 set_0.

Proof. $\bigcup_{j\leq\mu} L^j$ is low_0 by requirement (b) on INDEX; thus by Sep_{low} (10.1) so are its subsets L^j, $\bigcup_{j+1\leq k<\mu} L^k$, and L^μ. Thus by Rep_{low} (10.3) so are j-ancestor“$(\bigcup_{j+1\leq k<\mu} L^k)$ and j-rep“μ-rep$^{\leftarrow}$“L^μ; and hence also the j-coercion of L, by two applications of $Union_{low}$ (10.5). □

Lemma 17.2. If L is an index, then every member of j-coercion(L) is a j-rep.

Proof. By clause (d) of the definition of INDEX, every $member_0$ of L^j is a j-rep. By the definition of j-ancestor, so is every member of j-ancestor“$(\bigcup_{j+1\leq k<\mu} L^k)$. By the definition of “, so is every member of j-rep“μ-rep$^{\leftarrow}$“L^μ. □

Lemma 17.3 (Ancestorless Lemma)**.** Assume L is an index, $j < \mu$, $g \in_0 L^j$, but $\forall k < j.\, k\text{-ancestor}(g) \notin L^k$. Then
$\forall x.\, j\text{-rep}(x) = g \;\Rightarrow .\, x \in_2 {}^*L \vee j{+}1\text{-rep}(x) \in_0 j{+}1\text{-coercion}(L)$.

Informally, if g is a j-rep with no ancestors in L, then anything whose j-rep is g will be a member_2 of L unless its j+1-rep is in the j+1-coercion of L. (This was the purpose for the definition of j-coercion.)

Proof. For any index L, such j and g exist: for example, the first $j < \mu$ such that L^j is not empty_0, and any member_0 of such L^j.

Assume $g = j\text{-rep}(x)$ & $x \notin_2 {}^*L$. Show $j{+}1\text{-rep}(x) \in_0 j{+}1\text{-coercion}(L)$.

Consider sprig(L, x). It contains_0 $\langle j, g\rangle$, since $j\text{-rep}(x) = g \in_0 L^j$. It has an even number of members_0, since $x \notin_2 {}^*L$. Thus it must also contain_0 some $\langle k, h\rangle$, for $k \neq j$ & $k \leq \mu$, with $h = k\text{-rep}(x)$, $h \in_0 L^k$, and $k = \text{rank}(h)$. If k were less than j, then h would be the k-ancestor of g, which would contradict the hypothesis. Thus, since $k \leq \mu$, then $k > j$.

Claim: $k = j + 1 \vee j + 1 < k < \mu \vee k = \mu$. Since $j < k$, $k \not< j + 1$ by the Successor Lemma (10.12). By Required Properties of + (iv), $j + 1 \leq \mu$. If $k = \mu$, we are through, so (since $k \leq \mu$) assume $k < \mu$. If $j + 1 = \mu$, then $k < \mu = j + 1$, contradicting 10.12, since $k > j$. Thus $j + 1 < \mu$. Since μ is an ordinal, and also $k < \mu$, $j + 1 = k \vee j + 1 < k \vee k < j + 1$. The last disjunct is false, so $j + 1 = k \vee j + 1 < k$, as required.

Thus the following cases are exhaustive.

Case $k = j + 1$**:** Then $j{+}1\text{-rep}(x) = h \in_0 L^{j+1} \subseteq j{+}1\text{-coercion}(L)$.

Case $j + 1 < k < \mu$**:** Thus, since $k = \text{rank}(h)$, $j{+}1\text{-ancestor}(h) = j{+}1\text{-ancestor}(k\text{-rep}(x)) = j{+}1\text{-rep}(x)$, by (12.3). $h \in_0 L^k$, so $j{+}1\text{-rep}(x) = j{+}1\text{-ancestor}(h) \in_0 j{+}1\text{-ancestor``}L^k \subseteq j{+}1\text{-coercion}(L)$.

Case $k = \mu$**:** Then $\mu\text{-rep}(x) = h \in_0 L^\mu$, so $x \in_0 \mu\text{-rep}^{\leftarrow}\text{``}L^\mu$ and $j{+}1\text{-rep}(x) \in_0 j{+}1\text{-rep``}\mu\text{-rep}^{\leftarrow}\text{``}L^\mu \subseteq j{+}1\text{-coercion}(L)$. □

Lemma 17.4 (Nonemptiness_2 Lemma)**.** $\text{INDEX}(L) \Rightarrow \exists x.\, x \in_2 {}^*L$.

Proof. Assume not; so let L be an index with ${}^*L$ empty_2, with j the first $j < \mu$ such that L^j is not empty_0, and let g be a member_0 of L^j. Then by the preceding (since ${}^*L$ is empty_2) $\forall x.\, j\text{-rep}(x) = g \Rightarrow j{+}1\text{-rep}(x) \in_0 j{+}1\text{-coercion}(L)$.

Claim: This implies j+1-coercion(L) is a superset of the class of daughters of g. Let h be an arbitrary daughter of g; so $\exists y.\, j\text{-rep}(y) = g$ & $j{+}1\text{-rep}(y) = h$. Thus by the preceding $h \in_0 j{+}1\text{-coercion}(L)$.

But j-coercion(L) is a low_0 set_0; g is j-prolific by clause (d) of the definition of INDEX, contradiction. □

Theorem 17.5 (Interpretation of the Axiom of Extensionality for Sets (s))**.**
$\forall a \forall b.\, \text{nonempty}_2(a)$ & $\forall z.\, z \in_2 a \equiv z \in_2 b. \Rightarrow a = b$.

Proof. Assume not; then by the ∅-Theorem (s) (16.6), $\exists$c. INDEX(c) & empty_2(* c); but this contradicts the Nonemptiness_2 Lemma (17.4). □

Part III

j-Isomorphism, Foundation, Choice, the Interpretation, and Proof of the Axioms of CUSι

18 j-Isomorphism

In Part III, I define a specific sequence of restricted equivalence relations, $\eqsim^{j}$ (read "j-isomorphic"), and prove its two key properties: that the singleton function is the union of a small finite number (six in general, one in the current context) of 2-isomorphism classes (20.14), and that any non-degenerate j-isomorphism class is non-low (19.16).

After defining j-isomorphism, rather than proving the properties of a partially-specified membership relation (such as $\in_1$ or $\in_2$ in Parts I and II), I will instead define a specific relation $\in_3$; and I will assume for the base theory, in addition to the Basic Axioms, the Axioms of Foundation and Global Well-Ordering. Some of the uses of these axioms might be eliminable with sufficient care to relativization and the use of Scott's Trick [1955], but substantial use of Foundation seems necessary for the Replacing at Level*j construction, section 19.4 below.

18.1 Ordinals and Avoidance of Advanced Recursion

For the following subsection I will continue to avoid development of recursion on the finite ordinals beyond that used above. This may facilitate use of these techniques in other contexts, though whether this justifies the additional effort is by no means clear.

For $1 \leq j < \omega$, define $y \in^{\mathbf{j}} a \equiv_{df} \exists f.\ \exists c.\ \text{maps}(f, j{+}1, c)\ \ \&\ \ f`0 = a\ \ \&\ \ f`j = y\ \ \&\ \ \forall k{\in}j.\ f`k{+}1 \in f`k$. Read "y is a member at **level j** of a"

For convenience, define $y \in^{\mathbf{0}} a \equiv_{df} y = a$; this differs from Church's usage, but is convenient for usage with the j^{th} cumulative union, defined below. Repeated application of the Axiom of Pairs trivially shows that $y \in^1 a \equiv y \in a$.

Define $y \in^{<\mathbf{j}} a \equiv_{df} \exists k.\ 0 \leq k < j.\ y \in^k a$. (Note that this means $y \in^{<j} y$, for $j \geq 1$.)

Define $y \in^{\leq j} a \equiv_{df} \exists k.\ 0 \leq k \leq j.\ y \in^k a$.

Define $y \in^{*\mathbf{j}} a \equiv_{df} y \in^j a\ \ \&\ \ \neg\exists\, i < j.\ y \in^i a$. Read "y is a member at level*j *proper* of a"; **level*j** of a is the class of all members at level*j of a. Thus level*0 of a is {a}.

Define $\boldsymbol{\Xi}^{\mathbf{j}} a =_{df} \{y \mid y \in^{\leq j} a\}$, for $0 \leq j < \omega$. Read "the j^{th} *cumulative union* of a." This is a class abstract; it will have to be proved to be a set before making use of it. (Recall that a class abstract is never an urelement, so it is only necessary to eliminate the possibility that it is an ultimate class.) Note that $\Xi^0 a = \{a\}$, and that $\Xi^j a$ contains a for any j.

Define **TC**(a) $=_{df}$ {y | ∃ j ∈ ω. y $\in^j$ a }. Note that, because of my definition of $\in^0$, this differs slightly from the standard **transitive closure** of a, in that TC(a) also includes a. Informally, call a member of the transitive closure of x, a **constituent** of x.

18.2 Definition of j-Isomorphism

Define, for j ≤ 1 < ω, a $\bumpeq^j$ b $\equiv_{df}$ ∃F :

(1) SET(Ξ^ja) & SET(Ξ^jb) &
(2) $maps_{1\text{-}1}$(F, Ξ^ja, Ξ^jb) &
(3) F‘a = b &
(4) ∀y $\in^{<j}$ a. F‘y = F“y

Read "a is **j-isomorphic** to b." The first clause will be superfluous (by the Cumulative Union Lemma (19.1), below) in the presence of Foundation; it is only needed for ill-founded objects in the interpretation. When F is known, I will also write "F: a $\bumpeq^j$ b." For convenience, define $\bumpeq^0$ as the universal relation which holds between any two objects, and $\bumpeq^\omega$ as equality.

18.2.1 j-Isomorphism Notes

(1) Note that since $\in^0$ is equality, clauses (3) and (4) imply that b = F“a, for j≥1.
(2) By my definition of “, which maps empty objects to themselves, an empty object can be j-isomorphic only to itself, for j≥1.
(3) Conversely, for any empty a, the mapping {<a, a>} is a j-isomorphism for any j ≤ ω. So any empty object is j-isomorphic to itself, and, for j≥1, only to itself.
(4) Since b = F“a, if F|a (i.e., F restricted to a) is a set (which it will be in the presence of Foundation), then a is equinumerous to b.
(5) The intent is that the $\bumpeq^j$ are of increasing strictness, but proving this will require Foundation, which I assume below.
(6) Note that, despite my informal terminology, I have not yet proved that these are equivalence relations, nor even that they are reflexive. They won't necessarily be either for ill-founded sets, since the obvious proofs require Replacement.
(7) The state of j-isomorphism in the interpretation will be inelegant, especially the existence of set mappings witnessing j-isomorphism, rendering them merely restricted equivalence relations. (A similar difficulty arises with Church's theory, though he did not need to address it in his surviving writings.) It will be simple to show that j-isomorphism is an equivalence relation in the presence of Foundation, and j-isomorphism will be absolute for sets which are unaltered down to level j in the interpretation (see the j-Pure j-Isomorphism Absoluteness Theorem (20.10), below). This is somewhat short of showing that j-isomorphism will be a restricted equivalence relation in the interpretation, since some new set might be a mapping which witnesses a j-isomorphism for a new set, with no obvious guarantee of the existence of other mappings required for an equivalence relation.

(8) There will be further shortcomings of these equivalence relations in the interpretation. They will only provably be absolute for well-founded sets and those j-isomorphic to them, the j-pure sets (defined formally below); contemplating the requirements for absoluteness of such relations leads naturally to Oberschelp's existence criterion [Oberschelp 1973], [Sheridan 1990], which may loosely be described as mandating the set-hood of any predicate whose definition is absolute. The situation is inelegant even for $j = 1$. The intent is for two sets to be 1-isomorphic if they are equinumerous, and either both or neither are self-membered. No two urelements are 1-isomorphic in the base theory, but in the interpretation an old urelement might contain itself and be externally equinumerous to the universe (e.g., the urelement which represents the universal set itself), or not contain itself and be externally equinumerous to the universe (e.g., the set of all non-self-membered singletons). See also the discussion of the Bad Company problem in the philosophical introduction for further difficulties with j-isomorphism.

The crucial difference between my j-isomorphism and Church's j-equivalence (abbr: $\approx_j$) is that, in my definition, while the first two clauses deal with membership at level $\leq$j, the last deals with membership at level $<$ j, in order to enable the set-hood of the singleton function. A lesser differences is that I have a single mapping required to be one-one across all levels, while his sequence of mappings are only required to be individually one-one.

j-isomorphism classes do not seem to be closed under sum set, which is why my theory (unlike Church's) does not have an unrestricted axiom of sum set. The 2-isomorphism class of { {Ø, {Ø}}, {Ø} } will be a set in my theory, but its sum set does not seem to be. This union should be the set of all singletons plus the set of all pairs of the form {a, {a} }, but the latter does not seem to be a j-isomorphism class, nor a manageable combination thereof.

Church's equivalence relations have the property that if a class is roughly (i.e., modulo a well-founded set) closed under j-equivalence, its sum set is roughly closed under j-1-equivalence [Sheridan 1989], p. 75, 84; this would have been crucial in Church's consistency proof.

19 Foundation, Choice, j-isomorphism, and Less Generality

19.1 Foundation and Global Well Ordering

For the remainder of this work, I will drastically reduce the generality in which I have been working. I will work in a base theory which includes, in addition to RZFU, the Axiom of Foundation (sometimes merely three of its consequences—see below) and the Axiom Schema of Global Well-Ordering. This renders the Basic Axioms, some of them restricted to well-founded sets, equivalent to their standard counterparts in, for example, [Levy 1979]. (Since we are now assuming Choice, the usual proof will also go through that ω—here defined as the set of all Dedekind finite ordinals—is itself an ordinal.) This will allow the use of the standard results of ZFC, e.g., definition by recursion and Separation, and hence requires far less formality. It will also render unproblematic use of the (Cantor) cardinality of any set, with the standard definition as the least ordinal equinumerous to the given set.

As a further specialization, for the arbitrary ordinal μ, I will substitute ω; for the arbitrary sequence of relations $\sim^j$ (j≤μ) I substitute $\backsimeq^j$ (j≤ω), with $\backsimeq^\omega$ being equality and $\backsimeq^0$ being the universal relation. I will also substitute for the partially-specified relations $\in_1$ and $\in_2$, and predicate INDEX, a specific relation $\in_3$ and predicate INDEX3, defined below. For the partially-specified function "+" on the odd-or-even ordinals, I substitute the usual addition function on the finite ordinals. For convenience, I will reuse subsidiary terminology (e.g., j-rep, rank, *, and sprig) without explicitly distinguishing it, though the notations should henceforth be understood as defined in terms of $\backsimeq^j$ rather than $\sim^j$.

Some of the uses below of Foundation in the base theory are essential; the most extreme case is the Replacing at Level*j Construction, which is defined by recursion on the Cumulative Hierarchy. Some of the uses, however, are needed merely for three unrestricted consequences of the normal ZF axioms: Separation, pairwise union, and unrestricted sum set. Where appropriate, I will mark results which require only these consequences of Foundation.

The uses of unrestricted Separation are largely of one type, that a subclass of a set function (or of its domain or range) is also a set; I will call this the Function Subset Assumption: In set theories like Church's, this seems little, if any, weaker than full Separation, which needs to be restricted to well-founded sets. (Consider the identity function, which could plausibly be a set, and its subclass, the identity function restricted to non-self-membered sets, which is likely to lead to a paradox.) But the assumption recurs frequently enough in what follows that it seems worth calling attention to, for possible use of this construction in other theories; e.g., [Aczel 1985], which has finite self-membered sets but unrestricted Replacement.

Lemma 19.1 (Cumulative Union Lemma)**.** For $0 \le j < \omega$, ∀a. SET(Ξ^ja).

Proof. For j = 0, Ξ^ja = {a}, which exists by the Axiom of Pairs and is not an urelement. Given SET(Ξ^ja) for j ≥1, Ξ^{j+1}a is Ξ^ja ∪ ($\bigcup\Xi^j$a), by routine manipulation of finite sequences. This is also a set by two uses of Well-Founded Sum Set. The result follows by definition by recursion, which is unproblematic in the presence of Foundation. □

Similarly we have:

Lemma 19.2 (Transitive Closure Lemma)**.** ∀a. SET(TC(a)).

Lemma 19.3 (j-Isomorphism/Level j Lemma)**.** If F: a $\backsimeq^j$ b and i ≤ j, then ∀x. x $\in^i$ a ≡ F'x $\in^i$ b.

Proof.
Note that the result is trivial for j=0, so assume j≥1.

Part 1: Assume x $\in^{i}$ a; show F‘x $\in^{i}$ b. Since x $\in^{i}$ a, there is a function f such that f‘0 = a & f‘i = x & ∀k∈i. f‘k+1 ∈ f‘k.

Claim: F composed with f is a membership sequence of length i from b to F‘x. Clearly F‘f‘0 = F‘a = b, and F‘f‘i = F‘x. For arbitrary k in i, f‘k+1 ∈ f‘k, and (since f‘k $\in^{<i}$ f‘0 = a and F is a j-isomorphism) F‘f‘k = F“f‘k, which contains F‘f‘k+1, since f‘k+1 ∈ f‘k. Thus F‘f‘k+1 ∈ F‘f‘k.

Part 2: Assume F‘x $\in^{i}$ b and show x $\in^{i}$ a. Thus there is a function f such that f‘0 = b & f‘i = F‘x & ∀k∈i. f‘k+1 ∈ f‘k.

Claim: $F^{\leftarrow}$ composed with f is a membership sequence of length i from a to x. Clearly $F^{\leftarrow}$‘f‘0 = $F^{\leftarrow}$‘b = a, and $F^{\leftarrow}$‘f‘i = $F^{\leftarrow}$F‘x = x.

Subclaim: For arbitrary k in i, $F^{\leftarrow}$f‘k+1 ∈ $F^{\leftarrow}$f‘k and $F^{\leftarrow}$f‘k+1 $\in^{k+1}$ a, by induction. For k=0, this is $F^{\leftarrow}$f‘1 ∈ $F^{\leftarrow}$f‘0 = a and $F^{\leftarrow}$f‘1 $\in^{1}$ a; but f‘1 ∈ f‘0 = b, and b = F“a, so $F^{\leftarrow}$f‘1 ∈ a.

Assume $F^{\leftarrow}$f‘k+1 ∈ $F^{\leftarrow}$f‘k and $F^{\leftarrow}$f‘k+1 $\in^{k+1}$ a; show $F^{\leftarrow}$f‘k+2 ∈ $F^{\leftarrow}$f‘k+1 and $F^{\leftarrow}$f‘k+2 $\in^{k+2}$ a. But $F^{\leftarrow}$f‘k+1 $\in^{k+1}$ a, so F‘$F^{\leftarrow}$f‘k+1 = f‘k+1 = F“$F^{\leftarrow}$f‘k+1 by Clause (4) of the definition of j-isomorphism; thus, since f‘k+2 ∈ f‘k+1, by the nature of the “ operation, $F^{\leftarrow}$f‘k+2 ∈ $F^{\leftarrow}$f‘k+1, and hence $F^{\leftarrow}$f‘k+2 $\in^{k+2}$ a. This establishes the subclaim.

Thus ∀k∈i. $F^{\leftarrow}$f‘k+1 ∈ $F^{\leftarrow}$f‘k, so $F^{\leftarrow}$ composed with f is a membership sequence of length i from a to x, i.e., x $\in^{i}$ a, as required. □

The following result characterizes the first non-trivial j-isomorphism relation.

Lemma 19.4 (1-Isomorphism Lemma (Function Subset Assumption, Pairwise Union))**.** ∀a,b. non-empty(a) ⇒ a $≎^{1}$ b ≡ . a ≈ b & (a ∈ a ≡ b ∈ b).

(Recall that two empty objects are $≎^{j}$ iff they are equal, for j ≥ 1.) Note that the final conjunct is significant only in the absence of full Foundation; the main interest of this result is for CUSι and possible extensions, not the Base Theory, but such applications are beyond the scope of this paper.

Proof.

Part 1: Assume non-empty(a) & ∃F: a $≎^{1}$ b. Show a ≈ b & (a ∈ a ≡ b ∈ b).

As in j-Isomorphism Note (4), F|a: a ≈ b; by the Function Subset Assumption, F|a is a set. If a ∉ a, then a is not in the domain of F|a, so (since F is 1-1) b = F‘a is not in the range of F|a, which is b. Conversely, if a ∈ a, then a is in the domain of F|a, so b = F‘a is in the range of F|a, i.e. b.

Part 2: Assume ∃G: a ≈ b & (a ∈ a ≡ b ∈ b); show a $\Bumpeq^1$ b.

Case 1: a ∈ a, so b ∈ b.

Define F as follows, swapping elements so that F‘a will be b: G — {<a, G‘a>, <$G^{\leftarrow}$b, b>} ∪ {<a, b>, <$G^{\leftarrow}$b, G‘a>}, which exists by the Function Subset Assumption and Pairwise Union. ($G^{\leftarrow}$b is the G-inverse of b, which exists since G is one-one.) This is clearly still one-one, so F: a ≈ b.

Claim: F: a $\Bumpeq^1$ b. Since a ∈ a, then Ξ^1a = {a} ∪ a = a; likewise for b, which establishes clause (1) of the definition of $\Bumpeq^1$. This establishes clause (2) as well. (3) is true by fiat.

For clause (4), only the case j=0 is required; i.e., show F‘a = b = F“a. But G: a ≈ b, so b = G“a. G differs from F only in that their values for a and $G^{\leftarrow}$b are swapped, so b = F“a as well.

Case 2: a ∉ a, so b ∉ b.

Define F = G ∪ <a, b>, which exists by Pairwise Union.

Ξ^1a is a ∪ {a}, which exists by Pairwise Union; likewise Ξ^1b, which establishes clause (1) of the definition of $\Bumpeq^1$. Since a ∉ a and b ∉ b, F maps Ξ^1a to Ξ^1b, and is still one-one, which establishes clause (2). Clause (3) is true by fiat in this case as well. The only applicable case for clause (4) is j = 0, i.e., y = a. So it only remains to show F‘a = b = F“a. But F|a is G, and G: a ≈ b, so F“a is b, as required. □

19.2 1-Isomorphism and Paradox

This result, though it aids the consistency proof in this paper, would have disturbing consequences for the goal of extending my theory, CUSι, though apparently not Church's original theory: Natural extensions of CUSι (with unrestricted axioms of generalized Frege-Russell cardinals and some natural mappings as sets) lead to a variant of the Russell Paradox. (I would have hoped that some expansion of my theory could be useful for working mathematicians, for example, in category theory [Feferman 2006], but this seems to rule that out.)

Call a set *blasphemous* iff the universe is equinumerous to it via a set mapping; the formal definition is below. (This is a pun on the name Church and his conjecture about high sets ([1974a], p. 299), plus Cantor's notion of absolute infinity as presented in [Hallett 1984].) A sometimes helpful informal notion is being *weakly blasphemous,* via a class mapping rather than a set mapping. More formally, this is a definition schema: b is **weakly blasphemous via** φ iff$_{dfs}$ FUNCTION(φ) & ∀x. ∃!y ∈ b. φ(x,y). Often I will elide the formula in informal exposition; if I were to do so formally, there would be a risk of hidden quantification over virtual classes.

Informal Motivation: An easier, but not quite sufficient, version of this paradox is the 1-isomorphism class of the universe. Since the universe is a member of itself, this will be the set of all self-membered blasphemous sets. Does this set contain itself? If it does, then it does; if it doesn't, it doesn't. This isn't a paradox, but suggests a problem with such equivalence-class axioms, that in this case they say too little.

This does evoke a familiar route to a genuine paradox: Take a blasphemous set that isn't self-membered; the set of all singletons is a convenient one. The idea, which I work out in detail below, is that its 1-isomorphism class contains itself iff it doesn't.

Informally assume we are working in some partially-specified stronger theory than CUSι (call it CUSι#), which I will show inconsistent, with the *unrestricted* existence of 1-isomorphism classes, plus three additional properties, formally stated following the definitions.

(II), below, will mean simply that the 1-Isomorphism Lemma is still true in CUSι#, even for the new 1-isomorphism classes of ill-founded sets.

(III) is that being equinumerous to the set of all singletons (abbreviated $\mathbb{1}$) is equivalent to being equinumerous to the universe. To motivate this, note that the singleton function maps the universe one-one onto the set of all singletons. Thus any set equinumerous to $\mathbb{1}$ is at least weakly blasphemous, via the obvious composition map. Actually proving (III) in general would seem to require a fair amount of compositionality, which Church's technique does not seem to provide.

(IV) is that there is a set mapping from the universe one-one onto the 1-isomorphism class (abbreviated $\mathscr{I}$) of the set of all singletons. (Defined formally below; $\mathscr{I}$ will be a set by the unrestricted Axiom of Generalized Frege Cardinals.)

To motivate this, I will exhibit a class mapping which is one-one and might reasonably be hoped to map the universe into $\mathscr{I}$. (For readers worried about implicit quantification over ultimate classes, I stress that this motivational section is purely motivational: I am arguing that the desired properties of the hypothetical theory CUSι#, which turn out to lead to paradox, would have been reasonable to desire in the absence of paradox.)

Let $\mathbb{2}$ be the set of all pairs, which exists by a similar argument to that for $\mathbb{1}$. By the Axiom of Pairwise Union, $\mathbb{2} \cup \{\{x\}\}$, for arbitrary x, exists. It's non-self-membered, since it has more than two members, unlike any of its members. It's at least weakly blasphemous: Consider the mapping $z \rightarrow <z, z>$, which maps the universe one-one into $\mathbb{2}$, hence also into $\mathbb{2} \cup \{\{x\}\}$. This does not suffice to show that $\mathbb{2} \cup \{\{x\}\}$ is blasphemous, but does (I hope) make that seem a reasonable desideratum for CUSι#. If $\mathbb{2} \cup \{\{x\}\}$ is blasphemous and non-self-membered, it's a member of $\mathscr{I}$, the 1-isomorphism class of the set of all singletons.

Consider the class mapping $x \rightarrow \mathbb{2} \cup \{\{x\}\}$. It is obviously one-one. By the preceding, if $\mathbb{2} \cup \{\{x\}\}$ is blasphemous for each x, this mapping would inject the universe (abbr: $\mathbb{U}$) into $\mathscr{I}$. So $\mathscr{I}$ would also be weakly blasphemous, so it seems a reasonable desideratum that $\mathbb{U} \approx \mathscr{I}$.

More formally, define $\mathbb{U} =_{df}$ ɿu .∀x. x∈u. This will exist by the Unrestricted Axiom of Symmetric Difference.

Define **blasphemous**(b) iff_{df} $\mathbb{U} \approx b$, i.e. ∃f. $\text{maps}_{1\text{-}1}(f, \mathbb{U}, b)$.

Let $\mathbb{1}$ be the set of all singletons; this exists by the Unrestricted Axiom of Generalized Frege 1-Cardinals, as the 1-isomorphism class of $\{\emptyset\}$, unioned with the 1-isomorphism class of any self-membered singleton, if such exists. It has more than one member, hence does not contain itself.

$\mathscr{I}$ will be the class of things to which $\mathbb{1}$ is 1-isomorphic: $\mathscr{I} =_{df} \{x \mid \mathbb{1} \backsimeq^1 x\}$. This will be a set by the Unrestricted Axiom of Generalized Frege 1-Cardinals, (I) below.

Assumptions on CUSι#:

(I) Unrestricted Axiom of Generalized Frege 1-Cardinals: $\forall b.\ \exists F\ \forall x.\ x \in F \equiv b \backsimeq^1 x$

(II) The 1-Isomorphism Lemma still holds in CUSι#:

$\forall a,b.$ non-empty$(a) \Rightarrow a \backsimeq^1 b \equiv .\ a \approx b\ \ \&\ \ (a \in a \equiv b \in b)$

(III) $\forall x.\ \mathbb{1} \approx x$ iff $\mathbb{U} \approx x$

(IV) $\mathbb{U} \approx \mathscr{I}$

Thus, by (II) and the definition of $\mathscr{I}$, since $\mathbb{1} \notin \mathbb{1}$, we have $\forall x.\ x \in \mathscr{I} \equiv \mathbb{1} \approx x\ \ \&\ \ x \notin x$.

By (III), $\forall x.\ x \in \mathscr{I} \equiv \mathbb{U} \approx x\ \ \&\ \ x \notin x$.

Substituting $\mathscr{I}$ in the preceding, we have $\mathscr{I} \in \mathscr{I} \equiv \mathbb{U} \approx \mathscr{I}\ \&\ \ \mathscr{I} \notin \mathscr{I}$. But $\mathbb{U} \approx \mathscr{I}$ by (IV), so $\mathscr{I} \in \mathscr{I} \equiv \mathscr{I} \notin \mathscr{I}$, contradiction.

This could be interpreted as an example of the Bad Company Argument against equivalence sets ([Dummett 1991] pp. 188-9, [Boolos 1990], p. 214) or the Embarrassment of Riches Argument [Weir 2003], p. 28, or perhaps a confirmation of Forster's "Naturam expellas furca" argument [Forster 2006], p. 240. Cp. also Holmes' proof of the non-set-hood of the membership relation [Holmes 1998] p. 43, and Remark 7.7 on cardinalities and paradox in [Forster & Libert 2011].

I do not believe this is a counterexample to Heck's observation that "there are no set-theoretic paradoxes specifically concerning cardinal numbers" ([Heck 2013], p. 224), nor even evidence against Frege-Russell cardinals for ill-founded sets, but merely a hazard of a relation which can code enough information about membership to emulate the Russell Paradox.

This may also mean that extensions of Oberschelp's theory (which like CUSι, has the Singleton Function as a set, and which I believe also proves the existence of j-isomorphism classes for well-founded sets) cannot prove the set-hood of unrestricted generalized Frege cardinals and/or some of the preceding natural mappings, on pain of inconsistency.

19.3 Well-Founded Equivalence Relations

Theorem 19.5 (Well-Founded Equivalence Relation Theorem). $\forall j \in \omega$, $\backsimeq^j$ is an equivalence relation on the well-founded sets.

The result is actually slightly stronger; only one of the sets need be assumed well-founded. Note that we are now assuming Foundation, so both the assumption and the title of the theorem are redundant; but for possible use over other base theories, and to emphasize the nature of the result, I will limit my direct use of Foundation. (Explicitly calling out the indirect assumptions necessary for this theorem would be non-trivial, however, because of the use of recursion and the Cumulative Union Lemma.)

Proof.

Reflexive: ∀a. wf(a) → a $\stackrel{\sim}{\rightarrow}^{j}$ a.

By the Cumulative Union Lemma, SET(Ξ^ja). Define F as { <x,x> | x ∈ Ξ^ja }, which exists by Well-Founded Replacement. F obviously maps Ξ^ja one-one to itself, which establishes clause (2) of the definition of j-Isomorphism. (3) and (4) are both trivial for an identity mapping.

Symmetric: ∀a. wf(a) → a $\stackrel{\sim}{\rightarrow}^{j}$ b ≡ b $\stackrel{\sim}{\rightarrow}^{j}$ a.

Assume wf(a) and F: a $\stackrel{\sim}{\rightarrow}^{j}$ b. Construct G: b $\stackrel{\sim}{\rightarrow}^{j}$ a. (The other direction is the same after interchanging variables.) By clause 2 of the definition of j-Isomorphism, $\text{maps}_{1\text{-}1}$(F, Ξ^ja, Ξ^jb). Define G as the inverse of F, i.e., { <y,x> | x ∈ Ξ^ja & <x,y> ∈ F }, which exists by REP_{low}.

Claim: G is a j-isomorphism. Clauses (1), (2), and (3) of the definition of j-Isomorphism follow easily from the corresponding clauses for F. Clause (4) is ∀y $\in^{<j}$ b. G‘y = G“y, i.e., since G is $F^{\leftarrow}$, $F^{\leftarrow}$y = $F^{\leftarrow}$“y. Consider an arbitrary y and i with i<j and y $\in^{i}$ b; show $F^{\leftarrow}$y = $F^{\leftarrow}$“y.

Recall the j-Isomorphism/Level j Lemma (19.3): If F: a $\stackrel{\sim}{\rightarrow}^{j}$ b and i ≤ j, then ∀x. x $\in^{i}$ a ≡ F‘x $\in^{i}$ b. So since y = F‘$F^{\leftarrow}$y $\in^{i}$ b, then $F^{\leftarrow}$y $\in^{i}$ a. Thus by Clause (4) for F, F‘$F^{\leftarrow}$y = y = F“$F^{\leftarrow}$y.

So $F^{\leftarrow}$“y = $F^{\leftarrow}$“(F“$F^{\leftarrow}$y), which, since F is 1-1, is $F^{\leftarrow}$y, as required.

Transitive: ∀a. wf(a) & a $\stackrel{\sim}{\rightarrow}^{j}$ b & b $\stackrel{\sim}{\rightarrow}^{j}$ c → a $\stackrel{\sim}{\rightarrow}^{j}$ c.

Assume wf(a), F: a $\stackrel{\sim}{\rightarrow}^{j}$ b, and G: b $\stackrel{\sim}{\rightarrow}^{j}$ c. Construct H: a $\stackrel{\sim}{\rightarrow}^{j}$ c.

Define H as the composition of F and G, i.e., H = { <x,z> | x ∈ Ξ^ja & ∃y <x,y> ∈ F & <y,z> ∈ G }, which exists by REP_{low}. Claim: H is a j-isomorphism.

Clause (1) of the definition of j-Isomorphism follows from the first and second conjunct, respectively, of the corresponding clause in the definitions of a $\stackrel{\sim}{\rightarrow}^{j}$ b and b $\stackrel{\sim}{\rightarrow}^{j}$ c. Clauses (2) and (3) are obvious properties of composition of functions. Clause (4) is ∀y $\in^{<j}$ a. H‘y = H“y.

Consider an arbitrary y and i with i<j and y $\in^{i}$ a. Claim: H‘y = G‘F‘y = H“y. By clause (4) for F, F‘y = F“y. By clause (4) for G and the j-Isomorphism/Level j Lemma, G‘F‘y = G“F‘y. So H‘y = G‘F‘y = G“F‘y = G“F“y = H“y. □

Theorem 19.6 (Singleton Function/2-Isomorphism Theorem (Foundation for Finite Sets))**.** ∀b. <Ø, {Ø}> $\stackrel{\sim}{\rightarrow}^{2}$ b ≡ ∃d. b = <d, {d}>.

I.e., the singleton function is a 2-isomorphism equivalence class. The use of Foundation is only for the second part of the proof, is only needed for sets with three or fewer members, and could be avoided by explicitly cataloging the six possible failures of Foundation, as in [Sheridan 1989]. Note that <Ø, {Ø}> expands, by the definition of Kuratowski ordered pair, to { {Ø}, {Ø, {Ø}} }.

Proof.

Part 1: Assume F: <Ø, {Ø}> $\eqsim^2$ b; show **∃d.** b = <d, {d}>.

(Note that this direction of the proof makes no use of Foundation nor of F's being one-one.) Let d = F'Ø, which is defined, since Ø $\in^2$ <Ø, {Ø}>. Show b = <d, {d}>, i.e., <F'Ø, {F'Ø}> a.k.a. { {F'Ø}, {F'Ø, {F'Ø}} } or { {d}, {d, {d}} }. (By j-Isomorphism Note 1, b must be F"<Ø, {Ø}> = F"{ {Ø}, {Ø, {Ø}} } = { F'{Ø}, F'{Ø, {Ø}} }.)

By the instance for j = 1 of clause 4 of the definition of j-isomorphism, ∀y ∈ <Ø, {Ø}>. F'y = F"y. Since <Ø, {Ø}> is { {Ø}, {Ø, {Ø}} }, this amounts to F'{Ø} = F"{Ø} = {F'Ø} and F'{Ø, {Ø}} = F"{Ø, {Ø}} = {F'Ø, F'{Ø}} = {F'Ø, {F'Ø}}. Thus by the preceding expansion of j-Isomorphism Note 1, b = { F'{Ø}, F'{Ø, {Ø}} } = { {F'Ø}, {F'Ø, {F'Ø}} } = <d, {d}>, as required.

Part 2: Assume ∃d. b = <d, {d}>; show ∃F. F: <Ø, {Ø}> $\eqsim^2$ b.

Let F consist of the following ordered pairs: (level number prepended for clarity)

0: < <Ø, {Ø}>, b > (i.e., < { {Ø}, {Ø, {Ø}} }, { {d}, {d, {d}} } >)
1: < {Ø, {Ø}}, {d, {d}} >
1: < {Ø}, {d} >
2: < Ø, d >

Claim: F satisfies the definition of $\eqsim^2$. (The only interesting part will be demonstrating that F is one-one.)

Clauses 1 and 3 are trivial: F takes <Ø, {Ø}> to b by fiat, and maps(F, Ξ^2a, Ξ^2b) is true by inspection.

Clause 4 amounts to the cases below, each of which is true by inspection:

j = 0: F'{ {Ø}, {Ø, {Ø}} } = F"{ {Ø}, {Ø, {Ø}} }
j = 1: F'{Ø} = F"{Ø}
j = 1: F'{Ø, {Ø}} = F"{Ø, {Ø}}

All that remains is to show that F is one-one. Assume not; then two of the following must be equal: d; {d}; {d, {d}}; or { {d}, {d, {d}} }. (In a broader context, e.g., Aczel's ill-founded but small sets, the singleton function would have to be handled as a union of some or all of the following cases.) There are six cases:

(1) d = {d}
(2) d = {d, {d}}
(3) d = { {d}, {d, {d}} }
(4) {d} = {d, {d}}
(5) {d} = { {d}, {d, {d}} }
(6) {d, {d}} = { {d}, {d, {d}} }

In cases 1, 2, 4, 5, and 6, the set on the left is a member of that on the right, so we have a membership cycle of length one. In case 3, we have a membership cycle of length two. All violate Foundation, so we are done. □

This will help to show below that, in CUSι, the singleton function is a set; but there will be some non-obvious additional effort required, e.g., to show that an object which is 2-isomorphic to $<\emptyset, \{\emptyset\}>$ in the base theory remains so in the interpretation, and to verify that there are no new objects which are of the form $<x, \{x\}>$ but are not 2-isomorphic to $<\emptyset, \{\emptyset\}>$. If, for example, there were a Q such that $Q = \{Q\}$, then $<Q, \{Q\}> = \{ \{Q\}, \{Q, \{Q\}\} \} = \{ Q, \{Q\} \} = \{Q\} = Q$. This has only a single member, hence would not be 2-isomorphic to $<\emptyset, \{\emptyset\}>$. In the interpretation, all the new sets will be non-low, so this is not an issue. In a more general context, there are only a finite number of ways that this can go wrong, so it would not be hard to construct the singleton function as the union of a finite number of 2-isomorphism classes.

Lemma 19.7 (Level <j Equinumerosity Lemma (Function Subset Assumption))**.** If F: $x \Bumpeq^{j} y$ and $z \in^{<j} x$, then F‘z ≈ z.

Proof. By clause 4 of the definition of j-isomorphism, the required map is simply F restricted to z, which exists by the Function Subset Assumption, and is one-one by clause (2) of the definition of j-isomorphism. □

Lemma 19.8 (Increasing Strictness Lemma (Function Subset Assumption))**.** $\forall j, k \leq \omega\ \forall x, y.\ j \leq k\ \ \&\ \ x \Bumpeq^{k} y \Rightarrow x \Bumpeq^{j} y$. (Cp.$\sim^{j}$ Requirements (α), below.)

Assume $j \leq k \leq \omega\ \ \&\ \ F: x \Bumpeq^{k} y$; construct G: $x \Bumpeq^{j} y$. The required G will be $F|\Xi^{j}a$, which is a set by the Function Subset Assumption.

Proof. Clause (1) of the definition of j-isomorphism is true for G by the Function Subset Assumption. Clause (2), that G is 1-1 and that its domain is $\Xi^{j}a$, is true by its definition. Its range is therefore $\Xi^{j}b$ by the j-Isomorphism/Level j Lemma. Clause (3) is true by definition of G.

Clause (4) is true because G is a restriction of F, and $y \in^{<j} a \Rightarrow y \in^{<k} a$. Formally, assume $y \in^{<j} a$; show G‘y = G“y. Since $y \in^{<j} a$, $y \in^{<k} a$ and hence G‘y = F‘y = F“y. But for each member v of y, $v \in^{\leq k} a$; so for each member z of y, G‘z is defined and equal to F‘z. Thus F“y = G“y, which equals G‘y. □

19.4 Replacing at Level*j Construction

Given a well-founded set, a, not empty at level*j (for $j \in \omega$), an arbitrary object z, and an arbitrary infinite Cantor cardinal χ larger than the transitive closure of a, I will construct below a set b(z) such that $a \Bumpeq^{j} b(z)$ and $z \in^{j+2} b(z)$. Additionally, b(z) will have a member at level j of cardinality χ.

The function b(z) is one-one, so the universe can be injected into a's j-isomorphism class. In the interpretation below, the value of b( ) with a's j-isomorphism class (which will be a set_3) as argument, gives a $membership_3$ loop of length j+2; so every non-degenerate j-isomorphism class will be ill-founded.

Gandy's and my conjecture in [Sheridan 1989] that the following construction could be done by reverse recursion on membership depth seems to be false. A counterexample to the natural construction seems to be { {0}, {{Ø}} }, replacing 0 with $\aleph_1$ at level 2. The natural construction by reverse recursion on depth would leave {{Ø}} at level 1 unchanged, since {0} at level 2 would also be unchanged. But this would fail to preserve the level 1 graph edge from {{Ø}} to {0}, since {{Ø}} would be unchanged, but {0} would map to {$\aleph_1$}. (Part of the difficulty is that {0} is a member at level 1, hence not at level*2. Using maximal rather than minimal depth would not work, since, for instance, 0 is a member of ω at all finite levels.)

19.4.1 Preliminary Definitions

a is not empty at level*j, so it has a member at level*j, d. Choose an arbitrary object z; take χ as an arbitrary infinite (Cantor) cardinal larger than TC(a). Construct F and b(z), which will be a with d replaced at level*j by χ(z) [defined below], with F: a $\backsimeq^j$ b(z). F is constructed by transfinite recursion on a variant of the Cumulative Hierarchy, modified for urelements; thus this construction is essentially dependent on the Axiom of Foundation. (I had hoped that the consistency of the Axiom of Generalized Frege Cardinals would not be dependent on its restriction to well-founded sets, even though the available consistency proof is. But the above paradox of the set of all non-self-membered blasphemous sets mandates caution.) F will be constructed in stages, F^α for each ordinal α ≤ ρ(a), where ρ is the usual Cumulative Hierarchy rank function ([Levy 1979], §6.6); R(α) will be the usual α^{th} stage of the Cumulative Hierarchy, modified for the inclusion of the relevant urelements in R(0), as follows.

Let R(0) be the set of all urelements in TC(a); define, similarly to the usual cumulative hierarchy, $R(\alpha) = \bigcup_{\xi<\alpha} P(R(\xi))$, where P is power set. (Since TC(a) is a set, so will be the R(α)'s, and hence the F^α's defined on them below.) Let R*(α) be the collection of objects first appearing in stage R(α), i.e., R(α) — $\bigcup_{\xi<\alpha}$ R(ξ). (So R*(α) will be empty if α is a limit ordinal. R*(0) will be equal to R(0).) The union of the F^α's will be a mapping on the transitive closure of a; the desired F will be the restriction of the union of the F^α's to Ξ^ja. The desired b(z) will be F'a.

χ was taken above as an arbitrary infinite (Cantor) cardinal larger than TC(a); define χ(x) = χ — {{Ø}} ∪ {{x}}. Observe that χ(x) is one-one, and χ(x) contains x at level 2.

Let β be the first ordinal such that d ∈ R(β), i.e., β is unique such that d ∈ R*(β).

19.4.2 The Construction

Define F^α on R*(α), for ordinals α, as follows; the recursion will end at the first stage, γ, containing a (i.e., R*(ρ(a)+1). γ must be a successor ordinal, so γ - 1 exists. Observe that at each stage α ≤ β, F^α will be obviously one-one, since

χ(z) is distinct from—because it is larger than—any member of TC(a). Showing that later functions, and their union, are one-one will be more difficult.

If d is not an urelement, then F^0 will be the identity function on R*(0) (the urelements in TC(a)). Otherwise F^0 maps d to χ(z), and is the identity on the rest of R*(0); and β is 0. Formally,

F^0 =

$\{\langle d, \chi(z)\rangle\} \cup \{ \langle x,x\rangle \mid x \in R^*(0) \ \& \ x \neq d \}$, if urelement(d)

$\{ \langle x,x\rangle \mid x \in R^*(0)\}$, otherwise.

F^1, if d = Ø (and hence β = 1), maps Ø to χ(z), and is the identity on the rest of R*(1); otherwise F^1 is just the identity function on R*(1). Formally,

F^1 =

$\{\langle \emptyset, \chi(z)\rangle\} \cup \{ \langle x,x\rangle \mid x \in R^*(1) \ \& \ x \neq \emptyset \}$, if d = Ø,

$\{ \langle x,x\rangle \mid x \in R^*(1)\}$, otherwise.

For stages α between 1 and β (if any; if β is 0 or 1, this clause is vacuous, and the following clause coincides with clause 0 or 1), F^α is the identity on members of TC(a) in R*(α). I.e.,

$F^\alpha = \{ \langle x,x\rangle \mid x \in R^*(\alpha) \cap TC(a) \}$.

At stage β (where d $\in$ R*(β)), F^β maps d to χ(z), and is otherwise the identity on members of TC(a) in R*(β). I.e.,

$F^\beta = \{\langle d, \chi(z)\rangle\} \cup \{ \langle x,x\rangle \mid x \in R^*(\beta) \ \& \ x \in TC(a) \ \& \ x \neq d \}$.

Define $F^{\leq\alpha} = \bigcup_{\delta\leq\alpha} F^\delta$; this is a function, since the domains of the F^α are disjoint. (The continuation of the definition below maintains this disjointness; each F^α will be restricted to R*(α).)

For successor ordinals α+1 greater than β and less than γ, $F^{\alpha+1}$‘x is $F^{\leq\alpha}$“x, i.e.,

$F^{\alpha+1} = \{ \langle x, \{F^{\leq\alpha}\text{‘}w \mid w \in x\}\rangle \mid x \in R^*(\alpha+1) \cap TC(a) \}$.

Observe that each $F^{\leq\alpha}$‘w will be defined, since w $\in$ x, and hence w is earlier in the Cumulative Hierarchy.

The limit ordinal case is trivial, since R*(α) is empty for α a limit ordinal.

F^γ is defined only for a:

$F^\gamma = \{ \langle a, \{F^{\leq \gamma-1}\text{‘}y \mid y \in a\}\rangle \}$.

Let b(z) be $\{F^{\leq \gamma-1}\text{‘}w \mid w \in a\}$, i.e., F^γ‘a. (For brevity, in the rest of this proof, since z is fixed, abbreviate b(z) to b.) Let

$F^+ = \bigcup_{\delta \leq \gamma} F^\delta$; let F be F^+ restricted to Ξ^ja, i.e.,

$F = \{ \langle x,y\rangle \mid \langle x,y\rangle \in F^+ \ \& \ x \in \Xi^j a\}$.

Example: Let j = 2, a = 3 — {Ø} = { {Ø}, {{Ø}} } = { 1, {1} }, d = 0, γ = 4, β = 1, χ = ω, χ(z) = ω — {{Ø}} $\cup$ {{z}}.

R*(0) = Ø (Since there are no urelements in TC(a).)

R*(1) = { Ø }

R*(2) = { {Ø} }

R*(3) = { {{Ø}} ... }

R*(4) = { { {Ø}, {{Ø}} } ... }

F^0 is the empty function.
$F^1 = F^\beta = \{ <\emptyset, \chi(z)> \}$.
$F^2‘x = F^{\leq\beta}“x = \{ <x, \{F^{\leq 1}‘w \mid w \in x\}> \mid x \in R^*(2) \cap TC(a) \} = \{ <\{\emptyset\}, \{F^{\leq 1}‘\emptyset\}> \} = \{ <\{\emptyset\}, \{\chi(z)\}> \}$.
$F^3‘x = F^{2+1}‘x = F^{\leq 2}“x = \{ <x, \{F^{\leq 2}‘w \mid w \in x\}> \mid x \in R^*(3) \cap TC(a) \} = \{ <\{\{\emptyset\}\}, \{F^{\leq 2}‘w \mid w \in \{\{\emptyset\}\}> \} = \{ <\{\{\emptyset\}\}, \{ F^{\leq 2}‘\{\emptyset\} \} > \} = \{ <\{\{\emptyset\}\}, \{ \{\chi(z)\} \} > \}$.
$F^4 = \{ <a, \{F^{\leq 3}‘y \mid y \in a\}> \} = \{ <a, \{ F^{\leq 3}‘\{\emptyset\}, F^{\leq 3}‘\{\{\emptyset\}\} \}> \} = \{ <a, \{ \{\chi(z)\}, \{\{\chi(z)\}\} \}> \}$.
So $b(z) = \{ \{\chi(z)\}, \{\{\chi(z)\}\} \}$.

19.4.3 Properties of the Construction

Theorem 19.9 (Replacing Theorem)**.** F: a $\backsimeq^j$ b.

I will start by proving the two easy clauses of the definition of $\backsimeq^j$, then prove three results (leading to the crucial One-One Lemma), and then prove the remaining clauses; the order will be (3), (1), lemmata, (4), (2).
Proof.

Clause (3): F‘a = b. True by choice of b.

Clause (1): SET(Ξ^ja) & SET(Ξ^jb). True by the Cumulative Union Lemma.

Lemma 19.10 (Domain Lemma)**.** The domain of F is Ξ^ja.

Proof. By the definition of F, its domain is a subset of Ξ^ja; so it remains only to show that any member, x, of Ξ^ja is in the domain of F. Since x $\in \Xi^j$a, $\exists$k. $0 \leq k \leq j$ & $x \in^k a$. If k is 0, then x = a, and we are through. So there is a descending $\in$-chain of non-zero length from a to x; thus x's cumulative hierarchy rank is less than a's, say δ. So x is in the domain of $F^{\delta+1}$; it is in Ξ^ja, so it is in the domain of F, by the definition of F. □

Lemma 19.11 (Cardinality Lemma)**.** $\forall x \in$ domain(F). F‘x $\neq$ x iff TC(F‘x) $\geq \chi$.

(For this and the following lemma, ">" and "≥" will denote the usual cardinality inequalities; given the presence of Foundation and Choice, this is unproblematic.)

Proof.

Part (i): Assume TC(F‘x) $\geq \chi$; show F‘x $\neq$ x.
By choice of χ, $\chi >$ TC(a); but since x is in the domain of F, which is Ξ^ja, TC(x) $\subseteq$ TC(a). So $\chi >$ TC(x); thus F‘x $\neq$ x.

Part (ii): Show F‘x ≠ x → TC(F‘x) ≥ χ.

By induction on stages:

The antecedent is false for R*(α) with α < β.

At stage β, the consequent is true when x = d; otherwise the antecedent is false.

Assume true for α ≥ β; show for α+1. Assume F‘x = $F^{\alpha+1}$‘x ≠ x. Recall that $F^{\alpha+1}$‘x = $F^{\leq\alpha}$“x. Since $F^{\leq\alpha}$“x ≠ x, then by Extensionality ∃δ,y. δ≤α & y ∈ x & F^{δ}‘y ≠ y. (The case where x is an urelement does not arise; recall that, for any urelement u and any function ϕ, by my definition ϕ“u = u.) But δ ≤ α < α+1, so TC(y) ≥ χ, by the induction hypothesis. But y ∈ x; so TC(x) ≥ χ, as required.

The limit ordinal case is trivial, as usual. □

Corollary 19.12 (Cardinality Corollary)**.** ∀x ∈ Ξ^{j}b. x > TC(a) ↔ x = F‘d.

Proof. Up to and including stage β of the construction of b, F is an identity on everything except d, and obviously no member of Ξ^{j}a is larger than TC(a). F‘d is the only exception so far; it is of size χ, which is larger than TC(a).

For later stages, each $F^{\alpha+1}$‘w is $F^{\leq\alpha}$“w, where w ∈ R*(α+1) ∩ TC(a). Since w is in TC(a), by the nature of the “ operator, $F^{\alpha+1}$‘w will be no larger than TC(a).

For the final stage γ, b is the image of a function on a, and so is also no larger than TC(a). □

Lemma 19.13 (One-One Lemma)**.** F is one-one.

Proof. Assume F is not one-one; thus ∃η, κ, x, y. x ≠ y & F^{η}‘x = F^{κ}‘y. (Without loss of generality it may be assumed that η ≤ κ, renaming if necessary.) Choose η minimal satisfying the preceding. Note that η and κ are both successor ordinals, since F^{ξ} is empty for ξ a limit ordinal.

Note also that, since F‘x = F‘y, either both or neither have transitive closures whose cardinality is at least χ. Thus by the Cardinality Lemma, F‘x ≠ x iff F‘y ≠ y. So if F‘x = x ∨ F‘y = y, then F‘x = x & F‘y = y; so x = y, contradicting the hypothesis. Thus F‘x ≠ x and F‘y ≠ y, so TC(F‘x) = TC(F‘y) ≥ χ. Thus also η and κ are both at least β, since the F^{α}’s are all identity functions before stage β.

Case (i): η = β.

Note first that, since F^{β}‘x ≠ x, by the definition of F^{β}, x must be d, and hence F‘x = χ(z). (Observe that the proof for this case applies even if β = 0.)

Subcase (1): κ = β. This immediately leads to contradiction, since, as noted in its definition, F^{β} is one-one.

Subcase (2): κ > β. So F^{κ}‘y = $F^{\leq\kappa\text{-}1}$“y. But y is in the domain of F, hence in Ξ^{j}a; so its cardinality must be less than or equal to that of TC(a), hence less than χ; the cardinality of $F^{\leq\kappa\text{-}1}$“y will be no greater than that of y. But the cardinality of F‘x = χ(z) is χ, contradiction.

Case (ii): $\eta > \beta$. So $\kappa > \beta$ as well, since $\eta \leq \kappa$. Thus F‘x = $F^{\leq \eta\text{-}1}$“x and F‘y = $F^{\leq \kappa\text{-}1}$“y. But x ≠ y, so by Extensionality, ∃x' ∈ x. x' ∉ y or ∃y' ∈ y. y' ∉ x. (Neither x nor y can be an urelement, since η and κ are greater than β, hence greater than zero; but urelements are handled at stage F^0.)

Subcase (1): y' ∈ y & y' ∉ x. Since F‘x = $F^{\leq \eta\text{-}1}$“x = $F^{\leq \kappa\text{-}1}$“y = F‘y and y' ∈ y, $\exists\lambda\leq\kappa$-1. F^{λ}‘y' ∈ $F^{\leq \kappa\text{-}1}$“y = $F^{\leq \eta\text{-}1}$“x. So ∃x' ∈ x. $\exists\nu\leq\eta$-1. F^{ν}‘x' = F^{λ}‘y'. But y' ∉ x, so y' ≠ x'; $\nu < \eta$, so ν contradicts the minimality of η.

Subcase (2): x' ∈ x. x' ∉ y. Similarly, $\exists\nu\leq\eta$-1. F^{ν}‘x' ∈ $F^{\leq \eta\text{-}1}$“x = $F^{\leq \kappa\text{-}1}$“y. So ∃y' ∈ y. $\exists\lambda\leq\kappa$-1. F^{ν}‘x' = F^{λ}‘y', with ν also contradicting the minimality of η. □

Clause (4) of the Definition of $\backsimeq^{j}$: ∀y $\in^{<j}$ a. F‘y = F“y

Observe first that the definitions of each F^{α} divide into four types: (i) For argument d, which is chosen from level*j, so d $\notin^{<j}$ a, and hence is not relevant. (ii) For $\alpha<\beta$, or $\alpha=\beta$ and arguments other than d, F^{α} is an identity map. (iii) For successor ordinals $\alpha+1$ greater than β and less than γ, $F^{\alpha+1}$‘x is $F^{\leq\alpha}$“x, i.e., $F^{\alpha+1}$ = { <x, {$F^{\leq\alpha}$‘z | z ∈ x}> | x ∈ R*(α+1) ∩ TC(a) }. (iv) For argument a, which is similar to (iii).

Proof. Assume y $\in^{<j}$ a; show F‘y = F“y.

Since y $\in^{<j}$ a, by the Domain Lemma, F‘y exists. So $\exists\alpha\leq\gamma$. F‘y = F^{α}‘y.

Case $\alpha<\beta$: Then F‘y = y. So for each member w of y, w ∈ y $\in^{<j}$ a, so w $\in^{\leq j}$ a, so w is also in the domain of F. Thus ∃δ. F‘w = F^{δ}‘w. Since w ∈ y, w’s cumulative hierarchy rank is lower than y’s, so $\delta < \alpha < \beta$, so F‘w = F^{δ}‘w = w. So F“y = y = F‘y, as required.

Case $\alpha=\beta$: Since y ≠ d, as in the preceding case F‘y = F^{α}‘y = y, F“y = y.

Case $\alpha>\beta$, α a limit ordinal. R*(α) is empty for α a limit ordinal, so this case is vacuous.

Case $\alpha+1>\beta$: $F^{\alpha+1}$‘y is $F^{\leq\alpha}$“y by definition, so F‘y = F“y.

Case $\alpha=\gamma$: Recall the definition: F^{γ} is defined only for a: F^{γ} = { <a, {$F^{\leq\gamma\text{-}1}$‘w | w ∈ a}> }. I.e., F^{γ}‘a = $F^{\leq\gamma\text{-}1}$“a, so F‘a = F“a. □

Clause (2) of the Definition of $\backsimeq^{j}$: $maps_{1\text{-}1}$ (F, Ξ^{j}a, Ξ^{j}b):

Proofs of the three clauses of the definition of $maps_{1\text{-}1}$.

Clause (i) ∀p ∈ F ∃x ∈ Ξ^{j}a ∃y ∈ Ξ^{j}b. p = <x, y>

Let p be a member of F; by the definition of F, ∃x,y. <x,y> ∈ F^{+} & x ∈ Ξ^{j}a. It remains only to show that y ∈ Ξ^{j}b.

By the definition of F^{+}, <x, y> must be a member of F^{δ}, for some $\delta \leq \gamma$.

x ∈ Ξ^{j}a, so x $\in^{i}$ a, for some i with 0 ≤ i ≤ j. Show y = F‘x $\in^{i}$ b, by induction on i.

Case i = 0. Then x = a, so F‘x = b, which is in Ξ^{j}b.

Assume for i, show for i+1. Let x $\in^{i+1}$ a, so ∃w. x ∈ w $\in^{i}$ a. By the induction hypothesis, F‘w $\in^{i}$ b. By Clause (4) (proved above), since w $\in^{<j}$ a, F‘w = F“w. Since x ∈ w, F‘x ∈ F“w = F‘w $\in^{i}$ b. So F‘x $\in^{i+1}$ b, as required.

Clause (ii) $\forall x \in \Xi^j a\ \exists! y \in \Xi^j b\ \exists p \in F\ .\ p = <x, y>$

Let x be an arbitrary member of $\Xi^j a$. Its cumulative hierarchy rank is less than or equal to γ, so F‘x is uniquely defined in the definition of F, and $<x, F‘x>$ is a member of F, as required.

Clause (iii) $\forall y \in \Xi^j b\ \exists! x \in \Xi^j a\ \exists p \in F\ .\ p = <x, y>$

Let y be an arbitrary member of $\Xi^j b$. Note that x is unique if it exists, since F is one-one by the One-One Lemma. So it suffices to show $\exists x \in \Xi^j a$. F‘x = y.

By recursion on membership depth. $y \in \Xi^j b$, so $y \in^i b$, for some i with $0 \leq i \leq j$. Show $\exists x \in^i a$. F‘x = y.

Base Case: i = 0. Then y = b, so take x = a.

Inductive Case: Assume for i, show for i+1, with $i+1 \leq j$. Let $y \in^{i+1} b$, so $\exists u.\ y \in u \in^i b$. By the induction hypothesis, $\exists w \in^i a$. F‘w = u. Let α be the stage such that $w \in R^*(\alpha)$. Show $\exists x \in w$. F‘x = y. (Since $w \in^i a$, this will suffice.)

Subcase (a): $\alpha < \beta$. Then F‘w = w, so u = w, and (since y has lower Cumulative Hierarchy rank than u, which equals w), F‘y = y. So y is the required x: $y \in u = w$, and F‘y = y, as required.

Subcase (b): $\alpha = \beta$. So u = F‘w = F^{β}‘w. As before, the case w = d and u = $\chi(z)$ does not arise, since d is a member of a at level*j and $w \in^i a$, with $i < j$. So F‘w = w and u = w and $y \in u = w$, as before, and thus F‘y = y, as required.

Subcase (c): $\alpha > \beta$. Thus u = F‘w = F^{α}‘w = $F^{\leq \alpha\text{-}1}$“w. Thus $y \in u \in^i b$, so $\exists x \in w\ \exists \delta \leq \alpha\text{-}1$. F^{δ}‘x = y, as required. □

This concludes the proof of the Replacing Theorem. ∎

Observations: F‘d = $\chi(z)$: By the definition of F^0, if d is an urelement; by the definition of F^1, if d = ∅; otherwise by the definition of F^{β}.

$\chi(z) \in^j b$, since $\chi(z)$ = F‘d, and $d \in^j a$.

z is a member at level j+2 of b. (Recall that $\chi(z)$ contains z at level 2.)

Observation 19.14 (Cardinality Replacing Observation)**.** The above construction (considered now as a function of χ) provides an injection of the infinite cardinals larger than the transitive closure of a into a's j-isomorphism class.

Proof. By the preceding, $b(\chi)$ has a member at level j of cardinality χ. By the Cardinality Corollary (19.12), for any other infinite cardinal ξ larger than the transitive closure of a, $b(\chi)$ has no constituent of cardinality ξ, so $b(\chi) \neq b(\xi)$ for any such ξ. □

19.5 Definition of j-rep(x)

For arbitrary x, and j in ω, let r be the first object in the global well-ordering such that x $\Bumpeq^j$ r (if any, otherwise let r be x itself[23]); define **j-rep**(x) $=_{df}$ <j, r>. Say that <j, r> is a **j-rep** iff$_{df}$ there is an object x such that j-rep(x) is <j, r>.

Since ω-isomorphism is equality, define **ω-rep(x)** $=_{df}$ <ω, x>. Since Ø is the first object in the global well-ordering, 0-rep(x) = <0, Ø>. j-isomorphism on urelements will play little part in what follows, since they will be used for the new sets, but (as no two empty objects are j-isomorphic for j > 0) for 1<j<ω and u an empty object, j-rep(u) will be <j, u>.

Define **rank**(h) = j, for j≤ω, if ∃s. h = j-rep(s); undefined otherwise. Since any j-rep is an ordered pair with first component j, this will be single-valued.

19.6 Proof of $\sim^j$ Requirements

Lemma 19.15 ($\Bumpeq$j Requirements Lemma)**.** $\Bumpeq^j$, ω, rank, and j-rep satisfy the $\sim^j$ Requirements from Part II.11.

Substituting $\Bumpeq^j$ for $\sim^j$, ω for μ, and using the specific definitions of j-rep and rank, the $\sim^j$ Requirements are:

(α). ∀j, k ≤ ω ∀x, y. j ≤ k & x $\Bumpeq^k$ y ⇒ x $\Bumpeq^j$ y,
(β). ∀x, y. x $\Bumpeq^0$ y,
(γ). ∀x, y. x $\Bumpeq^\omega$ y ≡ x = y,
(δ). ∀j ≤ ω ∀b ∃r. r = j-rep(b),
(ε). For 0 ≤ j ≤ ω, x $\Bumpeq^j$ y iff j-rep(x) = j-rep(y),
(ζ). ∀h. rank(h) ≤ ω and ∀g. rank(g) = j ⇒ ∃x. g = j-rep(x),
(η). rank(0-rep(∅)) = 0 and ¬∃s: low(s) & ∀d. d ∈ s ↔ ∃x. 1-rep(x) = d.

Proofs: (α) is the Increasing Strictness Lemma, proven above.

(β) and (γ) are just the definitions of $\Bumpeq^0$ and x $\Bumpeq^\omega$ y.

(δ) The new definition of j-rep is defined for any argument for any j ≤ ω.

(ε). Let j be such that 0 ≤ j ≤ ω. (The result is trivial for j = 0 and for j = ω.)

Part 1: Assume x $\Bumpeq^j$ y; show j-rep(x) = j-rep(y). j-rep(x) = <j, r> and j-rep(y) = <j, s>, where r (respectively s) is the first object in the global well-ordering such that x $\Bumpeq^j$ r (respectively y $\Bumpeq^j$ s). By the Well-Founded Equivalence Relation Theorem (19.3), y $\Bumpeq^j$ r and x $\Bumpeq^j$ s, so by minimality, r = s.

Part 2: Assume j-rep(x) = j-rep(y); show x $\Bumpeq^j$ y. So j-rep(x) = j-rep(y) = <j, r>, where x $\Bumpeq^j$ r and y $\Bumpeq^j$ r. Thus by the Well-Founded Equivalence Relation Theorem again, x $\Bumpeq^j$ y.

(ζ). Both conjuncts are part of the definition of this formulation of rank.

(η) conjunct 1: rank(0-rep(Ø)) = 0. As noted above, 0-rep(x) = <0, Ø> because of the redefinition of the global well-ordering. Thus its rank is 0.

[23]This case does not arise in the base theory. The situation will be far more complicated in the interpretation, but the impact on the present consistency proof is limited.

The second conjunct of (η), ¬∃s: low(s) & ∀d. d ∈ s ↔ ∃x. 1-rep(x) = d, will be a consequence of the following lemma, as shown after its proof.

Lemma 19.16 (Non-Emptiness and Prolificity Lemma)**.** If <j, a> is a j-rep and a is not empty at level*j, then <j, a> is j-prolific. (This is a generalization of $\sim^{j}$ Requirement (η.2).)

Recall the definition: j-prolific(g) iff_{df} rank(g) = j & ¬∃s: low(s) & ∀d. d ∈ s ↔ daughter(d, g). Informally, something is j-prolific iff its rank is j and it has many daughters. Recall daughter(h, g) iff_{df} $\exists\, j < \omega$ ∃x. j = rank(g) & j-rep(x) = g & j+1-rep(x) = h.

Thus the lemma expands to: If <j, a> is a j-rep and a is not empty at level*j, then rank(<j, a>) = j & ¬∃s: low(s) & ∀d. d ∈ s ↔ daughter(d, <j, a>).

Proof. The first conjunct of the conclusion of the expansion of the lemma is immediate from the first conjunct of the hypothesis, by the new definition of rank.

To show the remaining conjunct of the lemma, apply the Replacing at Level*j Construction. a is not empty at level*j, with member d, say.

Recall that χ was an arbitrary infinite Cantor cardinal χ larger than the transitive closure of a. The construction was performed with z and χ arbitrary; for this lemma, z will be taken as 0, and the construction will be used as a function of χ mapping a to a(χ). The construction produces a mapping, called F above. The only two mappings it will be necessary to distinguish are for the cases a(χ) and a(ξ), for suitable cardinals χ and ξ; call these F_χ and F_ξ, respectively.

Replacing d with χ(Ø) gives a set a(χ) which is j-isomorphic to a, but has a member (χ(Ø)) at level j of cardinality χ. Likewise replacing d with ξ(Ø) gives a set a(ξ) which is also j-isomorphic to a, but has a member (ξ(Ø)) at level j of cardinality ξ.

Claim: The function (of χ) mapping χ to j+1-rep(a(χ)) is an injection of the class of infinite Cantor cardinals larger than TC(a) into the daughters of a. This will show there is no such low set, s, of daughters of a, and hence j-rep(a) is j-prolific.

Subclaim (1): j+1-rep(a(χ)) is a daughter of <j, a>. I.e., ∃x. j = rank(<j, a>) & j-rep(x) = <j, a> & j+1-rep(x) = j+1-rep(a(χ)). Substituting a(χ) for x, it suffices to show that j = rank(<j, a>) & j-rep(a(χ)) = <j, a> & j+1-rep(a(χ)) = j+1-rep(a(χ)). The last conjunct is trivial, as is the first. Thus it remains only to show that j-rep(a(χ)) = <j, a>.

But a $\simeq^{j}$ a(χ), by the Replacing at Level*j Construction. j-rep(a(χ)) is by definition <j, r> where r is the first object in the global well-ordering such that a(χ) $\simeq^{j}$ r.

But <j, a> is a j-rep, so there is an x such that a is the first object in the global well-ordering such that x $\simeq^{j}$ a. $\simeq^{j}$ is an equivalence relation by the Well-Founded Equivalence Relation Theorem (19.3), so a(χ) $\simeq^{j}$ x, so a is also first such that a(χ) $\simeq^{j}$ a; thus j-rep(a(χ)) is <j, a>.

Subclaim (2): The mapping which takes χ to j+1-rep(a(χ)) is an injection. (Cp. the Cardinality Replacing Observation above.) Let ξ be an infinite cardinal larger than TC(a) distinct from χ; show that j+1-rep(a(ξ)) $\neq$ j+1-rep(a(χ)). By the Replacing at Level*j Construction, a(χ) has a member at level j of cardinality χ; by the Cardinality Corollary to the Replacing at Level*j Construction (19.12), all the other members of a(χ) at level $\leq$ j are no larger than TC(a). Similarly for a(ξ).

If j+1-rep(a(ξ)) = j+1-rep(a(χ)), then by $\sim^{j}$ Requirements ($\in$), a(ξ) $\eqsim^{j+1}$ a(χ).

By the Cardinality Corollary, the only member of Ξ^{j}a(χ) larger than TC(a) is χ(0), and likewise the only member of Ξ^{j}a(ξ) larger than TC(a) is ξ(0).

So by the Level <j Equinumerosity Lemma, substituting j+1 for j, if there is a j+1-isomorphism from a(ξ) to a(χ), it must preserve cardinality for members at level $\leq$j, so the only possibility is to map ξ(0) to χ(0); but this would imply that ξ(0) is equinumerous to χ(0). But ξ(0) is of cardinality ξ, and χ(0) is of cardinality ξ, and ξ and χ were chosen to have distinct cardinalities, contradiction. This establishes the subclaim, claim, and lemma. □

Corollary 19.17 ($\sim$j Requirements (η))**.**

The first conjunct of (η) has been shown above. The second conjunct is ¬∃s: low(s) & ∀d. d ∈ s ↔ ∃x. 1-rep(x) = d.

Proof. First I will show that d is a daughter of the unique 0-rep iff it is a 1-rep. daughter(d, <0, Ø>) is equivalent to ∃x. 0-rep(x) = <0, Ø> & 1-rep(x) = d. Because $\eqsim^{0}$ is the universal relation, the first conjunct is trivially true, so daughter(d, <0, Ø>) is equivalent to ∃x. 1-rep(x) = d, i.e., d is a daughter of the unique 0-rep iff it is a 1-rep, as required.

The preceding lemma expands, as noted in its proof, to: if <j, a> is a j-rep and a is not empty at level*j, then rank(g) = j & ¬∃s: low(s) & ∀d. d ∈ s ↔ daughter(d, g). Substituting 0 for j and Ø for a, we get: if <0, Ø> is a 0-rep and Ø is not empty at level*0, then rank(<0, Ø>) = 0 & ¬∃s: low(s) & ∀d. d ∈ s ↔ daughter(d, <0, Ø>). All but the last major conjunct are trivial, which is ¬∃s: low(s) & ∀d. d ∈ s ↔ daughter(d, <0, Ø>). By the preceding, daughter(d, <0, Ø>) iff d is a 1-rep. So it suffices to show ¬∃s: low(s) & ∀d. d ∈ s ↔ d is a 1-rep.

But by the 1-Isomorphism Lemma (19.4), if there is such a set s, 1-rep(ξ) injects the sufficiently large Cantor cardinals into it. In our current Base Theory, with Foundation and Choice, this is an obvious contradiction by the usual argument. (Even in a different base theory without Foundation, that s is low would lead easily to a contradiction via Well-Founded Replacement.) □

Lemma 19.18 (j-Empty j-Isomorphism Lemma)**.** If a is empty at level j, then the only thing which is j-isomorphic to a is a itself.

Note that this is not true in a theory which violates Finsler Strong Extensionality [Aczel 1985].

Proof. Assume a is empty at level j, for j ∈ ω, and F: a $\approxeq^j$ b, with a≠b. (The result is trivial if either a or b is an urelement, so assume they are sets.)

Since a is empty at level j, Ξ^ja contains only members at level < j; so ∀y ∈ domain(F). F‘y = F“y.

Since a≠b, there exists an x $\in^i$ a with F‘x ≠ x and i maximal. (Since a is empty at level j, i < j.)

By my definition of “, which takes empty objects to themselves, this x cannot be empty.

But by the maximality of i, F is the identity on all the members of x. But this implies that F“x = x. But F‘x = F“x by the above; contradiction. □

Define **j-pure**(x) iff x has no members at less than level j which are urelements; more formally: **j-pure**(x) iff_{df} ∀y. y$\in^{<j}$x → ¬urelement(y). (For j=1, 1-pure(y) reduces to ¬urelement(y); 0-pure is vacuously true and will not be used.) (Note that this is simpler than the definition in [Sheridan 1993].) The intent is to exclude altered objects from being relevant to j-isomorphism, since this would make j-isomorphism different in the interpretation. This is manifested in the j-Isomorphism j-Purity Lemma, below. If I were doing this construction without using all urelements in the base theory as new sets, which would allow for urelements in the interpretation, it would make sense to define j-pure in terms of altered objects rather than urelements.

20 The Interpretation $\in_3$

20.1 Definition of INDEX3

As noted above in the introductory remarks for Foundation, Choice, j-Isomorphism, and Less Generality (III.19), the definition of INDEX3 will be similar to the earlier definition of INDEX, with the following differences:

- ω is substituted for the arbitrary ordinal μ.
- For the arbitrary sequence of relations $\sim^j$ (j≤μ) I substitute $\approxeq^j$ (j≤ω), with $\approxeq^\omega$ being equality and $\approxeq^0$ being the universal relation.
- One of the clauses (bracketed below) in the definition of INDEX is now redundant, given the assumption of Foundation in the Base Theory.
- Stricter conjuncts are substituted in clause (d), as noted in the original definition of INDEX (II.14).

Informally, an INDEX3 will be an ω+1-tuple, with at least one of its components (other than ω) non-empty, in which each L^j contains only j-reps of j-pure objects not empty at level*j.

Define **INDEX3**(L) $\equiv_{df}$

(a) ω+1-tuple(L) & ∀j ≤ μ. set(L_j),

[(b) low($\cup_{j\leq\omega}$ L^j)],

(c) ∃j < ω ∃x. x ∈ L^j ,

(d) ∀j < ω ∀a ∈ L^j . rank(a) = j & 2nd(a) is not empty at level*j & 2nd(a) is j-pure,

(e) ∀a ∈ L^ω ∃x. a = ω-rep(x),

(f) ∀x. odd-or-even(sprig(L, x)).

Note that, despite the specialization of the arbitrary family of relations $\sim^j$ (for j less than an arbitrary ordinal μ), to the $\backsimeq^j$ for finite j, (except for the trivial relation $\backsimeq^0$), conjunct (f) is still significant. The ω+1-tuple ({0-rep(ω)} {1-rep(ω)} ... {2-rep(ω)} ... {j-rep(ω)} ... {ω-rep(ω)}) is *not* an INDEX3, since its sprig for ω is neither odd nor even. The definition does not exclude all such unbounded ω+1-tuples, however: ({} {1-rep(1)} ... {2-rep(2)} ... {j-rep(j)} ... {}) is an INDEX3, since by the Increasing Strictness Lemma, its sprig for any object is of length either zero or one.

20.2 Excess Urelements

Previous uses of the Urelement Bijection Axiom have ignored urelements not used as indexes for new sets. If the class of unused urelements were equinumerous to the universe, this would cause problems with the Axiom of Generalized Frege Cardinals. The Frege 1-cardinal of the empty set is a set containing the empty set plus all urelements, and is well-founded. If this set can be mapped onto the universe, Well-Founded Replacement would then require the existence of the Russell Set, leading to a contradiction.

So rather than the mapping ϒ given by the Urelement Bijection Axiom, I will employ a mapping ϒ'' based on ϒ, but which is one-one from the class of INDEX3's onto the class of urelements. (Thus ϒ'' will also satisfy the ϒ' Injection Requirement.)

A Cantor-Schroeder-Bernstein-Dedekind construction will give a mapping ϒ'' from the class INDEX3 one-one onto the class of urelements. Note that the required definition for ϒ'' need merely be a particular definable formula; there is no need for a set mapping.

Since the class of all sets can be injected into the class INDEX3 (e.g., by the mapping from ξ to ({0-rep(Ø)} ... {ω-rep(ξ)})), the mapping defined in the standard proof of the Cantor-Schroeder-Bernstein-Dedekind Theorem, e.g., [Levy 1979] p. 85, gives a class mapping from the class INDEX3 one-one onto the class of all sets. The composition of this with the original bijection ϒ (from the sets one-one onto the urelements) gives the required bijection **ϒ''** from the class INDEX3 one-one onto the class of urelements.

Redefinition of *: Let ***** henceforth be an abbreviation for ϒ''; it will normally be used with parentheses omitted, as before. As noted above, I am reusing this terminology (as well as j-rep and rank); it is now being used in a more specific sense than in the more general proofs. For convenience, Υ'' will be abbreviated to $\boldsymbol{\Upsilon}$, since the original Υ will not be used again.

20.3 Definition of $\in_3$ & Interpretations of the Axioms of CUSι

Define x $\in_3$ y $\equiv_{df}$
(a) $\exists$L. y = *L & INDEX3(L) & odd(sprig(L, x))
$\vee$
(b) x $\in_0$ y.

As with $\in_1$ and $\in_2$, I will adopt the convention that a formula with subscript "$_3$" represents the formula with $\in_3$ substituted for the base theory's membership relation. As before, for convenience, "altered" will be redefined in terms of $\in_3$.

Discussion. As in Part II, the domain of the interpretation is the same as that of the ground model. Church's use of Compactness is again unnecessary, since the entire sequence of relations $\backsimeq^j$ is used, not merely a finite subsequence.

Since we are now assuming Foundation in the base theory, the sets$_0$ of the ground model will turn out to be definable as the low$_3$ sets$_3$ in the interpretation: after proving the Cardinal Injection Observation (20.1), the Ill-Foundedness Requirements, and the Unaltered Domain Lemma (20.3), below, it would not be hard to show, in the presence of a global well-ordering, that the altered objects are the non-low$_3$ sets$_3$. Observe that since the collection of old sets is definable in the interpretation, then so is the old membership relation, as the two relations differ only in that some old urelements are new sets.

It would be straightforward to alter this construction to use Church's j-equivalence (abbr: $\approx_j$) instead of $\backsimeq^j$. Chapter 7 of [Sheridan 1989] sketches a proof that any Church j-equivalence class is the union of a low number of j-isomorphism classes. For any two constructions using such related relation sequences, there is a natural embedding from the model with the looser relation into the model with the stricter. The embedding moves only the altered objects; substitute for each j-rep (in the sense of the looser relation) in the j^{th} component of the associated ω+1-tuple, the low (by hypothesis) collection of j-reps (in the sense of the stricter relation) whose second components bear the looser relation to the original j-rep's second component.

Provided that both relation sequences are absolute, the image of the embedded model is definable in the model with the stricter relation sequence: It is the unaltered objects, plus the altered objects which correspond to the combination of looser equivalence classes defined by the corresponding looser ω+1-tuple. E.g., define, for this section only, $\approx$j-rep(x) as the representative of the $\approx^j$ equivalence class of x, and INDEX4 and $\in_4$ as the INDEX predicate and membership relation, defined analogously to j-rep(x), INDEX3, and $\in_3$, but in terms of Church's j-equivalence in place of my j-isomorphism. Then x is in the image of the embedding of the looser model ($\in_4$) in the stricter ($\in_3$), iff it is either low$_3$ or there is an ω+1-tuple N which satisfies the requirements for INDEX4, such that membership in x (in terms of $\in_3$) satisfies the requirements specified by N in terms of $\approx^j$.

Somewhat more formally, this predicate is: $low_3(x) \vee \exists N.\ INDEX4(N)$ & $\forall z.\ z \in_3 x \equiv odd_3(\ \{ \langle j, \approx j\text{–rep}(z) \rangle \mid j \leq \mu$ & $\approx j\text{–rep}(z) \in_3 N^j \}_3\ )$. Note that the meaningfulness of this predicate depends heavily on the absoluteness of, among others, ≈j–rep and membership in low_3 $sets_3$.

20.3.1 Organization of the Verification of the Interpretations of the Axioms of CUSι

Verifying that the interpretation $\in_3$ satisfies the axioms of CUSι, which constitutes the rest of the body of the paper, will be organized as follows:

(1) ∈† Lemma: Verify that $\in_3$ satisfies the requirements for an ∈†-interpretation (§ I.9.1), i.e., the form of the definition, the Ill-Foundedness Requirements, and the ϒ' Injection Requirement. This, by the Basic Axioms Theorem (I.9.1), will establish the Basic Axioms except Extensionality.

(2) Verify that $\in_3$ satisfies the various assumptions of Part II. This will establish the Axioms of Extensionality (by Theorem II.17.5) and Symmetric Difference (by Theorem II.16.4). These assumptions are:

- (2.1) The definitions of INDEX3 and $\in_3$ are of the required form, with clause (d) of the former satisfying a strengthened requirement.
- (2.2) INDEX3 satisfies the Degeneracy/Diversity Properties (II.14.1).
- (2.3) Required Properties of + (II.10.2).
- (2.4) $\approx^j$ and j-rep satisfy the $\sim^j$ Requirements from section II.11. This was established in the $\approx^j$ Requirements Lemma (III.19.15), above.

(3) Verify the interpretation of the Unrestricted Axiom of Pairwise Union.

(4) Prove the j-Pure j-Isomorphism Absoluteness Lemma.

(5) This lemma, plus the Equivalence Class Observation (II.15.2), yields the interpretation of the Axiom of Generalized Frege Cardinals.

20.4 ∈† Lemma

I will show in the following that $\in_3$ (along with ϒ'') satisfies the requirements for an ∈†-interpretation. An immediate corollary will be, by the Basic Axioms Theorem (I.9.1), that $\in_3$ satisfies the Basic Axioms except Extensionality. The requirements on the membership relation for the Basic Axioms Theorem are (a) that the relation be defined in a certain form, which is true by inspection, (b) the ϒ' Injection Requirement, which is true for ϒ'', as noted in its construction, and (c) Ill-Foundedness Requirements (1)–(3), the proofs of which are after the following two results.

Observation 20.1 (Cardinal Injection Observation)**.** Given an altered set x, we can inject the sufficiently large (i.e. $infinite_0$ and larger than the transitive $closure_0$ of the index of x) Cantor $cardinals_0$ into the $members_3$ of x.

Note that the sense of Cantor cardinality used here is that of the Base Theory; the Unaltered Domain Lemma (20.3), below, will lessen this difficulty. Note also that the injection constructed here is a formula in the base theory, not necessarily a function in the interpretation.

Proof. Let x be an altered object; then membership in it must be determined by clause (2) of the definition of $\in_3$. So let L be its index. Let χ be an arbitrary infinite Cantor cardinal_0 larger than the transitive closure_0 of L. The Replacing at Level*j Construction will give a set $b(\chi)$, such that if ξ is a distinct infinite Cantor cardinal_0 larger_0 than the transitive closure_0 of L, then $b(\xi) \neq b(\chi)$, and $b(\chi)$ will be a member_3 of x. (Note that this notation differs from that in the construction; here the cardinal χ is displayed.)

Similarly to Corollary II.17.4 (Nonemptiness_2 Lemma), let j be the first ordinal $< \omega$ such that L^j is not empty_0, and let g be a member_0 of L^j. Let a be the second component of the j-rep g; by the Replacing at Level*j construction, since a is not empty_0 at level*j (by the definition of INDEX3) we have a $b(\chi)$ $\Bumpeq^j_0$ a. Since $b(\chi)$ has a member_0 at level j of cardinality χ, by the Level <j Equinumerosity Lemma it cannot be k-isomorphic_0, for any k with $j < k \leq \omega$, to any of the second components of the k-reps in L^k. Since $b(\chi) \Bumpeq^j_0$ a, by $\sim^j$ Requirement (ε), <j, j-rep($b(\chi)$)> will be equal to <j, a> which equals g; hence it is in L^j, hence also in sprig(L, $b(\chi)$). Since $b(\chi)$ is not k-isomorphic to any second component of any member of L^k for $k > j$, no <k, k-rep($b(\chi)$)> can be in L^k; thus there are no further elements of sprig(L, $b(\chi)$), which is therefore odd. So $b(\chi) \in_3$ x. By the Cardinality Replacing Observation, $b(\chi)$ is an injection. □

Corollary:

Observation 20.2 (Absolute Pairs Observation)**.** "x = {y, z}" and "x = <y, z>" (unordered and Kuratowski ordered pairs) are both absolute.

This result will be frequently used without comment.
Corollary:

Lemma 20.3 (Unaltered Domain Lemma)**.** (1) Any function_3 whose domain_3 is unaltered, is unaltered. (2) If a function_3 is unaltered, its domain_3 is unaltered. (3) If a function_3 is unaltered, its range_3 is unaltered.

Proof of (1): Assume, for the sake of a contradiction, maps_3(f, a, b), where f is altered and a is unaltered. Since by the Cardinal Injection Observation we can inject the sufficiently large Cantor cardinals_0 into f, this will also provide an injection_0 of all sufficiently large cardinals_0 into_0 a, giving a contradiction.

Let $\phi(\chi)$ be the injection into f given by the Cardinal Injection Observation. So by the first conjunct of the definition of "maps," for arbitrary sufficiently large χ, $\exists x \in_3 a\ \exists y \in_3 b.\ \phi(\chi) = \langle x, y\rangle_3$. By the Absolute Pairs Observation, the subscript "$_3$" to "<...>" may be omitted. By the second conjunct of the definition of "maps," this x is unique; it is a member of a. Therefore the mapping $\psi(\chi)$ (in the base theory) which takes χ to the unique x such that $\phi(\chi) = \langle x, y\rangle \in_3$ f, is an injection of the sufficiently large Cantor cardinals$_0$ into the members$_3$ of a. But a is unaltered, so (since the definition of $\in_3$ is an expression in the language of the base theory) $\psi(\chi)$ is an injection (in the base theory) of the sufficiently large Cantor cardinals into the members$_0$ of a, which is well-founded in the base theory, contradiction. ☐

Proofs of (2) and (3). These are similar but simpler. The mapping which takes a member$_3$ of f to its first (respectively second) component is a (class) mapping from f onto its domain$_3$ (respectively range$_3$). If either the domain or range were altered, this would provide a mapping from a set in the base theory onto a class as large as the class of all sufficiently large cardinals, contradiction. ☐

20.4.1 Verification of the Ill-Foundedness Requirements for $\in_3$

(1): $\forall$x. altered(x) $\to$ ill-founded$_3$(x)

Proof. Let x be an altered object; then membership in it must be determined by clause (2) of the definition of $\in_3$. So let L be its index; as in the Cardinal Injection Observation, let j be the first ordinal $< \omega$ such that L^j is not empty$_0$, and let g be a member$_0$ of L^j. Let a be the second component of the j-rep g; chose an arbitrary infinite Cantor cardinal$_0$ χ larger than the transitive closure$_0$ of L. By the Replacing at Level*j construction, since a is not empty at level*j there is a b $\eqsim^j{}_0$ a, with x $\in^{j+2}{}_0$ b and χ(x) $\in^j{}_0$ b. As before, sprig(L, b) has only the single member <j, j-rep(b)>, so b $\in_3$ x.

We have that x $\in^{j+2}{}_0$ b; so there is a finite sequence f (in the sense of $\in_0$) such that $\forall$i $\in$ j+2. f‘i+1 $\in$ f‘i, with f‘0 = b and f‘j+2 = x. Each f‘i (except x) is unaltered; so the range of f, which is a non-empty set in the base theory and hence also unaltered, is an unending chain in the sense of $\in_3$, since b $\in_3$ x. ☐

(2): $\forall$x$\forall$y. ill-founded$_3$(x) & x $\subseteq_3$ y $\to$ ill-founded$_3$(y)

Proof. Let x be ill-founded$_3$, with x belonging$_3$ to an unending-chain$_3$ c, and let x be a subset$_3$ of y. Show that y belongs$_3$ to an unending chain$_3$ e. (If c is unaltered, the result would be trivial, since c $\cup$ {y} in the sense of $\in_0$ would exist and be unaltered, and would be the required unending chain. But the following proof covers this case as well.) Working in the base theory, which has a strong form of Choice, and hence proves Dependent Choices, we have an ω-sequence d, starting with x, such that, for all i $\in \omega$, d‘i+1 $\in_3$ d‘i. (This use of Dependent Choices in the base theory, on something which might only be an unending chain in the sense of the interpretation, may seem odd; but it is legitimate, since $\in_3$ is an expression in the language of the base theory.) By Replacement in the base theory, the range of d is a set$_0$; by Sum Set in the base theory, e = range(d) $\cup_0$ {y} is also a set, and unaltered.

Claim: e is an unending chain$_3$. Since x $\subseteq_3$ y and x has a member$_3$ in$_3$ d, y also has a member$_3$ in$_3$ d, and hence in$_3$ e; so (since every other member$_3$ of e is also a member$_3$ of d) e is an unending chain in the sense of $\in_3$, as required. ☐

(3): $\forall x \forall y.$ ill-founded$_3$(x) & $x \in_3 y \rightarrow$ ill-founded$_3$(y)

Proof. Construct e as in the preceding case. The only difference is that, in this case, y's member$_3$ in$_3$ d (and hence e) is simply x. □

Corollary:

Corollary 20.4 (Basic Axioms $\in_3$ Corollary)**.** $\in_3$ satisfies the Basic Axioms except Extensionality.

20.5 Extensionality, Symmetric Difference, and the Application of Part II

This section verifies (as specified in §III.20.3.1) that $\in_3$ satisfies the assumptions in Part II required for the Symmetric Difference$_2$ Theorem (II.16.4) and the Interpretation of the Axiom of Extensionality for Sets (II.17.5). The requirements for the applicability of these results are as follows:

(I) The definition of INDEX3 (section III.20.1) is of the form required by the definition of $\in_2$ (II.15), with clause (d) of the definition of INDEX3 satisfying an additional requirement, as noted after the Degeneracy/Diversity Properties (II.14.1).
(II) The definition of $\in_3$ (III.20.3) is of the required form (II.15).
(III) INDEX3 satisfies the Degeneracy/Diversity Properties (II.14.1).
(IV) Addition on the natural numbers satisfies the Required Properties of + (II.10.2).
(V) $\backsimeq^j$, j-rep, and rank satisfy the $\sim^j$ Requirements from section II.11. This was established in the $\backsimeq^j$ Requirements Lemma (III.19.15), above.

Proof of (I): The definitions of INDEX3 in §III.20.1 matches the form of the definition of INDEX (II.14), except that (as discussed after the Degeneracy/Diversity Properties (II.14.1)) the second conjunct of clause (d) is replaced by a stronger condition than "j-prolific(a)", i.e. "2nd(a) is not empty at level*j and 2nd(a) is j-pure." The new condition is required to imply the old; this was established by the Non-Emptiness and Prolificity Lemma (III.19.16).

Proof of (II): By inspection, the definition of $\in_3$ (III.20.3) is an instance of the definition schema for $\in_2$ in §II.15, with INDEX3 substituted for INDEX, and * tacitly replaced with its new definition.

Proof of (III): INDEX3 satisfies the Degeneracy/Diversity Properties (II.14.1).

The proof will be abbreviated, since the main points are quite similar to the proof in the more general context, proposition II.14.1. The major difference is in the proofs for clause (d) of the definition of INDEX3, where j-prolific(a) has been replaced by a stronger conjunction.

Proof of Degeneracy/Diversity Property (g): INDEX3(L) & INDEX3(M) & diverse(L, M) ⇒ INDEX3(L ∂ M).

Assume INDEX3(L) & INDEX3(M) & diverse(L,M); show INDEX3(L ∂ M).

The proof is largely as before, though clause (b) of the definition of INDEX3 is trivial in the new context. The conclusion for clause (d) is different, but the reasoning is the same: Since every member of every component of L and M have the required property (formerly j-prolific(a), now rank(a) = j & 2nd(a) is not empty at level*j & 2nd(a) is j-pure), so will every member of every component of their componentwise symmetric difference.

Proof of Degeneracy/Diversity Property (h): INDEX3(L) & low(a) ⇒ INDEX3(L^0 L^1 . . . L^n . . . [L^ω δ ω-rep"a]).

Let L be an INDEX3 and let a be a set_0. ω-rep"a exists by Replacement in the base theory, so L^ω δ ω-rep"a exists by Union and Separation. So (L^0 L^1 . . . L^n . . . [L^ω δ ω-rep"a]) exists by another application of Replacement. Verify the clauses of the definition of INDEX3 for this ω+1-tuple:

(a) is trivial. (b) is trivial in the presence of Foundation. (c) and (d) are true, since they are true by hypothesis for L, and depend only on components other than ω. (e) is true because every member of L^ω δ ω-rep"a is a member either of L^ω (for which the claim is true by hypothesis) or of ω-rep"a, for which the claim is true by definition of ω-rep"a. (f) follows by the same argument as in the original proof; the old and new sprigs can differ only in one position, the ω component.

Proof of Degeneracy/Diversity Property (i): INDEX3({0-rep(Ø)} Ø . . . Ø).

As before, set υ= ({0-rep(∅)} ∅ … ∅). Clause (a) of the definition of INDEX3 is trivial. (b) is now trivial as before. (c) is true with 0 for j and {0-rep(Ø)} for x. Similarly for (d), which expands in this context to ∀j < ω ∀a ∈ L^j . rank(a) = j & 2nd(a) is not empty at level*j & 2nd(a) is j-pure, with 0 for j and {0-rep(Ø)} for a the only instance: The first conjunct is trivial. The second is true because nothing is empty at level*0. The third is true because the empty set is pure. (e) is vacuously true. (f) is true because the sprig for any object is of size 1, since any two objects are 0-isomorphic.

Proof of (IV): Verification of the Required Properties of + (II.10.2) on finite ordinals: (Recall that in the Part II, "+" was partially specified, with only the following properties used; it is now ordinary ordinal addition. The variables α and β are restricted to ordinals; variables a, b, and x are arbitrary.) The Required Properties were:

(i) α+0= α; α+1= α ∪ {α}; α+2= (α + 1) + 1.

(ii) ∀x. ¬odd(x) ∨ ¬even(x).

(iii) Parity Property: If odd-or-even(a) and odd-or-even(b) then odd(a δ b) ⇔ [odd(a) ≢ odd(b)], and even(a δ b) ⇔ [odd(a) ≡odd(b)].

(iv) $\forall\alpha, \beta$. ordinal(α) & ordinal(β) & $\alpha < \beta \Rightarrow \alpha + 1 \leq \beta$.

Verification:

The conjuncts of (i) are now merely well-known properties of ordinal addition.

(ii) is obvious in the current context: If x is not equinumerous to a natural number, then both disjuncts are true. If x is equinumerous to a natural number, this natural number is unique and cannot be both even and odd.

(iii) is an obvious property of finite sets.

(iv) is a well-known property of ordinals. □

The preceding establishes the applicability of the following two theorem schemas to the current interpretation, and hence the Axioms of Symmetric Difference and Extensionality in the interpretation:

Theorem II.16.4 (Symmetric Difference_2 Theorem)
$\forall a \forall b \exists z \forall w.\ w \in_3 z \equiv (w \in_3 a \not\equiv w \in_3 b)$.

Theorem II.17.5 (Interpretation of the Axiom of Extensionality for Sets)
$\forall a \forall b.\ \text{nonempty}_3(a)\ \&\ \forall z.\ z \in_3 a \equiv z \in_3 b.\ \Rightarrow a = b$.

20.6 Verification of the New Axioms

I turn now to the verfications of the interpretation of the new axioms, the first of which was just proven.

20.6.1 Unrestricted Axiom of Symmetric Difference

See above.

20.6.2 Unrestricted Axiom of Pairwise Union

Theorem 20.5 (Interpretation of the Unrestricted Axiom of Pairwise Union)**.**
$\forall x \forall y \exists z \forall w.\ w \in z \equiv (w \in x \vee w \in y)$

Define r "is **odd for component j of** L" iff_{df} sprig($(L^0 \ldots L^j\ \emptyset \ldots \emptyset)$, r) is odd; abbreviated r "is *odd for* L^j." (The abbreviation is slightly misleading, since the property depends on all of L, not just L^j.)

Two further abuses of notation, for special cases, will simplify the exposition: Define r "is *odd for* L^{-1}" as always false, and r "is *odd for* $L^{\omega-1}$" as sprig($(L^0 \ldots L^j \ldots \emptyset)$, r) is odd. r "is *odd for* L^{ω}" will simply be equivalent to sprig($(L^0 \ldots L^j \ldots L^{\omega})$, r) is odd, i.e., sprig(L, r) is odd.

Proof.

Case 1: x and y are altered, with x = *L and y = *M. I will construct the union_3, *N, of *L and *M, with motivation as follows. I will construct successively each component N^j of N, for $j \leq \omega$, which will be a subset of the union of L^j and M^j, and will exist by Separation in the Base Theory. For each j-rep <j, r> in L^j or M^j, put <j, r> in N^j if it is in the segments so far of either L or J, unless r is already in the union constructed so far; or it is not in the segments so far of either L or J, but *is* in the union constructed so far. I.e., according to the following:

Subcase 1: r is odd for either L^j or M^j: Then put <j, r> in N^j iff r is not already odd for N^{j-1}.

Subcase 2: r is odd for neither L^j nor M^j: Then put <j, r> in N^j iff r is already odd for N^{j-1}.

Formally, N is an $\omega+1$-tuple with $N^j =_{df} \{ <j, r> \in L^j \cup_0 M^j \mid (r \text{ is odd for } L^j \vee r \text{ is odd for } M^j) \not\equiv r \text{ is odd for } N^{j-1} \}$.

For the 0^{th} component, this simplifies to $N^0 =_{df} \{ <0, \emptyset> \}$ if either L^0 or M^0 is nonempty, $\emptyset$ otherwise.

For N^ω, by the special case in the notation, the last clause of the definition of N abbreviates sprig((N^0 ... N^j ... $\emptyset$), r) is odd, i.e.,

$N^\omega =_{df} \{ <\omega, r> \in L^\omega \cup_0 M^\omega \mid (r \text{ is odd for } L^\omega \vee r \text{ is odd for } M^\omega) \not\equiv \text{sprig}((N^0 \ldots N^j \ldots \emptyset), r) \text{ is odd} \}$. Substituting the definition of $\in_3$ twice, we have

$N^\omega =_{df} \{ <\omega, r> \in_0 L^\omega \cup_0 M^\omega \mid (r \in_3 {}^*L \vee r \in_3 {}^*M) \not\equiv \text{sprig}((N^0 \ldots N^j \ldots \emptyset), r) \text{ is odd} \}$.

Claim: N is an INDEX3.

Clause (a) of the definition of INDEX3 follows from the form of the definition of N.

Clause (b) is superfluous with Foundation in the base theory.

Clause (c): Let j be the first for which either L^j or M^j is non-empty. N^j will be nonempty as well. (In fact it will be $L^j \cup_0 M^j$.)

Clause (d) is true for N because it is true for both L and M, and everything in N^j is in L^j or M^j.

Similarly clause (e) is true because everything in N^ω is in L^ω or M^ω.

Clause (f) is true because, for any x, $\text{sprig}(N, x) \subseteq_0 \text{sprig}(L, x) \cup_0 \text{sprig}(M, x)$.

Claim: $\forall x.\ x \in_3 {}^*N \Leftrightarrow x \in_3 {}^*L \vee x \in_3 {}^*M$.

Part 1: Assume $x \in_3 {}^*L$; show $x \in_3 {}^*N$. (The proof is the same for the case where $x \in_3 {}^*M$.)

So by the definition of $\in_3$, sprig((L^0 ... L^j ... L^ω), x) is odd. Show sprig((N^0 ... N^j ... N^ω), x) is odd.

Informally, consider the last components of sprig(L, x) or sprig(M, x). sprig(L, x), at least, will have one or more components. There will be a last such component for sprig(L, x) by clause (f) of the definition of $\in_3$, and also for sprig(M, x) if it is nonempty. Hence sprig(N, x) will also have a last component if nonempty, which will be at most the maximum of the last component indexes of L and M, by the construction of N.

Formally, let <i, s> be the last component of sprig(L, x); this exists, because sprig(L, x) is odd. Let <k, t> be the last component of sprig(M, x), if the sprig is not empty.

Consider <q, v>, where q is the maximum of i and (if it exists) k, and <q, v> is q-rep(x). N^q will be

{ <q, r> $\in L^q \cup M^q$ | (r is odd for $L^q \vee$ r is odd for M^q) $\not\equiv$ r is odd for N^{q-1} }.

Subcase 1: <q, v> $\in_0 N^q$. Then (v is odd for $L^q \vee$ v is odd for M^q) $\not\equiv$ v is odd for N^{q-1}. Since x $\in_3$ *L, sprig(L, x) is odd.

Note that x $\Bumpeq^q$ v, since <q, v> is q-rep(x). Hence x is odd for $L^q \equiv$ v is odd for L^q, and likewise for M and N.

By the maximality of q, sprig(L, x)'s being odd implies that x is odd for L^q; hence v is also odd for L^q. Thus by the definition of N^q, v is not odd for N^{q-1}, hence it is even. So, since <q, v> $\in_0 N^q$, v is odd for N^q, and hence x is odd for N^q. Thus by the maximality of q, sprig(N, x) is odd. So x $\in_3$ *N as required.

Subcase 2: <q, v> $\notin_0 N^q$. This is largely the dual of the preceding subcase. <q, v> $\in L^q \cup M^q$, since <q, v> was chosen from either sprig(L, x) or sprig(M, x). Thus by the definition of N^q, and since <q, v> $\notin_0 N^q$, (v is odd for $L^q \vee$ v is odd for M^q) $\equiv$ v is odd for N^{q-1}. As before, sprig(L, x)'s being odd implies that x is odd for L^q, hence v is also odd for L^q; so v is odd for N^{q-1}. Since <q, v> $\notin_0 N^q$, v is odd for N^q as well. As before, x is thus odd for N^q and sprig(N, x) is odd. Thus x $\in_3$ *N.

Part 2: Assume x $\in_3$ *N; show x $\in_3$ *L $\vee$ x $\in_3$ *M. By the assumption, sprig((N^0 ... N^j N^ω), x) is odd.

As before, let <i, s> be the last component of sprig(L, x), if any; let <k, t> be the last component of sprig(M, x), if any. Since sprig((N^0 ... N^j N^ω), x) is nonempty, then either sprig(L, x) or sprig(M, x) must be non-empty. So either <i, s> or <k, t> exists; let q be the maximum of i and/or k, and with <q, v> equal to q-rep(x). As before, x $\Bumpeq^q$ v, so x is odd for $L^q \equiv$ v is odd for L^q, and likewise for M and N.

Subcase 1: <q, v> $\in_0 N^q$. Then by the definition of N^q, (v is odd for $L^q \vee$ v is odd for M^q) $\not\equiv$ v is odd for N^{q-1}. Since sprig((N^0 ... N^q ... N^ω), x) is odd, by the maximality of q, x is odd for N^q; thus v is odd for N^q. Since <q, v> $\in_0 N^q$, and <q, v> $\in_0$ sprig((N^0 ... N^q Ø ... Ø), v), then v is not odd for N^{q-1}. Thus v is odd for $L^q \vee$ v is odd for M^q, by the definition of N^q. Since x is odd for $L^q \equiv$ v is odd for L^q, and likewise for M, we have x is odd for $L^q \vee$ x is odd for M^q. Thus by the maximality of q, sprig(L, x) is odd or sprig(M, x) is odd; i.e., x $\in_3$ *L $\vee$ x $\in_3$ *M.

Subcase 2: <q, v> $\notin_0 N^q$. As before, <q, v> $\in L^q \cup M^q$, since <q, v> was chosen from either sprig(L, x) or sprig(M, x). Thus by the definition of N^q, (v is odd for $L^q \vee$ v is odd for M^q) $\equiv$ v is odd for N^{q-1}. As in the previous subcase, x is odd for N^q and hence v is odd for N^q. Since <q, v> $\notin_0 N^q$, v is also odd for

N^{q-1}. Thus v is odd for L^q ∨ v is odd for M^q; as in the previous subcase, this implies x $\in_3$ *L ∨ x $\in_3$ *M.

Case 2: x and y are unaltered; provide a z such that ∀w. w $\in_3$ z ≡ (w $\in_3$ x ∨ w $\in_3$ y). If x and y are both empty, take z as Ø, and the result is trivial. Otherwise take the z required by the axiom in the Base Theory: ∀x∀y∃z∀w. w $\in_0$ z ≡ (w $\in_0$ x ∨ w $\in_0$ y). Since either x or y is nonempty, z will be nonempty as well, hence unaltered. x and y are unaltered by hypothesis, so ∀w. w $\in_3$ z ≡ (w $\in_3$ x ∨ w $\in_3$ y), as required.

Case 3: x is altered and y is not, with x = *L. The result follows by a simplified version of Case 1.

Let N = (L^0 ... L^j ... N^ω), with N^ω = { <ω, z> $\in_0$ L^ω $\cup_0$ ω-rep“y | (z $\in_3$ *L ∨ z $\in_3$ y) ≢ sprig((N^0 ... N^j ... Ø), z) is odd }. Claim ∀z. z $\in_3$ *N ⇔ z $\in_3$ *L ∨ z $\in_3$ y. (That the definition gives the intended result may be intuitively obvious to the reader, by analogy with Case 1; the remaining portion of the proof may seem repetitive and worth skipping.) The proof will be in three parts: (1) Assume z $\in_3$ *L; show z $\in_3$ *N. (2) Assume z $\in_3$ y; show z $\in_3$ *N. (3) Assume z $\in_3$ *N; show z $\in_3$ *L ∨ z $\in_3$ y.

First observe that N is an INDEX3. Clauses (a) follows from the form of the definition of N; clause (b) is now superfluous. Clauses (c) and (d) are unaffected by the difference between L and N, which is only in the ω component. Clause (e) is true because everything in N^ω is either already in L^ω or a member_0 of ω-rep“y. Clause (f) follows from the corresponding clause for L, since the only difference between L and N is in N^ω, hence the size of any sprig can be changed by at most one.

Part 1: Assume z $\in_3$ *L; show z $\in_3$ *N.

Note that sprig(L, z) is odd by hypothesis, and that sprig(N, z) can only differ from sprig(L, z) in regard to <ω, z>.

Subcase 1: <ω, z> $\notin_0$ L^ω $\cup_0$ ω-rep“y. Thus <ω, z> $\notin_0$ N^ω, hence sprig((N^0 ... N^j ... Ø), z) is odd iff sprig((N^0 ... N^j ... N^ω), z) is odd. But z $\in_3$ *L, so sprig((L^0 ... L^j ... L^ω), z) is odd. <ω, z> $\notin_0$ L^ω, so sprig((L^0 ... L^j ... Ø), z) is odd as well. But (L^0 ... L^j ... Ø) = (N^0 ... N^j ... Ø), so sprig((N^0 ... N^j ... N^ω), z) is odd, i.e. z $\in_3$ *N.

Subcase 2: <ω, z> $\in_0$ L^ω $\cup_0$ ω-rep“y. Since z $\in_3$ *L, <ω, z> $\in_0$ N^ω iff sprig((N^0 ... N^j ... Ø), z) is not odd. So if sprig((N^0 ... N^j ... Ø), z) is not odd, then <ω, z> $\in_0$ N^ω, so sprig((N^0 ... N^j ... N^j), z) *is* odd. Conversely, if sprig((N^0 ... N^j ... Ø), z) is odd, then <ω, z> $\notin_0$ N^ω, so sprig((N^0 ... N^j ... N^ω), z) remains odd. In either case, z $\in_3$ *N.

Part 2: Assume z $\in_3$ y; show z $\in_3$ *N.

Note that, since y is unaltered, z $\in_0$ y, so <ω, z> $\in_0$ L^ω $\cup_0$ ω-rep“y. Thus <ω, z> $\in_0$ N^ω iff sprig((N^0 ... N^j ... Ø), z) is not odd, and so, as in the preceding subcase, z $\in_3$ *N.

Part 3: Assume z $\in_3$ *N; show z $\in_3$ *L $\vee$ z $\in_3$ y.

Subcase 1: <ω, z> $\notin_0$ L^ω $\cup_0$ ω-rep“y. Thus, as before, sprig((N^0 ... N^j ... Ø), z) is odd iff sprig((N^0 ... N^j ... N^ω), z) is odd, which is true by assumption. So sprig((N^0 ... N^j ... Ø), z) is odd, hence so is sprig((L^0 ... L^j ... Ø), z). But <ω, z> $\notin_0$ L^ω, so sprig((L^0 ... L^j ... L^ω), z) is odd as well; thus z $\in_3$ *L.

Subcase 2: <ω, z> $\in_0$ L^ω $\cup_0$ ω-rep“y. Thus by the definition of N^ω, <ω, z> $\in_0$ N^ω iff (z $\in_3$ *L $\vee$ z $\in_3$ y) $\not\equiv$ sprig((N^0 ... N^j ... Ø), z) is odd.

Subsubcase 1: <ω, z> $\in_0$ N^ω, so since sprig((N^0 ... N^j ... N^ω), z) is odd, sprig((N^0 ... N^j ... Ø), z) is not odd; thus z $\in_3$ *L $\vee$ z $\in_3$ y.

Subsubcase 2: <ω, z> $\notin_0$ N^ω, so dually sprig((N^0 ... N^j ... Ø), z) is odd and (z $\in_3$ *L $\vee$ z $\in_3$ y) $\equiv$ sprig((N^0 ... N^j ... Ø), z) is odd. Thus z $\in_3$ *L $\vee$ z $\in_3$ y, which concludes the proof of the interpretation of the Unrestricted Axiom of Pairwise Union. □

20.6.3 Purity Lemmata

For j $\in_0$ ω, define **j-unaltered**(x) iff_{df} ∀i≤j ∀u. u $\in^i_0$ x ≡ u $\in^i_3$ x. (Note that subscripts on the implicit “∈” in “≤” are unnecessary, since ordinals are unaltered.) Similarly to j-purity, the predicate 0-unaltered is vacuously true and will not be used; being 1-unaltered is equivalent to being unaltered.

Lemma 20.6 (j-Purity Chain Lemma)**.** If x is j-pure_0, then any membership_0 chain from x of length ≤ j is also a membership_3 chain, and conversely.

The result will obviously imply the following:

Corollary 20.7 (j-Unaltered j-Purity Corollary)**.** If x is j-pure_0, it is j-unaltered.

Proof of Lemma.

Part 1: Show that a membership_0 chain is also a membership_3 chain.

Assume x is j-pure_0 for 1 ≤ j < ω and i ≤ j and maps_0(f, i+1, c) & f‘0 = x & f‘i = u & ∀k$\in_0$i. f‘k+1 $\in_0$ f‘k. (The case for j=0 is vacuous.) Show that f is a membership_3 chain, i.e., maps_3(f, i+1, c) & f‘0 = x & f‘i = u & ∀k$\in_3$i. f‘k+1 $\in_3$ f‘k.

Note that the membership relation implicit in the notation “f‘” does not need to have $\in_0$ distinguished from $\in_3$, since f is nonempty_0 and hence unaltered, as is each of its members_0 (which are ordered pairs) and each of its members’$_0$ members_0 (which are unordered pairs or singletons). Note also that i and i+1 are also unaltered, because they are ordinals, and c is unaltered, by the Unaltered Domain Lemma.

Thus the first conjunct of the demonstrandum follows from the corresponding conjunct of the hypothesis, since the definition of “maps” depends only on equality of ordered pairs and membership in f, i+1, and c. The next two conjuncts follow from f’s being unaltered.

Show ∀k$\in_3$i. f‘k+1 $\in_3$ f‘k. Assume not; take m minimal_0 with m $\in_0$ i and f‘m+1 $\notin_3$ f‘m. i is unaltered, so m $\in_3$ i and (by hypothesis) f‘m+1 $\in_0$ f‘m. Thus f'm is altered, hence an urelement. But by hypothesis and the Function Subset Assumption, f restricted to m+1 is a membership$_0$ chain; it is of length less than j and ends in f'm, contradicting the j-purity$_0$ of x.

Part 2: Show that a membership$_3$ chain is also a membership$_0$ chain.

Assume x is j-pure$_0$ and maps$_3$(f, i+1, c) & f'0 = x & f'i = u & $\forall$k$\in_3$i. f'k+1 $\in_3$ f'k. Show that f is a membership$_0$ chain, i.e., maps$_0$(f, i+1, c) & f'0 = x & f'i = u & $\forall$k$\in_0$i. f'k+1 $\in_0$ f'k.

By the Unaltered Domain Lemma (1 & 3), f is unaltered and hence so is c. i+1 is of course unaltered, as it is an ordinal. Thus as before, the first three conjuncts of the demonstrandum follow from the corresponding conjuncts in the hypothesis.

Similarly to before, for the sake of a contradiction take m minimal with m $\in_0$i and f'm+1 $\notin_0$ f'm. Thus f'm is altered, hence an urelement. By the minimality of m, $\forall$k$\in_0$m. f'k+1 $\in_0$ f'k. Thus as before, f restricted to m+1 is a membership chain, contradicting the j-purity$_0$ of x. □

Lemma 20.8 (Ill-Founded Level j Lemma). $\forall$j$\in\omega$. ill-founded$_3$(x) & x $\in^j{}_3$ y → ill-founded$_3$(y)

Proof. Apply recursion to Ill-Foundedness Requirement (3), which was: $\forall$x$\forall$y. ill-founded$_1$(x) & x $\in_1$ y → ill-founded$_1$(y). The case j=1 is just Ill-Foundedness Requirement (3) itself. Assume true for j, show for j+1: i.e., show ill-founded$_3$(x) & x $\in^{j+1}{}_3$ y → ill-founded$_3$(y). So assume x $\in^{j+1}{}_3$ y and ill-founded$_3$(x); show ill-founded$_3$(y). x $\in^{j+1}{}_3$ y expands to $\exists$f. $\exists$c. maps(f, j+2, c) & f'0 = y & f'j+1 = x & $\forall$k$\in$j+1. f'k+1 $\in_3$ f'k. Substituting j for k, f'j+1 = x $\in_3$ f'j $\in^j{}_3$ y. By Ill-Foundedness Requirement (3), f'j will be ill-founded$_3$, since x is; so by the induction hypothesis, y must be ill-founded$_3$ as well. □

Lemma 20.9 (Well-Founded Purity Lemma). If b is well-founded$_3$, it is j-pure$_0$ for any j.

Proof. Assume x is not j-pure$_0$; show that it is ill-founded$_3$. So it has a member$_0$ at less than level j which is an altered urelement$_0$. Let i be the first such j, with u an urelement such that u $\in^i{}_0$ x. I.e., $\exists$f. $\exists$c. maps(f, i+1, c) & f'0 = x & f'i = u & $\forall$k$\in$i. f'k+1 $\in_0$ f'k. By the minimality of i, each of the f'k for k<j is unaltered, so u $\in^i{}_3$x as well. By Ill-Foundedness Requirement (1), u is ill-founded$_3$. So by the preceding, x is also ill-founded$_3$. □

Theorem 20.10 (j-Pure j-Isomorphism Absoluteness Theorem). $\forall$a, b. j-pure$_0$(a) ⇒ a $\simeq^j{}_0$ b ≡ a $\simeq^j{}_3$ b.

Proof. The consequent is trivially true for j = 0 or j = ω. Note that $\Xi^j{}_0$a is a set$_0$ by the Cumulative Union Lemma and is unaltered. By the j-Unaltered j-Purity Corollary (20.7), a is j-unaltered, and $\Xi^j{}_0$a is $\Xi^j{}_3$a by the j-Purity Chain Lemma (20.6).

Part 1: Assume j-pure$_0$(a) & F: a $\Bumpeq^{j}_{3}$ b; show F: a $\Bumpeq^{j}_{0}$ b.

Since the domain$_3$ of F, Ξ^{j}_{3}a (which equals Ξ^{j}_{0}a), is unaltered, by the Unaltered Domain Lemma (1), F itself is unaltered. So by the Unaltered Domain Lemma (3), Ξ^{j}_{3}b is unaltered as well.

Claim: $\Xi^{j}_{3}b = \Xi^{j}_{0}b$. Assume not; so by Extensionality and my definition of a class's coinciding with a set, $\exists i<j\ \exists x \in^{i}_{3} b.\ x \notin^{i}_{0} b$. (The other direction is impossible, since $\in_0$ is a subrelation of $\in_3$.) Choose i minimal; so b has a member$_3$, u, at level i-1 which is altered. By the Cardinal Injection Observation (20.1), the sufficiently large Cantor cardinals can be injected into the class of members$_3$ of u, and hence into Ξ^{j}_{3}b. But F is unaltered and has domain Ξ^{j}_{0}a, which is an obvious contradiction in the base theory.

Thus clauses (1) through (3) of the definition of j-isomorphic$_0$ for F, a, and b are true by inspection. (For clause (2), note that the definition of maps$_{1\text{-}1}$(f, x, y) only depends on membership in f, x, or y, and on equality of ordered pairs in f.)

To establish clause (4), let $y \in^{i}_{0} a$ for $i < j$. Show F'$_0$y = F''$_0$y.

Since a is j-unaltered, $y \in^{i}_{3} a$; and y is unaltered since $y \in^{i}_{0} a$.

F and its ordered pair members are unaltered, so it is only necessary to subscript explicitly the right side of the demonstrandum, so the equation may be expanded as F'y = F''$_0$y = $\{z \mid \exists x \in_0 y.\ z = F'x\ \}_0$.

Note that Replacement in the base theory shows the existence of F''$_0$y, which is (by my definition of "{...|...}") a set, unaltered, and unique.

By the definition of j-isomorphic$_3$, F'y = F''$_3$y = $\{z \mid \exists x \in_3 y.\ z = F'x\ \}_3$.

Since y is unaltered, this is $\{z \mid \exists x \in_0 y.\ z = F'x\ \}_3$; but this must be just F''$_0$y: By my definition of class abstract, the abstract corresponds to a set if some set exists which is coextensive with the class formula, and that set is unique. But in the Base Theory F''$_0$y is a set$_0$, unaltered and unique, so F'$_0$y = F''$_0$y as required.

Part 2: Assume j-pure$_0$(a) & F: a $\Bumpeq^{j}_{0}$ b; show F: a $\Bumpeq^{j}_{3}$ b.

Clauses (1) through (3) of the definition of j-isomorphic$_3$ for F are as in part (1).

Similarly to part (1), to establish clause (4), let $y \in^{i}_{3} a$ for $i < j$. Show F'y = F''$_3$y = $\{z \mid \exists x \in_3 y.\ z = F'x\ \}_3$. By the definition of j-isomorphic$_0$, F'y = F''$_0$y = $\{z \mid \exists x \in_0 y.\ z = F'x\ \}_0$. y is unaltered since a is j-unaltered, so $\{z \mid \exists x \in_0 y.\ z = F'x\ \}_0$ is $\{z \mid \exists x \in_3 y.\ z = F'x\ \}_0$, which is unaltered and hence coextensive to the required $\{z \mid \exists x \in_3 y.\ z = F'x\ \}_3$. By Extensionality in the Base Theory, nothing else is coextensive$_0$ to F''$_0$y. Uniqueness in the interpretation follows from the Cardinal Injection Observation (20.1). □

Corollary, by the Well-Founded Equivalence Relation Theorem (19.3), the j-Pure j-Isomorphism Absoluteness Theorem (20.10), and the Well-Founded Purity Lemma (20.9):

Remark 20.11 (Well-Founded$_3$ Equivalence Relation$_3$ Remark)**.** $\Bumpeq^{j}_{3}$ is a restricted equivalence relation on the well-founded$_3$ sets$_3$.

The following slightly stronger result follows from the above, but I will not use it; it serves only to justify the use of Frege's name and the use of the word "equivalence": If a is well-founded$_3$, then (i) $a \approx^{j}_{3} a$, (ii) $a \approx^{j}_{3} b \equiv b \approx^{j}_{3} a$, (iii) $a \approx^{j}_{3} b \ \ \& \ \ b \approx^{j}_{3} c \rightarrow a \approx^{j}_{3} c$.

Lemma 20.12 (j-Isomorphism j-Purity Lemma)**.** $\forall j{\in}\omega\ \forall a\, \forall b.$ j-pure$_0$(a) & $a \approx^{j}_{0} b \Rightarrow$ j-pure$_0$(b).

I.e., something to which something j-pure$_0$ is j-isomorphic$_0$, is also j-pure$_0$.

Proof. Assume that j-pure$_0$(a) and $F\colon a \approx^{j}_{0} b$; assume for the sake of a contradiction that $\neg$ j-pure$_0$(b). Then $\exists y {\in}^{<j}{}_{0} b.$ urelement$_0$(y), so $\exists i{<}j.\ y{\in}^{i}{}_{0} b$.

Since F is 1-1 by clause (2) of the definition of j-isomorphic, by the j-Isomorphism/Level j Lemma, $F^{\leftarrow}y \in^{i}{}_{0} a$. By clause (4), since $i<j$, $F`F^{\leftarrow}y = y = F``F^{\leftarrow}y$. By my definition of ``, the only way $F``F^{\leftarrow}y$ can be the urelement$_0$ y is if $F^{\leftarrow}y = y$. But this contradicts the j-purity$_0$ of a. ☐

20.6.4 Frege Cardinals and the Singleton Function

Theorem 20.13 (Interpretation of the Restricted Axiom of Generalized Frege Cardinals)**.** $\forall j \in_3 \omega\ \forall b.\ wf_3(b) \Rightarrow \exists F\ \forall x.\ x \in_3 F \equiv b \approx^{j}_{3} x$.

Proof. Let $j \in \omega$, and let b be well-founded$_3$, hence j-pure$_0$. Let R be (Ø ... {j-rep(b)} ... Ø), i.e., R is an $\omega{+}1$-sequence, with R^i empty for $i \neq j$, and R^j = {j-rep(b)}. In the non-degenerate case, the required Frege Cardinal will be *R. Let c be 2nd(j-rep(b)). By the definition of j-rep, $b \approx^{j}_{0} c$; so c is j-pure$_0$ by the j-Isomorphism j-Purity Lemma.

Case 1: c is not empty at level*j

Claim 1: R is an INDEX3. Clause (a) of the definition of INDEX3 is trivial, since R was chosen as an $\omega{+}1$-tuple. Clause (b) is automatic in the presence of Foundation. (c) is true with j as the witness. (d) is vacuously true except for j, in which case the first conjunct is true by the definition of j-rep, the second conjunct is simply the hypothesis for this case, and the third conjunct is that c is j-pure$_0$, which was noted above. Clause (e) is vacuously true by choice of R. Clause (f) is trivial since R has only one component.

Claim 2: The required Frege Cardinal is *R. I.e., $\forall x.\ x \in_3 {}^*R \equiv b \approx^{j}_{3} x$.

Subclaim (2a): Assume $x \in_3 {}^*R$; show $b \approx^{j}_{3} x$. By the definition of $\in_3$, $\forall x.$ $x \in_3 {}^*R$ iff odd(sprig(R, x), since *R is an urelement. Since R has only one non-empty component, the right-hand side implies j-rep(x) $\in R^j$ = {j-rep(b)}, i.e., j-rep(x) = j-rep(b).

So by the definition of j-rep, c = 2nd(j-rep(b)) = 2nd(j-rep(x)), and also by the definition of j-rep, $x \approx^{j}_{0}$ 2nd(j-rep(x)). So $x \approx^{j}_{0} c$ and $b \approx^{j}_{0} c$.

Thus by the symmetry and transitivity clauses of the Well-Founded Equivalence Relation Theorem (19.3), this implies b $\Bumpeq^{j}_{0}$ x. Since b is j-pure$_0$, by the j-Pure j-Isomorphism Absoluteness Theorem b $\Bumpeq^{j}_{3}$ x, as required.

Subclaim (2b): Assume b $\Bumpeq^{j}_{3}$ x; show x $\in_3$ *R.

Since b $\Bumpeq^{j}_{3}$ x, and b is j-pure$_0$, by the j-Pure j-Isomorphism Absoluteness Theorem, b $\Bumpeq^{j}_{0}$ x. Since sprig(R, x) can have at most one element, to show that it is odd, it will suffice to show that j-rep(x) $\in$ R^j = {j-rep(b)}, i.e., that j-rep(x) = j-rep(b).

As before, let c be 2nd(j-rep(b)). Thus it suffices to show that 2nd(j-rep(x)) = c. By the definition of j-rep, c is minimal in the global well-ordering such that b $\Bumpeq^{j}_{0}$ c.

Likewise, 2nd(j-rep(x)) is the minimal d in the global well-ordering such that x $\Bumpeq^{j}_{0}$ d.

We have b $\Bumpeq^{j}_{0}$ x & b $\Bumpeq^{j}_{0}$ c & x $\Bumpeq^{j}_{0}$ d. By symmetry and transitivity for $\Bumpeq^{j}_{0}$, we have x $\Bumpeq^{j}_{0}$ c & b $\Bumpeq^{j}_{0}$ d. But d is minimal such that x $\Bumpeq^{j}_{0}$ d, so d precedes, or is equal to, c. Similarly, c is minimal such that b $\Bumpeq^{j}_{0}$ c, and b $\Bumpeq^{j}_{0}$ d; so c precedes or is equal to d. Thus c = d, as required.

Case 2: c is empty at level*j. By the j-Empty J-Isomorphism Lemma (19.18), no other object is j-isomorphic to b. So the required equivalence class will be the singleton {b}, which exists by the Pair Set Axiom in the base theory, and is unaltered. □

Corollary:

Theorem 20.14 (Singleton Function Theorem)**.** The Singleton Function is a Set$_3$.

Proof. Let **ι** = *(Ø Ø {2-rep(<Ø, {Ø}>)} ... Ø). Claim: ∀x. x $\in_3$ ι ≡ ∃d. b = <d, {d}>.

As in Case 1 of the proof of the Restricted Axiom of Generalized Frege Cardinals, above, since <Ø, {Ø}> is 2-pure$_0$ and not empty$_0$ at level*2, we have ∀x. x $\in_3$ ι ≡ <Ø, {Ø}> $\Bumpeq^{2}_{3}$ x. By the j-Pure j-Isomorphism Absoluteness Theorem, since <Ø, {Ø}> is 2-pure$_0$, <Ø, {Ø}> $\Bumpeq^{2}_{3}$ x iff <Ø, {Ø}> $\Bumpeq^{2}_{0}$ x. By the Singleton Function/2-Isomorphism Theorem (19.6), <Ø, {Ø}> $\Bumpeq^{2}_{0}$ x ≡ ∃d. x = <d, {d}>. Thus ∀x. x $\in_3$ ι ≡ ∃d. x = <d, {d}>, as required. □

This concludes the proof of the relative consistency of CUSι, Q.E.D.

21 Conclusion and Future Work

The construction technique pioneered by Church and followed by Mitchell and myself suffices to rebut naive anti-Platonist arguments against the universal set and Frege-Russell cardinals, but in the long run it seems to be a dead end. Forster's Potemkin Village criticism fairly argues that the technique will not suffice for serious theories, and it is hardly clear that a serious set theory with a universal set must have consistency strength easily comparable to a theory based on the cumulative hierarchy. The approach seems an even worse dead end in terms of manpower; all three consistency proofs involve large amounts of unrewarding complexity without concomitant aesthetic or theoretical benefits.

The paradox involving my partially-specified theory CUS$\iota\#$ seems less profound: merely an instance of the obvious (in retrospect) point that while natural equivalence relations may have equivalence classes which are sets, a relation which can code enough information about the membership relation to emulate the Russell Paradox cannot.

The recent work by neo-Fregeans is to some extent a divergent method of rescuing Frege: Positing representatives for equinumerosity classes, rather than defining them as sets, suffices for much of arithmetic. This presumably would have been considerable consolation to Frege, who seemed willing to abandon set theory with a universal set as a foundation for mathematics, once an inconsistency was found.[24] But I like to think that he would have appreciated the benefit of honest toil in showing that something like his set theory could define Frege cardinals while avoiding the paradoxes.

To those considering doing further research in the field, I would advise against re-traversing Church's, Mitchell's, and my paths. Oberschelp's theory may repay verification and further investigation; perhaps his theory can place the singleton function on a firmer footing than my efforts. Constructions which alter the equality relation, such as Malitz's, and Church's abandoned construction, may allow theories of greater complexity to have their relative consistency proved. My concluding advice echoes and extends Gödel's and Malitz's: What is more important than relative consistency proofs is applying Platonistic intuition to develop new theories with new axioms.

[24] See, e.g., at the end of his career, "A New Attempt at a Foundation for Arithmetic," reprinted in *Posthumous Writings* pp. 278–281, in which he bases mathematics on the complex numbers and geometry rather than the natural numbers and sets.

Part IV
Appendices

Acknowledgments

Written in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Oxford. I am grateful to my late grandfather Joseph Gillis, my mother Nancyann Sheridan, my wife Olga Miroshnychenko, a United Kingdom Overseas Research Student Award, United Kingdom Alvey grant IKBS-047, Palm, Inc., and Bell Labs for support during the writing and rewriting of my dissertation. I would like to thank Thomas Forster, Randall Holmes, and Roger Janeway for comments on earlier versions of this paper.

In addition to the bibliography, I provide two appendices: an annotated listing of articles citing [Church 1974a], and of apparently relevant items in the Church Archives at Princeton. The latter is currently based primarily on the archive catalog, pending a change in the archives policy.

A Bibliography

Peter Aczel 1985. *Lecture Notes on Non-Well-Founded Sets,* Stanford Center for the Study of Language and Information, Lecture Notes #14, ISBN 0-937073-22-9.

Joseph Almog 2008. "Frege Puzzles?", *Journal of Philosophical Logic,* volume 37, number 6, pp. 549-574, DOI: 10.1007/s10992-008-9080-8.

C. Anthony Anderson 1998. "Alonzo Church's Contributions to Philosophy and Intensional Logic," *Bulletin of Symbolic Logic,* volume 4, number 2, June 1998.

Paul Benacerraf and Hilary Putnam (editors) 1983. *Philosophy of Mathematics: Selected Readings,* second edition, Cambridge University Press, ISBN: 978-0521296489.

George Boolos 1987. "Saving Frege from Contradiction," *Proceedings of the Aristotelian Society,* volume 87, pp. 137-151, http://www.jstor.org/pss/4545060.

George Boolos 1990. "The Standard of Equality of Numbers," reprinted in [Demopoulos 1995].

John Burgess 2005. *Fixing Frege*, Princeton University Press, ISBN: 978-0691122311.

Alonzo Church 1974a. "Set Theory with a Universal Set," *Proceedings of the Tarski Symposium,* Proceedings of Symposia in Pure Mathematics XXV, ed. Leon Henkin, American Mathematical Society, ISBN: 978-0821814253, pp. 297-308. (Delivered 24 June 1971.)

Alonzo Church 1974b. “Notes on a Relative Consistency Proof of Axioms A–K of Church’s Set Theory with a Universal Set,” unpublished lecture notes, Church Archives, box 47, Folder 5.

William Demopoulos, editor, *Frege’s Philosophy of Mathematics,* Harvard University Press 1995.

Werner Depauli-Schimanovich 2008. “ZFK: = ZFC with a Complement, or: Hegel And The Synto-Set-Theory,” http://arxiv.org/abs/0807.4938v1, citing his *Extension der Mengenlehre,* doctoral dissertation, Universität Wien, 1971.

Sir Michael Anthony Eardley Dummett 1991. *Frege: Philosophy of Mathematics,* Harvard University Press.

Herbert B. Enderton 2008(?). “Introduction: Alonzo Church: Life and work,” written for *Collected Works of Alonzo Church,* MIT Press, pp. xiii-xix. Preprint, creation date 06/26/08: http://www.math.ucla.edu/~hbe/church.pdf.[25]

Solomon Feferman 2006. “Enriched stratified systems for the foundations of category theory,” in *What is Category Theory?* (G. Sica, ed.) Polimetrica, Milano (2006), 185-203.

Thomas E. Forster 1983. *Quine’s New Foundations (An Introduction),* Cahiers du Centre de Logique, Université Catholique de Louvain, Institut Supérieur de Philosophie.

Thomas E. Forster 1985. “Status of the Axiom of Choice in Set Theory with a Universal Set”, *Journal of Symbolic Logic* v. L pp. 701-7.

Thomas E. Forster 1992. *Set Theory with a Universal Set: Exploring an Untyped Universe*, Oxford Logic Guides #20, Clarendon Press, ISBN 0-19-853395-0.

Thomas E. Forster 1995. *Set Theory with a Universal Set: Exploring an Untyped Universe,* second edition, Oxford Logic Guides #31, ISBN 978-0198514770.[26]

Thomas E. Forster 2001. “Church’s Set Theory with a Universal Set” in: *Logic, Meaning and Computation: Essays in Memory of Alonzo Church,* Synthese library 305, Kluwer. Corrected version at http://www.dpmms.cam.ac.uk/~tf/church2001.pdf.[27]

Thomas E. Forster 2006. “Permutations and Wellfoundedness: the True Meaning of the Bizarre Arithmetic of Quine’s NF,” *Journal of Symbolic Logic,* volume 71, March 2006, pp. 227-240.

Thomas Forster 2008, “The Iterative Conception of Set,” *Review of Symbolic Logic*, I pp 97-110.

Thomas Forster and Thierry Libert 2011. “An Order-Theoretic Account of Some Set-Theoretical Paradoxes,” *Notre Dame Journal of Formal Logic* volume 52, number 1, pp. 1-19.

[25]The publication of this book has apparently been blocked by Michael Zeleny: http://cs.nyu.edu/pipermail/fom/2006-September/010790.html.

[26]The second edition contains less detail.

[27]But note that I am a great-grand-student of Church’s (via Turing and Gandy), not a student as claimed on the first page.

Abraham Fraenkel, Yehoshua Bar-Hillel, and Azriel Lévy, *Foundations of Set Theory,* second edition, Elsevier 1973.

Gottlob Frege 1884, *Die Grundlagen der Arithmetik*, Koebner.

Gottlob Frege 1893 & 1903. *Grundgesetze der Arithmetik,* Georg Olms Verlagsbuchhandlung, Hildesheim.

Ulf Friedrichsdorf 1979. "Zur Mengenlehre über Klassen," *Zeitschrift f. Mathem. Logik* 25, pp. 379-383. Contains helpful summary of [Oberschelp 1964a & 1973].

Kurt Gödel, "What is Cantor's Continuum Problem," 1964, reprinted in [Benacerraf & Putnam 1983].

Theodore Hailperin 1944. "A set of axioms for logic," *Journal of Symbolic Logic* 9, pp. 1-19.

Michael Hallett 1984. *Cantorian Set Theory and Limitation of Size,* Oxford Logic Guides 10, Oxford: Clarendon Press

Richard G. Heck Jr. 2013, *Reading Frege's Grundgesetze*, Oxford University Press.

M. Randall Holmes 1998. *Elementary Set Theory with a Universal Set*, originally volume 10 of the Cahiers du Centre de Logique, Academia, Louvain-la-Neuve, Belgium; corrected online version at http://math.boisestate.edu/~holmes/holmes/head.ps.

M. Randall Holmes 2001 "Could ZFC be inconsistent?," unpublished, http://math.boisestate.edu/~holmes/holmes/sigma1slides.ps.

Andrej Kisielewicz 1989. "Double Extension Set Theory," Reports on Mathematical Logic 23, PWN-Warsaw, pp. 81-89, [Mathematical Reviews 92e:03084].

Azriel Levy 1969. "The Definability of Cardinal Numbers." In Jack J. Bulloff, Thomas C. Holyoke, S.W. Hahn, eds., *Foundations of Mathematics: Symposium Papers Commemorating the Sixtieth Birthday of Kurt Gödel*, Springer-Verlag.

Azriel Levy 1979. *Basic Set Theory.* **Ω** Perspectives in Mathematical Logic, Springer-Verlag.

Penelope Maddy 2011. *Defending the Axioms: On the Philosophical Foundations of Set Theory*, Oxford University Press 2011.

[Isaac] Richard Jay Malitz 1976, *Set Theory in which the Axiom of Foundation Fails,* unpublished doctoral thesis, University of California at Los Angeles, supervised by Alonzo Church.[28]

Dmitry Mirimanoff 1917b. "Remarques sur la Théorie des Ensembles et les Antinomies Cantoriennes—I," *L'Enseignement Mathématique* 19, pp. 208-217.

Emerson Mitchell 1976. *A Model of Set Theory with a Universal Set,* unpublished Ph.D. thesis, University of Wisconsin at Madison.

[28]Note that no complete copy of this thesis seems to be available, some copies purportedly of this thesis, e.g., the copy sent to me by Church, are early drafts, others are incomplete, and even the latest draft refers to itself as non-final.

Emerson Mitchell 1977. Abstract of [Mitchell 1976] in *Journal of Symbolic Logic,* March 1977, volume 42, number 1, p. 148, http://www.jstor.org/stable/2272332?seq=6.

Arnold Oberschelp 1964a. "Eigentliche Klasse als Urelemente in der Mengenlehre," *Mathematische Annalen* 157 pp. 234-260. [Mathematical Reviews 31#2136]. (Delivered 20 August 1962)

Arnold Oberschelp 1964b. "Sets and Non-Sets in Set Theory" (abstract), received 3 June 1964, *Journal of Symbolic Logic* XXIX p. 227

Arnold Oberschelp 1973. *Set Theory over Classes,* Dissertationes Mathematicæ (Rozprawy Mat.) 106. [Mathematical Reviews 42 #8300].[29]

Charles Parsons 1977. "What is the Iterative Concept of Set?" reprinted in [Benacerraf & Putnam 1983].

Sir Roger Penrose, "Response to Tony Dodd's 'Godel, Penrose and the Possibility of AI,'" *Artificial Intelligence Review* volume 5, number 4 (1991) p. 235.

Willard van Orman Quine 1937a. "New Foundations for Mathematical Logic," American Mathematical Monthly 44 pp. 70-80; reprinted in *From a Logical Point of View,* Harper & Row 1953.

Willard van Orman Quine 1937b. "On Cantor's Theorem," *Journal of Symbolic Logic* v. II, pp. 120-24

Willard van Orman Quine 1969. *Set Theory and Its Logic*, revised edition, Harvard: Belknap Press.

Dana Scott 1955. "Definitions by Abstraction" (abstract), *Bulletin of the American Mathematical Society* 61, p. 442.

Flash Sheridan 1982, "Against Half-Heartedly Platonistic Set Theory," senior essay, Yale University.

Flash Sheridan 1985, qualifying thesis, Institute of Mathematics, University of Oxford, examined by Robin Gandy and Sir Michael Dummett.

Flash Sheridan 1989. "The Singleton Function is a Set in a Mild Extension of Church's Set Theory," doctoral thesis, supervised by Robin Gandy, University of Oxford, resubmitted 2013.

Flash Sheridan 1989b. "Interpretations of ZF Restricted to Sets Equinumerous to a Well-Founded Set," unpublished.

Flash Sheridan 1990. "The Natural Generalization of Church's Set Theory with a Universal Set is Oberschelp's," unpublished.

Flash Sheridan 1991. "Churchish Oddities," unpublished.

Flash Sheridan 2005 (1980), "A Naive View of the Russell Paradox," unpublished.

Joseph R. Shoenfield 1967. *Mathematical Logic*, Addison-Wesley, Reading Massachusetts.

Helen Skala 1974. "An Alternative Way of Avoiding the Set-Theoretical Paradoxes," *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 20, pp. 233-237.

[29] Note that the crucial part of the consistency proof in both [Friedrichsdorf 1979] and [Oberschelp 1973] is merely a reference to [Oberschelp 1964a], which uses a significantly different formalism.

Ernst P. Specker 1953. "The Axiom of Choice in Quine's New Foundations for Mathematical Logic," *Proceedings of the National Academy of Sciences* 39, pp. 972-975.

Alan Weir 2003. "Neo-Fregeanism: An Embarrassment of Riches," *Notre Dame Journal of Formal Logic* volume 44, Number 1 (2003), pp. 13-48, http://projecteuclid.org/euclid.ndjfl/1082637613.

Reprintings of [Church 1974a]

International Logic Review 15, pp. 11–23, according to http://math.boisestate.edu/~holmes/holmes/setbiblio.html and http://www.dpmms.cam.ac.uk/~tf/finsler.ps.[30]

Philosophy of Quine 5: Logic, Modality, and Philosophy of Mathematics, pp. 215–226, edited by Dagfinn Føllesdal, Taylor & Francis, 2001, ISBN 0-8153-3742-6.

The Collected Works of Alonzo Church, MIT Press, if it ever appears, which seems uncertain.

B Works Incidentally Citing Church's Paper

The only subsequent writings substantially relying on Church's set theory seem to be [Mitchell 1976], my own work, and Thomas Forster's, chiefly the accounts in the two editions of Oxford Logic Guide ([1992] and [1995], with the first edition providing more detail), and [2001]. I list below other apparently relevant articles (excluding those by Forster and Mitchell) which are listed by Google Scholar as citing [Church 1974a],[31] or which I happen to have encountered and which mention Church's set theory or technique. Citation format and completeness varies dependent on the information available from Google. Where practical I append a brief summary or quotation of the apparently relevant passage(s). Articles on modal or paraconsistent logic are grouped together at the end.

C. Anthony Anderson 1987. Review of Bealer's *Quality and Concept, Journal of Philosophical Logic,* volume 16, issue 2, pp. 115-64. Review of a work on intensional logic.

Carsten Butz 2003. "Bernays–Gödel type theory," *Journal of Pure and Applied Algebra,* Volume 178, Issue 1, 15 February 2003, pp. 1-23. Preprint at http://euclid.math.mcgill.ca/butz/preprints/axiomatics.ps.gz.

> On category theory; also cites Aczel 1988. "our efforts to understand and apply the techniques and results of [8] [A. Joyal and I. Moerdijk, Algebraic Set Theory, Cambridge University Press, Cambridge 1995] to other set-theories, like for example Aczel's set-theory of non-well-founded

[30] I have been unable to verify this independently.

[31] Article URL: http://scholar.google.com/scholar?cluster=5427335978717460142; citations URL http://scholar.google.com/scholar?cites=5427335978717460142, as of 11 March 2011.

sets [1], or Church's set-theory with a universal set [4]." (But the paper is nonetheless on conventional Gödel-Bernays, not ill-founded set theory.)

M. Clavelli, M. Forti, and V.M. Tortorelli 1988. "A self-reference oriented theory for the foundations of mathematics," *Analyse Mathématique et Applications, 67-115.* No information available online.

Pui Kuen Fong 2008. "Privacy preservation for training datasets in database: application to decision tree learning," Masters thesis, University of Victoria, British Columbia, http://hdl.handle.net/1828/1291.

Computer science paper with references to multiple universal sets; it's not clear that this is an actual use of a universal set rather than merely a domain of discourse.

Ennio De Giorgi 1998. "Dal superamento del riduzionismo insiemistico alla ricerca di una più ampia e profonda comprensione tra matematici e studiosi di altre discipline scientifiche ed umanistiche" ("Overcoming set-theoretic reductionism in search of wider and deeper mutual understanding between mathematicians and scholars of different scientific and human disciplines."), *Rend. Mat. Acc. Lincei* s. 9, v. 9:71-80.

An "open-ended, non-reductionist axiomatic framework, grounded on the primitive notions of quality and relation." Lists Church but does not cite him, beyond a general allusion, "Tutte le teorie citate nella bibliografia..." ("All the theories mentioned in the bibliography...") after a discussion of V∈V, p. 78.

Michael Warren Jamieson 1994. *Set Theory with a Countable Universe,* unpublished doctoral thesis, University of Florida

From an email to the New Foundations mailing list, with LaTeX commands in the original, 16 August 1994:

(i) It satisfies the comprehension axiom

\exists y\forall x[xEy iff \phi(x)]

for any "strictly quantifier-free" formula (with parameters), i.e. any quantifier-free formula such that no atomic subformula of \phi contains a repeated variable. (So xEx and x=x are not permissible subformulas, and Russell's paradox is avoided.) Even though x=x does not have a comprehension axiom in this scheme, the model does have a universal set u via comprehension for the formula "xEp \vee \not(xEp)" for any parameter p.

(ii) The model satisfies the extensionality axiom.

(iii) The model satisfies closure under power set formation.

(iv) Though the model doesn't include a substructure rich enough to satisfy ZF, it does contain an element \omega whose elements (under the relation E) are isomorphic to the standard natural numbers.

(v) The model contains an element f such that the sentence "f maps the universe 1-1 into \omega" is satisfied.

Reinhard Kahle 2011, "The Universal Set and Diagonalization in Frege Structures," *The Review of Symbolic Logic,* 4, pp 205-218 doi:10.1017/ S1755020310000407. "*unrestricted* comprehension and a *partial* element-of relation.' Also cites Oberschelp and Aczel.

Martin Kühnrich 1986. "Untersuchungen zur Friedmanschen Theorie der Prädikate," *Mathematical Logic Quarterly,* volume 32 pp. 29 - 44, DOI: 10.1002/malq.19860320105. Apparently on a hierarchy of predicates.

John Lake 1973. "On an Ackermann-Type Set Theory," *The Journal of Symbolic Logic,* Vol. 38, No. 3 (December 1973), pp. 410-412.

Paper on conventional large cardinals. Notes on page 411 that a Hilbert-Infinite-Hotel-style interpretation resembles Church's construction. Paper received June 28, 1972; cites Church's paper as appearing in 1973, not 1974, and with slightly inaccurate page numbers.

John Lake 1979. "The approaches to set theory," *Notre Dame Journal of Formal Logic*, 20 (2):415-437.

"Another approach to set theories with a universal set has been made by Church [8]. Here the motivation is that the abstraction principle is desirable but (unfortunately?) it turns out to be inconsistent so that we must investigate all (formalistic) ways of approximating to it whilst remaining within the realms of consistency or, at least, relative consistency. This view also seems to be an assumption for the book by Frankel, Bar-Hillel and Levy [10]. We have little sympathy with such ideas as there does not seem to be any clear reason why we should have believed the abstraction principle in the first place."

Agustın Rayo and Timothy Williamson 2003. "A Completeness Theorem for Unrestricted First-Order Languages," chapter 15, *Liars and Heaps: New Essays on Paradox.*

"But one might be tempted to address the problem by adopting a set theory which allows for a universal set—Quine's New Foundations, the Church and Mitchell systems and positive set-theory all satisfy this requirement. * Unfortunately, set theories that allow for a universal set must impose restrictions on the axiom of separation to avoid paradox. So, as long as an MT-interpretation assigns a subset of its domain as the interpretation of a monadic predicate, some intuitive interpretations for monadic predicates will not be realized by any MT-interpretation.

*"Throughout the rest of the paper we will be working with ZFU plus choice principles, rather than a set theory which allows for a universal set."

Keith Simmons 2000. "Sets, Classes and Extensions: a Singularity Approach to Russell's paradox," *Philosophical Studies,* 2000 100: 109–149, http://www.springerlink.com/content/u78462464r4l6405/, DOI 10.1023/-A:1018666804035.

Dismisses Quine, Church, and Mitchell for their failure "to respect the intuition that every predicate has an extension," p. 112. About extensions of concepts rather than sets: "We resolve the Russell paradox as we

resolved our simple paradox: the term 'extension' is a context-sensitive expression that is minimally restricted on any occasion of use," p. 138.

Tim Storer 2010. *A Defence of Predicativism as a Philosophy of Mathematics,* doctoral thesis, University of Cambridge, http://www.dspace.cam.ac.uk/handle/1810/226320.

Parenthetical discussion of the universal set, citing Forster and Church: "In fact, no paradox results merely from the assumption that there is a universal set", in the context of indefinite extensibility.

Achille C. Varzi 2006. "The Universe Among Other Things," *Ratio*, 19: 107–120. doi: 10.1111/j.1467-9329.2006.00312.x.

A rebuttal to Simons, P. M., 2003, 'The Universe', *Ratio* 16: 237–250, largely on ontology and mereology, but refutes an argument against the objecthood of the universe with an account from [Forster 1995] citing Quine and Church.

"As for the second reason, it is certainly correct that under standard assumptions about the existence of sets there is no such thing as the universal set. And neither is the very idea of a universal set a popular one. This is not to say, however, that it is formally incoherent or otherwise unworkable. Examples of non-standard set theories in which the universe is a set among others can already be found in the works of Quine and Church,... and today the topic is gaining interest among logicians and mathematicians alike."

Paraconsistent and Modal Logics

NCA da Costa, D Krause, O Bueno 2004. "Paraconsistent logics and paraconsistency: Technical and philosophical developments," preprint: http://www.cfh.ufsc.br/~dkrause/pg/papers/CosKraBue2004.pdf

Brief comment about using Church's theory, apparently inessentially, p. 45 & 43. Reference to "A system of this type was already studied in" da Costa, N. C. A., 'On a set theory suggested by Ehresmann and Dedecker,' *Proceedings of the Japan Academy of Sciences* 45, 1969, pp. 880-888.

Other paraconsistent works citing Church by these authors:

NCA da Costa 1996, "Théories paraconsistantes des ensembles*," Logique et Analyse.*

NCA da Costa and O. Bueno. 2001. "Paraconsistency: towards a tentative interpretation" *Theoria* 16: 119-45.

"A classical set theory of the ZF kind with universal class was developed in Church (1974); in da Costa (1986) [da Costa, N.C.A.: 1986, 'On Paraconsistent Set Theory', Logique et Analyse 115, 361-371.] This was extended to a paraconsistent set theory."

NCA da Costa, D Krause, O Bueno 2007. "Paraconsistent Logics and Paraconsistency," *Philosophy of Logic.* Very similar to the above.

Modal Logic

M Oksanen 1999. "Russell-Kaplan paradox and other modal paradoxes: a new solution", *Nordic Journal of Philosophical Logic.*

Brief mention of Church and Forster; also mentions [Skala 1974], which suggests a lack of seriousness. "Indeed, the existence of a class of all classes with seventeen members is also consistent with the set theories of Church and Skala, though it does not follow from them as it follows from NFU." Obviously incorrect about Church. (He may well be right about Skala.)

C Alonzo Church Archives at Princeton

Below I list the papers in the Alonzo Church Archives (at the Department of Rare Books and Special Collections, Princeton University Library) which seemed from the catalog to be relevant to his set theory with a universal set, and his attempts to unify it with Quine's New Foundations. I have been unable to obtain some of these papers, pending procurement of scanning equipment by the archives. Its catalog is available at http://arks.princeton.edu/ark:/88435/fx719m49m; the archives were processed by Sylvia Yu and Laura Hildago, Princeton Class of 2006, in 2004, and the online catalog contains some obvious errors. (In particular, the capitalization of "box" and "Folder" are systematically incorrect, which I have sometimes followed, and "subseries" is repeatedly misspelled.) The catalog also lists a bibliography of Church's works compiled by Erin Zhu, December 22, 1993 (63 pp. box 9 folder 29), which I have not obtained. Italicized comments are my own.

C.1 Subseries 3D: Set Theory

Box 45, Apparently Relevant Folders

box 45, Folder 1 & 2: Notebook: Dec. 31 1970-Sept. 1971, Notebook: Set theory old notes, August-October 1971. *Possibly working notes for Church 1974a & b.*

box 45, Folder 5: Notebook: December 1973-July 1974, includes Summer 1974 Notes on Extensions of Set Theory

box 45, Folder 7 & 8: Notebook, part 1: Towards combining the Quine set theory with the basic axioms, June 1975. Notebook, part 2: Towards combining the Quine set theory with the basic axioms, June 1975-August 1976.

Box 46, Apparently Relevant Folders

box 46, Folder 2: Notebook: September 1975 "Sets of the Model Transfinitely Generated"; Notes prepared for the Illinois lectures, Notebook: September 1975 and some associated correspondence box 46, Folder 3: Notebook: December 1975-July 1976 "Continuing the typed and dittoed notes "Outline and Background Material, Arthur B. Coble Memorial Lectures"

Box 47, Apparently Relevant Folders

box 47, Folder 2: Notebook: Recursion clauses for $\mathrm{inv}^{\mathrm{m}}$ as revised Sept. 1980, Proof of the Main Lemma, cont.

Assumed relevant because of "inv^m*" (box 15, Folder 10), and secondarily "Proof of the Main Lemma" (note to box 15, Folder 11 below).*

box 47, Folder 5: Lecture notes, fall 1974 (Set Theory with a Universal Set). Church's note: "Probably not of much value - but possibly worth some reflection." Presumably [Church 1974b].

box 47, Folder 6: Set Theory material and notes: 1975-January 1983

box 47, Folder 7: Set Theory Notes as of June 1983: "incomplete set of originals"

box 47, Folder 8: Set Theory Notes 1981-June 1983: "Third copy, nearly complete"

box 47, Folder 9: Set Theory notes, September 1983: "To go back to L.A. January 1984" and "Left in G.B. Jan. 1984: Second and third copies of these set-theory notes, plus assorted left-over pages, plus some older set theory notes"

box 47, Folder 10: "Notes as to Set Theory with a Universal Set. Check completeness. Brought back from G.B. [Grand Bahama] 1989. These notes are old [1971] but might be reconsidered for the sake of ***some truth in it,*** which might guide a new approach." (photocopies); Ajdukiewicz's Paradox of the Name Relation.

I have only briefly been able to examine this document, and cannot obtain a copy from the archives, as their policy forbids making copies of copies. Page 1 is dated July 1971; it begins by stating that "As even the amended model of April 1971, ... is not yet satisfactory, we make a new start using the outline of June 1971." It seems to be an eventually-abandoned attempt at another consistency proof for the full CUS. It is roughly fifty unnumbered pages; the mathematics is quite complicated, and apparently not final: there is a Case 22 which has an amendment and a second amendment. The paper contains a definition of well-foundedness in terms of a predicate called "retrogressive," which is similar to my concept of unending chain. I provide Church's definition after my own, above. The photocopy, if not the manuscript, ends abruptly in the middle of a separate Case 8. It is followed by the Ajdukiewicz paper noted in the archives listing, and a paper entitled "Revised λ-δ-calculus," which is not.

(The Ajdukiewicz paper seems unrelated: "This is the now well-known problem of the failure of substitutivity of identity in intensional contexts." "Alonzo Church's Contributions to Philosophy and Intensional Logic," C. Anthony Anderson, The Bulletin of Symbolic Logic, volume 4, number 2, June 1998, p. 167 footnote 88.)

C.2 Other Subseries

Suberies [sic] 1D: Published Papers, 1966-1973

box 4, Folder 6 (Church 1974a): "Set Theory with a Universal Set" in Proceedings of Symposia in Pure Mathematics, vol. XXV, Proceedings of the Tarski Symposium (an international symposium held to honor Alfred Tarski on the occasion of his seventieth birthday), edited by Leon Henkin, John Addison, William Craig, C. C. Chang, Dana Scott, and Robert Vaught, American Mathematical Society, Providence, Rhode Island, 1974, pp. 297-308. 1 reprint and photocopy; proof and corrections; AMs, 23 pp., and photocopies.

Suberies [sic] 1H: Collected Works Projects and Bibliographies

box 9, Folder 20: "Set Theory with a Universal Set" (1975)

Suberies [sic] 1I: Lectures, Abstracts, Unpublished Papers, etc.

box 15, Folder 10: "Set Theory on a Universal Set" [sic] Background Material, Arthur B. Coble Memorial lectures, Sept. 23-25, 1975, Urbana, Illinois, AMs, 16 pp. of lecture notes: "Sets of the Model Transfinitely Generated" 7 pp.; "Conditions on the Relations invm," 1 p.; Introduction Equivalences of a Set a m+1," 1 p.; "Basic Axioms," 1 p.; "Set Existence," 1 p.; "Hailperin's Axioms for the Quine Set Theory," 1 p.; "Modified n-Equivalence," 1 p.; "The Recursion Order," 1 p.; "Invariance Relations Directed towards the Hailperin Axioms," 2 pp. Also 1 photocopy and typed, mimeographed copy, 18 pp.

The biggest item in the above is three versions of Church's notes for the Coble lectures:

- *Seventeen pages of handwritten notes, with an apparently later heading (apparently in Church's handwriting, but in lighter ink) "Lectures at Urbana, Ill., Sept 23–25, 1975". The notes ends in section headed "CORRECTION" which is crossed out, B15, F10, 17th page*
- *A second copy of the same notes, without the handwritten heading, but with instead a typewritten title page (the 18th page in the folder) with the incorrect title "Set Theory on a Universal Set". The CORRECTION is not crossed out in this copy (B15, F10, 34th page).*
- *A largely typewritten version of the preceding, ending in a different correction.*

The last item is the same as the mimeographed typescript which I obtained from Professor Enderton, and is apparently also in the University of California at Berkeley Logic Library, http://logic-library.berkeley.edu/catalog/detail/462.

The typescript seems to be a later version of the manuscript; the opening and ending seem the same, and p. 16 of the typescript is a hand-written table entitled "The Recursion Orders," identical to the 32nd page of the manuscript, *labeled 14. The final sections labeled "Correction" are different, however, and only the typewritten correction mentions the need for substantive change. The typewritten correction correlates p. 12 line 9 of the typewritten notes with p. 11 l. 9, and p. 17 l. 6 with p. 15 l. 6. The corrections are also added by hand (plus another correction to p. 11) to the Princeton typescript, but not the UCLA copy of Professor Enderton. The correction is not made on manuscript p. 11, and page numbering inconsistencies mean that there isn't a p. 15 of the manuscript.*

(Cp. box 46, Folder 3: 1620. Notebook: December 1975–July 1976 "Continuing the typed and dittoed notes "Outline and Background Material, Arthur B. Coble Memorial Lectures.")

box 15, Folder 11: Set Theory original manuscript, January-March 1979. AMs, 42 pp.; 1 photocopy. Includes "Hailperin's Axioms for the Quine Set Theory," 1 p.; "Analysis Directed Towards Proof of P6," 2 pp.; "The Main Lemma, 1 p.; "Lemmas Needed for the Proof of the Main Lemma," 4 pp.; "Proof of the Main Lemma (for P6) from Lemmas 1-5," 2 pp.

Page 41 does not seem to exist; the page numbered 40 ends abruptly, and page 42 begins two pages entitled "Proof of the Main Lemma (for P6) from Lemmas 1-5." Many of the other pages in this folder are unnumbered, but some are numbered, and most seem to follow the same sequence. The manuscript ends abruptly after the 43rd page, numbered 43.

This suggests the relevance of other lemmas with names of the form "P<n>", and possibly other papers with the notation "Proof of the Main Lemma," though the latter obviously need not be unique.

Subseries 3F: UCLA Courses and Miscellaneous Dated Notes

box 49, Folder 7, 2nd entry: Notes (4 pp.) for lecture in Finland, 1976. Shortened version of Urbana, Illinois lectures, "Set Theory on a Universal Set," [sic] at Abo Akademi in Turku, Finland, March 22, 1976.

box 49, Folder 8: (?) Misc. Notes (1980s), including "Corollaries of the Proof of P1, and other theorems about the model" (12 pp., Dec. 1982)

Subseries 3G: Miscellaneous Undated Notes

box 50, Folder 4: Notes (undated) titled "Proof of P1" (1 p.), "The sum set axiom" (10 pp.), and "P6" (4 pp.)

Series 5: Papers of Others

box 76, Folder 13: Sheridan, Flash. *(See also possibly [Correspondence] Sa-Si General 21 7).*

There seems to be no reference to [Mitchell 1976], except perhaps in [Correspondence] M General 19 14.

C.3 Relevance Unclear

box 15, Folder 8: "Frege on the Philosophy of Time," April 17, 1969. AMs, 11 pp.

Subseries 3D: Set Theory

box 45, Folder 3: Exercise book with notes on set theories from 1950s, Feb. 1972 addendum

box 45, Folder 6: Notebook: August-Sept. 1974; with July 1980 addendum

box 46, Folder 1: Notebook: July 1975-December 1977

box 46, Folder 4–9: Various dated but descriptionless notebooks.

box 47, Folder 1: Notebook: Set Theory Notes December 1979 and later (superseded pages preserved for reference)

box 47, Folder 3: Notebook: March-December 1981

box 47, Folder 4: Set Theory notes, Jan.-Dec. 1981 ("mostly secondary")

Subseries 3E: UCLA Courses

box 48, Folder 7: UCLA Philosophy 221B, Spring 1977: Set Theory Seminar Notes (85 pp.)

Subseries 3F: UCLA Courses and Miscellaneous Dated Notes

box 49, Folder 13–15 are probably not relevant: "Alternative (0)" probably refers to the Logic of Sense and Denotation; see folder 15 below.

box 49, Folder 13: Notes towards revision of the treatment of Alternative (0). First draft, Dec. 1974. Includes addenda dated March 1978 and August 1986.

box 49, Folder 14: Notes (July 1986). Refers to previous folder.

box 49, Folder 15: Notes (July-September 1986) from preliminary (working) notebook: "A Revised Formulation of the Logic of Sense and Denotation under Alternative (0)"

Subseries 4C: Academic Topics from Church's Files

box 53, Folder 12: Quine Set Theory

Box 53, folder 17 Set Theory (research and readings)

Subseries 4I: Publications (cont.) and Personal

box 60, Folder 1: Lecture Notes prepared in Connection with the Summer Institute on Axiomatic Set Theory, UCLA, 1967

60 2: Tarski Symposium, Berkeley, CA, June 1971 (abstracts)

box 60, Folder 2: Tarski Symposium, Berkeley, CA, June 1971

Series 5: Papers of Others

box 66, Folder 9 Frege, Gottlob (photocopies only)
box 68, Folder 9 Hailperin, Theodore
box 72, Folder 4–7: Malitz, Richard (includes correspondence)

C.4 Archive Errata

- *"Set Theory on a Universal Set" in box 49, Folder 7 and box 15, Folder 10 should be "Set Theory with a Universal Set" as in box 4 Folder 6, box 9 Folder 20, box 47 Folder 5, and box 47 Folder 10.*
- *"Suberies" in "Suberies 1A: Published Papers, 1924-1951" and eight other occurrences, through "Suberies 1I: Lectures, Abstracts, Unpublished Papers, etc." should be "Subseries".*
- *The capitalization in "box <N>, Folder <N>," passim, in unfortunate.*

Index