AXIOMATIZATION OF SET THEORY BY EXTENSIONALITY, SEPARATION, AND REDUCIBILITY

SEMINAR NOTES

by

Harvey M. Friedman
Department of Mathematics
Ohio State University
Columbus, Ohio 43210
friedman@math.ohio-state.edu
www.math.ohio-state.edu/~friedman/

October 5, 1997

We discuss several axiomatizations of set theory in first order predicate calculus with epsilon and a constant symbol W, starting with the simple system K(W) which has a strong equivalence with ZF without Foundation. The other systems correspond to various extensions of ZF by certain large cardinal hypotheses. These axiomatizations are unusually simple and uncluttered, and are highly suggestive of underlying philosophical principles that generate higher set theory.

The general philosophy is that at any time, one can attempt to conceive of the entire set theoretic universe. But according to Russell's Paradox, as one later conceives of the entire set theoretic universe, reflecting on the previous conception, one obtains a yet larger set theoretic universe. However, according to this general philosophy, nothing really new is obtained - the first set theoretic universe is, for all intents and purposes, the same as the second set theoretic universe, even though it is not literally the same. This principle takes the form of a Reducibility axiom that can be viewed as an extension of Russell's Reducibility.

We view the variables in these axiomatizations as ranging over today's set theoretic world, and the constant symbol W as representing yesterday's set theoretic world (the sub-world). In K(W), it can be proved that W is transitive - no element of yesterday's world picks up any new elements from today's world. In some of the stronger systems which generate large cardinals, W is provably non transitive - some elements of yesterday's world do pick up new elements from today's world.

This is supported by a natural weakening of separation, with a corresponding strengthening of Reducibility, which allows for the construction of a standard model of ZF + measurable cardinals and beyond. W gets interpreted as images of elementary embeddings associated with large cardinals.

The systems K(J),K*(J) give direct axiomatizations of an elementary embedding J, and are much more committal than K1(W)-K5(W). They postulate a “projection function” from today’s world into (onto) yesterday’s world, which is related to published ideas of Reinhardt. We include them because of their special simplicity.

Here is the organization of the results.

1. ZF\Foundation.
   Extensionality (EXT), Subworld Separation (SS), Reducibility (RED).
2. Indescribable and subtle cardinals.
   EXT, SS, RED+.
4. Measurable cardinals.
   EXT, SS-, RED, TRANS.
5. Elementary embeddings from V(□) into V(□).
   EXT, SS-, RED, RED’, TRANS.
6. ZF\Foundation again.
   EXT, SS-, RED, RED’’.
7. Elementary embeddings from V(□) into V(□).
   EXT, SS-, RED, RED’’, TRANS, Largeness.
8. Elementary embeddings from V into V.
   EXT, SS-, RED, RED’’, TRANS, Subworld Collection.

NOTE: SS- < SS; RED+ > RED; and RED’’> RED’.

Largeness says that every element of W can be mapped one-one into the intersection of W with an element of W. TRANS says that every transitive subset of W is a subset of a fixed element of W.

9. Axiomatic elementary embeddings: huge cardinals and V(□) into V(□).
   EXT, Separation, Elementarity, Fixed Points.
10. Axiomatic elementary embeddings: V into V.
   EXT, Replacement, Elementarity, Initial Fixed Points.

In each case, we prove a close relationship between these new systems and standard formal set theories without Choice, sometimes without Foundation. Foundation is inessential here in that the cumulative hierarchy provides a standard interpretation.

The absence of the axiom of Choice is a more delicate matter. Until we move significantly past the measurable cardinal level, we can use known inner model techniques going back to Gödel ablish the □ interpretability of the axiom of Choice. At the higher levels, in some cases we can invoke some unpublished work of Hugh Woodin that uses forcing models instead of inner models to establish □ interpretability of the axiom of Choice -- at the cost of weakening somewhat the large cardinal axioms in the forcing model. So, more or less, the absence of the axiom of Choice does not substantially weaken higher set theory with respect to □ interpretability. Also, of course, there is nothing to prevent you from simply adding the axiom of choice in any of its myriad forms to our systems. Then the standard set theories that they correspond to will simply be standard set theories with the axiom of Choice. We choose not to do this, since in doing so, there is no interaction with the rest of the axioms that we use.

We begin with a discussion of the system K(W) based on L(□,W), which is first order predicate calculus with the binary relation symbol □ and the constant symbol W. The axioms of K(W) are:

1. Extensionality (EXT). (□x_1)(x_1 □ x_2 □ x_1 □ x_3) □ (□x_1)(x_2 □ x_1 □ x_3 □ x_1).
2. Subworld Separation (SS). x_1 □ W □ (□x_2 □ W)(□x_3)(x_3 □ x_2 □ (x_3 □ x_1 & □)), where □ is a formula in L(□,W) in which x_2 is not free.
3. Reducibility (RED). (x_1,…,x_n □ W □ □) □ (□x_{n+1} □ W)(□), where n ≥ 0 and □ is a formula in L(□) whose free variables are among x_1,…,x_{n+1}.

Define x = y iff (□z)(z □ x □ z □ y). Using EXT:

LEMMA 1.1. Let □ be a formula in L(□,W) without x_2. K(W) proves: $x_1 = x_1$. $x_1 = x_2$ □ $x_2 = x_1$. ($x_1 = x_2$ & $x_2 = x_3$) □ $x_1 = x_3$. $x_1 = x_2$ □ (□ □ □[x_1/x_2]).

For any formula □ in L(□), let □^{W} be the result of relativizing all quantifiers to W.

LEMMA 1.2. Let φ be a formula in $L(\in)$ whose free variables are among $x_1,\ldots,x_n$, $n \geq 0$. The following is provable in K(W). $x_1,\ldots,x_n \in W \rightarrow (\varphi \leftrightarrow \varphi^W)$.

LEMMA 1.3. The following is provable in K(W). $(\exists x_1)(x_1 \in W)$. $x_1 \in x_2 \in W \rightarrow x_1 \in W$.

The idea is to form $y = \{b \in x_2: b \in W\} \in W$, by SS, and look at the statement $x_2 \subseteq y$. If a counterexample exists then there must a counterexample in W. This is a contradiction.

LEMMA 1.4. The following is provable in K(W). $(x_1 \in W \;\&\; x_2 \subseteq x_1) \rightarrow x_2 \in W$.

We now use an axiomatization of ZF\Foundation by extensionality, pairing, union, separation, power set, reflection, and infinity. In the presence of separation, pairing, union, power set take the form of the existence of a set containing all of the requisite elements, rather than being exact. Reflection takes the form:

$(\exists \text{ transitive } x_1)(x_2,\ldots,x_n \in x_1 \wedge (\forall x_{n+1},\ldots,x_m \in x_1)((\exists x_{m+1})(\varphi) \rightarrow (\exists x_{m+1} \in x_1)(\varphi)))$, where $m > n \geq 2$ and φ is a formula in $L(\in)$ whose free variables are among $x_2,\ldots,x_n$.

LEMMA 1.5. Pairing is provable in K(W).

LEMMA 1.6. Union is provable in K(W).

LEMMA 1.7. Separation is provable in K(W).

LEMMA 1.8. Power set is provable in K(W).

LEMMA 1.9. Reflection is provable in K(W).

LEMMA 1.10. Infinity is provable in K(W).

LEMMA 1.11. K(W) proves ZF(rfn)\Foundation.

We now want to discuss interpretations of K(W).

The most natural models of K(W) are of the form $(V(\kappa),\in,V(\lambda))$, where κ is a strongly inaccessible cardinal and $\lambda < \kappa$ is chosen with some care. Here it suffices to simply choose $\lambda < \kappa$ so that $V(\lambda)$ is an elementary submodel of $V(\kappa)$.

By adapting this construction:

LEMMA 1.12. Every sentence in L(□) that is provable in K(W) is provable in ZF(rfn)\F.

Here V(□) is replaced by V, and W is constructed via Reflection. Thus we have:

THEOREM. A sentence in L(□) is provable in K(W) iff it is provable in ZF(rfn)\F.

We now move on to K_1(W), which is connected to indescribable and subtle cardinals.

Let K_1(W) be the following theory in the language L(□,W).

1. EXT.
2. SS.
3. RED+. (x_1,…,x_n □ W & x_{n+1} □ W & □) □ (□x_{n+2},x_{n+3} □ W)(x_{n+3} □ x_{n+1} □ □[x_{n+1}/x_{n+3}]), where n ≥ 0 and □ is a formula in L(□) whose free variables are among x_1,…,x_{n+2}.

First of all, it is easy to see that RED+ implies RED by taking x_{n+1} = W. Hence K_1(W) extends K(W), and so we have ZF(rfn)\F at our disposal.

We now review the relevant large cardinals in the treatment of K_1(W).

In ZFC, let n ≥ 1 and □ be a cardinal. We say that □ is n-th order indescribable if and only if for all R □ V(□) and first order sentence □, if (V(□+n),□,R) satisfies □, then there is an □ < □ such that (V(□+n),□,R □ V(□)) satisfies □.

We say that □ is totally indescribable if and only if □ is n-th order indescribable for all n ≥ 2. This definition agrees with the one given in [Ka94], p. 59.

It is natural to go a bit further.

In ZFC, we say that □ is extremely indescribable if and only if for all R □ V(□) and first order sentence □, if (V(□+□),□,R) satisfies □, then there is an □ < □ such that (V(□+□),□,R □ V(□)) satisfies □.

The definition of extreme indescribability in ZF is identical to the definition given above in ZFC except that it is natural to simply drop the requirement that the ordinal be a cardinal. Thus we speak of extremely indescrib-able ordinals.

In ZF(rep)\F, we say that V(□) is strongly inaccessible iff if the range of every function from an element of V(□) into V(□) is an element of V(□).

LEMMA. ZF(rep)\F proves every strongly inaccessible rank satisfies ZF; and □ extremely indescribable □ V(□) is strongly inaccessible. □ an □-interpretation of ZFC + "□ an extremely indescribable cardinal" in ZF(rep)\F + "□ an extremely indescribable ordinal" using constructibility.

THEOREM. $K_1(W)$ proves: ZF(rfn)\F + "□ an extremely indescribable ordinal." □ a standard model of ZF + "□ an extremely indescribable ordinal." □ an □ model of ZFC + "□ an extremely indescrib-able cardinal."

Extremely indescribable cardinals fit into the standard large cardinal picture as quite a bit stronger than weakly compact cardinals, but quite a bit weaker than subtle or ineffable cardinals, which are still compatible with the axiom of constructibility. We now want to see what we need to build standard models of $K_1(W)$.

In ZFC, let □ be a cardinal. We say that f:□ □ S(□) is regressive if and only if for all □ < □, f(□) □ □. (Here S is the power set operation).

In ZFC, we say that □ is a subtle cardinal if and only if

i) □ is a cardinal;
ii) For all closed unbounded C □ □ and regressive f:□ □ S(□), there exists □ < □, □,□ □ C □ □, such that f(□) = f(□) □ □.

Subtle cardinals are strongly inaccessible, and in fact extremely indescribable.

THEOREM. ZFC + "there exists a subtle cardinal" proves the existence of a standard model of $K_1(W)$; i.e., of the form (V(□),□,V(□)).

Proof: Let be a subtle cardinal. We want to find < such that (V(),,V()) is a model of K_1(W). You can’t just pick any old elementary submodel V() of V() because Reducibility has been strengthened considerably to RED+. The idea is to define a map h from into V() such that for each < , h() is data sufficient for a counterexample to (V(),,V()) satisfying K_1(W); default otherwise. This setup can be seen to fit into the framework of the definition of subtle cardinal, so that one obtains an appropriate < < such that h() and h() cohere. But then one can show that in light of this coherence, the defaults must have been used, and so (V(),,V()) is seen to satisfy K_1(W).

Let K_2(W) be the following theory in L(,W).

1. EXT.
2. SS-. x_1 W (x_2 W)(x_3 W)(x_3 x_2 (x_3 x_1 &)), where is a formula in L(,W) in which x_2 is not free.
3. RED. (x_1,…,x_n W) (x_{n+1} W)(), where n ≥ 0 and is a formula in L() whose free variables are among x_1,…,x_n.
4. TRANS. (x1 W)(every transitive subset of W is a subset of x1).

Note that SS- is weaker than SS in that the quantifier x_2 is restricted to W. This is the natural formulation of Subworld Separation that accommodates the possibility that W is not transitive.

I.e., some elements of yesterday’s world may pick up new elements from today’s world.

This form of Subworld Separation asserts that every property of elements of any element of W has a representation, x_2, in W that at least represents the property with regard to elements of W. We actually prove that this representation x_2 W is unique.

Note that we can formally derive the nontransitivity of W. For suppose W is transitive. Let W x W. By SS-, let y W be such that for all z W, z x z z. Then x x x x.

We have proved that K_2(W) proves the existence of a standard model of ZF + "there is a measurable rank." And ZF + "there is a nontrivial elementary embedding from some rank into some rank" proves the existence of a standard model of K_2(W).

Measurable ranks are an appropriate formalization of measurable cardinals in the ZF context with Choice. In ZFC, these are equivalent: i.e., V(□) is a measurable rank if and only if □ is a measurable cardinal.

We say that V(□) is a measurable rank if and only if there is a complete measure m on V(□+1); i.e.,

i) m:V(□+1) □ {0,1};
ii) for all x □ V(□+1), if |x| □ 1 then m(x) = 0;
iii) for all x □ V(□+1), m(V(□)\x) = 1-m(x);
iv) if x □ V(□), f:x □ V(□+1), and for all b □ x, m(f(b)) = 0, then m(□rng(f)) = 0.

We use x = y for (□z)(z □ x □ z □ y).

LEMMA 4.1. Let □ be a formula in L(□,W) which does not mention x_2. The following is provable in K_2(W). $x_1 = x_1$. $x_1 = x_2$ □ $x_2 = x_1$. ($x_1 = x_2$ & $x_2 = x_3$) □ $x_1 = x_3$. $x_1 = x_2$ □ (□ □ □[x_1/x_2]).

For any formula □ in L(□), let $□^W$ be the result of relativizing all quantifiers to W.

LEMMA 4.2. Let □ be a formula in L(□) whose free variables are among $x_1,\ldots,x_n$, $n \geq 0$. The following is provable in K_2(W). $x_1,\ldots,x_n$ □ W □ (□ □ $□^W$).

LEMMA 4.3. K_2(W) proves Separation and Tupling.

Here Separation is as usual in set theory, except that we accommodate all formulas in our language L(□,W). Tupling is the scheme that allows us to form any $\{x_1,\ldots,x_n\}$.

To get Separation, it suffices to show that Separation holds in (W,□), because we can absorb W as a paremeter. But this is clear by SS-. So why didn't we use Separation instead of SS-? This is because SS- gives us something more: Separation in W with respect to arbitrary formulas in L(□,W).

LEMMA 4.4. Let □ be a formula in L(□,W) The following is provable in K_2(W). If x,y □ W and every common element of x and W lies in y, then x □ y. x_1 □ W □ [(□!x_2 □ W)(□x_3 □ W)(x_3 □ x_2 □ (x_3 □ x_1 & □)) & (□x_2 □ W □ x_1)(□x_3 □ W)(x_3 □ x_2 □ (x_3 □ x_1 & □))].

LEMMA 4.5. The following is provable in K_2(W). No set includes all ordinal as elements. The least ordinal, OW, that is not in W is a limit ordinal > □. There is an ordinal in W\OW.

Fix □ to be the first ordinal in W that is greater than OW. □ exists by Lemma 4.5.

LEMMA 4.6. The following is provable in K_2(W). □ is a limit ordinal. If V(□) exists then V(□) □ W □ □ □ W. For all □ < OW, V(□) □□ W. V(OW) exists and is a transitive subset of W. V(□) exists.

Let W′ = {x □ W: x □ V(□)}.

LEMMA 4.7. The following is provable in K_2(W). If x,y □ W′ and x □ V(OW) □ y, then x □ y. If x,y □ W′ and x □ V(OW) = y □ V(OW) then x = y.

LEMMA 4.8. The following is provable in K_2(W). The correspondence that sends x □ W′ to x □ V(OW) is one-one onto V(OW+1). Every element of W′\V(OW) has rank □+1. The correspondence is the identity on V(OW) and one-one from W′\V(OW) onto V(OW+1)\V(OW). This correspondence is an □-isomorphism.

We are not claiming that this correspondence exists as an actual function.

For x □ V(OW+1), we let H(x) be the y □ W′ such that x = y □ V(OW). Again, we emphasize that H does not necessarily exist as a function.

Now define m:V(OW+1) □ {0,1} by m(x) = 1 if V(OW) □ H(x); 0 otherwise.

LEMMA 4.10. The following is provable in K_2(W). V(OW) is strongly inaccessible.

LEMMA 4.11. K_2(W) proves: For all □ < OW there exists □ < □ < OW and a complete measure on some V(□+1).

THEOREM 4.12. K_2(W) proves that there is a standard model of ZF + “there exists arbitrarily large measurable ranks.”

We now wish to construct a standard model of K_2(W) using large cardinals.

In ZFC, we say that □ is an extendible cardinal if and only if for all □ > □ there is an elementary embedding from V(□) into some V(□) with critical point □.

THEOREM 4.13. ZFC + "there is an extendible cardinal" proves the existence of a standard model of K_2(W).

Here a standard model is of the form (V(□),□,A), where A □ V(□) is the interpretation of W; A cannot be a rank.

Let □ be an extendible cardinal. Let j:V(□+1) □ V(□+1) be an elementary embedding with critical point □. Then □ and □ are strongly inaccessible and that V(□) is an elementary submodel of V(□). Now let j':V(□+1) □ V(□+1) be an elementary embed-ding with critical point □. Then j'(□) = □, and since □ is strongly inaccessible, □ is also strongly inaccessible. We use j' below for j'|V(□).

Thus j' is an elementary embedding from V(□) into V(□) with critical point □. We prove that (V(□),□,rng(j')) is a model of K_2(W). Clearly this is a legitimate interpre-tation since rng(j') □ V(□), by the strong inaccessibility of □.

rng(j') is an elementary submodel of V(□) by general nonsense concerning elementary embeddings.

From this it follows that RED holds in (V(□),□,rng(j')).

Also the largest transitive subset of rng(j') is V(□). To see this, suppose A is a transitive subset of rng(j'). Then there are elements of A of every rank < rk(A). If rk(A) > □+1 then some element of A, and hence some element of rng(j'), has rank □+1. Therefore by RED, □ □ rng(j'), which contradicts that □ is the critical point of j'.

It remains only to verify SS-.

Let j'(x) be given, where x □ V(□). Let y □ j'(x). Consider j'(j'$^{-1}$[y]). Then j'(u) □ j'(j'$^{-1}$[y]) if u □ j'$^{-1}$[y] iff j'(u) □ y. Thus we have shown that for all v □ W, v □ j'(j'$^{-1}$[y]) □ v □ y.

We can actually use quite a bit less.

THEOREM. ZFC + “there exists a nontrivial elementary embedding from a rank into a rank” proves the existence of a standard model of $K_2(W)$.

The use of an extendible cardinal actually proves the standard model of an extension $K_3(W)$ of $K_2(W)$ by the additional axiom of Reducibility:

RED’’. □ □ (□x_1 □* W)(□), where □ is in L(□) and x1 is not free in □.

In $K_3(W)$, one can prove ZF(col)\F + “there exists a nontrivial elementary embedding of a rank into itself”, which Woodin has shown to imply Projective Determinacy.