

1. How NFU evades the paradoxes and what if anything this explains
2. Why we have confidence in NFU
3. The axiom of cantorian sets in NFU and n -Mahlo cardinals!

M. Randall Holmes

Boise State Math Department
Logic and Set Theory Seminar
10/29/2019 and onward

Abstract: I'll present the formal definition of the set theory NFU (New Foundations with urelements) proposed by R.B. Jensen in 1969 as a weakening of Quine's New Foundations (NF) and shown by Jensen to be consistent.

I'll present enough basic implementation of mathematics in NFU to explain why the usual arguments for the paradoxes of Russell, Cantor, and Burali-Forti do not go through in NFU.

The question then presents itself, in the light of the previous talk...in what sense and to what extent can we be taken to be thereby explaining why we can take NFU to be consistent (or NF, which avoids the paradoxes in the same way, and yet remains dubious)?

The exposition might include a discussion of why we actually do believe that NFU is consistent (Jensen did prove this) and why NF remains doubtful.

The definition of NFU

NFU is a first order unsorted theory whose primitive notions are a sethood predicate, equality, and membership. We provide in addition to the usual countable supply of variables a countable supply of variables x^i with a natural number superscript i for each $i \in \mathbb{N}$. An atomic formula will be said to be *well-typed* iff either one of the variables in it is unsuperscripted or if it is of a shape $\text{set}(x^{i+1})$, $x^i = y^i$, or $x^i \in y^{i+1}$. A formula will be well-typed iff each of its atomic subformulas is well-typed. Please note that a formula which is not well-typed is well-formed and has the same logical privileges as any other formula; we will see what this notion is for in a moment.

The axiom schemes of NFU are three:

sethood: $(\forall xy : x \in y \rightarrow \text{set}(y))$: objects with elements are sets.

extensionality:

$$(\forall xy : \text{set}(x) \wedge \text{set}(y) \wedge (\forall z : z \in x \leftrightarrow z \in y) \rightarrow x = y)$$

sets with the same elements are equal.

comprehension:

$$(\exists A : \text{set}(A) \wedge (\forall x^i : x^i \in A \leftrightarrow \phi)),$$

for any well-typed formula ϕ in which A does not appear and in which each bound variable is superscripted (note that x^i is stipulated to be superscripted by the form of the axiom): for each such formula ϕ , the set $\{x^i : \phi\}$ exists.

The original theory NF did not have a sethood predicate (and neither did NFU as originally proposed by Jensen), but it is very convenient. Notice that extensionality and comprehension together imply the existence of a unique empty set $\emptyset = \{x^i : x^i \neq x^i\}$, and that we could equally well (as has been done) taken the empty set as a primitive constant, provided an axiom asserting $(\forall x : x \notin \emptyset)$ and defined $\text{set}(x)$ as asserting $x = \emptyset \vee (\exists y : y \in x)$. We refer to empty objects which are not sets as urelements or atoms: in the original theory NF there were no atoms and the axioms were stated without mention of sethood. The presence or absence of atoms will not often make an obvious difference in our discussion (it does obtrude at one point in the discussion of Cantor's paradox).

The way in which we state comprehension is designed to minimize metamathematical machinery. Notice that as superscripted variables have no special logical status, we can assert the existence of any set $\{x : \phi\}$ for which we can replace x and each bound variable in ϕ with a superscripted variable in such a way as to make the formula well-typed. With a little thought, this can be seen to express exactly the usual criterion of “weak stratification” (weak because we do not require assignment of types to parameters), which might be thought metamathematically burdensome as it involves discussing the existence of a function from variables to natural numbers having a particular relationship to the syntax of ϕ .

Some mathematics in NFU

The existence of $\{x, y\} = \{z^0 : z^0 = x \vee z^0 = y\}$ is a theorem of NFU. Thus the ordered pair $(x, y) = \{\{x\}, \{x, y\}\}$ exists for any x, y . This can be proved to have the essential properties of an ordered pair exactly as in ordinary set theory.

We can then define functions and relations as sets of ordered pairs just as in ordinary set theory. Usual properties and classifications of relations and functions will be taken to be familiar in this talk.

Paradox the first: the paradox of Russell

It is a theorem of first order logic that

$$\neg(\exists R : (\forall x : x \in R \leftrightarrow x \notin x)).$$

$(\exists R : (\forall x : x \in R \leftrightarrow x \notin x))$ cannot be turned into an instance of the axiom of comprehension of NFU, because at the very least x must be replaced with x^i for some superscript i , and $(\exists R : (\forall x : x \in R \leftrightarrow x^i \notin x^i))$ has the would-be embedded subformula $\phi = x^i \notin x^i$ quite irretrievably ill-typed.

NFU cannot be convicted of paradox at the tribunal of Russell. The other paradoxes have more mathematical prerequisites.

Paradox the second: the Cantor paradox of the largest cardinal

The Cantor paradox of the largest cardinal goes as follows. The Cantor theorem asserts that $|A| < |\mathcal{P}(A)|$. The largest cardinal is of course the cardinality $|V|$ of the universal set. Now $\mathcal{P}(V) \subseteq V$ (there might be urelements!) so $|\mathcal{P}(V)| \leq |V|$, which contradicts the case $|V| < |\mathcal{P}(V)|$ of Cantor's theorem.

We need to do some work to set up the machinery to evaluate this argument in NFU.

We define $A^i \sim B^i$ as asserting the existence of a function f^{i+3} which is a bijection from A^i to B^i .

We define $|A^i|$ as $\{B^i : A^i \sim B^i\}$. This is the original Frege-Russell-Whitehead definition of cardinal number. We define $|A^i| \leq |B^i|$ as holding iff $(\exists C^i : C^i \subseteq B^i \wedge A^i \sim C^i)$. We define $|A^i| < |B^i|$ as holding iff $|A^i| \leq |B^i| \wedge |A^i| \neq |B^i|$.

We define $\iota^{\ulcorner A^i \urcorner}$ as $\{\{x^i\} : x^i \in A^{i+1}\}$. We define $T(|A^i|)$ as $|\iota^{\ulcorner A^i \urcorner}|$ (leaving it as an exercise to verify that this does not depend on the choice of A).

Theorem (Cantor): $T(|A|) < |\mathcal{P}(A)|$

Proof: Clearly $T(|A|) = |\iota^{\ulcorner A \urcorner}| \leq |\mathcal{P}(A)|$, since $\iota^{\ulcorner A \urcorner} \subseteq \mathcal{P}(A)$. Now suppose for the sake of a contradiction that $|\iota^{\ulcorner A \urcorner}| = |\mathcal{P}(A)|$, that is, $\iota^{\ulcorner A \urcorner} \sim \mathcal{P}(A)$, that is, there is a bijection f from $\iota^{\ulcorner A \urcorner}$ to $\mathcal{P}(A)$.

Consider the set $R = \{x^0 : x^0 \in A^1 \wedge x^0 \notin f^{\ulcorner \{x^0\}^1 \urcorner}\}$. This is a subset of A , so there is a singleton $\{r\}$ such that $f(\{r\}) = R$. Now $r \in R$ iff $r \in A \wedge r \notin f(\{r\})$, which is a contradiction, since $r \in A$ is true under the hypotheses, so we are asserting that $r \in R$ if and only if $r \notin f(\{r\}) = R$.

We provide a little explanation of the reason why R is a set.

$R = \{x^0 : x^0 \in A^1 \wedge x^0 \notin f^4(\{x^0\}^1)\}$ can be expressed at more length as

$\{x^0 : x^0 \in A^1 \wedge (\exists y^1 : x^0 \in y^1 \wedge (\exists z^1 : (y^1, z^1) \in f^4 \wedge y^1 \notin z^1))\}$

The subformula $(y^1, z^1) \in f^4$ can be expanded to

$(\exists w^3 \in f^4 : (\forall u^2 : u^2 \in w^3 \leftrightarrow (\forall v^1 : v^1 \in u^2 \leftrightarrow v^1 = y^1) \vee (\forall v^1 : v^1 \in u^2 \leftrightarrow v^1 = y^1 \vee v^1 = z^1)))$

And this only works because we have replaced A^1 itself with ιA^1 in the statement of the theorem and its proof. The general idea is that we can define sets and reason in general in NFU in the ways warranted in the typed theory of sets, and not much more.

The theorem we have proved is $T(|A|) < |\mathcal{P}(A)|$, and from this we get $T(|V|) < |\mathcal{P}(V)| \leq |V|$, so instead of $|V| < |V|$ we have obtained $T(|V|) < |V|$, the assertion that the cardinality of ιV , the set of all singletons, is less than the cardinality of the universe.

We seem to see a bijection $(x \mapsto \{x\})$ from V to ιV , but we must conclude that this is not a set, and further it is clearly not an object we can sensibly define in the typed theory of sets: there is no reason to believe that the singleton map can be defined as a set in a well-typed way.

It is worth observing that both V , the universal set, and $|V|$, the largest cardinal, actually exist in NFU. They are not paradoxical objects in the absence of additional assumptions.

Paradox the third: the Burali-Forti paradox of the largest ordinal

For any two binary relations R, S , we say that $R \approx S$, R and S are isomorphic, iff there is a bijection f from the field of R to the field of S such that for all x, y , $x R y \leftrightarrow f(x) S f(y)$. We define the isomorphism type $[R]$ as $\{S : S \approx R\}$.

In the particular case of a well-ordering \leq , we call $[\leq]$ the order type of \leq , and we refer to the set of order types of well-orderings as the set of ordinal numbers. This is the original Frege-Russell-Whitehead method of defining ordinal numbers.

It is a theorem about well-orderings that for any two well-orderings \leq_1 and \leq_2 , either \leq_1 is isomorphic to the restriction of \leq_2 to an initial segment, or vice versa. In addition, it is impossible for a well-ordering to be isomorphic to one of its proper initial segments. It follows (with more work not given here) that the relation $[\leq_1] \leq_\Omega [\leq_2]$ defined as holding iff \leq_1 is isomorphic to the restriction of \leq_2 to an initial segment is a well-ordering, and so has an order type Ω .

We can now state the Burali-Forti paradox. Ω is an ordinal and the order type of the natural order \leq_Ω on the ordinals below an ordinal α is obviously α , so the order type of the ordinals below Ω is Ω , whence the well-ordering on the ordinals is isomorphic to a proper initial segment of itself, which is impossible.

The problem with this is that the assertion that “the order type of the natural order \leq_{Ω} on the ordinals below an ordinal α is obviously α ” is not well-typed and is actually false. If $x \leq y$, the type of the well-ordering \leq is three higher than the common type of x and y (we have $\{\{x^0\}^1, \{x^0, y^0\}^1\}^2 \in \leq^3$) and the type of $[\leq]$ is one higher ($\leq^3 \in [\leq]^4$).

So the order type of the natural order \leq_{Ω} on the set of ordinals $< \alpha$ is in fact four types higher than α : to claim that it is α is to make an ill-typed assertion.

For any well-ordering \leq , the order $\leq' = \{(\{x\}, \{y\}) : x \leq y\}$ is one type higher than \leq and the order type $T([\leq]) = [\leq']$ is one type higher than $[\leq]$.

The assertion that “the order type of the natural order \leq_{Ω} on the ordinals below an ordinal α is $T^4(\alpha)$ ” is well-typed and provable in NFU.

We then discover that the order type of the natural order \leq_Ω on the ordinals below Ω is $T^4(\Omega)$, so $T^4(\Omega) < \Omega$. This should make us uncomfortable, since there is an obvious external isomorphism T^4 from Ω to $T^4(\Omega)$: but the way in which T^4 is defined should make it clear that we cannot expect it to be a set, just as we cannot expect the singleton map to be a set, and the resolution of the Burali-Forti paradox is that indeed it cannot be a set on pain of contradiction.

T (and T^4) certainly commute with order, so we have the alarming

$$\Omega > T^4(\Omega) > T^8(\Omega) > \dots > T^{4i}(\Omega) > \dots$$

but we do not have a descending sequence in the ordinals (which would be a contradiction) because this is not a set.

Note that the order type Ω of the ordinals exists in NFU. It is not, however, the largest ordinal. There can be a largest ordinal in NFU (another story) but only if Infinity fails to be true.

Is this an explanation of why we should trust NFU? I would tend to think not, because while it does show that NFU does not fall prey to the paradoxes, it also shows that NFU behaves in weird unexpected ways.

One slide might be thought to be indirectly an explanation of why we trust NFU: the careful exposition in support of the Cantor paradox resolution of the definition of the crucial set R is an illustration of the fact that our reasoning in NFU is closely tied to reasoning in the typed theory of sets, a theory in which we might have a lot more intuitive confidence.

A hard counterargument is that the entire strategy for resolving the paradoxes would proceed in the same way in the theory $NF + AC$ in which we assume that all objects are sets (there are no urelements) and the Axiom of Choice holds. And this theory is known to be inconsistent; it falls prey to an argument for a contradiction which might be construed as another, more elaborate paradox of set theory.

But this exposition of the resolution of the paradoxes does have an explanatory function in NF(U) studies. It exhibits how mathematics is done (in particular, the Frege-Russell-Whitehead cardinals and ordinals are presented). It teaches caution: a student of NFU needs to know that correspondences mediated through the singleton map are not likely to be set bijections. And there is the positive point that it exhibits the fact that reasoning in this kind of theory follows the pattern of reasoning in the typed theory of sets, in which we do have confidence: to this extent it might be taken as (partially) justifying confidence in this theory.

Part Two: why **do** we have confidence in NFU?

We do believe, since the publication of a paper in 1969 by R. B. Jensen, that NFU is consistent, and as a result of this and subsequent work we can form an intuition about what a model of NFU looks like and what is really happening in the resolutions of the paradoxes described above: one can look at the model and “see” the largest cardinal and the order type of the natural order on the ordinals at work.

Tuesday, November 5, 3:00 - 3:50 pm

Title: The Consistency of NFU (and what bearing it has on discourse about paradoxes)

Abstract: I will discuss the typed theory of sets TST, its variant TSTU (actually I tend to think of TST as a variant of TSTU :-)) and the consistency proof for R.B. Jensen's variant of Quine's set theory New Foundations.

I will take pains to talk about how the consistency proof with its presentation of a model of NFU (in one of its versions) might help us understand exactly what is going on in the curious ways that NF(U) avoids the paradoxes, the topic of our talk last time.

I shall continue to touch on the theme of mathematical explanation and the basis for our confidence in the reliability of mathematical theories, where appropriate.

Typed theory of sets

The typed theory of sets with urelements (TSTU) is the first order theory with a sethood predicate, equality and membership, with sorts indexed by the natural numbers, with formation rules for atomic formula summarized by the templates $\text{set}(x^{i+1}, x^i = y^i, x^i \in y^{i+1}$, and the axioms on the next slide:

sethood: $(\forall xy : x \in y \rightarrow \text{set}(y))$. Things with elements are sets.

extensionality:

$$(\forall xy : \text{set}(x) \wedge \text{set}(y) \wedge (\forall z : z \in x \leftrightarrow z \in y) \rightarrow x = y)$$

Sets with the same elements are the same.

comprehension:

$$(\exists A : \text{set}(A) \wedge (\forall x : x \in A \leftrightarrow \phi))$$

for each formula ϕ in which A does not appear. The witness to this axiom can be denoted by $\{x : \phi\}$ (and is of type one higher than that of x).

Note that we do not adorn every variable with a type superscript. Each variable x is understood to have a type $\text{type}(x)$, and any variable or term may be written with the appropriate type superscript for clarification when necessary.

Intuitive pictures and justification?

There is an intuitive idea behind this theory: we begin with type 0 (a population of featureless individuals), then construct all sets of type 0 objects and perhaps throw in some urelements to get type 1, then construct all sets of type 1 objects and perhaps throw in some urelements to get type 2, and so forth. The urelements may be unexpected: they are needed for Jensen's proof of consistency, and there is a general point that there is no particular reason to believe that everything is a set. The original proposals of this theory did not have a sethood predicate (in effect assuming that everything was a set) and satisfied strong extensionality.

Typical ambiguity

Assume that we have a bijection $(x \mapsto x^+)$ between variables in general and variables of positive type, with $\text{type}(x^+) = \text{type}(x) + 1$. For any formula ϕ , define ϕ^+ as the result of replacing every variable x in ϕ , free or bound, with x^+ .

Note the following facts:

1. ϕ is a formula iff ϕ^+ is a formula.
2. ϕ is an axiom iff ϕ^+ is an axiom.
3. ψ can be logically inferred from ϕ iff ψ^+ can be logically inferred from ϕ^+
4. Any object $\{x : \phi\}$ that we can define has an analogue $\{x^+ : \phi^+\}$ in the next higher type.

Motivation of NF(U) and the ambiguity scheme

The phenomena outlined on the previous slide are the motivation for NFU (well, originally for the strongly extensional version NF).

The idea is that we might be tempted to think not just that ϕ^+ is a theorem if ϕ is a theorem [a fact following from the observations above] but that ϕ is true iff ϕ^+ is true: the axiom scheme $\phi \leftrightarrow \phi^+$ will be termed the Ambiguity Scheme.

Further, we might be tempted to think (and the original proposal vaulted right over the suggestion of the ambiguity scheme to this more radical view) that analogous objects $\{x : \phi\}$ and $\{x^+ : \phi^+\}$ are simply to be viewed as the same object.

This leads to an unsorted set theory with the untyped versions of the axioms of TSTU, which we have already presented in the first talk, as NFU. The subtle point is that dropping the types does not give the inconsistent comprehension scheme of naive set theory: what one obtains are untyped versions of all the typable comprehension axioms.

Specker's ambiguity theorem

In 1962, Specker presented a formal justification for vaulting from the conjecture of the ambiguity scheme to the identification of the types: he proved that $\text{TST}(U) + \text{Ambiguity}$ is consistent if and only if $\text{NF}(U)$ is consistent.

In the case of TSTU + Choice + Ambiguity we can outline an argument for this. In TSTU + Choice, one can prove that each type is well-ordered and introduce a well-ordering \leq^i of each type. Define $(\nu x : \phi)$ as the first object x in the appropriate well-ordering (if there is one) such that ϕ and otherwise as the empty set (or a default object of type 0). This gives what is called a Hilbert symbol. Now it is a standard result that one can construct a model of TSTU + Choice with exactly the same theory as one's original model whose elements are exactly the closed Hilbert symbols. One can then safely identify each term $(\nu x : \phi)$ with $(\nu x^+ : \phi^+)$, because the ambiguity scheme ensures that all statements about the elements of the model will be unaffected by shifting type indices, and then one has a model of NFU + Choice.

It is an embarrassing fact that NF disproves Choice, and so of course does TST + Ambiguity. Specker's original proof of his theorem was intended to apply just to NF, and Specker knew that choice was false in NF already, so took a different tack. The original proof is rather more difficult.

We prove the consistency of NFU

We begin with a set model of TSTU (which can be supposed to exist in a model of TSTU: the types of the model will all be sets of the same type. For each set τ_i representing a type, we have an injective map $e_i : \mathcal{P}(\tau_i) \rightarrow \iota''\tau_{i+1}$ and we define the relation $x \in_i y$ as

$$x \in \tau_i \wedge y \in \tau_{i+1} \wedge x \in e^{-1}(\{y\}).$$

We pause to make this point because much nonsense has been said to the effect that alternative set theories cannot be taken seriously philosophically because we prove their consistency in the usual set theory. Here the arguments can be carried out quite sensibly in the system TSTU itself (the assumption that there is such a model of TSTU in TSTU is quite strong, of course, but if we believe TSTU to be consistent, we believe that it has models in TSTU (though not necessarily such rich ones as we describe here).

We now subvert the type system of our model creatively. Define $\iota_*(x)$ for each $x \in \tau_i$ as the element of $e(\{x\})$: this is just the implementation of the singleton operation in the model.

Define $x \in_{i,j} y$ for $x \in \tau_i$ and $y \in \tau_j$, $j > i$, as $\iota_*^{j-i-1}(x) \in_{j-1} y \wedge e^{-1}(\{y\}) \subseteq \iota_*^{j-i-1} \tau_i$. We are subverting the type system by supplying a membership relation for each type in each higher type.

For each strictly increasing sequence s , the structure in which type i is represented by $\tau_{s(i)}$ and the membership of type i in type $i + 1$ is represented by $\in_{s(i),s(i+1)}$ is readily seen to be a model of TSTU: each subset of type $s(i)$ is represented in type $s(i + 1)$ by its elementwise image under the appropriate iterated singleton map, and sets in type $s(i + 1)$ which are not such images are interpreted as urelements.

Now we can take any formula ϕ of the language of TSTU and transform it to ϕ^s by replacing each reference to type i to type $s(i)$ and suitably correcting membership relations. We are ready for (a version of) Jensen's proof of the consistency of NFU. Notice for application of the ambiguity theorem that if choice holds in our original model Choice^s also holds for every sequence s .

We will use the Ramsey partition theorem, which we state. For any infinite set I , we define $[I]^n$ as the collection of n -element subsets of I . If P is a partition of $[I]^n$, we say that $H \subseteq I$ is homogeneous for P iff all n element subsets of H belong to the same element of the partition P .

The Ramsey partition theorem asserts that each partition of $[I]^n$ for any infinite set I and any n has an infinite homogeneous set.

Let Σ be any finite set of formulas in the language of TSTU. Let n be greater than any type index appearing in any formula in Σ .

We can partition $[\mathbb{N}]^n$, placing each n -element A in one of $2^{|\Sigma|}$ compartments determined by the truth values of formulas ϕ^s for $\phi \in \Sigma$ and $s \in \{0, \dots, n-1\} = A$.

This partition has an infinite homogeneous set H , determining an increasing sequence h . The sentences ϕ^h describe the situation in a model of TSTU with types indexed by the elements of H , in which the instances of the Ambiguity Scheme $\phi \leftrightarrow \phi^+$ hold for each $\phi \in \Sigma$. So any finite subset of the Ambiguity Scheme is consistent with TSTU, from which it follows that the Ambiguity Scheme is consistent with TSTU by the Compactness Theorem, from which it follows by Specker's theorem that NFU is consistent.

It is important to notice that this procedure creates lots of urelements. Whenever a type is skipped in the construction, all sets which are not iterated singleton images of the type at the bottom of the skipped interval become urelements in the type at the top of the skipped interval.

So this procedure simply does not address whatever it is that is going on in the more mysterious NF.

It should also be clear that if Infinity and Choice hold in the original model of TSTU, they will continue to hold in the models with partial ambiguity schemes that are constructed by the procedure above, and so by compactness can be taken to hold in NFU.

An intuitive picture with the aid of the cumulative hierarchy

This approach can be embedded in the usual cumulative hierarchy. A model of TSTU is obtained from any sequence V_{α_i} where α_i 's are strictly increasing, with $x \in_i y$ defined as $x \in V_{\alpha_i} \wedge y \in V_{\alpha_{i+1}} \wedge x \in y$. Notice that this membership relation of elements of V_{α_i} in elements of $V_{\alpha_{i+1}}$ treats each element of $V_{\alpha_{i+1}} \setminus V_{\alpha_i}$ as an urelement.

Model theoretic magic not terribly different from the Jensen argument given above can arrange for a nonstandard model of an initial segment of the cumulative hierarchy with an external automorphism j sending each α_i to α_{i+1} . We can then get a model of NFU by letting its universe be any fixed V_{α_i} and defining $x \in_{\text{NFU}} y$ as $x \in j(y) \wedge j(y) \in V_{\alpha_{i+1}}$.

Third Talk

Tuesday November 12: 3-3:50 pm.

Title: The axiom of cantor sets in NFU and the existence of n -Mahlo cardinals

Abstract: We report on some observations of Robert Solovay, somewhat extended and refined by the speaker, regarding the relationship between the seemingly innocent Axiom of Cantorian Sets proposed by C Ward Henson for NF a long time ago, and the existence of n -Mahlo cardinals. A partition theorem of Schmerl will be described which handles the relationship in one direction.

It is important to remark that errors in this talk should not be imputed to Solovay, where they relate to anything he explained to me, but to my imperfect understanding.

This could be two talks.

Our base theory is NFU + Infinity + Choice, which we defined and showed to be consistent in earlier talks.

The talk relates two concepts.

The Axiom of Cantorian Sets was originally proposed by C. Ward Henson in the context of NF. It is a nice assumption regularizing the behavior of certain ill-typed notions which are natural to consider in NFU.

The other concept is the notion of inaccessible and (strongly) n -Mahlo cardinals, an initial segment of the large cardinal hierarchy. We will define these in TSTU or NFU because that is our metatheory, but of course these are usually discussed in ZFC.

Cantorian and strongly cantorion sets, cardinals, ordinals

A cantorion set is a set A such that $|A| = |\iota A|$. A cantorion cardinal is a cardinal which contains a cantorion set (it follows immediately that all sets of that cardinality are cantorion). A cantorion ordinal is the order type of a well-ordering on a cantorion set.

It is a consequence of Infinity that \mathbb{N} is cantorion (in fact, it is equivalent to Infinity). We can define a map f by induction which sends 0 to $\{0\}$ and which if it sends m to $\{n\}$ sends $m + 1$ to $\{n + 1\}$. It is easy to show by induction that this is a bijection from \mathbb{N} to $\iota \mathbb{N}$. Your itch to identify the variables m and n is understandable, but that would be ill-typed.

Rosser's axiom of counting

The assertion that $f(n) = \{n\}$, where f was defined on the previous slide, is equivalent to Rosser's Axiom of Counting, proposed in the context of NF by Rosser in his excellent book *Logic for Mathematicians*, the only book length treatment of foundations of mathematics in NF. The original form of Rosser's axiom is the entirely common sense

$$|\{x \in \mathbb{N} : 1 \leq x \leq n\}| = n.$$

This is ill-typed, but very natural. What we can prove, in entirely standard ways, are the assertions $f(n) = \{T^{-1}(n)\}$ and

$$|\{x \in \mathbb{N} : 1 \leq x \leq n\}| = T^2(n).$$

The axiom of counting is not a theorem of our base theory

The axiom of counting is not a theorem. In the models of NFU using an automorphism, the T operation coincides with the external automorphism j on the natural numbers of the interpreted NFU, and it is easy to build a model in which j moves a natural number. In fact, one needs to build quite a large model to allow it to be possible for j not to move a natural number. If the ordinal rank V_{α_i} moved by j is a $V_{\omega+n}$ for n a natural number (the immediate natural way to get NFU + Infinity) then n is certainly moved upward by j and the Axiom of Counting is false in the model. This is readily proved internally in NFU: if the cardinality of the universe is \beth_n for some natural number n , then $\beth_n = |V| > |\mathcal{P}(V)| = 2^{|V|} = 2^{\beth_{T(n)}} = \beth_{T(n)+1}$, so $n > T(n) > T^2(n)$ for this particular n , so $|\{x \in \mathbb{N} : 1 \leq x \leq n\}| = T^2(n) < n$. for this choice of n .

The axiom of counting is equivalent in consistency strength to the assertion that $\exists \omega_n$ exists for each standard natural number n . That $\exists \omega_n$ exists for each standard natural number (not necessarily for all natural numbers: the induction turns out to be on an unstratified condition) is provable in NFU with Counting by the same sort of computation exhibited on the previous slide. The result in the other direction is an amusing bit of model theory.

This should be surprising: the axiom of counting looks like a convenience for arithmetic, not a moderately strong set theoretic existence principle.

The difference between the theorem that \mathbb{N} is cantorion and the stronger assertion of the axiom of counting motivates a definition and an obviously convenient axiom, just a natural regularity principle.

We say that a set A is *strongly cantorion (s.c.)* iff $\iota \upharpoonright A$, the restriction of the singleton map to A , is a set. The Axiom of Counting asserts that our map f from above (which we knew was a set) is in fact $\iota \upharpoonright \mathbb{N}$. Strongly cantorion cardinals and ordinals are defined in the same way cantorion cardinals and ordinals were defined.

The Axiom of Cantorian Sets asserts that every cantorion set is strongly cantorion (and so every cantorion ordinal is s.c., every cantorion cardinal is s.c.). What could be more natural?

Large cardinals introduced

A regular cardinal is a cardinal κ such that a set of cardinality κ is not the union of any set of cardinality $< T(\kappa)$ whose elements are all of cardinality $< \kappa$. You may enjoy my type theoretical precision: I did!

A strong limit cardinal is a cardinal κ such that for no $\mu < \kappa$ do we have $2^\mu \geq \kappa$.

A subset C of the set of cardinals $< \kappa$ is a club in κ (a closed unbounded set) if its closure under the construction of least upper bounds of its subsets in the natural order on the cardinals is $C \cup \{\kappa\}$.

A inaccessible or 0-Mahlo cardinal is a regular strong limit cardinal. An $(n+1)$ -Mahlo cardinal is a cardinal any club in which contains an n -Mahlo cardinal. These are large cardinals: the existence of such cardinals is not proved by the usual set theory ZFC.

Our possibly shocking results

NFU + Infinity + Choice + Cantorian Sets (which we will call NFUA, following Solovay) proves the existence of n -Mahlo cardinals for each standard natural number n (not for all natural numbers: induction on an ill-typed condition is involved).

The consistency of NFUA is equivalent to the consistency of TSTU (or ZFC) plus the existence of n -Mahlo cardinals for each standard n (a scheme, not an assertion quantified over n).

A natural side remark

This has bearing on whether NFU is itself a plausible proposal for foundations of mathematics. We have argued that $\text{NFU} + \text{Infinity} + \text{Choice}$ is a plausible basis for foundations of mathematics, though propaganda for this is not our purpose in this series of talks.

There is a counterargument that NFU + Infinity + Choice is (to use Dr Ferrier's word from the comments after the last thought) "parasitic": that we only have confidence in it because we have confidence in ZFC. This is simply false. The proposal of NF historically was somewhat a stretch, but it was based on confidence in TST, not ZFC. The Jensen proof of consistency of NFU was in fact carried out in ZFC but is not intellectually dependent on ZFC: it could be carried out (as we have done) in TSTU + consistency of TSTU. We have confidence in TSTU: when we have justified the ambiguity scheme, we can bootstrap to NFU, as it were.

There is a counterargument that NFU + Infinity + Choice is not strong enough. Saunders Mac Lane has seriously proposed Zermelo set theory with separation restricted to bounded formulas as a general foundation, and this is exactly as strong as the base theory NFU + Infinity + Choice. And it appears that natural extensions of NFU with ill-typed principles (as Counting, Cantorian Sets) tend to give equivalents of principles which appear consistent with the usual set theory but are much stronger than expected. NFUA is much stronger than ZFC, and there are further natural extensions of NFU which are stronger yet.

I'm not engaged in a polemic for NFU as a foundational scheme. To underline this, I note that there is a third objection to NFU and its variants as a foundational scheme which I think either has genuine merit or necessitates some technical improvements in the way this set theory is presented. Sol Feferman has noted that NFU can lead to awkward presentation of almost any mathematical construction involving indexed families of sets. I got seriously burned by this in my published book on the subject, in which I fell into embarrassing error in the discussion of infinite products and sums of cardinals. This was fixed in the online version, but this sad history underlines Feferman's point. Sometimes one has to carefully revisit the roots of NFU in type theory to get things right.

Solovay's framework for proving the existence of inaccessibles (and n -Mahlos) in NFUA

We start doing some math. The framework of this argument is originally due to Solovay, but nothing about my actual presentation should be attributed to him without inquiry to me.

We will need as a tool a special function C . For each strong limit cardinal κ , choose a club $C(\{\kappa\})$ in κ which

1. is of minimal cardinality among clubs in κ ($T^{-1}(|C(\{\kappa\})|)$ is $\text{cf}(\kappa)$, the cofinality of κ in the usual sense).
2. does not contain n -Mahlos unless κ is $(n + 1)$ -Mahlo [in which case it must, by the definition of $(n + 1)$ -Mahlo].

We “cut down” C in a way which the Axiom of Cantorian Sets characteristically supports. Define C^ι as the function which sends $\{T(\kappa)\}$ to $T“C(\{\kappa\})$ for each κ .

This type-shifted version of C will agree with C for each cantorian (= s.c.) cardinal. T fixes a cardinal iff it is cantorian. The axiom of cantorian sets implies that every cardinal less than a cantorian cardinal is cantorian (this assertion is actually equivalent to the axiom, at least in the presence of choice), and so $T“A = A$ for any set A of cardinals less than κ .

Thus C and C^ι will agree for all cardinals less than some noncantorian $\chi \leq T(|V|)$; we restrict our attention to cardinals less than this χ , so that for all cardinals we consider we will have $C(\{T(\kappa)\}) = T“C(\{\kappa\})$. The fact that the cantorian cardinals/ ordinals are a non-set initial segment of the cardinals/ordinals in the natural order is a powerful consequence of the Axiom of Cantorian Sets.

A special function on pairs of cardinals

We define a partial function D acting on pairs of cardinals.

not strong limit: If κ, λ are distinct cardinals which are not strong limit, then $D(\kappa, \lambda) = (\kappa^-, \lambda^-)$, where for any cardinal μ , μ^- is the smallest cardinal ν such that $2^\nu \geq \mu$.

strong limit, not regular, different cofinality:

If κ, λ are distinct cardinals which are strong limit but not regular, and they have different cofinalities, then $D(\kappa, \lambda) = (\text{cf}(\kappa), \text{cf}(\lambda))$.

strong limit, not regular, same cofinality: If

κ, λ are distinct cardinals which are strong limit, but not regular, and they have the same cofinality, then $D(\kappa, \lambda)$ is the first pair (κ', λ') of distinct cardinals with corresponding positions in the natural orders on $C(\kappa)$ and $C(\lambda)$ respectively.

inaccessible: If κ, λ are distinct inaccessibles which are both n -Mahlo, but not $n + 1$ -Mahlo, for the same value of n , then $D(\kappa, \lambda)$ is the first pair (κ', λ') of distinct cardinals with corresponding positions in the natural orders on $C(\kappa)$ and $C(\lambda)$ respectively, if there is such a pair, and otherwise is undefined.

otherwise undefined: In all other cases, $D(\kappa, \lambda)$ is undefined.

Notice that $\kappa_i = \pi_1(D^i(\kappa, \lambda))$ and $\lambda_i = \pi_2(D^i(\kappa, \lambda))$ are strictly decreasing sequences of cardinals, so must be finite: D can only be iterated finitely many times on any given pair of cardinals.

We choose a cardinal κ such that $\kappa \neq T(\kappa)$ and all of $\kappa, T(\kappa), T^2(\kappa)$ are less than χ (which enforces nice behavior of the function C which chooses clubs).

We define $D^i(\kappa, T(\kappa)) = (\kappa_i, \lambda_i)$ and $D^i(T(\kappa), T^2(\kappa)) = (\mu_i, \nu_i)$

An unstratified induction

We would like to claim that $\nu_i = T(\mu_i) = T(\lambda_i) = T^2(\kappa_i)$ for each i . These conditions are clearly true for $i = 0$ and it is straightforward to argue that they hold for $i = k + 1$ if they hold for $i = k$. But these conditions are ill-typed (unstratified): we need to know that they define a set before we can be sure that induction applies.

They do, because the assertion that $(T(\kappa_i), T(\lambda_i)) = (\mu_{T(i)}, \nu_{T(i)})$ is well-typed and so defines a set, and since $i = T(i)$ by the axiom of counting (a consequence of NFUA), this assertion along with the well-typed $\lambda_i = \mu_i$ expresses the questionable condition, so it does define a set and can be proved to contain all relevant i by induction. It remains to actually present the argument.

What we prove by induction is the condition $\nu_i = T(\mu_i) = T(\lambda_i) = T^2(\kappa_i)$ for each i for which our sequences are defined, unless the cardinals involved are inaccessible, and that the sequences continue to index $i + 1$ except possibly in the case where κ_i and λ_i are distinct inaccessibles. This implies, since the sequence must terminate, that there are noncantorian inaccessibles.

With refinements to be discussed later, we can actually show that there are noncantorian n -Mahlos for each concrete n (not for all n , just for each standard natural number).

If κ_i, λ_i are distinct and not strong limit, the same is true of μ_i, ν_i (images under the T operation).

$D(\kappa_i, \lambda_i) = (\kappa_i^-, \lambda_i^-)$, and the $-$ operation certainly commutes with T so the same is true of μ_i, ν_i (and the relations $\nu_i = T(\mu_i) = T(\lambda_i) = T^2(\kappa_i)$ are preserved with subscript incremented).

We argue further that $\kappa_i^- \neq \lambda_i^-$. Otherwise we would have $\kappa_i, \lambda_i = T(\kappa_i)$ both less than $2^{\kappa_i^-} = 2^{\lambda_i^-}$: if these last two items are equal, they are cantorion, so s.c. and they dominate κ_i and $T(\kappa_i) = \lambda_i$ which thus must also be cantorion, and this is a contradiction, as κ_i and λ_i are distinct.

If κ_i and λ_i are strong limit, not regular, and have distinct cofinalities, things are straightforward. All that we have to observe is that the operation of taking cofinalities will commute with T , so all desired conditions will continue to hold.

If κ_i and λ_i are strong limit, not regular, and have the same cofinality, this cofinality is cantorion and the same as the common cofinality of μ_i, ν_i . Let (κ', λ') be the first pair of corresponding elements in the natural orders on $C(\kappa_i), C(\lambda_i)$ which differ (at ordinal index δ , say). Because all cardinals considered are below χ , the first pair (μ', ν') of corresponding elements in the natural orders on $C(\mu_i), C(\nu_i)$ is the componentwise image under T of (κ', λ') and occurs at index $T(\delta)$. But $T(\delta) = \delta$, so the changes happen at the same place, preserving the desired conditions. Note that actual work is done here to show that $\lambda' = \mu'$ [that $\lambda_{i+1} = \mu_{i+1}$ required no special comment in the first two cases].

If κ_i and λ_i are distinct n -Mahlos which are not $(n+1)$ -Mahlo, of course μ_i, ν_i satisfy the same condition.

Let (κ', λ') be the first pair of corresponding elements in the natural orders on $C(\kappa_i), C(\lambda_i)$ which differ (at ordinal index δ , say). Because all cardinals considered are below χ , the first pair of corresponding elements in the natural orders on $C(\mu_i), C(\nu_i)$ is the componentwise image under T of (κ', λ') and occurs at index $T(\delta)$. If $\delta = T(\delta)$, everything works nicely as above. If $\delta \neq T(\delta)$, failure occurs, and we will analyze this more carefully in further discussion at a later point. We do need to note that the point of difference must exist: if there was no first point of difference, then $C(\lambda_i)$ would be an initial segment of $C(\kappa_i)$ or vice versa, and the closure properties of a club would force λ_i into $C(\kappa_i)$ or vice versa, which cannot happen because $C(\kappa_i), C(\lambda_i)$ contain no n -Mahlos.

At this point, we have exhibited a procedure for constructing descending sequences of cardinals which must end at a pair of distinct inaccessible (the logical possibility of failing at a pair of distinct ω -Mahlos should not be ignored). A further refinement of the argument enables us to show that the failure can in fact only occur either at an ω -Mahlo or at an n -Mahlo with n not a standard natural number. This will be presented later if we have time.

Once more with feeling...

We reprise a final segment of the slides from the last talk with refinements allowing us to show that there are n -Mahlos for each concrete n .

We begin by specifying a special set \mathcal{A} of cardinals which has the properties that

1. If $\mu \in \mathcal{A}$ then the minimum of $T(\mu)$ and $T^{-1}(\mu)$ is in \mathcal{A} . A set with this property is said to be *semi-natural*.
2. \mathcal{A} has a noncantorian element.

For any cardinal μ which is not an upper bound for \mathcal{A} , we define $\mu_{\mathcal{A}}$ as the supremum of the set of cardinals in \mathcal{A} which are $\leq \mu$.

It must be noted that if $\mu \neq T(\mu)$ are not upper bounds of \mathcal{A} , we must have $T(\mu)_{\mathcal{A}} = T(\mu_{\mathcal{A}}) \neq \mu_{\mathcal{A}}$. That $\cdot_{\mathcal{A}}$ commutes with T should be evident. If $\mu_{\mathcal{A}} = T(\mu)_{\mathcal{A}}$ we then have this common value cantorian, and the next value in \mathcal{A} must also be cantorian and larger than both μ and $T(\mu)$, which are thus cantorian.

Note that we have demonstrated that the set of inaccessible cardinals has these characteristics.

A special function on pairs of cardinals

We define a partial function D acting on pairs of cardinals.

not strong limit: If κ, λ are distinct cardinals which are not strong limit, then $D(\kappa, \lambda) = (\kappa_{\mathcal{A}}^-, \lambda_{\mathcal{A}}^-)$, where for any cardinal μ , μ^- is the smallest cardinal ν such that $2^\nu \geq \mu$.

strong limit, not regular, different cofinality:

If κ, λ are distinct cardinals which are strong limit but not regular, and they have different cofinalities, then $D(\kappa, \lambda) = (\text{cf}(\kappa)_{\mathcal{A}}, \text{cf}(\lambda)_{\mathcal{A}})$.

strong limit, not regular, same cofinality: If

κ, λ are distinct cardinals which are strong limit, but not regular, and they have the same cofinality, then $D(\kappa, \lambda)$ is $(\kappa'_{\mathcal{A}}, \lambda'_{\mathcal{A}})$, where (κ', λ') is the first pair of distinct cardinals with corresponding positions in the natural orders on $C(\kappa)$ and $C(\lambda)$ respectively.

inaccessible: If κ, λ are distinct inaccessibles which are both n -Mahlo, but not $n + 1$ -Mahlo, for the same value of n , then $D(\kappa, \lambda)$ is $(\kappa'_{\mathcal{A}}, \lambda'_{\mathcal{A}})$, where (κ', λ') is the first pair of distinct cardinals with corresponding positions in the natural orders on $C(\kappa)$ and $C(\lambda)$ respectively, if there is such a pair, and otherwise is undefined.

otherwise undefined: In all other cases, $D(\kappa, \lambda)$ is undefined.

Notice that $\kappa_i = \pi_1(D^i(\kappa, \lambda))$ and $\lambda_i = \pi_2(D^i(\kappa, \lambda))$ are strictly decreasing sequences of cardinals, so must be finite: D can only be iterated finitely many times on any given pair of cardinals.

We choose a cardinal κ such that $\kappa \neq T(\kappa)$ and all of $\kappa, T(\kappa), T^2(\kappa)$ are less than χ (which enforces nice behavior of the function C which chooses clubs), and none of these cardinals are upper bounds for \mathcal{A} .

We define $D^i(\kappa, T(\kappa)) = (\kappa_i, \lambda_i)$ and $D^i(T(\kappa), T^2(\kappa)) = (\mu_i, \nu_i)$

An unstratified induction

We would like to claim that $\nu_i = T(\mu_i) = T(\lambda_i) = T^2(\kappa_i)$ for each i . These conditions are clearly true for $i = 0$ and it is straightforward to argue that they hold for $i = k + 1$ if they hold for $i = k$. But these conditions are ill-typed (unstratified): we need to know that they define a set before we can be sure that induction applies.

They do, because the assertion that $(T(\kappa_i), T(\lambda_i)) = (\mu_{T(i)}, \nu_{T(i)})$ is well-typed and so defines a set, and since $i = T(i)$ by the axiom of counting (a consequence of NFUA), this assertion along with the well-typed $\lambda_i = \mu_i$ expresses the questionable condition, so it does define a set and can be proved to contain all relevant i by induction. It remains to actually present the argument.

What we prove by induction is the condition $\nu_i = T(\mu_i) = T(\lambda_i) = T^2(\kappa_i)$ for each i for which our sequences are defined, unless the cardinals involved are inaccessible, and that the sequences continue to index $i + 1$ except possibly in the case where κ_i and λ_i are distinct inaccessibles. This implies, since the sequence must terminate, that there are noncantorian inaccessibles.

With refinements to be discussed later, we can actually show that there are noncantorian n -Mahlos for each concrete n (not for all n , just for each standard natural number).

If κ_i, λ_i are distinct and not strong limit, the same is true of μ_i, ν_i (images under the T operation).

$D(\kappa_i, \lambda_i) = ((\kappa_i)_{\mathcal{A}}^-, (\lambda_i)_{\mathcal{A}}^-)$, and the $-$ operation and the \mathcal{A} operation certainly commute with T so the same is true of μ_i, ν_i (and the relations $\nu_i = T(\mu_i) = T(\lambda_i) = T^2(\kappa_i)$ are preserved with subscript incremented).

We argue further that $\kappa_i^- \neq \lambda_i^-$. Otherwise we would have $\kappa_i, \lambda_i = T(\kappa_i)$ both less than $2^{\kappa_i^-} = 2^{\lambda_i^-}$: if these last two items are equal, they are cantorion, so s.c. and they dominate κ_i and $T(\kappa_i) = \lambda_i$ which thus must also be cantorion, and this is a contradiction, as κ_i and λ_i are distinct. It follows that $(\kappa_i)_{\mathcal{A}}^- \neq (\lambda_i)_{\mathcal{A}}^-$, due to the general fact that if $\mu_{\mathcal{A}} = T(\mu)_{\mathcal{A}}$ we must actually have $\mu = T(\mu)$.

If κ_i and λ_i are strong limit, not regular, and have distinct cofinalities, things are straightforward. All that we have to observe is that the operation of taking cofinalities will commute with T , so all desired conditions will continue to hold. Further, the operation $\cdot_{\mathcal{A}}$ commutes with T and preserves distinctness for the same reasons discussed at the end of the previous slide.

If κ_i and λ_i are strong limit, not regular, and have the same cofinality, this cofinality is cantorion and the same as the common cofinality of μ_i, ν_i . Let (κ', λ') be the first pair of corresponding elements in the natural orders on $C(\kappa_i), C(\lambda_i)$ which differ (at ordinal index δ , say). Because all cardinals considered are below χ , the first pair (μ', ν') of corresponding elements in the natural orders on $C(\mu_i), C(\nu_i)$ is the componentwise image under T of (κ', λ') and occurs at index $T(\delta)$. But $T(\delta) = \delta$, so the changes happen at the same place, preserving the desired conditions. Note that actual work is done here to show that $\lambda' = \mu'$ [that $\lambda_{i+1} = \mu_{i+1}$ required no special comment in the first two cases].

Further applying $\cdot_{\mathcal{A}}$ will preserve everything, because this operation commutes with T and preserves distinctness under these conditions for reasons described above.

If κ_i and λ_i are distinct n -Mahlos which are not $(n+1)$ -Mahlo, of course μ_i, ν_i satisfy the same condition.

Let (κ', λ') be the first pair of corresponding elements in the natural orders on $C(\kappa_i), C(\lambda_i)$ which differ (at ordinal index δ , say). Because all cardinals considered are below χ , the first pair of corresponding elements in the natural orders on $C(\mu_i), C(\nu_i)$ is the componentwise image under T of (κ', λ') and occurs at index $T(\delta)$. If $\delta = T(\delta)$, everything works nicely as above. If $\delta \neq T(\delta)$, failure occurs, and we will analyze this more carefully in further discussion at a later point. We do need to note that the point of difference must exist: if there was no first point of difference, then $C(\lambda_i)$ would be an initial segment of $C(\kappa_i)$ or vice versa, and the closure properties of a club would force λ_i into $C(\kappa_i)$ or vice versa, which cannot happen because $C(\kappa_i), C(\lambda_i)$ contain no n -Mahlos.

Applying $\cdot_{\mathcal{A}}$ will not break anything unless it is broken for the reason described.

Notice that the process must terminate with noncantorian inaccessible which are limits of the set \mathcal{A} .

Initially (stage 0), we show that any set eligible to be \mathcal{A} must contain noncantorian inaccessible = 0-Mahlos. This is just the original argument with the modified definition of D .

At stage $n + 1$, we show that there must be noncantorian $(n + 1)$ -Mahlos in any set eligible to be \mathcal{A} , on the assumption that there must be noncantorian n -Mahlos in any such set. To do this, we show that failure of our descending process cannot occur at an n -Mahlo, so must occur at a noncantorian $n + 1$ -Mahlo.

We reproduce the text for the failure at n -Mahlos.

If κ_i and λ_i are distinct n -Mahlos which are not $(n+1)$ -Mahlo, of course μ_i, ν_i satisfy the same condition.

Let (κ', λ') be the first pair of corresponding elements in the natural orders on $C(\kappa_i), C(\lambda_i)$ which differ (at ordinal index δ , say). Because all cardinals considered are below χ , the first pair of corresponding elements in the natural orders on $C(\mu_i), C(\nu_i)$ is the componentwise image under T of (κ', λ') and occurs at index $T(\delta)$. If $\delta = T(\delta)$, everything works nicely as above. If $\delta \neq T(\delta)$, failure occurs, and we will analyze this more carefully in further discussion at a later point. We do need to note that the point of difference must exist: if there was no first point of difference, then $C(\lambda_i)$ would be an initial segment of $C(\kappa_i)$ or vice versa, and the closure properties of a club would force λ_i into $C(\kappa_i)$ or vice versa, which cannot happen because $C(\kappa_i), C(\lambda_i)$ contain no n -Mahlos.

Applying $\cdot_{\mathcal{A}}$ will not break anything unless it is broken for the reason described.

The reason this cannot happen is that the common part of the clubs $C(\kappa_i), C(\lambda_i)$ if $\delta \neq T(\delta)$ would satisfy the conditions to be a set \mathcal{A} and so by ind hyp would have to contain an n -Mahlo, and these clubs do not contain n -Mahlos.

This argument works for each concrete natural number n but cannot be carried to a proof that there are n -Mahlos for each n , because the complicated inductive hypothesis does not in any obvious way define a set of natural numbers: the condition that any set eligible to be \mathcal{A} (an unstratified description) must contain a noncantorian n -Mahlo is ill-typed.

There is a converse result. If we suppose that there are n -Mahlo cardinals for each standard natural number n (a scheme, not a quantified statement) we can deduce the consistency of NFUA.

This relies on a partition relation established by Schmerl, which reads almost as if it were designed to prove this result (which it was not!)

We outline the argument from a partition theorem of Schmerl that NFUA is consistent if there are n -Mahlos for each n .

The Schmerl partition property $P(n, \alpha)$ asserts of a cardinal κ that if we have a well-ordered set X of order type $\text{init}(\kappa)$ and partitions C_ν of $[X]^n$ each of size $< T(\kappa)$ that we have a subset Y of X with order type α such that $Y - X_\nu$ is homogenous with respect to C_ν for each ν , where X_ν is the initial segment of X of order type ν .

The interesting theorem is that $P(n+2, n+5)$ holds for n -Mahlo cardinals (in fact, it characterizes n -Mahlo cardinals).

We use it as follows. Let Σ be a finite collection of formulas of the language of set theory containing $n + 2$ types, in a language which includes a countable supply of anonymous constants. Let X be the collection of ordinals less than the initial ordinal for an n -Mahlo cardinal. Let the partition C_ν be determined by the truth values of the formulas in Σ in the models determined by the levels of the hierarchy of isomorphism types of well-founded extensional relations with types taken from a given finite subset of size $n + 2$ of X , including versions of the formulas with every assignment of constant values of level $\leq \nu$ in the hierarchy to anonymous constants in the formulas. This partition will be of size less than the $n + 2$ -Mahlo in play. It then follows by the Schmerl property that there is an ambiguous model for these formulas with $n + 5$ types. Note further that a cantorion ordinal determined by a term $f(x_1, \dots, x_{n+2})$ and, because cantorion, equal

to $f(x_2, \dots, x_{n+3})$, will coerce any ordinal term known to be less than it to be similarly Cantorian, because the set of formulas will include a concrete assignment of a value to the f term and to the g term if these terms are mentioned (the assignment only operating in types with index higher than ν , but that is enough to make the point). TSTU + Ambiguity + Infinity is thereby modelled in an infinitary language with typically ambiguous names for a lot of ordinals, and passage to NFUA will yield Cantorian Sets in addition: a term will be Cantorian in the limiting theory iff it is equal to a typically ambiguous ordinal constant.