1996

Set theory without the axiom of foundation

Talia Leven
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Leven, Talia, "Set theory without the axiom of foundation" (1996). *Master's Theses*. 1237.
DOI: <https://doi.org/10.31979/etd.skh3-amf5>
https://scholarworks.sjsu.edu/etd_theses/1237

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

SET THEORY WITHOUT THE
AXIOM OF FOUNDATION

A Thesis

Presented to

The Faculty of the Department of Mathematics
San Jose State University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science

by

Talia Leven

May 1996

UMI Number: 1379352

UMI Microform 1379352
Copyright 1996, by UMI Company. All rights reserved.

**This microform edition is protected against unauthorized
copying under Title 17, United States Code.**

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

© 1996

Talia Leven

ALL RIGHTS RESERVED

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

Maurice C. Stanley

Professor Maurice Stanley

Jon Pearce

Professor Jon Pearce

M. Gitik

Professor Moti Gitik

MS.d

APPROVED FOR THE UNIVERSITY

M. Lou Lewandowski

ABSTRACT

SET THEORY WITHOUT THE AXIOM OF FOUNDATION

The consistency problem of NF is still an open problem, but there are a few related results concerning theories which are equiconsistent with NF and NFU. We present here the following results:

- (1) Specker's result about the equiconsistency of NF with an extension TT^* of the Theory of Types TT .
- (2) Jensen's result on the consistency of NFU.

The thesis is organized as follows. In Chapter 1 we give a formal definition of First Order Logic and introduce the notion of many sorts of variables. In Chapter 2 we present the topic of Saturated and Special models and give definitions and some theorems related to these kind of models. Chapter 3 presents definitions of ZFC, TT , NF and NFU. Chapter 4 presents Specker's result about the equiconsistency of NF with TT^* , an extension of TT . Chapter 5 presents Jensen's result on the consistency of NFU.

ACKNOWLEDGMENTS

I would like to thank my thesis advisor Professor Mark Stanley for all his help and guidance throughout this thesis.

I especially wish to thank Professor Moti Gitik of Tel-Aviv University who helped me complete the thesis after I moved back to Israel, for all his patience and time and for his excellent mentorship in making me understand every little detail of the material.

Special thanks also to Professor Pearce for all his encouragements throughout the writing of this thesis.

Finally, I would like to also thank my husband, Daniel, who helped type the thesis. Without his patience and that of our children, Yaniv, Itai and Oded, this thesis would not have been possible.

TABLE OF CONTENTS

<u>Chapter</u>	<u>Page</u>
1. Formalization of First Order Logic...	1
1.1 Definitions.....	1
1.2 First Order Logic with many sorts of variables.	8
2. Saturated and Special Models.....	16
3. ZFC, New Foundation and other sets of axioms	31
3.1 ZFC.....	31
3.2 Theory of Types (TT).....	40
3.3 New Foundation (NF).....	42
3.4 NFU.....	44
4. Proof of Equiconsistency of NF and TT*	46
5. On the Consistency of NFU	52
REFERENCES.....	59

CHAPTER 1: FORMALIZATION OF FIRST ORDER LOGIC

In this chapter we give a formalization of first order logic which is required for the understanding of the material in this thesis. We follow this by presenting the notion of first order logic with many sorts of variables.

1.1 Definitions

Definition: A language $\mathcal{L}$ is a set of symbols, namely, relation symbols, function symbols, and constant symbols:

$$\mathcal{L} = \{R, \dots, F, \dots, C, \dots\}.$$

Each R is assumed to be an n -ary (n -placed) relation symbol for some integer $n \geq 1$, each F is an m -ary (m -placed) function symbol for some $m \geq 1$.

Definition: Let $\mathcal{L}$ be a fixed language. Any finite sequence, each element of which is either one of the symbols $(,)$, $,$ (comma), $\wedge$, $\vee$, $\neg$, $\longrightarrow$, $=$, $\forall$, $\exists$, x , y , z , $\dots$, or an element of $\mathcal{L}$, is called an expression of $\mathcal{L}$.

Definition: The terms of $\mathcal{L}$ form the smallest set of expressions containing variables x , y , z , $\dots$, all constant symbols of $\mathcal{L}$ and closed under the formation rule: if $t_1, \dots, t_n$ are terms of $\mathcal{L}$ and if $f \in \mathcal{L}$ is an n -ary function symbol, then the expression $f(t_1, \dots, t_n)$ is a term of $\mathcal{L}$.

Definition: An atomic formula of $\mathcal{L}$ is an expression of either of two forms: $(t_1=t_2), R(t_1, \dots, t_n)$ where, in the first case, t_1 and t_2 are terms of $\mathcal{L}$ and in the second case $R \in \mathcal{L}$ is any n -ary relation symbol and $t_1, \dots, t_n$ are terms of $\mathcal{L}$.

Definition: The formulas of $\mathcal{L}$ form the smallest set of expressions containing the atomic formulas and closed under the following formation rules:

- (i) If φ, ψ are formulas so are the expressions $\neg\varphi, (\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \longrightarrow \psi)$
- (ii) If φ is a formula and v is a variable, then $(\exists v\varphi)$ and $(\forall v\varphi)$ are formulas.

Definition: The set $FV(\varphi)$ of free variables of a formula φ is defined as follows:

- (i) If φ is an atomic formula, then $FV(\varphi)$ is the set of variables appearing in φ .
- (ii) $FV(\neg\varphi) = FV(\varphi)$.
- (iii) $FV(\varphi \wedge \psi) = FV(\varphi \vee \psi) = FV(\varphi \longrightarrow \psi) = FV(\varphi) \cup FV(\psi)$.
- (iv) $FV(\exists v\varphi) = FV(\forall v\varphi) = FV(\varphi) - \{v\}$.

Definition: A sentence of $\mathcal{L}$ is a formula without any free variables.

Definition: A structure or model for $\mathcal{L}$ is a pair $\mathcal{U}=\langle A, f \rangle$ where A is a non-empty set and f is an operation with domain $\mathcal{L}$ such that writing $X^{\mathcal{U}}$ for $f(X)$,

- (i) If $R \in \mathcal{L}$ is an n -ary relation symbol, then $R^{\mathcal{U}} \subseteq A^n$.
- (ii) If $F \in \mathcal{L}$ is an n -ary function symbol then $F^{\mathcal{U}}: A^n \rightarrow A$.
- (iii) If $c \in \mathcal{L}$ is a constant symbol then $c^{\mathcal{U}} \in A$.

Definition: A model $\mathcal{U}'=\langle A', f' \rangle$ is a submodel of a model $\mathcal{U}=\langle A, f \rangle$, denoted by $\mathcal{U}' \subset \mathcal{U}$, if $A' \subset A$ and

- i) Each n -placed relation R' of $\mathcal{U}'$ is the restriction to A' of the corresponding relation R of $\mathcal{U}$, i.e., $R' = R \cap (A')^n$.
- ii) Each m -placed function G' of $\mathcal{U}'$ is the restriction to A' of the corresponding function G of $\mathcal{U}$, i.e., $G' = G| (A')^m$.
- iii) Each constant of $\mathcal{U}'$ is the corresponding constant of $\mathcal{U}$.

Definition: Let $\mathcal{U}=\langle A, f \rangle$ be a model for a language $\mathcal{L}$. An assignment in $\mathcal{U}$ is a function s with domain the set of variables of $\mathcal{L}$ and range a subset of A . We think of s as assigning a meaning $s(v)$ to the variable v . We can then identify each term $t^{\mathcal{U}}$ with its interpretation in the model.

Example: Let $\mathcal{L}$ be the language of rings and let t be the term, or polynomial, $x^2+2\cdot x+1$. If $s(x)=a$ then $t^u(s)=a^2+2\cdot a+1$, the operation $+$ and $\cdot$ being those of the ring u .

Definition: Let u be an $\mathcal{L}$ -structure. We define a relation $u \models \varphi[s]$ for all assignments s and formulas φ as follows:

- (i) $u \models (t_1=t_2)[s]$ iff $t_1^u[s]=t_2^u[s]$.
- (ii) $u \models R(t_1, \dots, t_n)[s]$ iff $(t_1^u[s], \dots, t_n^u[s]) \in R^u$.
- (iii) $u \models \neg \varphi[s]$ if not $u \models \varphi[s]$.
- (iv) $u \models (\varphi \wedge \psi)[s]$ iff $u \models \varphi[s]$ and $u \models \psi[s]$.
- (v) $u \models (\varphi \vee \psi)[s]$ iff $u \models \varphi[s]$ or $u \models \psi[s]$ or both.
- (vi) $u \models (\varphi \rightarrow \psi)[s]$ if either not $u \models \varphi[s]$ or else $u \models \psi[s]$.
- (vii) $u \models (\exists v \varphi)[s]$ iff there is an $a \in A$ such that $u \models \varphi[s(v/a)]$
where $s(v/a)$ is defined as

$$s(v/a)(x) = \begin{cases} s(x) & \text{if } x \neq v, \\ a & \text{if } x = v. \end{cases}$$
- (viii) $u \models (\forall v \varphi)[s]$ iff for all $a \in A$ $u \models \varphi[s(v/a)]$.

Definition: Let Σ be a set of sentences. $\Sigma \models \varphi$ means $\forall u$ $u \models \Sigma \Rightarrow u \models \varphi$. $u \models \Sigma$ means $u \models \psi$ for each $\psi \in \Sigma$.

Definition: Fix a set $\{T, F\}$ of truth values consisting of two distinct points

T , called truth,

F , called falsity

then a truth assignment s for a set $\mathcal{L}$ of sentences is a

function

$$s: \mathcal{L} \rightarrow \{T, F\}$$

assigning either T or F to each sentence in $\mathcal{L}$.

Definition: Let Σ be a set of formulas, then Σ tautologically implies φ iff every truth assignment for symbols in Σ and φ which satisfies every member of Σ also satisfies φ .

Definition: $\mathcal{U}'$ is an elementary submodel of $\mathcal{U}$, denoted by $\mathcal{U}' \prec \mathcal{U}$, if $\mathcal{U}' \subset \mathcal{U}$ and for all formulas φ of $\mathcal{L}$ in the variables

$x_1, \dots, x_n$, and all elements $a_1, \dots, a_n$ of A' , we have

$$\mathcal{U}' \models \varphi[a_1, \dots, a_n] \text{ iff } \mathcal{U} \models \varphi[a_1, \dots, a_n].$$

Definition: The rule of inference known as modus ponens is defined as follows: From the formulas α and $\alpha \rightarrow \beta$ we may infer β .

Now we give the set Λ of logical axioms. These are arranged in six groups. Say that a formula φ is a generalization of ψ iff for some $n \geq 0$ and some variables $x_1, \dots, x_n$,

$$\varphi = \forall x_1 \dots \forall x_n \psi.$$

We include the case $n=0$; any formula is a generalization of itself. The logical axioms are then all generalizations of formula of the following forms, where x

and y are variables and α and β are formulas:

1. Tautologies;
2. $\forall x \alpha \rightarrow \alpha_t^x$, where t is substitutable for x in α ;
3. $\forall x (\alpha \rightarrow \beta) \rightarrow (\forall x \alpha \rightarrow \forall x \beta)$;
4. $\alpha \rightarrow \forall x \alpha$, where x does not occur free in α ;
5. $x = x$;
6. $x = y \rightarrow (\alpha \rightarrow \alpha')$, where α is atomic and α' is obtained from α by replacing x in zero or more (but not necessarily all) places by y .

Definition: A sentence φ is deducible from Σ , denoted by $\Sigma \vdash \varphi$, if there is a finite sequence $\varphi_1, \varphi_2, \dots, \varphi_n$ of sentences such that $\varphi = \varphi_n$ and each sentence φ_m either belongs to $\Sigma \cup \Lambda$ or is inferred from two earlier sentences of the sequence by modus ponens.

Theorem 1.1: (Completeness/Soundness) (Goedel)

$$\Sigma \vdash \varphi \iff \Sigma \models \varphi.$$

Example: Peano Arithmetic (PA).

Language of PA

Let $\mathcal{L} = \{+, \cdot, S, 0\}$, where $+, \cdot$ are 2-placed function symbols, s is a 1-placed function symbol (called the successor function) and 0 is constant symbol.

An example of a term, $S(0)$.

An example of a formula $x+y=S(0)$.

Axioms of PA:

- (i) $0 \neq S(x)$.
- (ii) $S(x) = S(y) \rightarrow x = y$.
- (iii) $x + 0 = x$.
- (iv) $x + S(y) = S(x + y)$.
- (v) $x \cdot 0 = 0$.
- (vi) $x \cdot S(y) = (x \cdot y) + x$.
- (vii) For each formula $\varphi(v_0, \dots, v_n)$ of $\mathcal{L}$ where v_0 is free in φ we have
$$\varphi(0, v_1, \dots, v_n) \wedge (\forall v_0) (\varphi(v_0, \dots, v_n) \rightarrow \varphi(S(v_0), v_1, \dots, v_n)) \rightarrow (\forall v_0) \varphi(v_0, \dots, v_n).$$

Example of a model of PA:

$$\mathcal{U} = \langle \mathbb{N}, +, S, 0, \cdot \rangle.$$

Definition: A set of formulas Σ is consistent, denoted by $\text{con}(\Sigma)$, if there is no β such that both β and $\neg\beta$ are theorems of Σ .

Second Incompleteness Theorem

Assume PA is consistent. Then consistency of PA is not provable in PA.

1.2 First order logic with many sorts of variables

Many Sorted Language

We will present now a first-order language, with many sorts of variables, running over different universes. In informal mathematics we sometimes say things like, "We use Greek letters for ordinals, capital script letters for sets of integers,...." In effect, one thereby adopts several sorts of variables, each sort having its own universe. We will now carefully examine this situation.

Assume that we have a non-empty set I , whose members are called sorts, and symbols arranged as follows:

Variables: For each sort i , there are variables $v_1^i, v_2^i, \dots$ of sort i .

Predicate Symbols: For each $n > 0$ and each n -tuple $\langle i_1 \dots i_n \rangle$ of sorts, there is a set (possibly empty) of n -place predicate symbols, each of which is said to be of sort $\langle i_1, \dots, i_n \rangle$.

Equality Symbols: The symbol $=_i$ is a predicate symbol of sort $\langle i, i \rangle$.

Quantifier Symbols: For each sort i there is a universal quantifier symbol $\forall_i$.

Constant Symbols: For each sort i there is a set (possibly empty) of constant symbols each of which is said to be of sort i .

Function Symbols: For each $n > 0$ and each $(n+1)$ -tuple $\langle i_1 \dots i_n, i_{n+1} \rangle$ of sorts, there is a set (possibly empty) of n -place function symbols, each of which is said to be of sort $\langle i_1, \dots, i_n, i_{n+1} \rangle$.

We assume that these categories of symbols are disjoint, and further that no symbol is a finite sequence of other symbols.

Each term will be assigned a unique sort. We define the set of terms of sort i inductively and simultaneously for all i :

- 1) Any variable of sort i or constant symbol of sort i is a term of sort i .
- 2) If $t_1 \dots t_n$ are terms of sort $i_1 \dots i_n$, respectively, and f is a function symbol of sort $\langle i_1 \dots i_n, i_{n+1} \rangle$, then $f(t_1, \dots, t_n)$ is a term of sort i_{n+1} .

An atomic formula is a sequence $Pt_1 \dots t_n$ consisting of a predicate symbol of sort $\langle i_1 \dots i_n \rangle$ and terms $t_1 \dots t_n$ of sort $i_1 \dots i_n$, respectively. The non-atomic formulas are then formed using the connectives $\neg, \rightarrow, \wedge, \vee$ and the quantifiers $\forall_i \forall_n^i, \exists \forall_n^i$.

Many Sorted Structure.

The many sorted structure M^* is a function on the set of parameters which assigns to each parameter the appropriate type of object:

- 1) To the quantifier symbol $\forall_i, M^*$ assigns a non-empty set $M^*_{i_1}$, called the universe of M^* of sort i .
- 2) To each predicate symbol P of sort $\langle i_1 \dots i_n \rangle, M^*$ assigns the relation $P^{M^*} \subseteq M^*_{i_1} \times \dots \times M^*_{i_n}$.
- 3) To each constant symbol c of sort i , M^* assigns a point c^M in $M^*_{i_1}$.
- 4) To each function symbol f of sort $\langle i_1, \dots, i_n, i_{n+1} \rangle, M^*$ assigns a function: $f^{M^*}: M^*_{i_1} \times \dots \times M^*_{i_n} \rightarrow M^*_{i_{n+1}}$.

The definitions of truth and satisfaction are the obvious ones, given that $\forall_i$ is to mean "for all members of the universe $M^*_{i_1}$ of sort i ."

In many-sorted structures, the universes of the various sorts may or may not be disjoint. Since the equality

relation $=_i$ holds only between elements of a single sort, any non-disjointedness of different sorts can be regarded as accidental. Without loss of generality, we may assume different sorts in a many sorted structure to be disjoint.

Converting a many-sorted language to a one-sorted language

Many-sorted language may at times be convenient, but there is nothing essential that can be done with many sorted language that cannot be done without it. This is because we can always define a one-to-one translation function from a many-sorted language to (an equivalent) one-sorted language. An example of such a translation function is as follows:

Consider the language $\mathcal{L} = \{\epsilon, =\}$ and suppose $(m, n) \mapsto \langle m, n \rangle$ is a one to one function from $\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$. The many-sorted language here has $=_i$ of type $\langle i, i \rangle$ and ϵ_{i+1} of type $\langle i, i+1 \rangle$ for each $i \geq 0$.

Define $\cdot^b$, a function from many sorted formulas to one sorted formulas, by recursion as follows:

$$\begin{aligned} (v_{j \ i+1}^i v_k^{i+1})^b &= v_{\langle i, j \rangle}^{\epsilon} v_{\langle i+1, k \rangle} \\ (v_{j \ i}^i = v_k^i)^b &= (v_{\langle i, j \rangle} = v_{\langle i, k \rangle}) \\ (\neg \phi)^b &= (\neg \phi^b) \end{aligned}$$

$$(\varphi \rightarrow \psi)^b = (\varphi^b \rightarrow \psi^b).$$

$$(\forall v_j^i \varphi)^b = (\forall v_{\langle i, j \rangle} \varphi^b).$$

That is, φ^b is the result of replacing the many sorted variable v_j^i with the one sorted variable $v_{\langle i, j \rangle}$ in φ . Note that the definition of the $\cdot^b$ function depends on the function $(m, n) \mapsto \langle m, n \rangle$.

Definition: A formula ψ of the one-sorted language $\{=, \in\}$ is stratified when there exists a one-to-one function $(m, n) \mapsto \langle m, n \rangle$ such that $\psi = \varphi^b$ for some many sorted formula φ .

Examples:

The formulas $(x \in y)$ and $(x \in z \rightarrow \neg(y \in z))$ are stratified, whereas the formula $x \in x$ is not.

Converting a many-sorted structure M^* to one sorted structure M .

Turning now to semantics we can convert a many sorted structure M^* into a structure M for the above one-sorted language. The universe M is the union $\bigcup_{i \in I} M_i^*$ of all the universes of M^* . M agrees with M^* on the predicates and constant symbols. For a function symbol f , the function f^M is an arbitrary extension of f^{M^*} .

An example of converting a many sorted structure to a one sorted structure which will be required later in the thesis is as follows:

M^* , a many sorted structure, is defined as follows:

Let M be a structure of the language $\mathcal{L}=\{\in,=\}$. Then

- $M^*=\omega \times M$.
- $M^*_i=\{i\} \times M$.
- $(i,x) \in^{M^*} (j,y)$ iff $j=i+1$ and $x \in^M_{i+1} y$.

Let $s:\{\text{many-sorted variables}\} \rightarrow M^*$ be such that $s(v^i_j) \in M^*_i$.

Define: $s^b:\{\text{one-sorted variables}\} \rightarrow M$ by

$$s^b(v_k) = \begin{cases} s(v^i_j) & \text{if } k=\langle i,j \rangle \\ a_0 & \text{if } k \neq \langle i,j \rangle \text{ for all } i,j \end{cases}$$

where $(m,n) \rightarrow \langle m,n \rangle$ is a one-to-one function and a_0 is some fixed element of M .

Claim 1.2.1: Suppose $(m,n) \rightarrow \langle m,n \rangle$, s , s^b are as above.

Then $M^* \models \varphi[s]$ iff $M \models \varphi^b[s^b]$ for all typed formulas φ .

Proof:

The proof is by induction on formulas. Consider the atomic

formulas $\varphi_1=(v^i_{j+1} \in^{M^*} v^i_k)$ and $\varphi_2=(v^i_j = v^i_k)$.

For φ_1 :

$M^* \models \varphi_1[s]$ iff $M^* \models v_j^i \in_{i+1} v_k^{i+1}[s]$ iff $s(v_j^i) \in_{i+1} s(v_k^{i+1})$
iff $s^b(v_t) \in^M s^b(v_r)$ where $t = \langle i, j \rangle$ and $r = \langle i+1, k \rangle$
iff $M \models \varphi_1^b[s^b]$.

For φ_2 :

$M^* \models \varphi_2[s]$ iff $M^* \models v_j^i =_i v_k^i[s]$ iff $s(v_j^i) =_i s(v_k^i)$
iff $s^b(v_t) =^M s^b(v_r)$ where $t = \langle i, j \rangle$ and $r = \langle i, k \rangle$
iff $M \models \varphi_2^b[s^b]$.

For $\neg\varphi$:

$M^* \not\models \varphi[s]$ iff $M^* \models \neg\varphi[s]$.

Assume $M^* \models \neg\varphi[s]$ but $M \models \varphi^b[s^b]$. We get a contradiction from the above statement and induction. Therefore
 $M^* \models \neg\varphi[s]$ iff $M \models \neg\varphi^b[s^b]$.

For $\varphi \rightarrow \psi$:

$M^* \models (\varphi \rightarrow \psi)[s]$ iff $M^* \not\models \varphi[s]$ or $M^* \models \psi[s]$
iff $M \not\models \varphi^b[s^b]$ or $M \models \psi^b[s^b]$ iff $M \models (\varphi \rightarrow \psi)^b[s^b]$.

For $\forall v_i \varphi$:

$M^* \models \forall v_i \varphi[s]$ iff $M^* \models \varphi[s(v_j^i/a)]$, for all a in the universe of the i 'th sort of M^*
iff $M \models \varphi^b[s^b(v_{\langle i, j \rangle}/a)]$, for all a in the universe of M
(for all i we have $\{i\} \times M = M^*_i$ and therefore
 $(i, a) \in M^*_i \iff a \in M$)

iff $M \models \varphi^b[s^b(v_j^i/a)]$, for all a in the universe of M .

iff $M \models \forall v_{\langle i,j \rangle} \varphi^b[s^b]$.

■

CHAPTER 2: SATURATED AND SPECIAL MODELS

If we start with a model $M = \langle A, f \rangle$ for the many-sorted language $\mathcal{L}$ we can always expand M to a model $M_x = (M, a)_{a \in x}$ where $x \subseteq A$ for the many sorted language $\mathcal{L} \cup \{c_a^k : a \in x \cap A_k\}$, where $\{c_a^k : a \in x \cap A_k\}$ is disjoint from $\mathcal{L}$, by choosing an interpretation f' such that $f'(c_a^k) = a$ and $a \in x \cap A_k$. We say that M_x is an expansion of M to $\mathcal{L}_x$.

Definition:

- (a) The cardinal, or power, of the model M is the cardinal $|\bigcup_{i \in I} |A_i||$. M is said to be finite, countable or not countable if $|\bigcup_{i \in I} |A_i||$ is finite, countable or uncountable respectively.
- (b) The cardinal, or power, of the many sorted language $\mathcal{L}$, $\|\mathcal{L}\|$, is defined as $\|\mathcal{L}\| = \omega \cup |\mathcal{L}|$. $\mathcal{L}$ is said to be countable or uncountable if $\|\mathcal{L}\|$ is countable or uncountable respectively.

Definition: Let M and N be two models of a language $\mathcal{L}$.

We say that

- (a) N is an extension of M if M is a submodel of N .
- (b) N is an elementary extension of M , $M \equiv N$, if N is an extension of M and if for any formula $\varphi(v_1, \dots, v_n)$ of $\mathcal{L}$ and any $a_1, \dots, a_n$ such that $a_i \in A$ and $a_i \in A$, $a_1, \dots, a_n$ satisfies φ in M iff it satisfies φ in N .

Definition: By a type $\Sigma(v_1^i, \dots, v_n^k)$ in the variables $v_1^i, \dots, v_n^k$ we mean a maximal consistent set of formulas of $\mathcal{L}$ in those variables.

Definition: We say that M realizes Σ iff some n -tuple of elements of A satisfies Σ in M .

Definition:

- a) Let α, β be limit ordinals. If $f: \alpha \rightarrow \beta$, then f maps α cofinally iff $\text{range}(f)$ is unbounded in β .
- b) The cofinality of β , denoted by $\text{cf}(\beta)$, is the least α such that there is a map from α cofinally into β .

Definition:

- a) A chain of models is an increasing sequence of models

$$\mathcal{U}_0 \subset \mathcal{U}_1 \subset \dots \subset \mathcal{U}_\beta \subset \dots, \quad \beta < \alpha,$$

whose length is an ordinal α .

- b) An elementary chain is a chain of models

$$\mathcal{U}_0 \overset{e}{\subset} \mathcal{U}_1 \overset{e}{\subset} \dots \overset{e}{\subset} \mathcal{U}_\beta \overset{e}{\subset} \dots, \quad \beta < \alpha,$$

such that $\mathcal{U}_\gamma \overset{e}{\subset} \mathcal{U}_\beta$ whenever $\gamma < \beta < \alpha$.

Definition: For a model $\mathcal{U}$, define the theory of $\mathcal{U}$, $\text{Th}(\mathcal{U})$, to be the set of all sentences of $\mathcal{L}$ which hold in the model $\mathcal{U}$.

Definitions:

- a) A model A is α -saturated iff for every subset $X \subset A$ with fewer than α elements, the expansion M_x realizes $\forall i, j$ every type $\Sigma(v_j^i)$ of the language $\mathcal{L}_x$ which is consistent with the theory of M_x .
- b) A model M is said to be special iff M is the union of an elementary chain of models M_β , $\beta < |A|$, where β is a cardinal and each M_β is β^+ -saturated.

Note: At first sight it may appear that we would obtain a stronger notion by considering sets of formulas with finitely many free variables. The next claim shows that we do not obtain a stronger notion that way.

Claim 2.1: M is α -saturated if and only if for every ordinal $\zeta < \alpha$ and every $a \in {}^\zeta A$, the model $(M, a_\eta)_{\eta < \zeta}$ realizes all types $\Sigma(v_1^m, \dots, v_n^k)$ consistent with its theory.

Proof:

The proof is by induction on n .

($\Leftarrow$) Assume that for every ordinal $\zeta < \alpha$ and every $a \in {}^\zeta A$ the model $(M, a_\eta)_{\eta < \zeta}$ realizes all types $\Sigma(v_1^m, \dots, v_n^k)$ consistent with its theory, so definitely $(M, a_\eta)_{\eta < \zeta}$ realizes every type $\Sigma(v_j^i)$ consistent with its theory. By definition M is α -saturated.

($\Rightarrow$) For $n=1$, by definition of α -saturated.

Assume the claim is true for n we want to prove it

for $n+1$. Suppose now that $\Sigma(v_1^m, \dots, v_{n+1}^k)$ is a type consistent with the theory $\text{Th}((M, a_\eta)_{\eta < \zeta})$, where $\zeta < \alpha$.

Note that then

- (i) If $\varphi_1, \dots, \varphi_k \in \Sigma$ then $\varphi_1 \wedge \dots \wedge \varphi_k \in \Sigma$ and
- (ii) If $\varphi \in \Sigma$, then $\exists v_1^m \dots \exists v_{n+1}^l \varphi \in \text{Th}((M, a_\eta)_{\eta < \zeta})$
 (since Σ is consistent with $\text{Th}((M, a_\eta)_{\eta < \zeta})$ there is $a_1, \dots, a_{n+1} \in \langle a_\eta : \eta < \zeta \rangle$ such that $\varphi(a_1, \dots, a_{n+1})$ is true in $(M, a_\eta)_{\eta < \zeta}$)

Set $\Gamma = \{\exists v_{n+1}^l \varphi : \varphi \in \Sigma\}$. To see that Γ is consistent with $\text{Th}((M, a_\eta)_{\eta < \zeta})$, suppose that $\exists v_1^m \varphi_1, \dots, \exists v_{n+1}^l \varphi_k \in \Gamma$. Then the formulas $\varphi_1, \dots, \varphi_k \in \Sigma$ and therefore $\varphi_1 \wedge \dots \wedge \varphi_k \in \Sigma$. Thus $\exists v_1^m \dots \exists v_{n+1}^l (\varphi_1 \wedge \dots \wedge \varphi_k) \in \text{Th}((M, a_\eta)_{\eta < \zeta})$. Now $\exists v_1^m \dots \exists v_{n+1}^l (\varphi_1 \wedge \dots \wedge \varphi_k) \vdash \exists v_1^m \dots \exists v_n^l (\exists v_{n+1}^l \varphi_1 \wedge \dots \wedge \exists v_{n+1}^l \varphi_k)$ and thus Σ is consistent with $\text{Th}((M, a_\eta)_{\eta < \zeta})$. By induction, say $b_1, \dots, b_n$ realize Γ in $(M, a_\eta)_{\eta < \zeta}$.

Set $\Delta(v_{n+1}^l) = \{\varphi(c_{b_1}^m, \dots, c_{b_n}^t, v_{n+1}^l) : \varphi \in \Sigma\}$. Suppose that we have the formulas

$$\varphi_1(c_{b_1}^m, \dots, c_{b_n}^t, v_{n+1}^l), \dots, \varphi_k(c_{b_1}^m, \dots, c_{b_n}^t, v_{n+1}^l) \in \Delta.$$

Then $\varphi_1(v_1^m, \dots, v_{n+1}^l), \dots, \varphi_k(v_1^m, \dots, v_{n+1}^l) \in \Sigma$, consequently the single formula

$$\varphi_1(v_1^m, \dots, v_{n+1}^l) \wedge \dots \wedge \varphi_k(v_1^m, \dots, v_{n+1}^l) \in \Sigma.$$

It follows that the formula

$$\exists v_{n+1}^l (\varphi_1(v_1^m, \dots, v_{n+1}^l) \wedge \dots \wedge \varphi_k(v_1^m, \dots, v_{n+1}^l)) \in \Gamma.$$

Because $b_1, \dots, b_n$ realize Γ , we have

$$\exists v_{n+1}^l \varphi_1(c_{b_1}^m, \dots, c_{b_n}^t, v_{n+1}^l) \wedge \dots \wedge \exists v_{n+1}^l \varphi_k(c_{b_1}^m, \dots, c_{b_n}^t, v_{n+1}^l) \in$$

$\text{Th}(M, b_1 \dots b_n, a_\eta)_{\eta < \zeta}$. Of course $n + \zeta < \alpha$ since $\zeta < \alpha$ and α

is an infinite cardinal. Choose b_{n+1} realizing Γ in $(M, b_1 \dots b_n, a_\eta)_{\eta < \zeta}$. Then $b_1, \dots, b_{n+1}$ realizes Σ in $(M, a_\eta)_{\eta < \zeta}$.

■

The existence of many sorted special models

We shall prove that many-sorted special models exist in each power $\alpha = 2^{<\alpha}$, α a limit cardinal.

Proposition 2.1: Suppose that $\alpha = 2^{<\alpha}$, α a limit cardinal, $\|\mathcal{L}\| < \alpha$, and $\omega \leq |A| < \alpha$. Then there is a many-sorted special elementary extension $\mathcal{B}$ of M of power α .

Proof:

Let $\alpha = 2^{<\alpha}$. α is a limit cardinal; therefore, if $\beta < \alpha$ then $\beta^+ < \alpha$ and $2^\beta \leq \alpha$. Let $\gamma = \|\mathcal{L}\| \cup |A|$. We want to construct an elementary chain of models $\mathcal{B}_\beta$, β a cardinal, $\beta < \alpha$ and $\mathcal{B}_\beta$ is an extension of M .

to prove the proposition we need the following claim:

Claim 2.2: Suppose $\|\mathcal{L}\| \leq \gamma$ and $\omega \leq |A| \leq 2^\gamma$, then there is a many-sorted γ^+ -saturated elementary extension of M of power 2^γ .

Proof:

We construct an elementary chain $\mathcal{B}_\xi$, $\xi < 2^\gamma$, of length 2^γ such that:

- (1) Each $\mathcal{B}_\xi$ is an elementary extension of M of power 2^γ .
- (2) For every $x \in \mathcal{B}_\xi$ of power γ , $(\mathcal{B}_{\xi+1}, a)_{a \in x}$ realizes every type $\Sigma(v_\zeta^1)$ of $(\mathcal{B}_\xi, a)_{a \in x}$.

We first construct $\mathcal{B}_0$. Since $|A| \leq 2^\gamma$, the number of subsets $x \subset A$ of power γ is at most 2^γ . The language $\mathcal{L}_x = \mathcal{L} \cup \{c_a^k : a \in x \cap |A|_k\}$ still has at most γ symbols. Hence the total number of types $\Sigma(v_j^1)$ of $\mathcal{L}_x$ is at most 2^γ . (Since the cardinal number of all subset of $\mathcal{L}_x$ is 2^γ). For each $x \subset A$ of power γ and each type $\Sigma(v_j^1)$ of $(M, a)_{a \in x}$ we introduce a constant $c_{x\Sigma}^1$. Let T be the theory $\text{Th}(M_A)$ of all sentences of $\mathcal{L}_A$ which hold in $M_A = (M, a)_{a \in A}$ plus all the sentences $\Sigma(c_{x\Sigma}^1)$ for every $x \subset A$ of power γ and every type $\Sigma(v_j^1)$ of $(M, a)_{a \in x}$. T is consistent (by compactness T has a model if and only if every $T' \leq T$ $|T'| < \aleph_0$ has a model. Let $T' \leq T$ be finite. Then $T' = T'_1 \cup T'_2$ where $T'_1 \leq \text{Th}(\mathcal{U}_A)$, T'_2 the rest. T'_1 is consistent. Let $\sigma = \bigwedge_{1 \leq k \leq n} \Sigma_k(v_j^1)$. By definition of T , T'_2 is the union of the finite Σ_k 's. Consider any formula $\varphi \in \Sigma_k(v_j^1)$. $\exists a$ such that a satisfies φ . Therefore every type is satisfied in the model M_A with the new constants. So T'_2 is consistent. Therefore T has a model.) The model of T is infinite because it contains an infinite number of constants $c_{x\Sigma}^1$. Since T has at most 2^γ symbols, T has a model of power 2^γ . The reduction of this model to $\mathcal{L}$ is the desired model $\mathcal{B}_0$.

If η is a limit ordinal, $0 < \eta < 2^\gamma$, we let $\mathcal{B}_\eta = \bigcup_{\xi < \eta} \mathcal{B}_\xi$. If $\eta = \xi + 1$, we let $\mathcal{B}_{\xi+1}$ be the elementary extension of $\mathcal{B}_\xi$ as we did in the case of $\mathcal{B}_0$ with M replaced by $\mathcal{B}_\xi$. Let $\mathcal{B} = \bigcup_{\eta < 2^\gamma} \mathcal{B}_\eta$. Clearly $\mathcal{B}$ is an elementary extension of M of power 2^γ . Let $x \subset \mathcal{B}$ be of power γ , and let $\Sigma(v_j^1)$ be a type of $(\mathcal{B}, a)_{a \in x}$. Since 2^γ has cofinality greater than γ , there is some $\xi < 2^\gamma$ such that $x \subset \mathcal{B}_\xi$. Now $\mathcal{B}_\xi$ is an elementary submodel of $\mathcal{B}$ so $\Sigma(v_j^k)$ is also a type of $(\mathcal{B}_\xi, a)_{a \in x}$. By our construction, some $b \in \mathcal{B}_{\xi+1}$ realizes Σ . That is $(\mathcal{B}_{\xi+1}, a)_{a \in x} \models \Sigma(b)$. Since $\mathcal{B}$ is an elementary extension of $\mathcal{B}_{\xi+1}$ we get $(\mathcal{B}, a)_{a \in x} \models \Sigma(b)$. ■

We can now go back to the proof of proposition 2.1. Let $\mathcal{B}_\gamma$ be any γ^+ -saturated elementary extension of M of power 2^γ . For any cardinal $\delta < \gamma$, let $\mathcal{B}_\delta = \mathcal{B}_\gamma$. Let $\gamma \leq \beta < \alpha$ where α is a limit cardinal. Since we assume that α is a limit cardinal, if $\beta < \alpha$, then also $\beta^+ < \alpha$ and $2^{\beta^+} \leq \alpha$ since $2^{<\alpha} = \alpha$. Given $\mathcal{B}_\beta$ of power 2^β , let $\mathcal{B}_{\beta^+}$ be any β^{++} -saturated elementary extension of power 2^{β^+} . Given $\mathcal{B}_\delta, \delta < \beta$, where β is a limit cardinal less than α , we first take the union $\mathcal{B}' = \bigcup_{\delta < \beta} \mathcal{B}_\delta$, which is an infinite model of power at most 2^β ; then we let $\mathcal{B}_\beta$ be any β^+ -saturated elementary extension of $\mathcal{B}'$ of power 2^β . Now the chain $\mathcal{B}_\beta, \beta < \alpha$, is a specializing chain of the model $\bigcup_{\beta < \alpha} \mathcal{B}_\beta$ with power α . ■

Definition: Two many-sorted models M and N for $\mathcal{L}$ are equivalent if every sentence that is true in M is true in N , and vice versa. We express this relationship between models by $\equiv$. Note that the definition of equivalence is the same as for one-sorted models.

Definition: Two models $M=\langle A, \epsilon_1 \rangle$ and $N=\langle B, \epsilon_1 \rangle$ for $\mathcal{L}$ are isomorphic iff there is a 1-1 function f mapping $|A|_1$ onto $|B|_1$ $\forall i \in \omega$ satisfying:

- (i) for each n -placed relation R of M and the corresponding relation R' of N

$$R(a_1, \dots, a_n) \text{ iff } R'(f(a_1), \dots, f(a_n))$$

for all $a_1, \dots, a_n$ in A .

- (ii) for all m -placed functions G of M and the corresponding function G' of N

$$f(G(a_1, \dots, a_m)) = G'(f(a_1), \dots, f(a_m))$$

for all $a_1, \dots, a_m$ in A .

- (iii) for each constant a_k of M and the corresponding constant b_1 of N we have $f(a_k) = b_1$.

We use the notation $M \cong N$ for " M is isomorphic to N ".

Lemma 2.1: Let M be α -saturated and for every i M_1 is

infinite then for every i , $\alpha \leq |M_i|$.

Proof:

Assume that there is k such that $|M_k| = \beta < \alpha$. Let $M_k = \{a_\zeta^k \mid \zeta < \beta\}$ and let $\Sigma(v_0^k)$ be the type $\Sigma(v_0^k) = \{(v_0^k \neq a_\eta^k) \mid \eta < \beta\}$. $\Sigma(v_0^k)$ is a type of M_k but cannot be realized in M_k (since M_k is infinite and M is α -saturated), a contradiction to our assumption. ■

Theorem 2.1: If $M = \langle A, \epsilon_{i+1} \rangle$ and $N = \langle B, \epsilon_{i+1} \rangle$ are equivalent saturated models of the same power, then $M \cong N$.

Proof:

Let $|A| = |B| = \alpha$ and let $\langle a_\zeta^1; \zeta < \alpha \rangle$ be an enumeration of A_i and $\langle b_\zeta^1; \zeta < \alpha \rangle$ and enumeration B_i , $i < \omega$. By Lemma 2.3 below there are $\langle \bar{a}_\eta^1; \eta < \alpha \rangle$, $\langle \bar{b}_\eta^1; \eta < \alpha \rangle$, $i < \omega$ such that $(M, \bar{a}_\eta^1)_{\eta < \alpha} \equiv (N, \bar{b}_\eta^1)_{\eta < \alpha}$. Since $\langle \bar{a}_\eta^1; \eta < \alpha \rangle$ and $\langle \bar{b}_\eta^1; \eta < \alpha \rangle$, $i < \omega$ still enumerate A_i and B_i the isomorphism of M and N follows. ■

Lemma 2.2: Suppose that α is infinite, M and N are both α -saturated, and $M \cong N$. Let $\langle a_\zeta^1; \zeta < \alpha \rangle$, $\langle b_\zeta^1; \zeta < \alpha \rangle$, $i_\zeta < \omega$ be sequences consisting of elements of M and N . Then, there

are $\langle \bar{a}_\xi^1; \xi < \alpha \rangle$, $\langle \bar{b}_\xi^1; \xi < \alpha \rangle$, $i_\xi < \omega$ such that:

$$\{a_\zeta^1; \zeta < \alpha\} \subset \{\bar{a}_\xi^1; \xi < \alpha\}, \quad i_\zeta < \omega,$$

$$\{b_\zeta^1; \zeta < \alpha\} \subset \{\bar{b}_\xi^1; \xi < \alpha\}, \quad i_\zeta < \omega,$$

$$(M, \bar{a}_\xi^1)_{\xi < \alpha} \equiv (N, \bar{b}_\xi^1)_{\xi < \alpha}.$$

Proof:

We represent each ordinal $\xi < \alpha$ uniquely as a sum $\xi = \lambda + n$, where $\lambda \leq \xi$ is a limit ordinal and $n \in \omega$. ξ is said to be even if n is even; otherwise ξ is odd. We wish to find two sequences $\langle \bar{a}_\xi^i; \xi < \alpha \rangle$, $\langle \bar{b}_\xi^i; \xi < \alpha \rangle$, $i_\xi < \omega$ such that:

- (a) For all $\eta < \xi$, if $\eta = \lambda + 2n$ is even, then $\bar{a}_\eta^i = a_{\lambda+n}^i$.
- (b) For all $\eta < \xi$, if $\eta = \lambda + (2n+1)$ is odd, then $\bar{b}_\eta^i = b_{\lambda+n}^i$.
- (c) $(M, \bar{a}_\eta^i)_{\eta < \xi} \equiv (N, \bar{b}_\eta^i)_{\eta < \xi}$.

Suppose that $\xi < \alpha$ and $\bar{a}_\eta^i, \bar{b}_\eta^i, \eta < \xi$, have been found so that

(a), (b), (c) hold for ξ . We now find $\bar{a}_\xi^i, \bar{b}_\xi^i$. Assume η is even, i.e. $\eta = \lambda + 2n$. Let $\bar{a}_\eta^i = a_{\lambda+n}^i$ and $\Sigma(v_k^i)$ be the type of $\bar{a}_{\lambda+n}^i$ in $(M, \bar{a}_\eta^i)_{\eta < \xi}$. By (c), $\Sigma(v_k^i)$ is a type of $(N, \bar{b}_\eta^i)_{\eta < \xi}$ with a's replaced by b's. So Σ is realized by some element $\bar{b}_\xi^i$. Therefore, $(M, \bar{a}_\eta^i)_{\eta \leq \xi} \equiv (N, \bar{b}_\eta^i)_{\eta \leq \xi}$.

The case where ξ is odd is treated similarly. ■

Lemma 2.3: Suppose M is special and M_i is infinite for all i then $|M_i| = |M|$.

Proof:

Let M be a special model. By definition $M = \bigcup_{\beta < |M|} M_\beta$ where every M_β is β^+ -saturated. According to Lemma 2.1 for every i $\beta^+ \leq |M_{\beta_i}|$. $M_i = \bigcup_{\beta < |M|} M_{\beta_i}$ but $|M_i| \geq \bigcup_{\beta < |M|} \beta^+ = |M|$. Therefore $|M_i| = |M|$. ■

Theorem 2.2: If $M = \langle A, \epsilon_{i+1} \rangle$ and $N = \langle B, \epsilon_{i+1} \rangle$ are equivalent many sorted special models of the same power, such that M_i, N_i are infinite for all i , then $M \cong N$.

Proof:

Suppose M and N are both of cardinality α and have specializing chains $\langle M_\beta : \beta < \alpha \rangle$ and $\langle N_\beta : \beta < \alpha \rangle$ respectively, then $|M_i| = |N_i| = \alpha$. Let $\langle a_\zeta^i; \zeta < \alpha \rangle, \langle b_\zeta^i; \zeta < \alpha \rangle, i_\zeta < \omega$ be enumerations of A and B respectively. Let $A_\beta = \bigcup_{i \in \omega} A_{\beta i}$, $B_\beta = \bigcup_{i \in \omega} B_{\beta i}$ $\beta < \alpha$ be the universe of M_β, N_β respectively and let $a_\eta^i \in A_{\beta i}, b_\eta^i \in B_{\beta i}$ $\eta < \zeta$ such that the cardinality of η is β . We are going to define by recursion two sequences $\langle \bar{a}_\gamma^i; \gamma < \alpha \rangle, \langle \bar{b}_\gamma^i; \gamma < \alpha \rangle, i_\gamma < \omega$ such that for each $\delta < \gamma$, with cardinality of γ being β ,

- (a) $\bar{a}_\gamma^i \in A_{\beta i_\gamma}$ and $\bar{b}_\gamma^i \in B_{\beta i_\gamma}$.
- (b) $(M, \bar{a}_\delta^i)_{\delta < \gamma} \equiv (N, \bar{b}_\delta^i)_{\delta < \gamma}$.
- (c) $\bar{a}_\delta^i = a_{\eta+n}^i$, if $\delta = \eta + 2n$ with η a limit ordinal and $n < \omega$
and $\bar{b}_\delta^i = b_{\eta+n}^i$, if $\delta = \eta + 2n + 1$ with η a limit ordinal
and $n < \omega$.

Suppose $\mu < \alpha$, the cardinality of μ is β , and (a), (b), (c) hold for all $\gamma < \mu$, then

$$\text{by (b)} \quad (M, \bar{a}_\gamma^i)_{\gamma < \mu} \equiv (N, \bar{b}_\gamma^i)_{\gamma < \mu}, \quad (1)$$

$$\text{by (a)} \quad \{\bar{a}_\gamma^i; \gamma < \mu\} \subseteq A_\beta, \quad \{\bar{b}_\gamma^i; \gamma < \mu\} \subseteq B_\beta. \quad (2)$$

Therefore since $M_\beta \overset{c}{\subseteq} A$ and $N_\beta \overset{c}{\subseteq} B$ it follows from (1) and (2) that

$$(M_\beta, \bar{a}_\gamma^i)_{\gamma < \mu} \equiv (N_\beta, \bar{b}_\gamma^i)_{\gamma < \mu} \dots\dots\dots (3)$$

Since M_β and N_β are both β^+ -saturated and $\mu < \beta^+$ we can conclude from (3) that we can find $\bar{a}_\mu^i \in A_\beta$, $\bar{b}_\mu^i \in B_\beta$ so that (a), (b), (c) are satisfied when γ is replaced by μ . This completes the recursive definition of $\bar{a}_\gamma^i$, $\bar{b}_\gamma^i$. By (c) $\langle \bar{a}_\mu^i, \mu < \alpha \rangle$, $\langle \bar{b}_\mu^i, \mu < \alpha \rangle$ are enumerations of A , B respectively and by (b) $(M, \bar{a}_\mu^i)_{\mu < \alpha} \equiv (N, \bar{b}_\mu^i)_{\mu < \alpha}$. Therefore by the definition of $\equiv$ and $\cong$ we conclude that M is isomorphic to N .

■

Suppose that $\mathcal{L}$ is the many sorted language given by $\mathcal{L} = \{\epsilon_{i+1} : i \geq 0\}$. Let $\mathcal{L}^*$ be the result of increasing the types of every symbol in $\mathcal{L}$ by 1. That is

$$\mathcal{L}^* = \{\epsilon_{i+1} : i \geq 1\}.$$

If φ is a formula of the language $\mathcal{L}$, define φ^* to be $\mathcal{L}^*$ -formula resulting from increasing each type index in φ by 1. That is, letting t^i and u^i denote terms of type i (either v_j^i , for some j , or a constant symbol of type i), set

$$(t^i \epsilon_{i+1} u^{i+1})^* = (t^{i+1} \epsilon_{i+2} u^{i+2}).$$

$$(t^i = u^i)^* = (t^{i+1} = u^{i+1}).$$

$$(\neg \varphi)^* = (\neg \varphi^*).$$

$$(\varphi \rightarrow \psi)^* = (\varphi^* \rightarrow \psi^*).$$

$$(\forall_i v_j^i \varphi)^* = (\forall_{i+1} v_j^{i+1} \varphi^*).$$

Now suppose that M is a (typed) structure for $\mathcal{L}^+$. Define a structure M^+ for $\mathcal{L}$ as follows:

$$M_i^+ = M_{i+1}, \text{ for all } i \geq 0$$

$$(i, a) \in_{i+1}^{M^+} (i+1, b) \text{ iff } (i+1, a) \in_{i+2}^M (i+2, b).$$

Let s be an assignment of individuals in M to typed variables, define an assignment s^+ of individuals in M^+ to typed variables by setting $s^+(v_j^i) = s(v_j^{i+1})$, for all i, j .

Lemma 2.4: Suppose M is a typed structure for $\mathcal{L}^+$ and that s is an assignment of individuals in M to typed variables, then:

$$M^+ \models \varphi[s^+] \text{ iff } M \models \varphi^+[s] \text{ for all formulas } \varphi \text{ of } \mathcal{L}.$$

Proof:

The proof is by induction on formulas. Let $\varphi_1 = (v_{j_{i+1}}^i v_k^{i+1})$ and $\varphi_2 = (v_j^i v_k^i)$ be the atomic formulas.

For φ_1

$$\begin{aligned} M^+ \models \varphi_1[s^+] &\text{ iff } M^+ \models v_{j_{i+1}}^i v_k^{i+1}[s^+] \\ &\text{ iff } s^+(v_{j_{i+1}}^i) \in_{i+1}^{M^+} s^+(v_k^{i+1}) \\ &\text{ iff } s(v_j^{i+1}) \in_{i+2}^M s(v_k^{i+2}) \\ &\text{ iff } M \models (v_{j_{i+1}}^i v_k^{i+1})^+[s] \\ &\text{ iff } M \models \varphi_1^+[s]. \end{aligned}$$

For φ_2

$$M^+ \models \varphi_2[s^+] \text{ iff } M^+ \models v_j^1 = v_k^1[s^+]$$

$$\text{iff } s^+(v_j^1) =_i^{M^+} s^+(v_k^{i+1})$$

$$\text{iff } s(v_j^{i+1}) =_{i+1}^M s(v_k^{i+1})$$

$$\text{iff } M \models (v_j^1 = v_k^1)^+[s]$$

$$\text{iff } M \models \varphi_2^+[s].$$

For $\neg\varphi$

$$M^+ \models \neg\varphi[s^+] \text{ iff } M^+ \not\models \varphi[s^+]$$

$$\text{iff } M \not\models \varphi^+[s]$$

$$\text{iff } M \models \neg\varphi^+[s].$$

For $\varphi \rightarrow \psi$:

$$M^+ \models (\varphi \rightarrow \psi)[s] \text{ iff } M^+ \not\models \varphi[s] \text{ or } M^+ \models \psi[s]$$

$$\text{iff } M \not\models \varphi^+[s] \text{ or } M \models \psi^+[s] \text{ iff } M \models (\varphi \rightarrow \psi)^+[s].$$

For $\forall v_j^1 \varphi$:

$$M^+ \models \forall v_j^1 \varphi[s^+] \text{ iff } M^+ \models \varphi[s^+(v_j^1/a)], \text{ for all } a \text{ in the universe of the } i\text{'th sort of } M^+$$

$$\text{iff } M \models \varphi^+[s(v_j^{i+1}/a)], \text{ for all } a \text{ in the universe of the } (i+1)\text{st sort of } M,$$

$$\text{iff } M \models \forall v_j^{i+1} \varphi^+[s].$$

■

Claim 2.3: $M \text{ is special} \Rightarrow M^+ \text{ is special.}$

Proof:

M is special, therefore M is a union of an elementary chain $M_\beta, \beta < |A|$ where β is a cardinal and each M_β is

β^+ -saturated. Take M_β^+ . We want to show that M_β^+ is β^+ -saturated. Define Σ^+ to be the result of increasing the types of every formula in Σ by 1. By Lemma 2.4 every $\Sigma(v_j^i)$, a type of M_β^+ , has a type $\Sigma^+(v_j^{i+1})$ where $\Sigma^+(v_j^{i+1})$ is a type of M_β . Since M_β is β^+ -saturated, $\Sigma^+(v_j^{i+1})$ is a type of an element $b \in M_\beta$. Obviously $b \in |M_\beta^+|_1$ and realizes $\Sigma(v_j^i)$. Therefore M_β^+ is β^+ -saturated. So M^+ is a union of M_β^+ , $\beta < |A| = |A^+|$ and each M_β^+ is β^+ -saturated.

■

CHAPTER 3: ZFC, NEW FOUNDATIONS AND OTHER SETS OF AXIOMS

3.1: ZFC

Language: $\mathcal{L} = \{\in, =\}$.

Atomic formulas: $v_i = v_j, v_i \in v_j$.

Axioms of ZFC

Axiom 0: Axiom of (Set) Existence

$\exists x (x = x)$.

Axiom 1: Axiom of Extensionality

$\forall x \forall y \forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y$.

A set is entirely determined by its members.

Axiom 2: Axiom of Comprehension (called the Separation axiom in some books).

$\exists y \forall x (x \in y \leftrightarrow x \in z \wedge \varphi(x))$ for each formula $\varphi(x)$ without y free.

Note that the Comprehension Axiom is not a single axiom, but an infinite set of axioms, one for each formula $\varphi(x)$.

Axiom 3: Pairing Axiom

$\forall x \forall y \exists z (x \in z \wedge y \in z)$.

Axiom 4: Union Axiom

$\forall z \exists A \forall y \forall x (x \in y \wedge y \in z \rightarrow x \in A).$

The union of all the elements of a set z is a set.

Definition: $z \subseteq x \leftrightarrow \forall y (y \in z \rightarrow y \in x)$

Axiom 5: Power Set Axiom

$\forall x \exists y \forall z (z \subseteq x \rightarrow z \in y)$

If x is any set, the collection of all subsets of x is a set.

Definition: $p(x) = \{z : z \subseteq x\}.$

Axiom 6: Replacement Axiom

For each formula φ without y free

$\forall x \in a \exists ! y \varphi(x, y) \rightarrow \exists z \forall x \in a \exists y \in z \varphi(x, y).$

Axiom 7: Axiom of Infinity

$\exists x (\emptyset \in x \wedge \forall y \in x (y \cup \{y\} \in x)).$

Axiom 8: Axiom of Foundation

$\forall x (\exists y (y \in x) \rightarrow \exists y (y \in x \wedge \neg \exists z (z \in x \wedge z \in y))).$

Definition:

- 1) A binary relation $<$ over a set z is a partial ordering of z if:

(i) $x \not< x$ for any $x \in z$.

(ii) if $x < y$ and $y < r$, then $x < r$.

$(z, <)$ is called a partially ordered set.

2) A partial ordering $<$ of z is a linear ordering if $x < y$ or $x = y$ or $y < x$ for every $x, y \in z$.

If $<$ is a partial (linear) ordering, then the relation $\leq$ is also called a partial (linear) ordering (and $<$ is sometimes called a strict ordering).

Definition:

If z is a partially ordered set and y is a non-empty subset of z and $a \in y$, then:

1) a is the least element of y if

$$a \in y \wedge (\forall x \in y) [a \leq x].$$

2) A linear ordering $<$ of a set P is a well ordering if every non-empty subset of P has a least element.

Axiom 9: Axiom of Choice

$\forall a \exists R (R \text{ is well ordering } a).$

Note: The statement above is not the axiom of choice but is equivalent to it.

Model of ZFC

We define ordinal numbers (see below) so that

$$\alpha < \beta \text{ iff } \alpha \in \beta. \text{ Then } \alpha = \{\beta : \beta < \alpha\}.$$

Definition: A set z is transitive if

$$\forall x(x \in z \rightarrow x \subseteq z).$$

Definition: A set is an ordinal number (an ordinal) if it is transitive and well ordered by $\in$.

Ordinals in ZFC

- (i) $0 = \emptyset$ is an ordinal.
- (ii) If β is an ordinal then $\alpha = \beta + 1$ is a successor ordinal.
- (iii) If α is not a successor ordinal then $\alpha = \sup\{\beta : \beta < \alpha\}$; α is called a limit ordinal.

Definition:

The set of all level α well-founded sets, V_α , are developed by transfinite recursion as follows:

- (i) $V_0 = \emptyset$.
- (ii) $V_\alpha = \cup_{\beta < \alpha} V_\beta$ if α is a limit ordinal.
- (iii) $V_{\alpha+1} = p(V_\alpha)$.

The V_α 's have the following properties (by induction):

- (i) Each V_α is transitive.
- (ii) If $\alpha < \beta$ then $V_\alpha \subset V_\beta$.
- (iii) $\alpha \leq V_\alpha$.

Definition: For every x in some V_α , we may define the rank of x , denoted by $\text{rank}(x)$, as the least α such that $x \in V_{\alpha+1}$.

Definition: $P = \bigcup_{\alpha \in \text{ord}} V_\alpha$.

Claim 3.1.1:

The class P is a model of ZFC.

To prove this claim we need the following lemmas.

Lemma 3.1.1:

$\langle P, \epsilon \rangle$ is a transitive model.

Proof:

In order to show that P is a transitive class, it is enough to check that each V_α is transitive since $x \in P \implies x \in V_\alpha$ for some α . Thus, if $x \in V_\alpha$ then $x \in P$. The proof is by induction.

Case 1: Successor ordinals

Assume α is a successor ordinal.

For $n=0$:

$V_0 = \emptyset \quad x \in V_0 \implies x \in V_0$ because $\forall x (x \notin V_0)$.

For successor ordinal such that $\alpha \geq 1$:

Assume now that V_α is transitive, we want to prove $V_{\alpha+1}$ is

transitive. That is, we want to prove: If V_α is transitive, then $P(V_\alpha)$ is transitive. $V_\alpha \in P(V_\alpha)$ induces $V_\alpha \subset P(V_\alpha)$. Let $x \in P(V_\alpha)$, then $x \subset V_\alpha$. V_α is transitive and therefore $x \in V_\alpha \Rightarrow x \subset P(V_\alpha)$ by definition of Power set.

Case 2: Limit ordinals

Assume α is a limit ordinal.

Let $x \in V_\alpha \Rightarrow x \in V_\gamma$ for some $\gamma < \alpha$ (since $V_\alpha = \bigcup_{\gamma < \alpha} V_\gamma$)
 $\Rightarrow (x \in V_\gamma \Rightarrow x \in V_\alpha)$.

■

By φ^P we mean the meaning of φ in P . Note that the formula φ^P is obtained from φ by replacing all $\exists x$ and $\forall x$ by $\exists x \in P$ and $\forall x \in P$. If φ is quantifier free, φ^P is the same as φ .

Definition:

- a) $\exists x \in P$ and $\forall x \in P$ are called restricted quantifiers.
- b) A formula is called a restricted formula if it has no quantifiers other than restricted quantifiers.

Lemma 3.1.2: If P is a transitive class and φ a restricted formula, then for all $x_1 \dots x_n \in P$

$$(*) \quad \varphi^P(x, \dots x_n) \text{ iff } \varphi(x, \dots x_n).$$

Proof:

- (a) If φ is an atomic formula, then $\varphi \equiv (a=b)$ or $\varphi \equiv (a \in b)$.

By definition of φ^P , $(a=b) \Leftrightarrow (a=b)^P$ and
 $(a \in b) \Leftrightarrow (a \in b)^P$.

- (b) If (*) holds for φ and ψ , then it holds for $\neg\varphi$, $\varphi \wedge \psi$,
 $\varphi \vee \psi$, $\varphi \rightarrow \psi$, $\varphi \Leftrightarrow \psi$, since,

$$\begin{aligned} P \models \neg\varphi &\Leftrightarrow P \models \neg\varphi^P \Leftrightarrow (\neg\varphi)^P \\ P \models \varphi \wedge \psi &\Leftrightarrow (P \models \varphi \text{ and } P \models \psi) \Leftrightarrow \varphi^P \wedge \psi^P \\ P \models \varphi \vee \psi &\Leftrightarrow P \models (\varphi \text{ or } \psi) \Leftrightarrow \varphi^P \vee \psi^P \\ P \models \varphi \rightarrow \psi &\Leftrightarrow (P \models \varphi \text{ implies } P \models \psi) \Leftrightarrow \varphi^P \rightarrow \psi^P \Leftrightarrow (\varphi \rightarrow \psi)^P \\ P \models (\psi \Leftrightarrow \varphi) &\Leftrightarrow P \models \psi \text{ iff } P \models \varphi \Leftrightarrow (\psi^P \Leftrightarrow \varphi^P) \Leftrightarrow (\psi \Leftrightarrow \varphi)^P \end{aligned}$$

- (c) Let φ be the formula $(\exists u)\psi(u, y)$ and assume that
 (*) is true for ψ . We want to show that (*) is true
 for φ .

($\Rightarrow$) If $\exists u \in P \psi^P(u, y)$ then $\exists a \in P$ such that $\psi^P(a, y)$.

By the induction assumption it is equivalent to
 $P \models \psi(a, y)$. Therefore $P \models \exists u \psi(u, y)$. The proof for
 $\forall u \in x$ is similar.

($\Leftarrow$) $P \models (\exists u)\psi(u, y) \Rightarrow \exists a \in P P \models \psi(a, y)$. By the induction
 assumption $\psi^P(a, y)$. Then $\exists u \in P \psi^P(a, y)$.

Therefore $(\exists u \psi(a, y))^P$

■

We are now ready for the proof of claim 3.1.1.

I. $\langle P, \epsilon \rangle \models$ Extensionality.

$[(\forall u \in x) (u \in y) \wedge (\forall u \in y) (u \in x)] \rightarrow x = y$ is a restricted formula.

II. $\langle P, \epsilon \rangle \models$ Pairing

Given $a, b \in P$ let $c = \{a, b\}$. Therefore, $\exists V_\alpha$ and V_β such that $a \in V_\alpha$ and $b \in V_\beta$. Therefore, $\exists V_\gamma$ such that $c \in V_\gamma \Rightarrow c \in P$. Since each V_α is transitive $c \in V$ and $P \models [a \in c \wedge b \in c]$ since the formula [...] is restricted and holds.

III. $\langle P, \epsilon \rangle \models$ Separation

Given $x, y \in P$, let $z = \{w \in x : P \models \varphi(w, y)\}$. Since $z \subseteq x$ and $x \in P$ we have $z \in P$ and $P \models \forall w (w \in z \longleftrightarrow w \in x \wedge \varphi(w, y))$.

IV. $\langle P, \epsilon \rangle \models$ Union

Given $x \in P$, let $y = \cup x$. $y \in V_\alpha$ for some $\alpha \in \text{ord}$ (since if $x \in V_\alpha$ then $\cup x \in V_\alpha$). Therefore, $y \in P$ and since each V_α is transitive $y \in P$. We get, $P \models (\forall z \in x) (\forall u \in z) u \in y$ since [...] is a restricted formula, saying $y = \cup x$.

V. $\langle P, \epsilon \rangle \models$ Power Set

Given $x \in P$, let $y = p(x)$. Let, $y \in P$. The formula $\forall u [u \in y \longleftrightarrow u \subseteq x]$ is not restricted because the quantifier $\forall u$ is not a restricted quantifier. But $[u \in y \longleftrightarrow (\forall z \in u) z \in x]$ is restricted because the only quantifier in it is $\forall z \in u$ which is a restricted quantifier, and for every $u \in P$, $P \models [u \in y \longleftrightarrow (\forall z \in u) z \in x]$ since [...] is true for all u . We get $\forall u [u \in y \longleftrightarrow u \subseteq x]$.

VI. $\langle P, \epsilon \rangle \models$ Infinity

The formula $P \models \exists y [\emptyset \in y \wedge (\forall x \in y) x \cup \{x\} \in y]$ contains defined notions $\{\}$, $\cup$, $\emptyset$ so first we should eliminate these symbols and replace the formula by an equivalent one that uses only ϵ and $=$. But we have already proved that $\{a, b\}^P = \{a, b\}$, $\cup^P = \cup$, $\emptyset^P = \emptyset$ where $\{a, b\}^P$, $\cup^P$, $\emptyset^P$ denote Pairing, Union and empty set are in the model P and therefore they can be written using only ϵ and $=$. Since $\omega \in P$, we let $y = \omega$. By the definition of ω and by the fact that ω is ordinal we get that $P \models \exists y [\emptyset \in y \wedge (\forall x \in y) x \cup \{x\} \in y]$ holds for this y .

VII. $\langle P, \epsilon \rangle$ Replacement

Given $x \in P$, assume that $P \models \forall x \forall y \forall z (\phi(x, y, p) \wedge \phi(x, z, p) \longrightarrow y = z]$. Thus $F = \{(x, y) \in P : P \models \phi(x, y, p)\}$ is a function and we let $y = F[x]$. Since $y \in P$ and P is transitive we have $y \in P$, and for every $y \in P$, $P \models y \leftrightarrow (\exists x) \phi(x, y, p)$.

VIII. $\langle P, \epsilon \rangle \models$ Foundation

Note: The axiom of Foundation is used to show that

$$P = \bigcup_{\alpha \in \text{ord}} V_\alpha.$$

If $s \in P$ and $s \neq \emptyset$, then let $x \in s$ be of minimum rank; then $s \cap x = \emptyset$, therefore $(\exists x \in s) s \cap x = \emptyset$. Since $[s \neq \emptyset \longrightarrow (\exists x \in s) s \cap x = \emptyset]$ has only restricted quantifiers we get $P \models [s \neq \emptyset \longrightarrow (\exists x \in s) s \cap x = \emptyset]$.

VIII. $\langle P, \epsilon \rangle \models$ Choice

What we actually want to prove is: If $A \in P$, then A can be well ordered iff $(A \text{ can be well ordered})^P$.

If A can be well ordered, let $R \subset A \times A$ well order A . $A \times A \subset P$ so $R \in P$ and $(R \text{ well orders } A)^P$, therefore $(A \text{ can be well-ordered})^P$.

Conversely, if $(A \text{ can be well-ordered})^P$ fix $R \in P$ such that $(R \text{ well-ordered } A)^P$. R totally orders A and every non zero subset of A in P has an R -least member. But every subset of A is in P , so R well orders A .

(Claim 3.1.1) ■

3.2 Theory of Types (TT)

Definition of TT - Theory of Types

Language: $\mathcal{L} = \{ \epsilon_i, =_i \}$.

Logical symbols: $\forall_i, \exists_i, \neg, \rightarrow, \wedge, \vee$.
 v_j^i for all $i, j \geq 0$ natural numbers.

Atomic formula: $v_j^i =_i v_k^i$ for all i, j, k .

$v_j^i \epsilon_{i+1} v_k^{i+1}$ for all i, j, k .

where the superscript $i \in I$ means the sort of the variable

Formulas:

- (1) Any atomic formula is a formula.
- (2) If α, β are formulas, then so are $(\neg \alpha), (\alpha \rightarrow \beta)$.
- (3) If α is a formula, $\forall_i v_j^i \alpha$ is a formula.

Axioms of TT

Axiom 1: Axiom of Extensionality

$$\forall v_j^{i+1} \forall v_k^{i+1} (v_j^{i+1} =_{i+1} v_k^{i+1} \leftrightarrow \forall v_m^i (v_m^i \in_{i+1} v_j^{i+1} \leftrightarrow v_m^i \in_{i+1} v_k^{i+1})).$$

Axiom 2: Axiom of Abstraction

$$\exists v_j^{i+1} \forall v_k^i (v_k^i \in_{i+1} v_j^{i+1} \leftrightarrow \varphi)$$

Where φ is a formula of the language, v_j^{i+1} does not occur in φ .

Model of TT

Let x be a set in the model of ZFC. Define P' to be

$$P' = \langle x, p(x), p^2(x), \dots \rangle.$$

Claim 3.2.1:

P' is a model for TT.

Proof:

$P' \models \text{Extensionality}$

Let $A_1, A_2 \in p^{i+1}(x)$. Since $p^{i+1}(x) = p(p^i(x))$ we have $A_1, A_2 \subseteq p^i(x)$. Assume that for every $a \in p^i(x)$ we have $a \in A_1 \leftrightarrow a \in A_2$. We get $A_1 =_{i+1} A_2$. Therefore $P' \models \text{Extensionality}$.

$P' \models \text{Abstraction}$

Given $a \in p^i(x)$ we let $y = \{a : P' \models \varphi[a]\}$. Since $y \subseteq p^i(x)$ we get $y \in p^{i+1}(x)$. Therefore $\exists z^{i+1} \forall x^i (\varphi(x^i) \leftrightarrow x^i \in z^{i+1})$, and

$P' \models \text{Abstraction.}$

■

3.3 New Foundation (NF)

Definition of NF

In NF, we allow the use of unstratified formulas in the language but limit the axiom of Abstraction to a stratified formula. NF is therefore defined as follows:

Language: $\{\in, =\}.$

Atomic formula: $v_i = v_j$ for all i, j ,
 $v_i \in v_j$ for all i, j .

Formulas:

- (1) Any atomic formula is a formula.
- (2) If α, β are formulas, then so are $(\neg\alpha), (\alpha \rightarrow \beta).$
- (3) If α is a formula, $\forall x\alpha$ is a formula, when x is any variable v_i .

Axioms of NF

Axiom 1: Axiom of Extensionality

(Ext) $\forall x \forall y \forall z (z \in x \leftrightarrow z \in y) \rightarrow x = y.$

Axiom 2: Axiom of Abstraction

(Abs') $(\exists x)(\forall y)(y \in x \leftrightarrow \phi)$ where ϕ is stratified and does not contain x .

Note: A model for NF has not been found assuming only the consistency of ZFC or some extension of ZFC by strong axioms of Foundation.

Adding the axiom of Infinity to NF

Obviously, the axiom of Foundation is inconsistent with NF since $\exists x(x \in x)$ is a consequence of NF (but not of ZFC) since if we take $\exists x \forall y (y \in x)$ in NF and replace x by v we get $v \in v$. This is why in NF it is natural to add the following axiom of Infinity.

Axiom of Infinity:

(Inf) There exists a non-empty linear ordering $\mathbb{R}$ which has no minimum element.

An alternative version of this axiom of Infinity can be given in terms of cardinal numbers.

Definition:

$\bar{x}$, the cardinal number of a set x , is the collection of sets which are in one to one correspondence with x .

Claim 3.3.1: $\bar{x}$ is a set.

Proof:

By Abstraction. Let ϕ formalize " $\exists f$ (f is a one to one correspondence between x and y)". Then:

$$\exists \bar{x} \forall y (y \in \bar{x} \leftrightarrow \phi).$$

■

Definitions:

- 1) R is a finite total ordering if R and its inverse R^{-1} are both well orderings.
- 2) If, for some finite R , n is the cardinal of R then n is called a finite cardinal.

The axiom of Infinity is then equivalent to:

$$\exists x (\bar{x} \text{ is not finite}).$$

Note that $\neg \text{Inf}$ implies the existence of a greatest finite cardinal, which we denote by $\bar{v}$. In particular, it follows (from $\neg \text{Inf}$) that V is well orderable and hence $\neg \text{Inf}$ implies the axiom of Choice. But Specker, in his article "The Axiom of Choice in NF" [11] proved that AC is refutable in NF. Therefore, Inf is provable in NF.

3.4 Definition of NFU

NFU is NF with the axiom of Extensionality being weakened to hold for non-empty sets only as follows:

(Ext') $\forall x \forall y \exists z (z \in x) \wedge \forall z (z \in x \leftrightarrow z \in y) \longrightarrow x = y.$

Note that Jensen [10] proved that NFU is essentially weaker than NF and not just a trivial variation of NF as Quine himself thought [7]. "U" is for urelement, which is what individuals with no member are called.

CHAPTER 4. PROOF OF EQUICONSISTENCY OF NF AND TT*

We prove the equiconsistency of NF and TT*, an extension of TT defined as follows:

Definition of TT*

Let φ be a sentence in a many sorted language with many sorted (ϵ_{i+1} and $=_i$) relations. A new sentence φ^+ can be obtained by raising all sort indices by 1. A sentence of the form $\varphi \longleftrightarrow \varphi^+$ is called an axiom of ambiguity. TT* is obtained by adding to TT all axioms of ambiguity.

Let M be a structure of the language $\mathcal{L} = \{\epsilon, =\}$.

Define: M^* (as it was defined in chapter 1) a many sorted structure such that:

- $M^* = \omega \times M$.
- $M^*_i = \{i\} \times M$.
- $(i, x) \epsilon_{i+1}^{M^*}(j, y)$ iff $j = i+1$ and $x \epsilon^M y$.

Claim 4.1: $\langle M, \epsilon \rangle \models \text{NF}$ iff $\langle M, \epsilon \rangle^* \models \text{TT}$.

Proof:

TT is written in many-sorted language while NF is written in one-sorted language. Both are based on the axioms of Extensionality and Comprehension. The axiom of Extensionality does not have free variables. In the axiom of Comprehension Φ is already stratified. Therefore we

have a translation functions from one set of axioms to another when we do not have to worry about the free variables. Assuming the translation function $.^b$ defined in Ch. 1 and by claim 1.2.1, given any axiom ψ of NF it is possible to find a pairing $(m,n) \longrightarrow \langle m,n \rangle$ that gives a $.^b$ function under which $\psi = \varphi^b$ for some axiom φ of TT and vice versa.

Therefore, $\langle M, \epsilon \rangle \models \text{NF}$ iff $\langle M, \epsilon \rangle^* \models \text{TT}$.

■

Claim 4.2: Let φ be a sentence of $\mathcal{L}$ then

$$M^* \models \varphi^+ \text{ iff } (M^*)^+ \models \varphi.$$

Proof:

The proof is an immediate consequence of Lemma 2.5.

■

Claim 4.3: $\langle M, \epsilon \rangle^* \models \text{TT}$ iff $\langle M, \epsilon \rangle^* \models \text{TT}^*$.

Proof:

$(\Leftarrow)$ Trivial ($\text{TT} \subset \text{TT}^*$).

$(\Rightarrow)$ Let $(M^*) \models \text{TT}$. Let φ be a sentence of $\mathcal{L}$.

First we show that $(M^*) \cong (M^*)^+$ (recall that $M^* = \langle M_0, M_1, \dots \rangle$ and $(M^*)^+ = \langle M_1, M_2, \dots \rangle$). Define $h: (M^*)^+ \longrightarrow (M^*)$ by setting $h(i+1, a) = (i, a)$ (h exists since $M^*_i = M^*_{i+1}$). We now prove that h is an isomorphism.

onto: $\forall (i,a) \in M^*$ there is $(i+1,a) \in (M^*)^+$ such that $h(i+1,a) = (i,a)$.

1-1: Let $(i,a) \stackrel{M^*}{=} (i,b)$. By definition of h there are $(i+1,a)$ and $(i+1,b)$ such that $h(i+1,a) = (i,a)$ and $h(i+1,b) = (i,b)$. From $(i,a) \stackrel{M^*}{=} (i,b)$ we get $a=b$ and therefore $h(i+1,b) = (i,a)$. By definition of h we get $(i+1,a) = (i+1,b)$.

homomorphism: $(i+1,a) \in (M^*)^+ (i+2,b)$ so we have (i,a) and $(i+1,b)$ in M^* such that $h(i+1,a) = (i,a)$ and $h(i+2,b) = (i+1,b)$. From the definition of M^* and $(M^*)^+$ we get $(i,a) \in^{M^*} (i+1,b)$.

Since onto 1-1 and homomorphism exists we get $(M^*) \cong (M^*)^+$.

By Claim 4.2 we have $M^* \models \varphi^+ \longleftrightarrow (M^*)^+ \models \varphi$.

$(\implies)$ $M^* \models \varphi^+ \implies (M^*)^+ \models \varphi$. But $M^* \cong (M^*)^+$ and therefore $M^* \models \varphi$.

$(\impliedby)$ Suppose $M^* \not\models \varphi$. Now $M^* \cong (M^*)^+$, so $(M^*)^+ \not\models \varphi$, so $M^* \not\models \varphi^+$. Therefore we get $M^* \not\models (\varphi \longleftrightarrow \varphi^+)$.

■

Let N be a model of NF , then $N^* = \langle N, N, N, \dots \rangle$.

Theorem 4.1 (Specker's Theorem): Suppose ψ is a stratified sentence of the one-sorted language of NF . Fix a one-to-one pairing $(m, n) \rightarrow \langle m, n \rangle$, hence a $\cdot^b$ -function $\psi = \varphi^b$ as defined on p.11, where φ is a (many-sorted) sentence in the language of TT . Then $NF \vdash \psi$ iff $TT^* \vdash \varphi$.

Proof:

($\Leftarrow$) Suppose $TT^* \vdash \varphi$ then by soundness $TT^* \models \varphi$. Suppose $N \not\models NF$. By claims 4.1, 4.2, 4.3 $N^* \not\models TT^*$, so $N^* \not\models \varphi$. Then $N \not\models \psi$. Since N was arbitrary, by completeness $NF \not\vdash \psi$.

($\Rightarrow$) Suppose $NF \vdash \psi$ then by soundness $NF \models \psi$. Suppose $M \not\models TT^*$. By Lemma 4.1 (Specker's Lemma) below there exists $N \models NF$ such that $M \equiv N^*$. Now $N \models \psi$, so $M \equiv N^* \models \varphi$. Since M was arbitrary then by completeness $TT^* \vdash \varphi$.

■

Lemma 4.1 (Specker's Lemma): $M \models TT^*$ iff there exists a structure N such that $N \models NF$ and $M \equiv N^*$.

Proof:

($\Leftarrow$) Suppose $N \models NF$ then (by Claims 4.1, 4.2, 4.3) $N^* \models TT^*$ so by setting $M = N^*$ we get $M \models TT^*$.

($\Rightarrow$) Suppose $\langle M, \epsilon^M \rangle \models TT^*$.

By Proposition 2.1 we can construct $\langle M', \epsilon^{M'} \rangle$, such

that $\langle M, \epsilon^M \rangle \leq \langle M', \epsilon^{M'} \rangle$ is special.

Let $\bar{M} = (M')^+$.

Claim 4.4: $\bar{M} \equiv M'$.

Proof:

($\Rightarrow$) Let $M' \models \varphi$. Since $M \leq M'$ and $M \models TT^*$ we get $M' \models \varphi^+$. By Claim 4.2 $\bar{M} \models \varphi$.

($\Leftarrow$) Let $\bar{M} \models \varphi$. By Claim 4.2 we get $M' \models \varphi^+$. But $M \models (\varphi \leftrightarrow \varphi^+)$ and $M \leq M'$ so $M' \models \varphi$.

■

Claim 4.5: $\bar{M}$ is special.

Proof:

An immediate consequence of Claim 2.3.

■

$\bar{M}$ and M' are special and $\bar{M} \equiv M'$ so by Theorem 2.2 $\bar{M} \cong M'$.

Let f be an isomorphism $f: \langle \bar{M}', \epsilon^{\bar{M}'} \rangle \rightarrow \langle M', \epsilon^{M'} \rangle$. Define a structure $\langle N, \epsilon^N \rangle$ as follows:

$N = M'_0$ where $x \epsilon^N y$ iff $x \epsilon^{M'} f(y)$.

Define $g: \omega \times N \rightarrow M'$ by $g(i, x) = f^i(x)$ where $f^0(x) = x$ and $f^{i+1}(x) = f(f^i(x))$.

Claim 4.6: g is an isomorphism of $\langle N, \epsilon \rangle^*$ and $\langle M', \epsilon \rangle$.

Proof:

Onto: We want to show that for each $x \in M'_i$ there exists $y \in N^*_i$ such that $g(i, y) = x$. Let $x \in M'_i$ then $\exists y \in \bar{M}$ such that $f^i(y) = x$. By definition of g we get $g(i, y) = f^i(y) = x$.

1-1 Let $g(i, x) = g(i, y)$. By definition of g , $f^i(x) = f^i(y)$ so $x = y$ because f is one-to-one.

Homomorphism: We must see that $(i, x) \in^{N^*} (i+1, y)$ iff $f^i(x) \in^{M'} f^{i+1}(y)$. Indeed,
 $(i, x) \in^{N^*} (i+1, y) \Leftrightarrow x \in^N y \Leftrightarrow x \in^{M'} f(y) \Leftrightarrow f^i(x) \in^{M'} f^{i+1}(y)$.

■

We can now complete the proof of Specker's Lemma. By claims 4.4 - 4.6 we have $N^* \cong M' \cong M$, therefore since $N^* \equiv M \models \text{TT}$ we have that $N \models \text{NF}$.

(Lemma 4.1 (Specker's Lemma)) ■

We have proved Specker's Lemma from which Specker's Theorem (Theorem 4.1) follows, thus proving the equiconsistency of NF and TT*.

CHAPTER 5: ON THE CONSISTENCY OF NFU

Theorem 5.1: If TT (respectively, TT+Inf, TT+Inf+AC) is consistent so is NFU (respectively, NFU+Inf, NFU+Inf+AC).

Proof:

Let TTU be obtained by restricting the axiom of Extensionality (Ext) to the non-empty sets.

$$(Ext') \quad \exists z^i (z^i \in x^{i+1}) \wedge \forall z^i (z^i \in x^{i+1} \leftrightarrow z^i \in y^{i+1}) \rightarrow x^{i+1} \equiv y^{i+1}.$$

Let $M = \langle M_0, M_1, \dots \rangle$ be a model of TTU. For each increasing sequence of natural numbers $i_0 < i_1 < \dots$ let us define a new model $N = \langle M_{i_0}, M_{i_1}, \dots \rangle$ as follows: For $x^k \in M_{i_k}, y^{k+1} \in M_{i_{k+1}}$ we have

$$N \models (x^k \in y^{k+1}) \text{ iff}$$

$$M \models (\exists z^{i_{k+1}}) (x^k \in z^{i_{k+1}} \wedge y^{k+1} = \{\dots \{z^{i_{k+1}}\} \dots\}), \text{ where the}$$

singleton operation is iterated $i_{k+1} - i_k - 1$ times.

To prove Theorem 5.1 we need the following Lemmas:

Lemma 5.1: N is a model of TTU.

Proof:

1) $N \models Ext'$

We want to see that

$$\exists x^k (x^k \in s^{k+1}) \wedge \forall x^k (x^k \in s^{k+1} \leftrightarrow x^k \in t^{k+1}) \rightarrow s^{k+1} = t^{k+1}.$$

Assume $\exists x^k (x^k \in s^{k+1}) \wedge \forall x^k (x^k \in s^{k+1} \leftrightarrow x^k \in t^{k+1})$. Pulling

$\exists x^k (x^k \in s^{k+1}) \wedge \exists x^k (x^k \in t^{k+1})$ back to M, we have $y^{i_{k+1}}$ and

$z^{i_{k+1}}$ such that $s^{i_{k+1}} = \{\dots\{y^{i_{k+1}}\}\dots\}$ and $t^{i_{k+1}} = \{\dots\{z^{i_{k+1}}\}\dots\}$ and $y^{i_{k+1}} \neq \emptyset$, $z^{i_{k+1}} \neq \emptyset$ where the singleton operation is iterated $i_{k+1} - i_k - 1$ times. So in M we get:

$$\forall x^{i_k} (x^{i_k} \in z^{i_{k+1}} \leftrightarrow x^{i_k} \in y^{i_{k+1}}).$$

By Ext' in M we get $z^{i_{k+1}} = y^{i_{k+1}}$. Therefore by induction and definition of ϵ^N we get $s^{k+1} = t^{k+1}$ in N and Ext' holds in N .

2) $N \models$ Abstraction

We want to see that

$$\exists y^{k+1} \forall x^k (x^k \in y^{k+1} \leftrightarrow \varphi) \text{ holds in } N.$$

Pulling this back to M , we are required to verify the formula

$$(*) \quad (\exists y^{i_{k+1}} \forall x^{i_k} (\exists z^{i_{k+1}} (x^{i_k} \in z^{i_{k+1}} \wedge y^{i_{k+1}} = \{\dots\{z^{i_{k+1}}\}\dots\}) \leftrightarrow \varphi^N))$$

where φ^N is defined by recursion on φ as follows:

$$\begin{aligned} \text{Let } (x^k \in y^{k+1})^N, \\ (\neg \varphi)^N &= \neg \varphi^N, \\ (\varphi \rightarrow \psi)^N &= \varphi^N \rightarrow \psi^N, \\ (\forall v_j^k \varphi)^N &= \forall v_j^{i_k} \varphi^N. \end{aligned}$$

Then $M \models \varphi^N[a_1, \dots, a_n]$ iff $N \models \varphi[a_1, \dots, a_n]$ for all $a_1, \dots, a_n$ chosen from the appropriate types in M .

By abstraction in M, there exists $z^{i_{k+1}}$ such that $\forall x^{i_k} (x^{i_k} \in z^{i_{k+1}} \leftrightarrow \varphi^N)$. Since $z^{i_{k+1}}$ is a set, by Abstraction there exists a set $y^{i_{k+1}} = \{.. \{z^{i_{k+1}}\} ..\}$ in M and this set witnesses (*). By definition of N $\exists y^{i_{k+1}} \forall x^k (x^k \in y^{i_{k+1}} \leftrightarrow \varphi)$ holds in N.

■

Such a model N which is defined by the assumption of the existence of M is called "a model extracted from M". We shall write $N \leq M$ to denote that N is extracted from M. $\leq$ is reflexive and transitive.

Definition: For any model M of TTU and any stratified sentence φ we say that

- 1) M forces φ when φ is true in every model extracted from M.
- 2) M decides φ when M forces either φ or $\neg\varphi$.

Lemma 5.2 (Extracted Lemma): Given any model M of TTU and any sentence φ , there is a model $N \leq M$ which decides φ .

Proof:

Let k be greater than all the types appearing in φ . Define a partition G_1, G_2 of $[\omega]^{k+1}$ as follows:

$$G_1 = \{i_0 < i_1 \dots < i_k \mid \langle M_{i_0}, M_{i_1}, \dots, M_{i_k} \rangle \models \varphi\},$$

$$G_2 = \{i_0 < i_1 \dots < i_k \mid \langle M_{i_0}, M_{i_1}, \dots, M_{i_k} \rangle \models \neg\varphi\}.$$

By Ramsey's theorem (below), take an infinite set x of natural numbers $i_0 < i_1 < \dots < i_n < \dots$ such that $[x]^{k+1} \subseteq G_1$ or $[x]^{k+1} \subseteq G_2$ and set $N = \langle M_{i_0}, M_{i_1}, \dots \rangle$. In the first case $([x]^{k+1} \subseteq G_1)$ N forces φ and in the second case N forces $\neg\varphi$. ■

TTU is formulated in many-sorted language, NFU in a one-sorted language. Let Σ' be the set of sentences in many-sorted language translated to one-sorted language by fixing one-to-one pairing $(m, n) \rightarrow \langle m, n \rangle$ to define a function $\cdot^b$ as defined in Chapter 1.

Lemma 5.3: Let M be a model of TTU and let $\Sigma \subseteq \Sigma'$ be the set of all sentences which are forced by M . Then $\text{NFU} + \Sigma^b$ is consistent.

Proof:

Let TTU^* be defined by adding to TTU the axioms of ambiguity. An immediate generalization of Specker's result is that a many-sorted theory T^* generated by some set of axioms plus the axioms of Ambiguity and the one-sorted theory T_1 generated by the same axioms have the same stratified theorems. TTU and NFU are obtained by restricting the axiom of extensionality to the non empty sets. Therefore TTU and NFU have the same stratified consequences. Hence it is enough to prove the consistency of $\text{TTU}^* + \Sigma$. By compactness, it suffices to prove this for

any finite set of stratified sentences. Let $A \subseteq \text{TTU} * \cup \Sigma$ be a finite set of sentences such that $A = A_1 \cup A_2$ where $A_1 \subseteq \text{TTU} \cup \Sigma$ and $A_2 = \{\varphi_i \leftrightarrow \varphi_i^+ \mid i < n\}$. Since $A_1 \subseteq \text{TTU} \cup \Sigma$, for every model $N \models M$ $N \models A_1$. For A_2 , by the extraction lemma, since M is a model of TTU there is a model N_0 , $N_0 \models M$ such that N_0 decides φ_0 . Assume that N_0 forces φ_0 . Then, for every model $N_1 \models N_0$ $N_1 \models \varphi_0$. Since $N_1^+ \models N_1$ $N_1^+ \models \varphi_0$. Then by Lemma 2.3 $N_1 \models \varphi_0^+$. Similarly, if N_0 forces $\neg \varphi_0$ then by the same argument $N_1 \models \neg \varphi_0^+$. Using extraction n times we get the consistency of $\text{TTU} * + \Sigma + A$.

■

At this point we can conclude that if TT is consistent then NFU is consistent as follows:

By definition of TTU it follows that
 $\text{con}(\text{TT}) \Rightarrow \text{con}(\text{TTU})$. By lemma 5.3 we get
 $\text{con}(\text{TTU}) \Rightarrow \text{con}(\text{NFU})$. Therefore $\text{con}(\text{TT}) \Rightarrow \text{con}(\text{NFU})$.

We can now complete the proof of Theorem 5.1.

The axioms of Infinity and Choice are stratified.

Therefore from Lemma 5.3 we can conclude:

If N exists such that $N \models \text{TTU} + \text{Inf}$ and $N \models \text{TTU} + \text{Inf} + \text{AC}$ then $\text{NFU} + \text{Inf}$ and $\text{NFU} + \text{Inf} + \text{AC}$ are consistent.

(Theorem 5.1) ■

Theorem 5.2 - Ramsey's Theorem:

Let I be an infinite set and let $n < \omega$. Suppose $f: [I]^n \rightarrow \{0,1\}$ then there is an infinite subset $J \subseteq I$ such that f is constant on $[J]^n$.

Proof:

We can assume without loss of generality that I is countable (and therefore well ordered) since otherwise we can take an infinite countable subset of I and use it in our proof. Denote by $<$ the order in I .

The proof is by induction on n . For $n=1$ the theorem is trivial since either the set $K_0 = \{a \in I : f(a) = 0\}$ or the set $K_1 = \{a \in I : f(a) = 1\}$ must be infinite and we can take J to be that infinite set.

Assuming now that the theorem holds for n we prove it for $n+1$. We do so by constructing the subset J of I as follows: We start by defining for each $i < \omega$ an infinite subset of I , K_i , and an element $b_i \in K_i$ such that for all i , $K_{i+1} \subset K_i$ and $b_i < b_{i+1}$. We start with $K_0 = I$. Obviously K_0 is infinite. Given K_i we define b_i to be the least element of K_i . Let $L_i = \{b \in K_i : b_i < b\}$. L_i is infinite since K_i is infinite and obviously $L_i \subset K_i$. For each n -tuple $a_{i_1} < \dots < a_{i_n}$ of elements in L_i we define

$$f_i(a_{i_1}, \dots, a_{i_n}) = f(b_i, a_{i_1}, \dots, a_{i_n}). \quad (f_i \text{ is well defined})$$

since f is defined for all $(n+1)$ -tuples in I and $b_i \in L_i$. By the induction hypothesis there is an infinite subset K_{i+1} of L_i such that $f_i([K_i]^n) = l_i \in \{0,1\}$. By repeating this construction we build the sequence K_i, b_i, f_i, l_i for all $i < \omega$.

We now define the subset J of I as follows. Since for all $i < \omega$, $f_i([K_i]^n) \in \{0,1\}$ then (by the same argument as for $n=1$ above) there is an infinite sequence $i_0 < i_1 < \dots$ of natural numbers such that for all $j < \omega$ $f_{i_j}([K_{i_j}]^n)$ is a constant $l \in \{0,1\}$. We define J to be the set $\{b_i : f_i([K_i]^n) = 1\}$.

To complete the proof we need to show that $f([J]^{n+1}) = 1$. Indeed, let $i_0 < \dots < i_n$ be any $(n+1)$ -tuple of natural numbers such that $b_{i_j} \in J$, $j=0, \dots, n$. By the construction $b_{i_0} < \dots < b_{i_n}$. Also since $K_{i+1} \subset K_i$ for all $i < \omega$ we have $b_{i_1}, \dots, b_{i_n} \in K_{i_0+1}$. But by the construction $f(b_{i_0}, b_{i_1}, \dots, b_{i_n}) = f_{i_0+1}(b_{i_1}, \dots, b_{i_n}) = 1$.

■

REFERENCES

1. Aczel, P. (1988). CSLI Lecture Notes No. 14, Leland Stanford Junior University. Palo Alto.
2. Barwise, J. (1977). Handbook of Mathematical Logic. New York: North-Holand Publishing Company.
3. Chang, C.C. & Keisler, H.J. (1973). Model Theory. New York: North-Holand Publishing Company.
4. Kunen, K. (1980). Set Theory - An Introduction to Independence Proofs. New York: North-Holand Publishing Company.
5. Jech, Thomas (1978). Set Theory. San Diego: Academic Press, Inc., Harcourt Brace Jovanovich Publishers.
6. Nagel, Ernest (1958). Godel Proof. New York: University Press.
7. Quine, W. (1953). From a Logical Point of View. Boston: Harvard University Press.
8. Tieszen, Richard (1992). Symbolic logic: A computational approach. San Jose.
9. Boffa, M. (1977). The consistency problem for NF. The Journal of Symbolic Logic, 42, 215-220.
10. Jensen, R. (1968/69). On the consistency of a slight (?) modification of Quine's new foundation. Synthese, 19, 215-220.
11. Specker, E.P. (1953), "The axiom of Choice in Quine's New Foundation for Mathematical Logic", proceedings of the National Academy of Science, 39, 972-975.
12. J.L. Bell, A.B. Slomson (1971), Models and Ultraproducts, North Holland Publ. Co.