AXIOMATIC SET THEORY

by Saunders Mac Lane

University of Chicago 1964

Notes prepared by

Murray Schacher and Richard Shaker

PROPERTY OF
R. Schultz

ACKNOWLEDGEMENT

The following assisted in the compilation of these notes: Thomas Heise, Richard Larson, Gerson Levin, James McNamara, Clark Overley, Max Plager and Lance Small.

R. Schultz

ERRATA

Page 2, line 4: Should read: 2) A set C of constants of type I and type II.

Page 3, line 10: (Include constants here instead of in (d) below.)

Page 3, line 23: 1) should read: If γ is an n- place predicate of type I (II) and $\tau_1, \ldots, \tau_n$ wft of type I (II), then $\gamma(\tau_1, \ldots, \tau_n)$ is a wff.

Page 8: Add the statement: The reader should verify that $\{x, y\}$ as defined has exactly x and y as elements.

Page 10, line 10: A3) should read: (power class) ...

Page 14, line 13: Should read: computing ...... string $\gamma\alpha\gamma\alpha\gamma\alpha A$.)

Page 14, line 15: Should read: .... $\langle x_{i_1}, \ldots, x_{i_{n-1}} \rangle\} \in A$.

Page 15, line 15: Reads: Using 3.5... . Should read: Using 4.2 .....

Page 15, line 19: Reads: If α is a primitive wff... . Should read: If α is a ppf .... .

Page 15, line 25: Should read: ...... the class of .....

Page 16: Change all τ's to ζ .

Page 16, line 14: Reads .... reduced to a term ... . Should read ... reduced to a formula ..... .

Page 16, line 17: Reads; Theorem 3.6. Should read: Theorem 4.3.

Page 16, line 22: Reads $(\exists x)\beta$. Should read: $(\exists x_1)\beta$.

Page 17, line 1: Should read: .... on α in 4.3.

Page 21, line 5: Reads $A \subset B$. Should read: $A \in B$.

Page 23, line 19: (5) should read: If $u \subset \omega$, and $\emptyset \in u$ .....

Page 31, line 4 : Should read: .... on complete classes of ordinals.

next page Page 36, line 15: Reads: .... $\alpha = \mathcal{P}u$. Should read: $\alpha = \overline{\overline{\mathcal{P}u}}$

Page 34, line 3: Should read: ... ~~We define F ar~~ $F'\alpha = C'(a - R(F|\alpha))$ if $a - R(F|\alpha) \neq \emptyset$. If $a - R(F|\alpha) = \emptyset$ (i.e., $R(F|\alpha) \supset a$), define $F'\alpha = a$. There...

Page 35, line 17: Should read: ... maps on well-ordered sets, Page 37, line 9: Should read: .... there is a least ordinal, ...

Page 37, line 22: Should read: .... $\alpha + \beta = \overline{\overline{\alpha \cup \beta}}$ and $\alpha \cdot \beta = \overline{\overline{\alpha \times \beta}}$.

Page 38, line 17: "df" means "by definition" .

Page 39, line 7 : Should read : function with domain x ....

Page 40, line 7: Amend to read: 5.1 and the class theorem, and ......

Page 40, line 20: Should read: $t = \{\omega, \mathcal{P}\omega, \mathcal{P}(\mathcal{P}\omega), \ldots\}$......

Page 44, line 6 : Should read: $R(\alpha) = \bigcup_{\beta < \alpha} \mathcal{P}R(\beta)$; ...

Page 45, line 9: Should read: ... the whole of Ω follows easily from the fact that Ω is a proper class (by 9.5 and 9.6).

Table of Contents

Chapter 1. Introduction: The Language of the Theory

In our axiomatic set theory we will consider certain undefined objects and present a list of axioms ascribing certain relations between the objects. Because of the imprecision of ordinary English, we will use a specialized language for our theory. As a first step we will describe this language and the grammar of its usage. In our description of the language we will utilize the reader's understanding of the words "set" and "function" merely to shorten the wordage. Suffice it to say that it is possible to describe the language without using these undefined terms, but we will follow the shorter path and use them.

We assume that the reader knows the basic facts of the propositional calculus. (A proposition is a sentence or statement with a truth value.) By convention, we adopt the following symbols:

(1).	$\sim$	"not"	$\sim p$	not p
(2).	$\&$	"and"	$p \& q$	p and q
(3).	$\vee$	"or"	$p \vee q$	p or q
(4).	$\Longrightarrow$	"implies"	$p \Longrightarrow q$	p implies q
(5).	(x)	"all"	$(x) \ldots x \ldots$	For all x such that ...
(6).	$\exists(x)$	"there exists"	$\exists(x) \ldots x \ldots$	There exists x such that ...

We label the operations (1) - (4) with the term "basic logical connectives."

The language we will use is a particular example of a class of languages known as <u>first order predicate calculi</u>. We will now characterize the structure of this class of languages. A <u>first order predicate calculus</u> has the following "elements" :

1) There are two "sets" of variables a and A. The elements of a will be called variables of type I ; those of A will be called variables of type II.

2) A set C of constants of type I and type II.

3) Two lists of predicates , one list for predicates of type I, and another for those of type II.

4) Two lists of operations, the lists separating those of type I from those of type II.

It may happen that there are no variables of type II, no predicates of type II, or no operations of type II.

We will now give an intuitive description of the notions "predicate" and "operation."

If r is a predicate of type I, it can be associated with an integer n called the degree of r. In this case we will say r is an n-place predicate of type I. For any n-place predicate r of type I and any variables $x_1, \ldots, x_n$ with $x_i \in a$, $r(x_1, \ldots, x_n)$ is a proposition in our predicate calculus.

Every predicate R of type II can be associated with a pair of integers (n, m). We then say R is an (n, m)-place predicate of type II. For any $x_1, \ldots, x_n$, $X_1, \ldots, X_m$ with $x_i \in a$, $X_j \in A$, $R(x_1, \ldots, x_n, X_1, \ldots, X_m)$ is a proposition in our system.

If f is an operation of type I, f is associated with an integer n called the degree of f. For any type I variables $x_1, \ldots, x_n$, $f(x_1, \ldots, x_n)$ will be a type I variable. Similarly, if F is an operation of type II, then F has a degree (n, m), and for any $(x_1, \ldots, x_n, X_1, \ldots, X_m)$ with $x_i \in a$, $X_j \in A$, we have $f(x_1, \ldots, x_n, X_1, \ldots, X_m)$ is a variable of type II.

We conclude the description of the language with the notions of well-formed formulas (wff) and well-formed terms (wft). We can view wft's as the "words" of our language, which build up meaningful "sentences," or wff's . A punctuated string of symbols in our language will make sense precisely when the string is a wff.

A well-formed term of type I (wft_I) is defined by the conditions:

1) Each variable of type I is to be considered a wft_I .

2) If f is an n-place operation from type I to type I, and $\tau_1, \ldots, \tau_n$ are wft_I , then $f(\tau_1, \ldots, \tau_n)$ is a wft_I .

3) Every wft_I is obtained (and cannot be otherwise obtained) from 1) and 2) .

A term is said to be well-formed term of type II (wft_{II}) if the three conditions above are satisfied with I everywhere replaced by II .

We now have the wherewithal to define a general wft:

A wft (of type I and II) is given by the following four conditions:

(a) Each variable of type I is wft_I

(b) Each variable of type II is wft_{II}

(c) If f is an operation having $\left\{\begin{array}{l}\text{n type I places} \\ \text{m type II places}\end{array}\right\}$ where $I_1, \ldots, I_n$ are wft_I and $\sum_1, \ldots, \sum_m$ wft_{II} , then $f(I_1, \ldots, I_n, \sum_1, \ldots, \sum_m)$ is wft of type I and II .

(d) Each constant of type I (II) is wft_I (II). (Up to 3)

Finally, we define well-formed formula. Every wff is given by:

1) If γ is an n-place predicate of type I (II) and $\tau_1, \ldots, \tau_n$ wft of type I (II), then $\gamma(\tau_1, \ldots, \tau_n)$ is a wff.

2a) If α is wff so is $\sim\alpha$. If α, β wff so are

$$\alpha \,\&\, \beta$$

$$\alpha \vee \beta$$

$$\alpha \Longrightarrow \beta$$

2b) If α is wff, so are $(x)\alpha$, $(\exists x)\alpha$, $(A)\alpha$, $(\exists A)\alpha$.

3) All wff are so obtained.

Note that our building blocks are wft's, or "words" making up wff's, or "sentences."

In 2b) we declare that if α is a wff, so are $(x)\alpha$ and $(\exists x)\alpha$. In this instance α is the scope of (x) or $(\exists x)$. An occurence of a variable x in a wff α is said to be bound if it is

1) in a quantifier (i.e., in $(\exists x)$, x is bound).
2) in the scope of a quantifier (x) or $(\exists x)$.

If x at a particular occurence is not bound, it is said to be free at that occurence. We will say x is free in a wff α precisely when every occurence of x is free in α.

Chapter 2. A Fragment of Set Theory - S

We are ready to describe a particular application of the first order predicate calculus that we will use. The following list of the "elements" of our calculus is not complete. But it will suffice for the fragment S.

1) Variables of type I, denoted by small letters $x, y, z, \ldots$, and tagged with the name "set variables" or simply "sets".

2) Two constants denoted by ϕ and ω.

3) A two place predicate $p(x,y)$ which we will write as $x \in y$. If x and y are set variables, then $x \in y$ is a proposition of the system.

4) Three operations, a binary operation $J(x,y)$ and two unary operations $\mathcal{P}$ and $\sum$. (Note: The fact that $\mathcal{P}$, $\sum$, and J appear in our list of elements says the following: If x and y are sets, so are $\mathcal{P}x$, $\sum x$, and $J(x,y)$.) Of course we have said nothing by this without a description of what $\mathcal{P}$, $\sum$, and J are. We may revise our initial statement to read: If x and y are sets and if $\mathcal{P}$, $\sum$, and J are defined as in axioms S3, S4, and S5 [to follow], then $\mathcal{P}x$, $\sum x$, and $J(x,y)$ are sets.

Our intention is to consider $\mathcal{P}$ the power set operator and $\sum$ the "sum" or union operator. ϕ, of course, will represent the empty set, and ω the set of integers. J will be the operator which gives us the "successor" of any integer.

Before listing axioms for the system S, we should first define the notion of a special wff in this predicate calculus. This will be called a primitive-propositional formula (ppf). A ppf is a wff which contains no operations or constants (i.e., $\mathcal{P}$, $\sum$, J, ϕ, and ω do not appear).

As a consequence of this definition, a pff involves only set variables tied together by the $\in$-relation and our previous logical connectives.

<u>Examples of pff</u>:

1) $x \in y$

2) $(x \in y) \vee (y \in z)$; $(x \in y) \,\&\, (y \in z)$; $\sim(x \in y)$;

$(\exists x)\ (x \in y) \ldots$; $(x)(x \in z) \ldots$.

We will be unable to list the axioms for S without a working definition for the idea of equality. As usual two sets will be equal if they are "subsets" of each other. Hence the following two definitions:

$$x \subset y \Longleftrightarrow (t)(t \in x \Longrightarrow t \in y)$$

$$x = y \Longleftrightarrow (x \subset y \,\&\, y \subset x)$$

The first of the axioms for system S offers an alternative criterion for equality.

S1) (Extensionality) $x = y \Longleftrightarrow [(z)(x \in z \Longrightarrow y \in z)]$

S2) $\sim(x \in \emptyset)$

S3) $x \in \sum u \Longleftrightarrow (\exists t)\ (x \in t \,\&\, t \in u)$

S4) $(x)(x \in \mathcal{P} y \Longleftrightarrow x \subset y)$

S5) $t \in J(u, x) \Longleftrightarrow (t \in u \vee t = x)$

<u>Theorem 2.1.</u> <u>1) $x = x$ and 2) $x = y \Longrightarrow y = x$</u> .

<u>Theorem 2.2.</u> <u>If</u> α <u>is a ppf, and</u> $x = y$, <u>then</u> $\alpha(x) \Longleftrightarrow \alpha(y)$.

The proof proceeds by induction on the number of "steps" involved in the construction of α :

Case 1: α has the form $x \in z$. Then $x = y \Longrightarrow (x \in Z \Longleftrightarrow y \in z)$ by S1.

Case 2: α has the form $z \in x$. Then $x = y \Longrightarrow (z \in x \Longleftrightarrow z \in y)$ by the definition of equality.

Case 3: α has the form $\alpha' \vee \beta'$. Then $\alpha'(x) \Longleftrightarrow \alpha'(y)$, $\beta'(x) \Longleftrightarrow \beta'(y)$ whence $[\alpha'(x) \vee \beta'(x)] \Longleftrightarrow [\alpha'(y) \vee \beta'(y)]$, so $\alpha(x) \Longleftrightarrow \alpha(y)$. The cases $(\exists x)\alpha'$ and $\sim\alpha'$ go similarly.

Theorem 2.2 follows.

Before writing S6, we require a brief development of the integers. We shall define: $0 = \emptyset$, $1 = \{\emptyset\}$, $2 = \{\emptyset, 1\}$, and in general $n+1 = J(n,n)$.

S6) $x \in \omega \Longleftrightarrow (z)\ [\emptyset \in z \,\&\, (t)\ [(t \in z \Longrightarrow J(t,t) \in z) \Longrightarrow x \in z\ ]$

A well-formed s-term, (wfst) will be a wft in the particular predicate calculus S. Similarly, a wfsf (well-formed s-formula) will be a wff in this calculus.

<u>Theorem 2.3</u>: If α is a wfsf, then there is a p p f with the same free variables as α and $\alpha \Longleftrightarrow \beta$ (i.e., we can eliminate $\emptyset$, J, $\sum$, ω, $\mathcal{P}$ from α).

The proof proceeds by induction on the "complexity" of α. We define the complexity of α as:

2(number of symbols $\emptyset, \omega$, J, $\sum$, $\mathcal{P}$ to the left of $\in$) >< (number of symbols $\emptyset, \omega$, J, $\sum$, $\mathcal{P}$ to the right of $\in$).

Now note that replacing any $\emptyset$, ω, J, $\sum$, $\mathcal{P}$ in α by its definition in S2-S6 reduces the complexity of α. It follows by induction that we can eliminate all $\sum$, ω, J, $\mathcal{P}$, $\emptyset$.

We may define the singleton x (written $\{x\}$) with the help of S5:

$$\{x\} = J(\emptyset, x).$$

The reader should verify that $\{x, y\}$ as defined has exactly x and y as elements.

Exercises : 1. $\{x, y\} = J(\{x\}, y)$

2. $\{x, y\} = \{x', y'\} \Longleftrightarrow (x = x' \,\&\, y = y') \vee (x = y' \,\&\, y = x')$

Our final concern in axiom system S will be the construction of <u>ordered pairs</u>. We define

$$\langle x, y \rangle \text{ (the ordered pair) } = \{\{x\}, \{x, y\}\}$$

<u>Theorem 2.4</u>: $\langle x, y \rangle = \langle x', y' \rangle \Longleftrightarrow x = x', y = y'$.

We can easily generalize to the construction of <u>ordered n-tuples</u> with the convention: $\langle x \rangle = x$

$\langle x, y \rangle$ is defined as above

$\langle x_1, \ldots, x_n \rangle = \langle x_1, \langle x_2, \ldots, x_n \rangle\rangle$ for each n.

As an immediate consequence of Theorem 2.4:

<u>Corollary 2.4.1</u>:

$$\langle x_1, \ldots, x_n \rangle = \langle y_1, \ldots, y_n \rangle \Longleftrightarrow x_1 = y_1, x_2 = y_2, \ldots, x_n = y_n.$$

Chapter 3. Axiom Group - A

One of the chief contributions of axiom system S is the construction of ordered pairs (and so ordered n-tuples). We will now enlarge S into two larger systems (called axiom systems A and B). Axiom set A will keep all intact except for adding "classes" to the repertoire of "sets" we have already constructed. Axiom set B will allow us to obtain the class theorem, that is, the procedure for constructing the collection of all $\langle x_1, \ldots, x_n \rangle$ which satisfy a given property.

We begin with axiom set A. The following are primitive in this system.

Variables:	(called classes) $A, B, C \ldots$ (upper case letters)
Constants:	$\mathcal{U}$, ω
Predicates:	$\in$
Terms	$A \in B$
Operations	(unary): $\mathcal{P}$, $\sum$; (binary): J

(We continue to write A0, A1 ... for axioms of system A.)

A0: (the universe) $\quad A \in B \Longrightarrow A \in \mathcal{U}$

We shall say A is a set iff $A \in \mathcal{U}$. $\mathcal{U}$ is hence the universe of all sets. Axiom A0 says that only sets can be members of either sets or classes. A special notation, use of lower case letters, will be used to denote sets and quantification of set variables. Examples of this convention are

$$(\forall x)(\alpha(x)) \overset{\text{def.}}{\Longleftrightarrow} (\forall A)(A \in \mathcal{U} \Longrightarrow \alpha(A))$$
$$(\exists x)(\alpha(x)) \overset{\text{def.}}{\Longleftrightarrow} (\exists A)(A \in \mathcal{U} \;\&\; \alpha(A));$$

x above is referred to as a set variable.

The inclusion relation is defined just as before:

$$A \subset B \Longleftrightarrow (\forall C)(C \in A \Longrightarrow C \in B).$$

We then define $A = B$ precisely when $A \subset B$ and $B \subset A$.

<u>Exercise</u>: Equality is an equivalence relation.

Equal classes should be equal as elements. They should belong to precisely the same things. This is guaranteed by the axiom:

A1) (extensionality): $A = B \Longrightarrow (\forall C)(A \in C \Longrightarrow B \in C)$.

The following axioms bring system A up to date with what was accomplished in $\mathbf{S}$:

A2) (null set): $x \notin \emptyset$

A3) (power ~~sets~~ class) $x \in \mathcal{P}A \Longrightarrow x \subset A$

A4) (union) $x \in \sum A \Longleftrightarrow (\exists y)(x \in y \; \& \; y \in A)$

A5) (the successor J): $x \in J(A, B) \Longleftrightarrow x \in A \vee x = B$

A6) (the natural numbers):

$$x \in \omega \Longleftrightarrow (\forall z)[\emptyset \in z \; \& \; (y)(y \in z \Longrightarrow J(y,y) \in z) \Longrightarrow x \in z]$$

It may seem that defining ω, $\emptyset$, $\mathcal{P}$, $\sum$, and J again amounts to formal nonsense. The reader should note, however, that they were previously defined only in terms of "set variables". A2-A6 incorporate our old definitions into the larger system A.

In the same manner as before we can proceed to construct ordered pairs: The <u>singleton A</u> (written $\{A\}$) will be given by

$$\{A\} = J(\emptyset, A)$$

<u>Theorem 3.1:</u> $B \in \{A\} \Longleftrightarrow B \in \mathcal{U} \;\&\; B = A$.

$\{A, B\}$, the <u>unordered pair</u> A and B , obeys the rule

$$\{A, B\} = J(\{A\}, B)$$

<u>Theorem 3.2:</u> $\{A, B\} = \{A', B'\} \Longleftrightarrow (A = A' \;\&\; B = B'$ or $A = B' \;\&\; B = A')$

The <u>ordered pair</u> $\langle A, B\rangle$ is now, as it was before:

$$\langle A, B\rangle = \{\{A\}, \{A, B\}\}$$

<u>Theorem 3.3:</u> $\langle A, B\rangle = \langle A', B'\rangle \Longleftrightarrow A = A', B = B'$.

We proceed to the ordered n-tuple. For the case $n = 1$, we define

$$\langle A \rangle = A$$

If $n = 2$:

$$\langle A, B\rangle = \{\{A\}, \{A, B\}\}$$

In general:

$$\langle A_1, \ldots, A_r\rangle = \langle A_1, \langle A_2, \ldots, A_r\rangle\rangle ,$$

where the right-hand side defines the left.

We will write $t \in \mathcal{U}^n$ precisely when t is an n-tuple, i.e., $t = \langle t_1, \ldots, t_n\rangle$ with $t_1, \ldots, t_n \in \mathcal{U}$.

We now move on to Axiom Group B. What is sought is to define the "classifier" $\{x \mid \phi(x)\}$. The most direct route is to adjoin an infinite set of axioms (an <u>axiom schema</u>) to those of system A. We would like the following statement to apply:

<u>Axiom Schema B</u>: If α is a wff containing free set variables $x_1, \ldots, x_n$, and free class variables $A_1, \ldots, A_m$, all of whose bound variables are set variables, then $\exists$ a class C such that

1) $C \subset \mathcal{U}^n$ (i.e., $t \in C \Longrightarrow (\exists t_1)(\exists t_2)\ldots(\exists t_n)(t = \langle t_1, \ldots, t_n \rangle)$)

2) $\langle x_1, \ldots, x_n \rangle \in C \Longleftrightarrow \alpha(x_1, \ldots, x_n \;,\; A_1, \ldots, A_n)$.

The following are immediate consequences of Axiom Schema B:

$$\emptyset = \{ x \mid x \neq x \}$$
$$\mathcal{P} A = \{ x \mid x \subset A \}$$
$$\sum A = \{ x \mid (\exists y)(x \in y \,\&\, y \in A) \}$$
$$J(A,B) = \{ x \mid x \in A \vee x = B \}$$
$$\omega = \{ x \mid (\forall z)(\emptyset \in z \,\&\, (\forall y)(y \in z \Longrightarrow J(y,y) \in z) \Longrightarrow x \in z) \}$$
$$\mathcal{U} = \{ x \mid x = x \}$$

The following definitions are possible by virtue of Axiom Schema B:

$$A \cap B = \{ x \mid x \in A \,\&\, x \in B \}$$
$$-A = \{ x \mid x \notin A \}$$
$$A + B = \{ x \mid x \in A \,\&\, x \notin B) \vee (x \notin A \,\&\, x \in B) \}$$
$$A \times B = \{ \langle x_1, x_2 \rangle \mid x_1 \in A \,\&\, x_2 \in B \}$$
$$A^{-1} = \{ \langle x_1, x_2 \rangle \mid \langle x_2, x_1 \rangle \in A \}$$
$$\mathcal{I} = \{ \langle x, y \rangle \mid x = y \}$$
$$\mathcal{E} = \{ \langle x_1, x_2 \rangle \mid x_1 \in x_2 \} \quad .$$

Chapter 4. Axiom Group B

It is desirable to replace Axiom Schema B, leading as it does to an infinite set of axioms, with a finite list of axioms from which the schema will be a consequence. For this purpose, a new system, Axiom System B, will be introduced and shown to be equivalent to system A with axiom schema B added.

The following are primitives for B:

Variables (called classes): $A, B, C, \ldots$

Constants: $\mathcal{U}, \mathcal{E}, \tau$ ($\mathcal{E}, \tau$ defined as below)

Predicates: $\in$

Terms: $A \in B$

Operations: (unary) γ, α, π; (binary) $< \ >$, $\cap$, $+$, $\times$

(all defined as below).

We now list a complete set of axioms for system B. Sets, set variables, equality, and extensionality are all carried over from system A.

B0) (The universe): $A \in B \Longrightarrow A \in \mathcal{U}$

B1) (Ordered pairs): $<A, B> = <A', B'> \Longleftrightarrow A = A', B = B'$

B2) (Ascension): $u \in \mathcal{E} \Longleftrightarrow (\exists x)(\exists y)(u = <x,y> \,\&\, x \in y)$

B3) (The diagonal identity): $u \in \tau \Longleftrightarrow (\exists x)(u = <x, x>)$

B4) (Intersection): $u \in A \cap B \Longleftrightarrow u \in A \,\&\, u \in B$

B5) (Symmetric difference): $u \in A + B \Longleftrightarrow (u \in A \,\&\, u \notin B) \vee (u \notin A \,\&\, u \in B)$

B6) (Cross product) $u \in A \times B \Longleftrightarrow (\exists x)(\exists y)(u = <x,y> \,\&\, (x \in A \,\&\, y \in B))$

B7) (Commutativity): $u \in \gamma A \Longleftrightarrow (\exists x)(\exists y)(u = <y,x> \,\&\, <x,y> \in A)$

B8) (Associativity): $u \in \alpha A \Longleftrightarrow (\exists x)(\exists y)(\exists z)(u = \langle\langle x,y\rangle, z\rangle \,\&\, \langle x,y,z\rangle \in A)$

B9) (Projection): $u \in \pi A \Longleftrightarrow (\exists x)(<x, u> \in A)$

α in B8) is called the <u>associator</u> because it allows us to construct the class of all $\langle\langle x,y\rangle, z\rangle$ with $\langle x, \langle y,z\rangle\rangle$ in A . γ of B7) is called the <u>converse operator</u> because it allows us to "flip" the ordered pairs of A; i.e., γ allows us to construct the class of all $\langle y, x\rangle$ where $\langle x,y\rangle$ is an ordered pair in A .

Note that B5) gives rise to our familiar notions of union and complement, for: $A \cap B = ((A + (A \cup B)) + (A \cup B)) + ((A \cup B) + B)$. The complement of A , written $-A$ is defined by : $-A = A + \mathcal{U}$.

The equivalence of the system given by Axiom Schema B with system B will now be shown. Obviously $B \Longrightarrow$ B1) - B9). To prove the converse, some preliminaries are needed:

<u>Lemma 4.1</u> $(\alpha\gamma)^3 A = A \cap \mathcal{U}^3$. (The reader should verify this by computing out the string $\alpha\gamma\ \alpha\gamma\ \alpha\gamma A$.)

It is now necessary to define classes corresponding to $\{\langle x_1,\ldots,x_n\rangle \mid \langle x_{i_1},\ldots, x_{i_{n-1}}\rangle\} \in A$. That is, we must establish rules for inserting new variables within n-tuples. First we make some definitions:

$$\alpha^{-1} = \gamma\alpha\gamma\alpha\gamma \quad \text{(the associative inverse)}$$

$$j_2 A = A \times \mathcal{U}$$

$$j_3 A = \alpha^{-1}(\gamma^2 A \times \mathcal{U})$$

$$i_1 A = \mathcal{U} \times A$$

$$i_2 A = \gamma\alpha(\mathcal{U} \times \gamma A)$$

$$i_3 A = \alpha^{-2}\gamma\,\alpha^{-1}(\gamma\alpha A \times \mathcal{U})$$

<u>Theorem 4.2:</u>

$$j_2 A = \{\langle x,t\rangle \mid x \in A\}$$

$$j_3 A = \{\langle x,y,t\rangle \mid \langle x,y\rangle \in A\}$$

$$i_1 A = \{\langle t, x\rangle \mid x \in A\}$$
$$i_2 A = \{\langle x, t, y\rangle \mid \langle x, y\rangle \in A\}$$
$$i_3 A = \{\langle x, y, t, z\rangle \mid \langle x, y, z\rangle \in A\}$$

As a sample we prove the last of these statements:

$$\begin{aligned}
\langle x_1, x_2, x_3, x_4\rangle \in i_3 A &\Longleftrightarrow \langle x_1, x_2, x_3, x_4\rangle \in \alpha^{-2}\gamma\alpha^{-1}(\gamma\alpha A \times \mathcal{U})\\
&\Longleftrightarrow \langle x_1, \langle x_2, \langle x_3, x_4\rangle\rangle\rangle \in \alpha^{-2}\gamma\alpha^{-1}(\gamma\alpha A \times \mathcal{U})\\
&\Longleftrightarrow \langle\langle x_1, x_2\rangle, \langle x_3, x_4\rangle\rangle \in \alpha^{-1}\gamma\,\alpha^{-1}(\gamma\,\alpha A \times \mathcal{U})\\
&\Longleftrightarrow \langle\langle\langle x_1, x_2\rangle, x_3\rangle, x_4\rangle \in \gamma\alpha^{-1}(\gamma\alpha A \times \mathcal{U})\\
&\Longleftrightarrow \langle x_4, \langle\langle x_1, x_2\rangle, x_3\rangle\rangle \in \alpha^{-1}(\gamma\,\alpha A \times \mathcal{U})\\
&\Longleftrightarrow \langle\langle x_4, \langle x_1, x_2\rangle\rangle, x_3\rangle \in \gamma\,\alpha A \times \mathcal{U}\\
&\Longleftrightarrow \langle x_4, \langle x_1, x_2\rangle\rangle \in \gamma\,\alpha A\\
&\Longleftrightarrow \langle\langle x_1, x_2\rangle, x_4\rangle \in \alpha A\\
&\Longleftrightarrow \langle x_1, \langle x_2, x_4\rangle\rangle \in A\\
&\Longleftrightarrow \langle x_1, x_2, x_4\rangle \in A
\end{aligned}$$

Using 4.2 it is possible to insert a variable anywhere in any ordered n-tuple.

We now proceed to show Axiom System B is strong enough to prove all that we wanted.

Theorem 4.3: (The Class Theorem): If α is a ~~primitive~~ wff containing free set variables among $x_1, x_2, \ldots, x_n$ and free class variables among $A_1, \ldots, A_n$, all of whose bound variables are set variables, then there exists a wft τ, all of whose variables occur among $A_1, \ldots, A_n$ such that:

1. $\tau \subseteq \mathcal{U}^n$

2. $\langle x_1, x_2, \ldots, x_n\rangle \in \tau \Longleftrightarrow \alpha(x_1, \ldots, x_n, A_1, \ldots, A_n)$ (i.e., it is possible to form $\{\langle x_1, \ldots, x_n\rangle \mid \alpha\}$, the ~~set~~ class of all "so-and-so's" with property α.

It is possible to move all class variables in α to the right of the " $\in$ " in α (α is required to be primitive); assume that this has been done.

Case I: α is atomic. Then α is of the form $x_r \in A$ or $x_r \in x_s$, $r, s \leq n$. If α is "$x_r \in A$", choose τ to be $i_1^{r-1} j_2^{n-r} A$. Otherwise for $n \leq 2$:

If n is	and α is	take τ as
1	$x_1 \in x_1$	$\mathcal{E} \cap \tau$
2	$x_1 \in x_1$	$j_2(\mathcal{E} \cap \tau)$
↓	$x_1 \in x_2$	$\mathcal{E}$
↓	$x_2 \in x_1$	$\gamma \mathcal{E}$
↓	$x_2 \in x_2$	$i_1(\mathcal{E} \cap \tau)$

For $n > 3$, define $\tau_2 = \{< x_1 , x_s > \mid \alpha\}$ which is possible by the above. Then take $\tau = i_1^{r-1} i_2^{s-r} i_3^{n-s-1} j_3 \tau_2$.

Case II: α is not atomic. Let α be reduced to a ~~term~~ formula involving only the logical operators " $\&$ ", " $\sim$ ", and " $\exists$ ". (This is again possible since α is primitive). We use induction on the number of symbols in α. If β has fewer symbols than α, let τ be the term satisfying Theorem 4.3. Then:

If α is of form	take τ as
$\beta \& \psi$	$\tau_\beta \cap \tau_\psi$
$\sim \beta$	$\tau_\beta + \mathcal{V}$
$(\exists x_1) \beta$	$\pi \tau_\beta (x_1, \ldots, x_n)$

This completes the proof.

It is desirable to reduce some of the restrictions on α in 43.

A wff α is normal if

1) α is primitive or

2) every constant, operation, or predicate in α can be expanded in terms of its definition, and every constant, predicate, or operation thus resulting can also be replaced by its definition, and so on so as to obtain a primitive statement in a finite number of steps.

Then the qualification "primitive α" of the class theorem may be weakened to "normal α."

Our complete system to date, then, consists of:

Variables: (called classes): A, B, C, ...

Constants: $\mathcal{U}$, $\emptyset$, ω, J, $\mathcal{E}$

Predicates: $\in$

Terms: $A \in B$

Operations (unary): $\mathcal{P}$, $\sum$, γ, α, π, ;

(binary): J, $\cap$, +, $\times$

Axioms: A0–A6, B1–B9

A class will be called proper iff it is not a set.

Theorem 4.4: System A, System B, and System AB are all absolutely consistent. For take $\mathcal{U} = \emptyset = \omega = \mathcal{E} = J$. There is only one class, and it is proper. The axioms are obviously satisfied.

Taking stock, we have now evolved a system in which we have sets and classes, which admit of n-tuples, and in which the Class Theorem provides machinery for constructing specified classes.

Chapter 5. Relations and Functions

From an axiomatic point of view, the idea of function will add nothing new to our theory; it is already included as a consequence of the construction of ordered pairs. Nevertheless, any attempt to axiomatize set theory must single out functions and relations as objects of detailed study.

A relation is simply a class whose elements are ordered pairs (i.e., $R \subset \mathcal{U}^2$). We can think of x being related to y if and only if $\langle x, y \rangle \in R$.

The image of a relation R will be $\pi \gamma R$, that is, $\text{im } R = \{ y \mid (\exists x) \langle y, x \rangle \in R \}$. The domain of R is $\pi R = \{ x \mid (\exists y) \langle y, x \rangle \in R \}$. As an analogue to the idea of composition of functions, we define the relative product $R \circ S$ of two relations R and S as:

$$R \circ S = \{ \langle x, z \rangle \mid (\exists y) \langle x, y \rangle \in R \mathbin{\&} \langle y, z \rangle \in S \}.$$

Exercises:
1. $(R \circ S) \circ T = R \circ (S \circ T)$.
2. $\gamma(R \circ S) = \gamma S \circ \gamma R$.

The restriction (in domain) of a relation R to B will be denoted by $R \restriction B$, and:

$$R \restriction B = R \cap (\mathcal{U} \times B).$$

A function F is a relation R which is "single-valued", i.e.,

$$\langle y_1, x \rangle, \langle y_2, x \rangle \in R \Longrightarrow y_1 = y_2.$$

The domain, range, restriction, and relative product of functions can now be obtained as special cases of relations. We can also define the restriction (in range) of a relation (function) R to A:

$$R \restriction A = R \cap (A \times \mathcal{U}).$$

We will write $F: A \longrightarrow B$ for a function F with domain A and image $F \subset B$. We would now like to define F(x) so as to obtain the familiar notion of "the result of acting on x by F."

Consider the formula $\alpha(t, x): (\exists y) t \in y \;\&\; \langle y, x\rangle \in F$. By the class theorem we can construct the class

$$C = \{ t \mid \alpha(t, x)\}.$$

For $x \in \mathrm{dom}(F)$, there is a unique y $(\exists! y)$ with $\langle y, x\rangle \in F$, so $t \in y \Longleftrightarrow t \in C$. For $x \notin \mathrm{dom}(F)$, there is no y with $\langle y, x\rangle \in F$ so no $t \in$ (such a y), and $C = \emptyset = y$.

Therefore we define: $F'x = C$. We say F is defined at x if $\langle x, F'x\rangle \in F$ and undefined otherwise. F is undefined at x if and only if $F'x = \emptyset$ and $\langle x, \emptyset\rangle \notin F$.

We will define $F''u$, the image of u under the function F, by

$$F''u = \{ y \mid F \text{ is a function } \& \ (\exists x) \ \langle y, x\rangle \in F \ \& \ x \in u\}.$$

There is as yet no reason why the image of a set under a function might not be a proper class; we would like to rule out this possibility. We will therefore introduce an axiom, and derive from it a useful theorem.

Axiom C1: If u is a set and F is a function, then $F''u$ is a set.

Theorem 5.1: If u is a set and $A \subset u$, then A is a set, (that is, subclasses of sets are sets). $I \cap (\mathcal{U} \times A)$, the restriction of I to A, is a function, so $(I \cap \mathcal{U} \times A)''u$ is a set, but this is precisely A.

It is 5.1, rather than Axiom C itself, which will be utilized in the construction of ordinals; however, Axiom C is necessary in other set-theoretic considerations, and cannot be derived from 5.1 and preceding axioms.

At this point we would like to know that certain familiar classes are actually sets. For this purpose we need two axioms:

<u>Axiom C2</u>: If u is a set, then $\sum u$ is a set.

<u>Axiom C3</u>: If u is a set, then $\mathcal{P} u$ is a set.

Chapter 6: <u>The Complete (Gödel-Bernays) Axiom System</u>

The complete axiom system will be, essentially, an accumulation of everything we have done up to now.

The primitive terms are, as usual, variable classes, written $A, B, C \ldots$ There is a primitive binary relation between classes, $A \in B$. To this skeleton we add all constants defined heretofore, all of Axiom Group A, Axiom Group B, and Axioms C1, C2, C3.

Temporarily, we define $\sigma(x) = J(x, x)$. By Axioms C2) and C3), we can construct $\sum u$ and $\mathcal{P} u$.

We now list some results of Axioms C1, 2, 3.

<u>Theorem 6.1:</u>

(a) u, v sets $\Longrightarrow u \cup v$ is a set

(b) u, v sets $\Longrightarrow u \times v$ is a set

(c) u, v sets $\Longrightarrow u \cap v$, $u + v$, $\gamma v, \alpha u, \pi u$ are sets.

(d) u, v sets $\Longrightarrow \overset{v^u}{\text{map}}(u, v) =$ [class of all functions: $u \longrightarrow v$)] is a set.

(a) is immediate, for $u \cup v = \sum \{u, v\}$ which is a set. To prove (b), note $\{x\} \in \mathcal{P} u$, $\{y\} \in \mathcal{P} v \Longrightarrow \{x, y\} \in \mathcal{P}(u \cup v)$, a set. Therefore, $\{\{x\}, \{x, y\}\} \in \mathcal{P}\mathcal{P}(u \cup v)$, so $\langle x, y\rangle \in \mathcal{P}\mathcal{P}(u \cup v)$. Hence:

$$u \times v \subset \mathcal{P}\mathcal{P}(u \cup v)$$

which is a set.

We skip the proof of (c). To prove (d) note: $f : u \longrightarrow v \Longrightarrow f \in \mathcal{P}(u \times v)$, which is a set, so $\text{map}(u, v) \subset \mathcal{P}(u \times v)$. If $f : u \longrightarrow v$, then $f \in \text{map}(u, v)$ so f is a set by 6.1.

Axiom D: (Axiom of Foundations): $A \neq \emptyset \Longrightarrow (\exists x)\ x \in A \ \& \ x \cap A = \emptyset$.

Theorem 6.2: $\sim(x \in x)$.

Let $A = \{x\}$. Clearly $A \neq \emptyset$, so $(\exists t)(t \in A$ and $t \cap A = \emptyset)$. But $t = x$, so $x \in x \Longrightarrow x \in x \cap A \Longrightarrow t \cap A \neq \emptyset$, which is a contradiction.

Theorem 6.3: $\sim(y \in x \ \& \ x \in y)$.

Let $A = \{x,y\} \neq \emptyset$. Then $(\exists t)(t \in A \ \& \ t \cap A = \emptyset)$ But $t = x$ or $t = y$. Let us suppose $t = x$. Suppose $y \in x$ and $x \in y$, then $y \in x = t \Longrightarrow y \in t \cap A \neq \emptyset$, which is a contradiction.

Similarly, it is clear that for any string of x_i, $\sim(x_1 \in x_n \in x_{n-1} \in \ldots \in x_1)$. We would like to show that the infinite analogue holds; i.e., that there can be no infinite sequence of sets $\{x_n\}$ such that $x_{n+1} \in x_n$ for all n. A precise formulation of this notion is:

Theorem 6.4: There is no function f on ω with $f'(n+1) \in f'n$ for all $n \in \omega$.

Let $A = f''\omega$. $f'0 \in A$, so $A \neq \emptyset$; $\exists t \in A$ such that $t \cap A = \emptyset$, therefore $(\exists n)\ t = f'n$, but $f'n{+}1 \in f'n = t$ and $f'n{+}1 \in A$, so $t \cap A \neq \emptyset$, which is a contradiction.

Axiom E (Axiom of Choice): There is a function C such that if $x \neq \emptyset$ then $(\exists y)\ \langle y,x \rangle \in C$ and $y \in x$ (i.e., $x \neq \emptyset \Rightarrow C'x \in x$).

Theorem 6.5: Assuming all preceding axioms (except D) and axiom E, then Axiom D $\Longleftrightarrow$ Theorem 6.4.

The $\Longrightarrow$ part of the biconditional has already been demonstrated. We prove only $\Longleftarrow$. Assume, then, that D is false, i.e.,

$$(\exists A)(A \neq \emptyset) \text{ and } (x)(x \in A \Longrightarrow x \cap A \neq \emptyset).$$

We define recursively a function f on ω:

$$f'0 = C'A = x_0 \quad \ldots$$

$$f'n{+}1 = C'(f'n \cap A) = x_{n+1}$$

Then $f'n \in A$, $f'n \in f'n{-}1$, which is a contradiction.

The proof of 6.5 is, of course, not valid until we can justify the mechanism of defining functions by recursion. This will be done in Chapter 9.

Before we go on to the construction of ordinals, a detailed study of the integers is in order. The usual tactic is to start with the <u>Peano Postulates</u> for ω . Each of these postulates is a deducible statement in our theory. For the remainder of this chapter all quantification will be restricted to ω ; we shall interpret $(n) \ldots$ to mean: $(n)(n \in \omega \Longrightarrow \ldots)$. Set $\sigma n = J(n, n)$.

<u>Theorem 6.6</u>: (1). $\emptyset \in \omega$

(2). $(x)(x \in \omega \Longrightarrow \sigma x \in \omega)$

(3). $(n, m)(\sigma n = \sigma m \Longrightarrow n = m)$

(4). $(n)(\sigma n \neq \emptyset)$

(5). If $u \subset \omega$, and $\emptyset \in u$, and u is σ-closed, then $u = \omega$.

σ is called the <u>successor function</u>, and σn is the <u>successor</u> of n .

We shall prove only statement (3) of 6.6; the others are all immediate from the definition of ω and 6.3.

Suppose, then, $\sigma n = \sigma m$. Then $n \in n \cup \{n\} = m \cup \{m\}$, so $n \in m \cup \{m\}$.
We conclude either $n = m$ or $n \in m$. By the same reasoning either $m = n$ or $m \in n$. If $m \neq n$, we have $m \in n$ and $n \in m$, contradicting 6.3.

Better to use "defining technique of defining functions by recursion".

We would now like to have the notion of the sum of two integers. We will show that there is a function S from ω^2 to ω with the properties:

(1^*) $S'(m, 0) = m$

(2^*) $S'(m, \sigma n) = \sigma(S'(m, n))$

We will show by induction on n that S is the unique function having these properties. We can then call $S'(m, n)$ the sum of m and n (written $m + n$).

Call $u \subset \omega^3$ a <u>summation</u> if

(1^0) $\langle m, m, 0\rangle \in u \quad \forall m \in \omega$

(2^0) $\langle k, m, n\rangle \in u \Longrightarrow \langle \sigma k, m, \sigma n\rangle \in u$

Let $S = \bigcap \{\text{all summations}\}$. Note that S satisfies (1^0) and (2^0), and is thus a summation.

<u>Lemma 6.7</u>: <u>S has properties (1^*) and (2^*)</u>. The proof is immediate. We have only to prove now that S is a function. To prove this we first define:

$t = \{\langle k, m, n\rangle \in S \mid$

1) $n = 0$ and $k = m$ or

2) $n \neq 0$ and $\exists k', n'$ such that $k = \sigma k'$, $n = \sigma n'$,

and $\langle k', m, n'\rangle \in S\}$.

<u>Lemma 6.8</u>: <u>t is a summation</u>. For $\langle k, m, n\rangle \in t \Longrightarrow \langle k, m, n\rangle \in S$. We conclude that $\langle \sigma k, m, \sigma n\rangle \in t$, and so t is a summation. Hence $t = S$.

<u>Theorem 6.9</u>: <u>S is a function.</u>

Assume $\langle k, m, n\rangle$, $\langle k^*, m, n\rangle \in S$. We can prove that $k = k^*$ by induction on n since $\langle k, m, n\rangle, \langle k^*, m, n\rangle \in S \Longrightarrow \langle k', m, n'\rangle$, $\langle k^{*\prime}, m, n'\rangle \in S$, $\sigma(k') = k$, $\sigma(k^{*\prime}) = k^*$, and $\sigma(n') = n$; by induction $k' = k^{*\prime}$, and $k = k^*$ since σ is a function.

Chapter 7. Ordinal Numbers

The axioms allow a very convenient definition of the ordinal numbers. The finite ordinal numbers are just the non-negative integers n already defined. The set ω can be regarded as the first infinite ordinal number; it is the set of all finite ordinal numbers. The next ordinal numbers are $\omega + 1$, $\omega + 2$, ... The first ordinal beyond these, 2ω, may be defined as the set of all finite ordinals and all ordinals of the form $\omega + n$, n a non-negative integer. The general principle will be that of defining each ordinal number as the set of all preceding ordinals.

Specifically, we will say that a set u is complete if $x \in u$ implies $x \subset u$; equivalently, u is complete if $t \in x$ and $x \in u$ implies $t \in u$. We then define an ordinal number α to be a complete set with the additional property that if $x, y \in \alpha$, either $x \in y$, $y \in x$ or $x = y$.

It follows at once from Axiom D that only one of these alternatives can hold. Furthermore, if x, y, z are elements of α with $x \in y$ and $y \in z$, then $x \in z$ by completeness. Thus the set α is (linearly) ordered by the $\in$-relation E. We may therefore say that an ordinal number is a complete set ordered by the $\in$-relation. It is then immediate that every $n \in \omega$ is an ordinal and ω is an ordinal. The following theorem effectively demonstrates that the elements of an ordinal α are the preceding ordinals, thus confirming that our definition corresponds to the general principle of the first paragraph.

Theorem 7.1. The elements of an ordinal α are precisely the complete proper subsets of α.

Let $x \in \alpha$. Since α is complete, $x \subset \alpha$, and the inclusion must be proper, for otherwise $\alpha \in \alpha$, violating 6.2. To show that x is complete, take any $y \in x$. Then $y \in x \subset \alpha$, hence $y \in \alpha$, so $y \subset \alpha$ by the completeness of α.

Thus, for any $t \in y$, we have $t \in \alpha$, so either $t \in x$, $t = x$ or $x \in t$. In the third case $x \in t$, $t \in y$ and $y \in x$, contradicting 6.3. In the second case, $t = x$, we have $x = t \in y$ and $y \in x$, a similar contradiction. Hence for every $t \in y$ we have $t \in x$, so $y \subset x$.

Conversely, let x be any proper complete subset of α. Denote the complement of x in α by z; then, by Axiom D there exists in z an element d with $d \cap z = 0$. Since z is a subset of α, it suffices to show that $x = d$, for then $x \in \alpha$ as asserted by the theorem.

Let t be an element of d. Then since $d \in z$ and $z \subset \alpha$, we have $d \in \alpha$ and thus $d \subset \alpha$ by the completeness of α, so $t \in \alpha$. Since $t \in d$, $t \notin z$, so t is in the complement of z in α, but this is precisely x, so $t \in x$ and hence $d \subset x$.

For the opposite inclusion, let r be any element of x. Since r and d are elements of α, either $r \in d$, $r = d$ or $d \in r$. If $r = d$, then $d \in x$, hence $d \notin z$, a contradiction. If $d \in r$, since $r \in x$ and x is complete, it again follows that $d \in x$, which leads to the same contradiction. Therefore $r \in d$, but since this is true for arbitrary r in x we have $x \subset d$ and thus $x = d$.

Corollary 7.1.1. If α is an ordinal, every element x of α is an ordinal.

By 7.1, x is a complete set and a subset of α; hence x, like α, is linearly ordered by $\in$, thus is an ordinal.

Corollary 7.1.2. If α and γ are ordinals, $\alpha \in \gamma$ if and only if $\alpha \subset \gamma$ and $\alpha \neq \gamma$.

If $\alpha \subset \gamma$ and $\alpha \neq \gamma$, then $\alpha \in \gamma$ by 7.1. Conversely, if $\alpha \in \gamma$, then $\alpha \subset \gamma$, and $\alpha = \gamma$ is impossible by 6.2. Henceforth, we will write $\alpha < \gamma$ to mean $\alpha \in \gamma$ $(\alpha \subset \gamma,\ \alpha \neq \gamma)$.

The following theorem is primarily a tool for Chapter 8. Its corollary here is a result that can be immediately derived from 7.1.2.

<u>Theorem 7.2.</u> <u>If</u> A <u>is any non-empty class of ordinals, then</u> $\beta = \bigcap_{\alpha \in A} \alpha$ <u>is an ordinal and an element of the class</u> A.

Clearly β is a set. To prove it is complete, let $x \in \beta$ and $t \in x$. Then for every α in A, $t \in x \in \beta \subset \alpha$, so $t \in x \in \alpha$, hence $t \in \alpha$. Thus $x \subset \alpha$ for every α in A, hence $x \subset \beta$ by the definition of the intersection, so β is complete. Since β is contained in some α_o in A, and α_o is ordered by membership, so is β. Thus β is an ordinal.

If β is not in the class A, then $\beta \neq \alpha$ for every $\alpha \in A$ but $\beta \subset \alpha$, so by 7.1 $\beta \in \alpha$ for all α. Therefore, $\beta \in \beta$, contradicting 6.2.

<u>Corollary 7.2.1.</u> <u>If</u> α <u>and</u> γ <u>are ordinals, either</u> $\alpha \subset \gamma$ <u>or</u> $\gamma \subset \alpha$ <u>holds (or both, of course, if</u> $\alpha = \gamma$<u>)</u>.

Apply 7.2 to $A = \{\alpha, \gamma\}$. It follows that $\alpha \cap \gamma$ is either α or γ.

Chapter 8. Construction of Ordinals

A class A is said to be well-ordered by a relation R if A is linearly ordered by R and if every non-void subclass of A has a first element in this order. The results of the previous chapter will enable us to show that the class of all ordinals (whose existence is guaranteed by 4.3.) has this important property.

Theorem 8.1. The class Ω of all ordinals is well-ordered by the relation $\alpha < \gamma$.

Let A be any subclass of Ω, and apply 7.2 to get an ordinal $\beta \in A$, with $\beta \subset \alpha$ for every α in A. This ordinal is the desired first element of A in the given order.

Corollary 8.1.1. If α is an ordinal, α is well-ordered by the relation $\beta < \gamma$ for its elements.

For any subset of α is a subclass of the class of all ordinals, hence has a first element.

We are also able to improve somewhat on the characterization of the elements of an ordinal α given by 7.1.

Theorem 8.2. The elements of an ordinal α are exactly those ordinals smaller than α.

Since if $x \in \alpha$, then x is an ordinal by 7.1.1, and $x \in \alpha$, hence $x \subset \alpha$. Conversely, by 7.1.2, if β is an ordinal and $\beta < \alpha$, then $\beta \in \alpha$.

The final theorem of this chapter (in particular, the corollaries derivable from it) provides us with a means of constructing large ordinals from a given one, say α. From the first corollary and the principle of finite induction it follows that an element resulting from reiterated application of the successor operation to α is an ordinal larger than α. The second corollary implies that the union of all such ordinals is an even larger one.

Theorem 8.3. Every complete set of ordinals is an ordinal.

For if u is a complete set whose elements are ordinals, then the elements of u, being ordinals, are linearly ordered by $\in$; hence u is an ordinal by definition.

Corollary 8.3.1. If α is an ordinal, so is $\sigma(\alpha)$, and $\sigma(\alpha)$ is the first ordinal larger than α.

For $\sigma(\alpha)$ has as elements α and the elements of α, hence $\sigma(\alpha)$ is a set of ordinals. If $x \in \sigma(\alpha)$, then either $x \in \alpha$, in which case $x \subset \alpha \subset \sigma(\alpha)$ by the completeness of α, or $x = \alpha$, in which case $x \subset \sigma(\alpha)$ by definition of $\sigma(\alpha)$. So $\sigma(\alpha)$ is a complete set of ordinals, thus an ordinal by the theorem.

Since $\alpha \in \sigma(\alpha)$, $\sigma(\alpha)$ is larger than α. If β is any ordinal larger than α, then $\alpha \in \beta$ and $\alpha \subset \beta$, hence $\sigma(\alpha) \subset \beta$. Therefore, $\sigma(\alpha)$ is the smallest such ordinal.

Corollary 8.3.2. If u is any set of ordinals, then $\beta = \bigcup_{\alpha \in u} \alpha$ is an ordinal. If the given set u has no largest ordinal, β is larger than every $\alpha \in u$.

By definition, β is a set whose elements are ordinals. It is also complete, for if $x \in \beta$, then $x \in \alpha$ for some α, so $x \subset \alpha$ by the completeness of α. Take any $t \in x$. Then $t \in \alpha$, so $t \in \beta$. Hence $x \subset \beta$, and β is a complete set of ordinals, thus an ordinal.

Finally, suppose that u has no largest member. Clearly, $\beta \supset \alpha$ for every $\alpha \in u$. If $\beta = \alpha$ for some α, then there is a second $\alpha_1 \in u$ with $\beta = \alpha < \alpha_1$, contradicting $\beta \supset \alpha_1$. Thus β is larger than every α.

9. Well-Ordering of Sets

The goal of this chapter is to demonstrate that for sets the property of being well-ordered is effectively (up to order-isomorphism) a characterization of the ordinal numbers. It can then be shown, by use of the axiom-of-choice, that a well-ordering can be imposed on an arbitrary set, from which it follows that any set is order-isomorphic to some ordinal number (if the order on the set is properly defined).

In order to begin the study of well-ordering, we must distinguish between two "kinds" of ordinals. We will say that an ordinal is of type I if it is 0 or the successor of some other ordinal; otherwise, we will say it is of type II. The ordinal ω is an example of the second sort, which are often called "limit" ordinals because they can be characterized according to the following theorem.

Theorem 9.1. If α is an ordinal, it is either of type I or $\alpha = \bigcup_{\beta \in A} \beta$, where $A \subset \Omega$ and A has no largest element.

For let A be the class of all ordinals smaller than α, and let $\alpha_0 = \bigcup_{\beta \in A} \beta$. If A has a largest element, that element is α_0; then $\alpha = \alpha_0 \cup \{\alpha_0\}$, so α is of type I. If A has no largest element, α_0 is larger than every element of A by 8.3.2, so $\alpha_0 = \alpha$.

We shall now use the well-ordered property of the ordinals as machinery to justify the procedure of defining function by recursion, thus completing the proof of 6.5. We will prove the theorem first for functions defined on ordinals, and then generalize to functions defined on complete ~~sets~~ classes of ordinals.

Let A be a class. We define an A-sequence to be a function f with domain an ordinal α and range a subclass of A. If f is an A-sequence defined on α and $\beta < \alpha$, it follows immediately from the definition that $f \upharpoonright \beta$ is an A-sequence.

Theorem 9.2. Let γ be an ordinal, A a class and G a function defined from all A-sequences to A. Then there exists exactly one function f_γ from γ to A with $f_\gamma{}'\alpha = G'(f_\gamma \upharpoonright \alpha)$ for all $\alpha \in \gamma$.

By 8.1, if a class of ordinals, C, has the property that for all γ $(\beta)(\beta < \gamma \Rightarrow \beta \in C) \Rightarrow \gamma \in C$, then $C = \Omega$, for if it were not, a contradiction could be derived by considering the least ordinal not contained in the class C. This justifies our assuming inductively that the theorem is true for all $\beta < \gamma$.

If $\gamma = 0$, the theorem is trivially true.

Suppose $\gamma \neq 0$ and is of type I. Then there exists an ordinal, δ, so that $\gamma = \delta \cup \{\delta\}$. Let $f_\gamma = f_\delta \cup \{\langle G'f_\delta, \delta\rangle\}$; then f_γ has the desired property.

Suppose γ is of type II. By 9.1, $\gamma = \bigcup_{\beta < \gamma} \beta$, and we can define inductively f_β on β for all $\beta < \gamma$. Since $\beta_1 < \beta_2$ implies that $f_{\beta_2} \upharpoonright \beta_1 = f_{\beta_1}$, $f_\gamma = \bigcup_{\beta < \gamma} f_\beta$ is a single-valued function and has the desired property.

It is clear that in either case only one function has the desired property, for if two distinct functions have the property, consideration of the least ordinal on which the functions disagree will lead to a contradiction.

Theorem 9.3. Let M be a complete class of ordinals. Given a class A and a function G from all A-sequences to A, then there exists exactly one function F on M to A with $F'\alpha = G'(F \upharpoonright \alpha)$ for all $\alpha \in M$.

The theorem is true if M is an ordinal, by 9.2. Otherwise, let $F = \bigcup_{\gamma \in M} f_\gamma$. If M contained some greatest ordinal, then M would be an ordinal by 8.3.2, so if $\alpha \in M$, for some $\beta \in M$, $\alpha \in \beta$, and $f_\beta' \alpha = G'(f_\beta \upharpoonright \alpha)$, so $F'\alpha = G'(F \upharpoonright \alpha)$.

We now employ the technique of defining functions by recursion to prove three fundamental theorems of well-ordering. We first need to define the notion of an R-section (where R is an arbitrary relation) of a class Y; we say that X is such an object if every R-predecessor in Y of an element of X also belongs to X. X is a proper R-section of Y if it is an R-section of Y but not the whole of Y.

Theorem 9.4. If A is a proper class well-ordered by W such that any proper W-section is a set, then A is order-isomorphic to Ω.

Let Rx denote the range of x. Let G be the relation with domain all A-sequences and range A such that $\langle y, x\rangle \in G$ if and only if $y \in (A - Rx)$ and $(A - Rx) \cap (W''\{y\}) = \emptyset$. G is, in fact, a function; for if $\langle u, x\rangle$ is also contained in G, $u \neq y$, we have uWy (or yWu, treated similarly), so $u \in A - Rx$ and $u \in W''\{y\}$, a contradiction. The function G is defined on all A-sequences, x, for since x is a set, Rx is a set, so $A - Rx$ is not empty. By 9.3, we can define F on Ω so that $F'\alpha = G'(F \wedge \alpha)$; $F'\alpha$ will then be the first element in $A - R(F \wedge \alpha)$.

F is one-to-one. For if not, we have $\alpha < \beta$ and $F'\alpha = F'\beta$. Since $\langle G'(F \wedge \beta), F \wedge \beta \rangle \in G$, we have $G'(F \wedge \beta) = F'\beta \in A - R(F \wedge \beta)$, but $F'\beta = F'\alpha \in R(F \wedge \beta)$, a contradiction.

F maps Ω onto A. It is clear from the definition of F that $RF \subset A$. Since F is one-to-one, it follows from Axiom C that RF is a proper class. It therefore suffices to show that RF is a W-section. We will show that $\beta W F'\alpha$ implies that $\beta \in R(F \wedge \alpha)$. If not, we have $\beta \in W''\{F'\alpha\}$ and $\beta \in A - R(F \wedge \alpha)$, but since $\langle F'\alpha, F \wedge \alpha \rangle \in G$ it follows that $W''\{F'\alpha\} \cap A - R(F \wedge \alpha) = \emptyset$, a contradiction.

It remains to prove that F is order preserving, that $\alpha < \beta$ implies $(F'\alpha)W(F'\beta)$. If not, since F is one-to-one, we would have $(F'\beta)W(F'\alpha)$, or $F'\beta \in W''\{F'\alpha\}$. But $A - R(F \wedge \beta) \subset A - R(F \wedge \alpha)$ and $F'\beta \in A - R(F \wedge \beta)$, so $F'\beta \in A - R(F \wedge \alpha) \cap W''\{F'\alpha\}$, a contradiction.

Theorem 9.5. <u>If</u> a <u>is a set well-ordered by</u> W, <u>then</u> a <u>is order-isomorphic to some ordinal</u>.

Define F as in 9.4. Since a is a set, for some α we have that $A - R(F \wedge \alpha) = \emptyset$, or else RF, a proper class, is a subclass of a. Let β be the least such α; then $F \wedge \beta$ is a one-to-one order-preserving map from β onto a.

Theorem 9.6. <u>If</u> a <u>is a set, there exists an ordinal</u> α <u>and a 1-1 onto function</u> F <u>from</u> α <u>to</u> a. (This function can be thought of as transmitting the well-ordering of α to a, so the theorem can be phrased "any set can be well-ordered.")

The proof relies heavily on Axiom E. There exists by that axiom a function C so that $x \neq \emptyset$ implies $C'x \in x$. We define F, as permitted by 9.3, so that $F'\alpha = C'(a - R(F \upharpoonright \alpha))$. ‡ There exists ordinals β so that $R(F \upharpoonright \beta) \supset a$, for otherwise F would be a one-to-one (verify as in 9.5) function from Ω to the set a. The least β, γ, has the property that $R(F \upharpoonright \gamma) = a$. ‡ if $a - R(F|\alpha) \neq \emptyset$. If $a - R(F|\alpha) = \emptyset$ (i.e. $R(F|\alpha) \supseteq a$, define $F'\alpha = a (= C'\{a\})$.

There is one thing fishy about this proof although the basic idea is right.

Chapter 10. Cardinal Numbers

We have seen in Chapter 9 that sets are characterized by the ordinal numbers up to order-isomorphism. It is natural to inquire about the characterization that results when the requirement of order preservation is dropped. This motivates us to define the notion of class equivalence so that A and B are said to be equivalent if and only if there exists a one-to-one function C with the property that the domain of C is A and the range of C is B.

By 9.6, we know that an arbitrary set x is equivalent to some ordinal. We define the cardinal of a set x, $\bar{\bar{x}}$, to be the smallest ordinal with which x is equivalent. We define the cardinal of a proper class to be Ω and denote the class of all cardinals by N. As might be expected, the map from the class of all sets (ordered by inclusion) which takes a set into its cardinal is order preserving.

Theorem 10.1. If $\alpha \in N$ and $x \subseteq \alpha$, then $\bar{\bar{x}} \le \alpha$.

Since x is well-ordered, by 9.5 there exists an ordinal β and a function $f : \beta \longrightarrow x$ that is one-one, onto and order preserving. If $t \in \beta$, $t \le f(t)$ (this is a general property of one-one order preserving maps on well-ordered sets, the proof of which is left as an exercise), and $f(t) < \alpha$, so $t < \alpha$ and $\beta \subset \alpha$. Since $\bar{\bar{x}} \le \beta$, we conclude that $\bar{\bar{x}} \le \alpha$.

The following three corollaries are easy consequences of Theorem 10.1. Their proofs are left as exercises.

Corollary 10.1.1. $x \subset y$ implies $\bar{\bar{x}} \le \bar{\bar{y}}$.

Corollary 10.1.2. $\alpha > 1$ implies $\overline{\overline{\sigma(\alpha)}} \le \overline{\overline{\alpha \times \alpha}}$.

Corollary 10.1.3. If f is a function, $\overline{\overline{f''x}} \le \bar{\bar{x}}$.

It is clear from 10.1.1 that the cardinality of $\mathcal{P}A$ is $\geq$ the cardinality of A, since there exists a natural equivalence map between the class A and the class of singletons of elements of A, a subclass of $\mathcal{P}A$. A famous theorem, due to Cantor, asserts that for sets this inequality is strict.

Theorem 10.2. $\overline{\overline{x}} < \overline{\overline{\mathcal{P}x}}$

Assume $\overline{\overline{x}} = \overline{\overline{\mathcal{P}x}}$. Then there exists a one-one onto mapping f from x to $\mathcal{P}x$. Let u be the subset of x so that $t \in u$ if and only if $t \notin f(t)$. Then some element of x, say v, has the property that $f(v) = u$. If $v \in u$, by definition of u we have that $v \notin f(v) = u$, so we conclude that $v \notin u$. But if $v \notin u = f(v)$, then $v \in u$. This contradiction demonstrates the impossibility of the existence of a function satisfying the condition imposed on f, so $\overline{\overline{x}} < \overline{\overline{\mathcal{P}x}}$.

Corollary 10.2.1. <u>If</u> $x \subset N$, <u>there exists some</u> $\alpha \in N$ <u>so that</u> $\beta \in x$ <u>implies that</u> $\beta < \alpha$.

Let $u = \sum x$, $\alpha = \overline{\overline{\mathcal{P}u}}$. Then since $\beta \in u$, we have $\beta = \overline{\overline{\beta}} \leq \overline{\overline{u}} < \alpha$.

Corollary 10.2.2. <u>N is a proper class.</u>

If not, by 10.2.1 there exists some $\alpha \in N$ greater than all the cardinals, which is absurd, since α is a cardinal.

Corollary 10.2.3. <u>There exists a function</u> $\aleph$ <u>mapping</u> Ω <u>onto</u> $N - \omega$ <u>such that</u> $\aleph$ <u>is one-one and order preserving.</u>

$N - \omega$ is a proper class, well-ordered and each of its proper sections is a set; therefore, the corollary follows from 9.4.

In particular, $\aleph(0) = \omega$ and $\aleph(1)$ is the first uncountable ordinal. Henceforth, $\aleph(\alpha)$ will be written $\aleph_\alpha$.

It is possible to well-order $\Omega \times \Omega$ by the following definition: $\langle \alpha, \beta \rangle < \langle \gamma, \delta \rangle$ if $\max\{\alpha, \beta\} < \max\{\gamma, \delta\}$; or $\max\{\alpha, \beta\} = \max\{\gamma, \delta\}$ and $\beta < \delta$, or $\beta = \delta$ and $\alpha < \gamma$. With this ordering, $\Omega \times \Omega$ is a well-ordered proper class having each proper section a set (verify). Therefore, there exists a function P which maps $\Omega \times \Omega$ onto Ω that is one-one and order-preserving. This order is employed in the proof of the following theorem.

Theorem 10.3. $\underline{\aleph_\alpha = \overline{\overline{\aleph_\alpha \times \aleph_\alpha}}}$.

If the theorem is false, there is a least ~~cardinal~~ ordinal, say γ, for which the equality does not hold. Suppose $\langle a, b \rangle \in \aleph_\gamma \times \aleph_\gamma$, and let x be the set of predecessors of $\langle a, b \rangle$ in the ordering $\Omega \times \Omega$. It suffices to show that $\bar{\bar{x}} < \aleph_\gamma$.

Let u be the maximum of a and b; then $u \subset \aleph_\gamma$ and $x \subset$ the predecessors of $\langle u, u \rangle$, so $x \subset \sigma(u) \times \sigma(u)$. If u is finite, $\bar{\bar{x}} \leq (u+1)^2 < \aleph_\gamma$. If u <u>is infinite,</u> $\bar{\bar{u}} = \aleph_\delta$ for some $\delta \in \Omega$, $\delta < \gamma$; by choice of γ, $\overline{\overline{\aleph_\delta \times \aleph_\delta}} = \aleph_\delta$, so $\overline{\overline{u \times u}} = \bar{\bar{u}}$. By 10.1.2, $\overline{\overline{\sigma(u)}} \leq \overline{\overline{u \times u}} = \bar{\bar{u}}$, so $\bar{\bar{x}} \leq \overline{\overline{\sigma(u) \times \sigma(u)}} = \bar{\bar{u}} < \aleph_\gamma$, completing the proof.

Corollary 10.3.1. x <u>infinite implies</u> x <u>is equivalent to</u> $x \times x$.

The proof is immediate, since $\bar{\bar{x}} = \aleph_\alpha$ for some α.

Corollary 10.3.2. <u>Let</u> $\alpha, \beta \in N$, $\alpha \neq 0$, $\alpha \leq \beta$, β <u>infinite.</u> <u>Then</u> $\alpha + \beta = \alpha \cdot \beta = \beta$, <u>where</u> $\alpha + \beta = \overline{\overline{\alpha \cup \beta}}$ (disjoint union) not necessarily <u>and</u> $\alpha \cdot \beta = \overline{\overline{\alpha \times \beta}}$.

It is clearly true if $\alpha = 1$; if $\alpha \geq 2$, we have $\beta \leq \alpha + \beta \leq \alpha \cdot \beta$, the latter from an easily constructible explicit correspondence between $\alpha \cup \beta$ and $\alpha \times \beta$. Since $\alpha \times \beta \subset \beta \times \beta$, we have $\overline{\overline{\alpha \cdot \beta}} \leq \bar{\bar{\beta}}$, by 10.1.1.

Chapter 11. Zermelo and Zermelo-Fraenkel Axiomatizations

The first careful axiomatization of set theory, authored by Gottlob Frege in 1884, proved unsatisfactory because certain of the entities recognized as sets (such as "the set of all sets not members of themselves") led to logical inconsistencies. In an attempt to sidestep this difficulty, Ernst Zermelo (1908) proposed a system of axioms (Z) which admitted as sets only the results of specific operations on sets known previously to exist (either from the axioms or subsequent derivations) and certain subsets of these sets. Zermelo's axiomatization was also considered to be somewhat unsatisfactory, however, because the number of axioms required was infinite. Therefore, later mathematicians introduced the notion of "class" as an extra-logical primitive distinct from "set" and succeeded in reducing the axioms to a finite number, as in the Goedel-Bernays (GB) system which has been presented in the previous chapters.

Zermelo's system will now be described in order that it may be compared with GB. The primitive terms are "set" and the binary set-relation "$\in$"; the binary set-relation "$=$" is defined by $x = y \overset{\text{by definition}}{\Longleftrightarrow} \text{df}(t)\ (t \in x \Longleftrightarrow t \in y)$. The axioms follow; the fifth, which is not really an axiom but rather an axiom schema, effectively corresponds to 4.3 but is more restrictive in that it allows only the construction of the set of all elements of a given set which satisfy a given property, as opposed to the construction of the class of all elements satisfying a given property.

Z - I. Extensionality: $(x = y \;\&\; y \in z) \Longrightarrow x \in z$.

II. Pairing : for all sets u, v, $\exists$ a set $\{u,v\}$ such that $(t)\ (t \in \{u, v\} \Longleftrightarrow (t = u \vee t = v))$.

III. Union : If u is a set, $\exists$ a set $\sum u$, with

$$t \in \sum u \iff (\exists x)(t \in x \,\&\, x \in u).$$

IV. Power set: If u is a set, $\exists$ a set $\mathcal{P}u$, with

$$t \in \mathcal{P}u \iff t \subset u.$$

V. Separation: given a wff $\phi(t)$, then $(u)(\exists v)(x)(x \in v \iff x \in u \,\&\, \phi(x))$

VI. Axiom of Choice: (local version) $(x)(\exists f)$ so that f is a function ~~of~~ with domain x and $y \in x \,\&\, y \neq \emptyset \implies f(y) \in y$.

VII. Axiom of Infinity: the existence of ω as described in A6 .

VIII. Axiom of Foundations: Axiom D.

The subject of "comparative axiomatics" is said to be "metamathematical" in that the proofs are not carried through within an axiomatic system but deal with the systems themselves. In this chapter and the next we shall metamathematically compare systems as to their relative consistency and their relative strength.

An axiom system is said to be consistent if it is not possible to prove both a formula and its denial in the system. In case the notion of denial (negation) is not present in the system (which is not the case in any of the systems presented in these notes, since all contain the propositional caluculus in their logical base), an alternative definition is available: a system is consistent if some well-formed formula is not provable. That these are equivalent (when the notion of denial is present) follows from the fact that $(p \,\&\, \sim p) \implies q$ is a theorem in the propositional calculus.

We shall say that a system A is weaker than a system B if every proposition demonstrable in A is also demonstrable in B, properly weaker if, in addition, there is a well-formed formula in A which is not demonstrable in A but is demonstrable in B. A and B will be said to be of <u>equal strength</u> relative to A if the demonstrable propositions in A are precisely the demonstrable propositions in B which are well-formed in A.

Theorem 11.1. Z <u>is properly weaker than</u> GB.

All the axioms of Z are axioms or theorems of GB (V can be derived from 5.1, and the class theorem, ~~and VI follows from~~ Axiom E), so Z is certainly weaker than GB.

Within the Zermelo system, it is possible to construct a number theory which is sufficiently adequate to justify the notational use of recursive functions defined on the integers (all of whose values are sets). The details as to how such a construction is carried out may be found in Paul Bernays' <u>Axiomatic Set Theory</u>, Amsterdam (North-Holland Publishing Company, 1950, Ch. 3 (section 4)). As a result, if we define F by

$$F'0 = \omega$$
$$F'(n+1) = P(F'n),$$

the statement $(\exists z)\,(x \in z \Longleftrightarrow (\exists n)\,(x \in F'n))$ is translatable into a well-formed formula, S, of the Zermelo system, even though F need not be a set in Z.

However, F is a function (class) in GB, so by Axiom C1 the image of this function, $t = \{\omega, P\omega, P(P\omega), \ldots\}$, is a set in GB, as is $\sum t$ by Axiom C2. It is an easy matter to verify that $\sum t$ has as elements precisely those that z was required to contain, so S is demonstrable in GB.

On the other hand, it can be shown that S is not demonstrable in Z by assuming that the elements of the universe are precisely the subsets of elements of t. Were this the case, the axioms of Z would be satisfied (verify), but S would not be true, since $\sum t$ is not a subset of $P^n\omega$ for any n. Therefore, the formula S is not a logical consequence of the axioms of Z.

We pass to a modification of the Zermelo system by Adolf Fraenkel (1925), which we will denote by ZF. The primitive terms are the same; the axioms are:

ZF - I. Extensionality.

III. Union.

IV. Power set.

VI. Axiom of Choice either that of Z or

$$(t)((x, y \in t \Longrightarrow x \cap y = \emptyset) \Longrightarrow (\exists u)(x \in t \& x \neq \emptyset$$
$$\Longrightarrow x \cap u \quad \text{is a singleton})).$$

VII. Axiom of Infinity.

VIII. Replacement given a wff $\emptyset(s, t)$ so that

$$\emptyset(s_1, t_1) \& \emptyset(s_1, t_2) \Longrightarrow t_1 = t_2, \qquad \text{then}$$
$$(x)(\exists y)(t \in y \Longleftrightarrow (\exists s)(s \in x \& \emptyset(s, t)).$$

IX. Axiom of Foundations .

The conspicuous omission of II and V is due to the fact that they are consequences of the other axioms. To derive II it is merely necessary to show that there exists a set with two elements (verify!); a function $\emptyset$ can then be defined so that one of the elements is mapped to x and all the remaining elements are mapped to y. By application of VIII (which is an axiom schema that effectively says the image of a function is a set) we have II. V also follows from VIII by letting the $\emptyset(s, t)$ of VIII be $\emptyset(s) \& (s = t)$, the latter $\emptyset$ being the $\emptyset$ of V. We have thus proved that ZF is stronger than Z; the complete result follows.

Theorem 11.2. Z is properly weaker than ZF.

The preceding has shown that all theorems derivable from Z are derivable from ZF. The converse is false, since in Z the image of a set need not be a set, as was demonstrated in the proof of 11.1.

It is natural to ask how ZF compares with GB. We present the result without proof. The latter may be found, however, in an article by Joseph R. Schoenfield entitled "A Relative Consistency Proof" (The Journal of Symbolic Logic, 1954, 19: 21-28).

Theorem 11.3. ZF and GB are of equal strength relative to ZF. (In other words, the provable formulas of ZF are exactly the provable formulas of GB which involve no class terms.)

Corollary 11.3.1. If ZF is consistent, so is GB.

If GB were not consistent, two contradictory formulas, p and $\sim p$, could be derived from its axioms. Since contradictory formulas imply all well-formed formulas, we have that all formulas of GB, including "class-free" ones, are deducible from the axioms of GB. By the theorem, all formulas of ZF would be deducible from the axioms of ZF. (The corollary, due to Novak and Mostowski, is historically prior to the theorem.)

In Z we can adopt in place of Axiom VI a stronger version of the Axiom of Choice: "The primitive symbol C has the property that if t is a well-formed set term, then so is Ct, and $x \neq 0 \Longrightarrow Cx \in x$." We call the modified system ZC. The following result compares the strength of ZC and ZF; it was first presented by Azriel Levy in "Comparing the Axioms of Local and Universal Choice" (Essays on the Foundations of Mathematics, Jerusalem (Magnus Press), 1961, pp. 83-90) and is stated here without proof.

Theorem 11.4. If a theorem not involving the term "C" is provable in ZC, it is provable in ZF.

Chapter 12. Inner Models.

We are sometimes able to show that the formation of a new axiomatic system by the addition of a particular axiom does not affect the consistency of the system, even though the appended axiom cannot be deduced from the others. Such a relative consistency proof is often carried out by the construction of an inner model.

Suppose we have an axiomatic system A, consisting of axioms $A_1, A_2, \ldots, A_n$. It is often possible to reinterpret certain of the notions of A so that the reinterpreted axioms of A, denoted by $A_1^*, A_2^*, \ldots, A_n^*$, are theorems of A. If an additional proposition, B, has the property that its reinterpreted formulation, B^*, is a theorem of A, then it follows that A with B appended is as consistent an axiomatic system as A itself. For if the assumption that $A_1, A_2, \ldots, A_n$ and B are true leads to a contradiction, then this contradiction can also be derived from $A_1^*, A_2^*, \ldots, A_n^*$ and B^*, hence follows from A.

Goedel was able to construct from the Goedel-Bernays system minus Axiom E a model in which all of the Goedel-Bernays axioms, including E, were satisfied. His demonstration is often referred to as a proof of the "consistency of the Axiom of Choice", but it actually testifies only to the relative consistency of the axiom. Goedel's model also has the property that it satisfies the generalized Continuum Hypothesis, a mathematical statement whose proof has been unattainable from the Goedel-Bernays axioms. This celebrated hypothesis states that $\mathcal{P}(\aleph_\alpha) = \aleph_{\alpha+1}$ for all ordinals α. So Goedel succeeded in proving that the Goedel-Bernays system lacking Axiom E is no more consistent than the apparently stronger system in which Axiom E and the generalized Continuum Hypothesis are added.

As an illustration of the use of an inner model, we will sketch a proof that demonstrates that the removal of Axiom D from the Goedel-Bernays system does not affect its consistency. A detailed proof may be found in an article by J.C. Shepherdson ("Inner Models for Set Theory", The Journal of Symbolic Logic, 1951, 16: 161-190), although the theorem was originally proved by von Neumann.

The essential notion we will use to construct our model is that of rank. We set $R(\alpha) = \sum_{\beta < \alpha} \mathcal{P}\{R(\beta)\}$; this defines R recursively as a function on Ω. For a set x, we define the rank of x to be the first γ so that $x \in R(\gamma)$; this makes $R(\alpha)$ the class of all sets of rank $\leq \alpha$. We have that

$R(0) = \emptyset$

$R(1) = \{\emptyset\}$

$R(2) = \{\{\emptyset\}, \emptyset\}$

$R(3) = \{\{\{\emptyset\}, \emptyset\}, \{\{\emptyset\}\}, \{\emptyset\}, \emptyset\}$, etcetera.

It is not *a priori* true that all sets will have a rank.

Theorem 12.1. *R has the following properties:*

1. $\beta \leq \alpha \Longrightarrow R(\beta) \subset R(\alpha)$
2. $t \in R\alpha \Longrightarrow t \subset R\alpha$
3. $\mathcal{P}R\alpha = R(\alpha + 1)$
4. $x \in R\alpha \Longrightarrow \mathcal{P}x \in R(\alpha + 2)$
5. if α is a limit ordinal, $R(\alpha) = \bigcup_{\beta < \alpha} R(\beta)$
6. $\alpha < \beta \Longrightarrow R\alpha < R\beta$
7. $x \in y \Longrightarrow$ rank of x < rank of y.

The proof of these statements are obvious; we will prove 7, leaving 1-6 for the reader.

If $y \in R(\alpha)$, then $y \in \bigcup_{\beta < \alpha} \mathcal{P}R(\beta)$, so $y \in \mathcal{P}R(\beta)$ for some $\beta < \alpha$. Hence $x \in y \subset R(\beta)$ for some $\beta < \alpha$.

We now define the class π to be the union of all $R(\alpha)$ over all ordinals α; that is, the class of all sets with a rank. As a consequence of this definition we have the following:

Theorem 12.2. $u \subset \pi \iff u \in \pi$.

Since $x \in u \implies x \in \pi$, the function Rk which corresponds to any set its rank is defined on u. The image of u under Rk is a set of ordinals, follows easily from the fact that Ω is a proper class (by 9.5, 9.6). but cannot be the whole of Ω ~~by 9.4~~. If $\beta = \bigcup_{\alpha \in Rk'' u} \alpha$, we have by 8.3.2 that β is as large as any member of $Rk''u$. Then $x \in u$ implies that $Rk\,x < \beta$, so $x \in R(\beta)$, hence $u \in R(\beta + 1)$.

The next theorem is of interest but is somewhat of a sidetrack since it is more like the converse of the theorem we are seeking.

Theorem 12.3. <u>Axiom D implies that</u> $\pi = \mathcal{U}$.

For suppose not. Then $\mathcal{U} - \pi \neq \emptyset$, so by D there exists some $y \in \mathcal{U} - \pi$ with the property that $y \cap (\mathcal{U} - \pi) = \emptyset$. But then $y \subset \pi$, and by 12.2 $y \in \pi$, a contradiction.

We are now ready to outline the proof of von Neumann's result. We begin with the axioms A, B, C, and E, and denote this system by S. Axiom D was relied upon heavily in the construction of the ordinals and the verification of their properties, but its need can be eliminated by incorporating Axiom D into the definition of the ordinal numbers. In other words, we can in the system S define an ordinal to be a set α such that

1. α is complete

2. α is linearly ordered by $\in$

3. $\emptyset \neq x \subset \alpha \Longrightarrow (\exists y)(y \in x \,\&\, y \cap x = \emptyset)$.

The ordinals defined in this way will satisfy in S all of the properties of the ordinals in GB.

We can therefore define rank in S and introduce π, the class of sets with a rank. We reinterpret four notions:

for	"$(x)\ldots$"	read	"$(x)\, x \in \pi \Longrightarrow \ldots$"
for	"$(\exists x)\ldots$"	read	"$(\exists x)\, x \in \pi \,\&\, \ldots$"
for	"$(A)\ldots$"	read	"$(A)\, A \subset \pi \Longrightarrow \ldots$"
for	"$(\exists A)\ldots$"	read	"$(\exists A)\, A \subset \pi \,\&\, \ldots$" .

Of course, all derivative notions will have to be reinterpreted in light of this. We will denote reinterpreted propositions by the superscript * .

Theorem 12.4. <u>A^*, B^*, C^* and E^* are theorems in S.</u>

That A^*, B^*, and C^* are theorems is easy to prove and is left to the reader. To prove E^*, it is necessary to show that there exists a choice function on π that is a subclass of π. By Axiom E, there exists a choice function C defined on $\mathcal{U}$; it will be demonstrated that $C \wedge \pi$ is a subclass of π, hence the required function.

The image of π under C is a subclass of π, for if $y \in \pi$ and $x \in y$ it follows from 12.1, 7) that $x \in \pi$. Since $C \wedge \pi \subset \pi \times \pi$, it suffices to show that $\pi \times \pi \subset \pi$. But if $\alpha = \max(\text{rank } x, \text{rank } y)$, then $x, y \in R(\alpha)$, so $\langle x, y \rangle \in \mathcal{P}^2(R(\alpha))$, so $\langle x, y \rangle \in R(\alpha + 2)$, proving the theorem.

Theorem 12.5. <u>D^* is a theorem in S.</u>

If not, it can be shown as in 6.5 that there exists a function f on ω to π with $\ldots f'(n+1) \in f'n \in \ldots \in f'2 \in f'1$. Since by 12.1 7) we have that $x \in y \Longrightarrow \text{rank } x < \text{rank } y$, $\{f'n\}$ represents an infinite descending chain of ordinals. But this is impossible by 8.1.

From our previous remarks, it follows that S and GB are relatively consistent.