PROPERTY OF R. Schultz

MATHEMATICS 261

THE AXIOMS FOR SET THEORY

by
Saunders MacLane

AXIOMS FOR SET THEORY

1. Introduction. Since the standard foundations of Mathematics are based upon the concepts of a set (and the elementary facts about symbolic logic), and since the naive notion of a set leads to paradoxes, it is important to have an adequate axiomatic foundation of set theory. We present below the von Neumann-Godel-Bernays axioms for set theory. Further details appear in

K. Godel, The Consistency of the continuum hypothesis, Annals of Mathematics Studies, No. 3, Princeton, 1940.

P. Bernays, A system of axiomatic set theory, Journal for Symbolic Logic, vol. 2 (1937), pp. 65-77; vol. 6 (1941), pp. 1-17; vol. 7 (1942), pp. 65-89 and pp. 133-148; vol. 8 (1942), pp. 89-106; vol. 13 (1948), pp. 65-79.

The paradoxes depend upon the construction of "bad" sets which are "too big" or which contain themselves as elements. The essential device here adopted to avoid these paradoxes is that of distinguishing between sets which may be "bad" and those which are legitimate. To this end, we distinguish sets and classes. Every "collection" of things, in the naive sense, either good or bad, is a class. Only the legitimately constructed classes rate as sets; in particular, it will be arranged that classes which are not sets can never be members of other sets or classes. This arrangement for meeting the paradoxes is essentially simpler than the theory of types.

Classes and sets are the only mathematical objects we consider. In particular, the elements of a class or a set are themselves always sets; this is more convenient than the alternative of starting with "atomic" elements from which the sets are built up.

2. The axiom of extensionality. The basic primitive notions are those of class and membership. We write capital letters A, B, C,... for classes, and $\in$ for membership so that $A \in B$ means that the class A is a member of the class B. By definition, a class A is a set if and only if $A \in B$ for some class B. We write the lower case letters x, y, z,... for (classes known to be) sets. Thus the formulas $x \in y$, $x \in A$ mean that the set x is a member of the set y or the class A, respectively.

It is convenient to use logistic symbols as follows:

$\forall x$: "for every set x"

$\exists x$: "there exists a set x"

$\longrightarrow$: "implies"

$\longleftrightarrow$: "if and only if"

&: "and"

$\vee$: "or"

$\neg$: "not"

Class inclusion and class equality are defined as usual.

(2.1) $A \subset B$ for "$(\forall x)(x \in A \longrightarrow x \in B)$",

(2.2) $A = B$ for "$A \subset B$ and $B \subset A$."

Thus classes (or sets) are equal if and only if they have the same members. It follows readily that class equality is reflexive, symmetric, and transitive, and that class inclusion is reflexive and transitive.

The first axiom is the axiom of extensionality.

A1. If $A = B$, and $A \in C$, then $B \in C$.

In words, equal classes are members of the same things. In the hypothesis of A1 we have $A \in C$, hence A must be a set. It follows that any class B equal to a set is itself a set. Axiom A1 may be replaced by this statement, and the statement

A1' If $x = y$ and $x \in C$, then $y \in C$.

The general principle of substitution for equality states that, in any valid formula involving the set x, one may replace x by any equal set y. The statement A1' is a special case of this general principle, for the formula "$x \in C$". From this special case the general principle may be deduced.

3. The null set and ordered pairs. A basic feature of set theory is the construction of new sets from given ones. The next axioms assert the possibility of two basic constructions of this sort. We must have one set as a starter; we adopt the null set 0. We wish also to be able to adjoin a new element x to a given set u; that is, form the set $\alpha(u, x)$ which has as elements x and the elements of the set u. The corresponding axioms are

A2. 0 is a set with no members.

A3. If u and x are sets, then $\alpha(u, x)$ is a set and

$z \in \alpha(u, x) \longleftrightarrow z \in u$ or $z = x$.

In particular, $\alpha(0, x)$ will be the set whose only element is x. This set is usually written $\{x\}$, and called the unit set of x. Also, we define

$$\{x, y\} = \alpha(\{x\}, y) = \alpha(\alpha(0, x), y);$$

it is the set whose only elements are x and y; that is

$$(3.1) \qquad z \in \{x,y\} \longleftrightarrow z = x \text{ or } z = y.$$

Thus $\{x,y\}$ acts like the unordered pair of the sets x and y.

The notion of an ordered pair is essential to the definition of functions and relations. We define it as

$$(3.2) \qquad < x, y > = \{\{x\}, \{x,y\}\},$$

the set whose only elements are the unit set of x and the set with exactly x and y as elements. This definition can also be written in terms of α directly. Because the expression on the right of (3.2) is not symmetric in x and y, it has the essential property that the ordered pair $< x, y >$ is different from the ordered pair $< y, x >$ unless $x = y$. More explicitly, one proves the

THEOREM 3.3. $< x, y > = < u, v >$ if and only if $x = u$ and $y = v$.

This states that the symbol $< x, y >$ has the basic property requisite for an ordered pair.

An ordered triple of sets may now be defined as

$$(3.4) \qquad < x, y, z > = < x, < y, z > > .$$

Note that this is not the same as $< < x, y > z >$. More generally, one defines an ordered n-tuple by induction on the positive integer n as follows:

$< x > = x$; $< x, y >$ as above; and for $n \geqq 2$,

$$(3.5) \qquad < x_1, \ldots, x_n, x_{n+1} > = < x_1, < x_2, \ldots, x_{n+1} > > .$$

Induction on n and Theorem 3.3 then yield

(3.6) $< x_1, \ldots, x_n > = < y_1, \ldots, y_n >$ if and only if $x_1 = y_1, \ldots, x_n = y_n$.

One may also show by induction on n, for fixed m that

$$(3.7) \qquad < x_1, \ldots, x_n, < x_{n+1}, \ldots, x_{n+m} > > = < x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+m} >.$$

Indeed, the case n = 1 is exactly the definition (3.5).

It is possible to replace Axiom A3 by the two stronger but more natural axioms:

A3a. If x is a set, $\{x\}$ is a set and $y \in \{x\} \longleftrightarrow y = x$.

A3b. If x and y are sets, $x \cup y$ is a set, and $z \in (x \cup y) \longleftrightarrow z \in x$ or $z \in y$.

The set $\{x, y\}$ with exactly x and y as elements may then be defined as

$$\{x, y\} = \{x\} \cup \{y\} ,$$

and the ordered pairs, etc. are constructed as before. One may also define α by

$$\alpha(u, x) = u \cup \{x\} .$$

In the definition (3.5) we apparently use the ordinary integer n. However, it is better to regard (3.5) as a set of instructions for writing down an infinite sequence of definitions of $< x_1, \ldots, x_n >$, one for each n. In any one argument, we use only a finite number of these definitions. Similarly (3.6) and (3.7) are descriptions of an infinite sequence of theorems, one for each n (resp. n and m).

Problems: Prove (3.5), (3.6), and (3.7).

(3.8) Express the ordered pair directly in terms of α .

4. Formulas. Many constructions of sets depend on the description of a new set u as the set of all elements having such and such a property. In order to formulate this and similar principles, we interpret "property" to mean anything which can be formulated in terms of the primitive notions of our axiomatics and the usual logical connectives. The connectives are

& (and), $\vee$ (or), - (not), $\longrightarrow$ (implies),

together with the "quantifiers" $(\forall x)$ (for all sets x) and $(\exists x)$ (there exists a set x such that...). It is important that we do not use these quantifiers for classes; that is, we shall not consider properties defined by phrases such as "for all classes A, ...".

More precisely, we shall define the notion of a <u>formula</u> W of set theory. Such a formula will be a sequence of symbols; constructed in grammatical fashion from the primitives and the connectives and quantifiers. Those set variables x, y, ... which appear in the formula in a context $(\forall x)$ or $(\exists y)$ will be said to be <u>bound letters</u> in the formula. Otherwise they are said to be <u>free</u> in the formula. The related notions "formula" and "bound" or "free" letters involved in a formula are defined by induction on the length of the formula as follows:

(4.1) If x, y are set letters and A is a class letter then $x \in y$, $x \in A$ are formulas in which x and y (respectively x and A) are free. Formulas of this sort are termed <u>atomic formulas</u>.

(4.2) If W is a formula, so is -W; the free and bound letters of -W are those of W.

(4.3) If W_1 and W_2 are formulas such that no letter free in W_1 is bound in W_2, or vice versa, then

$$W_1 \mathbin{\&} W_2, \quad W_1 \vee W_2, \quad W_1 \longrightarrow W_2$$

are formulas. A letter is free (bound) in one of these formulas if and only if it is free (bound) in either W_1 or W_2.

(4.4) If W is a formula and x is free in W, then

$$(\forall x)\, W, \qquad (\exists x)\, W$$

are formulas. In each of them x is bound. Any other letter is free (bound) in one of the formulas if and only if it is free (bound) in W.

The hypothesis of (4.3) is designed to avoid such troublesome formulas as

$$[(\forall x) x = x] \ \& \ x \in y,$$

which can alternatively be expressed as

$$[(\forall z)\ z = z] \ \& \ x \in y.$$

Experience shows that any "mathematical" property of a set x, defined in terms of our own primitive notions, can be expressed by means of a formula with x as its only free letter. The same applies to properties of two sets, x, y, of a set x and a class A, etc.

It will be convenient to transform the formula $z \in u$ into a different form, not involving the statement "z is a member of a class u."

$$z \in u \ \longleftrightarrow \ (\forall x)\ (x = z \longrightarrow x \in u).$$

Now replace the equality symbol by its definition, to get

(4.5) $$z \in u \longleftrightarrow (\forall x)\ [(\forall y)(y \in x \longleftrightarrow y \in z)] \longrightarrow x \in u.$$

Note that on the right z appears only in the atomic formula $y \in z$, as the <u>second</u> argument of this formula.

This elimination of = illustrates the fact that the terms which are defined by means of the primitive terms of set theorems can be eliminated from a formula not involving them; thus we can always replace formulas involving $x \subset y$ and $x = y$ by formulas involving only the $\in$. In the same fashion, we might consider formulas built up from atomic formulas which involve our basic symbols 0 and $\alpha(u, x)$. From every such formula, these primitive symbols may be eliminated by the appropriate axiom, for each axiom states exactly which sets belong to it;

$$z \in 0 \longleftrightarrow -(z = z) \qquad \text{by (A2)},$$

$$z \in \alpha(u, x) \longleftrightarrow (z \in u) \vee z = x \qquad \text{by (A3)},$$

and similarly in other cases. If one of these primitive symbols occurs in front of an $\in$-sign, we first apply the transformation (4.5) to move it behind the $\in$-sign, and then the appropriate axiom.

For this reason, we have restricted our definition of formula to cover only formulas built up from the $\in$-symbols, and not formulas built up from other concepts of set theory.

In terms of the notion of formula, we may give an exact statement of the principle of substitution for equality

<u>THEOREM 4.6</u>. If $W(x)$ is a formula in which x is free, and $W(y)$ is the formula obtained from $W(x)$ by replacing x (at some or all of its occurrences in $W(x)$) by a letter y not previously occurring in $W(x)$, then

$$(x = y) \longrightarrow [W(x) \longleftrightarrow W(y)] .$$

The proof is by induction on the length of the formula $W(x)$. The initial step of the induction involves for $W(x)$ only the three possible atomic formulas

$$x \in z, \qquad x \in A, \qquad u \in x.$$

Hence the theorem in this case requires only the three following assertions

$$(x = y) \longrightarrow [x \in z \longleftrightarrow y \in z],$$

$$(x = y) \longrightarrow [x \in A \longleftrightarrow y \in A],$$

$$(x = y) \longrightarrow [u \in x \longleftrightarrow u \in y].$$

The first and second of these are immediate consequences of the axiom of extensionality (A1). The third is a consequence of the definition of the equality of sets.

In the remainder of the induction, when $W(x)$ is not an atomic formula, the proof is straightforward and depends upon the basic properties of the logical connectives. These we do not discuss in detail here, although their properties are already necessary for a systematic completion of our foundations.

Problem (4.7). Prove that the "set" of formulas is denumerable ("set" means intuitive set here).

5. The existence of classes. A binary relation R is by definition a class each of whose elements is an ordered pair. Similarly, for each particular positive integer n, an n-adic relation R is a class each of whose elements is an ordered n-ad. The basic axiom on the existence of classes is

AXIOM B. Let W be a formula involving the free set letters $x_1, \ldots, x_n$, and the (free) class letters $A_1, \ldots, A_p$. Then there exists an n-adic relation R such that

$$< x_1, \ldots, x_n > \in R \text{ if and only if } W.$$

For example, let W be the formula $-(x \in A)$. There is then a 1-adic relation (i.e., a class B) such that

(5.1) $\qquad x \in B \longleftrightarrow \text{not } (x \in A).$

By the definition of equality of classes, B is completely determined by A; indeed, it is exactly the complement A' of A.

The formula $x = x$ similarly defines a class U such that $x \in U \longleftrightarrow x = x$. Thus U is the <u>universal class</u>, whose elements are <u>all</u> sets.

If axiom B is applied to the formulas $x \in A \;\&\; x \notin B$, it yields the existence of the intersection $A \cap B$ of the given classes A and B. Thus on the basis of axiom B we can obtain the usual Boolean algebra of classes, from the operations $A \cap B$ and A'.

From the formula $x = y$ we obtain by the axiom the identity relation I, which is the class of all pairs $< x,x >$. From the formula $x \in y$ we obtain the class E of all pairs $< x,y >$ for which $x \in y$. From the formula $x \in A \;\&\; y \in B$ we obtain the cartesian product $A \times B$. From the formula $(\exists y)(< y, x > \in A)$ we obtain the class Proj(A), consisting of those elements x for which $< y,x >$ is a member of A, for some set y. If A is a relation, Proj(A) is its domain. From the formula $< x_2, x_1 > \in A$ we obtain the class Conv(A) of all pairs $< x_1, x_2 >$ such that $< x_2, x_1 > \in A$. If A is a relation, Conv(A) is its converse.

Thus axiom B can be used to construct a large number of classes. From this point of view, Axiom B is not a single axiom, but an infinite family of axioms, one for each possible formula W. It is possible to replace axiom B by 8 ordinary axioms, which assume the existence of the following seven classes:

A', $A \times B$, E, I, Proj(A), [redacted] U, Conv(A),

as described above. The eighth axiom concerns the existence of the set $\gamma(A)$, whose elements are all $<<zy>x>$ with $<z<yx>> \in A$.

(This is a sort of re-association). Given these eight axioms, Axiom B becomes a Theorem, the "class Theorem" of Bernays, loc. cit.

Infinite intersections and unions also may be obtained from Axiom B. If A is a given class, the formulas

$$(\exists y)(x \in y \ \& \ y \in A) \quad \text{and} \quad (\forall y)(y \in A \longrightarrow x \in y)$$

each yield the class of all elements x satisfying the stated condition. These classes (which, of course, depend on A) are respectively

$$\bigcup_{y \in A} y \quad , \qquad \bigcap_{y \in A} y \quad .$$

If A is a class, the formula

$$(\forall x) \ [\ x \in u \longrightarrow x \in A\ ]$$

defines the class P(A) of all <u>sets</u> u which are subsets of A. This is the usual "power class" of A. Note that it has as elements the <u>subsets</u> of A, but not the subclasses of A which are not sets.

6. <u>Functions</u>. A function may be described as usual as a suitable class of ordered pairs. Explicitly, a function F is a class with the following two properties

(6.1) $\qquad z \in F \longrightarrow (\exists x,y)\ [\ z = \langle y, x \rangle\ ]$,

(every element of F is an ordered pair), and

(6.2) $\qquad \langle y_1, x \rangle \in F \ \& \ \langle y_2, x \rangle \in F \longrightarrow y_1 = y_2$,

(the value y = F(x), when defined, is unique). The <u>domain</u> of the function F is the class of all those x's for which there exists a set y with $\langle y, x \rangle \in F$; the domain is thus exactly the set Proj (F), as defined above.

The ordinary symbol $F(x)$ for the value of the function x at the argument x may be defined formally. Consider the formula

(6.3) $$(\exists y)\ [t \in y \quad \langle y, x \rangle \in F]$$

with a free set variable t. By Axiom B, we may introduce the class C consisting of all t's with this property. If F is a function and x is in its domain, then there exists exactly one y with $\langle y, x \rangle \in F$, so that the class of these t's is exactly the class of all elements of y. Hence $C = y$, and C is a class. If x is not in the domain of F, there is no y with $\langle y, x \rangle \in F$, so that there can be no t's with the property (6.2). Thus C is the null set. Since C is determined by F and x, we set $C = F(x)$. We have proved

<u>THEOREM 6.4</u>. If F is a function and x is a set, then the class $F(x)$ of all t's with property (6.3) is a set. If x is in the domain of F, then $\langle F(x), x \rangle \in F$, and $F(x)$ is the (ordinary) value of F for the argument x. If x is not in the domain of F, then $F(x) = 0$.

Our notation then has the following artificiality: $F(x)$ is defined for every x, by setting $F(x) = 0$ when the (usual) $F(x)$ would be undefined.

If F is a function and A a class, the property

(6.5) $$(\exists x)\ [x \in A \ \& \ \langle y, x \rangle \in F]$$

holds for exactly those sets which are images of elements of A under the function F. By axiom B, this formula (6.5) thus defines exactly the class of all such y. Since this class is the (ordinary) image of the class A under the function F, we write it as $F_*(A)$. Thus

(6.6) $$y \in F_*(A) \longleftrightarrow (\exists x)[x \in A \;\&\; \langle y, x \rangle \in F] \quad .$$

In particular, if u is a set (and hence a class) we may construct the image $F_*(u)$ of the set u under the function. This is to be distinguished from the previous notation F(u) for the value of F at the argument u.

If A and B are two classes, the formula

$$(\exists y)\,[\langle z, y \rangle \in B \;\&\; \langle y, x \rangle \in A]$$

defines a class of pairs $\langle z, x \rangle$, which we denote by $B \circ A$. In particular, if F and G are functions, $F \circ G$ is their composite. One may establish the usual properties of the composite, such as

(6.7) $$(F \circ G)_*(A) = F_*(G_*(A)).$$

If A and B are classes, a function F en A to B (notation $F: A \longrightarrow B$) is a function F with domain A and with $F_*(A) \subset B$. We cannot here introduce the <u>class</u> of all functions on A to B, because these functions F are classes and not necessarily sets, so that they may not be themselves members of a class.

Any subclass of a function F is clearly a function. In particular, if A is any class and U the universal, the cartesian product $U \times A$ consists of all pairs $\langle y, x \rangle$ with $x \in A$, hence the intersection $F \cap (U \times A)$ consists of all pairs $\langle y, x \rangle$ in F with $x \in A$. Hence

$$F \cap (U \times A) = F|A$$

is the function F with domain "cut down" to A. If F has domain D, then $F|A$ has domain $D \cap A$.

Problem: Prove (6.7).

(6.8) If F is a function and A a class, then $y \in F_*(A)$ if and only if there is an x in $A \cap$ (domain F), with $y = F(x)$.

7. The existence of sets. The next axioms provide for the existence of suitable sets. They are

C1. If u is a set, P(u) is a set,

(the class of all subsets of u is a set)

C2. If u is a set, then $\bigcup_{x \in u} x$ is a set.

(the union of all sets x contained in a set u is a set)

C3. If u is a set and F a function, then $F_*(u)$ is a set

(any image of a set under a function is a set).

The third axiom is especially powerful in producing sets.

THEOREM 7.1. If $A \subset u$, where u is a set, then A is a set.

PROOF: The identity function I cut down to A yields the function $I|A$ which is the identity on A (i.e., consists of all pairs $\langle x, x \rangle$ with $x \in A$). The image of the set u under this function is precisely the class A. Hence A is a set, by C3.

One repeatedly wishes to construct sets as the set of all elements x which are in a given set u and have some specified property. This property will be given by a formula W(x) in which x is free. This construction is always legitimate, as follows:

THEOREM 7.2. If u is a set and W(x) a formula with x free, then there is a set w with

$$x \in w \longleftrightarrow [x \in u \; \& \; W(x)].$$

PROOF: By axiom B, the formula $[x \in u \; \& \; W(x)]$ determines the class A of all the sets x satisfying this formula. This class is clearly a subclass of u, hence is a set, say w, by Theorem 7.1. We write w in the usual notation

(7.3) $$w = [x \mid x \in u \; \& \; W(x)] .$$

The axioms C1-C3 enable us to construct the Boolean algebra of all subsets x of a set w. Indeed, the class P(w) of all subsets of w is a set, by Axiom C1. Hence by Theorem 7.1, any class of subsets of w is a set u. The union of all sets x in u is a set by C2. The intersection

$$\bigcap_{x \in u} x$$

of all sets x in u is also a set, since it is itself a subset of w. If $x \subset w$, then the usual complement of x in w ($= w \cap x'$) is a subclass of w, hence a set. The usual properties of this Boolean algebra follow.

In particular, the intersection $u \cap v$ of any two sets u and v is a set (by Theorem 7.1) and the union $u \cup v$ of two sets is also a set. For, we may construct the set $\{u, v\} = z$ with exactly u and v as elements, and hence by C2 the set

$$w = \bigcup x, \qquad x \in \{u, v\}$$

which is exactly the ordinary union $w = u \cup v$.

THEOREM 7.4. If u and v are sets, so is their cartesian product $u \times v$.

PROOF: We have already defined the cartesian product as the class $u \times v$ of all ordered pairs $< x, y >$ with $x \in u$ and $y \in v$; it remains to prove this class a set. The proof depends upon our definition of ordered pairs, as

$$< x, y > = \{ \{x\} , \{ x, y \} \} .$$

Since u and v are sets, so is $u \cup v$ and $P(u \cup v)$, the latter by Axiom C1. Since $x \in u$ and $y \in v$, both $\{x\}$ and $\{x, y\}$ are subsets of $u \cup v$, so that

$$\{ x \} ; \{ x, y \} \in P(u \cup v) .$$

Thus $< x, y >$ is a subset of $P(u \cup v)$, hence an element of $PP(u \cup v)$. The class $u \times v$ of all such $< x, y >$ for $x \in u$, $y \in v$ is thus a subclass of $PP(u \cup v)$. The latter is a set, by Axiom C1, hence $u \times v$ is a set, by Theorem 7.1.

COROLLARY 7.5. If u and v are sets, any function $F: u \longrightarrow v$ is a set.

PROOF: By definition, F is a subset of $v \times u$, hence a set, by Theorems 7.4 and 7.1.

COROLLARY 7.6. If u and v are sets, the class v^u of all functions $F: u \longrightarrow v$ is a set.

PROOF: Let f be a set variable. The formula F which states that the set f is a function and that $f: u \longrightarrow v$ defines, by Axiom B, a class A whose elements are (by Corollary 7.5) <u>all</u> functions on u to v. Since each such f is a subset of $v \times u$, it is an element of $P(v \times u)$. Thus A is a subset of the set $P(v \times u)$, hence is a set. It is the desired set v^u.

Problems (7.7). If u is a set, prove that proj(u) and Conv(u) are sets.

(7.8) If the functions f and g are sets, prove that $f \circ g$ is a set.

8. The restrictive axiom. It is convenient but not necessary to adopt the following axiom.

D. Every non-empty set u has an element x with $u \cap x = 0$.

Its force is indicated by the following Corollaries.

COROLLARY 8.1. If u is a set, $u \in u$ is false; i.e., no set can be a member of itself.

PROOF: Form the unit set $\{u\}$ of u. This is non-empty, hence has an element x with $\{u\} \cap x = 0$. But x can only be the set u; hence $\{u\} \cap u = 0$. If now $u \in u$, then both $u \in u$ and $u \in \{u\}$, so $\{u\} \cap u \neq 0$, a contradiction.

COROLLARY 8.2. If u and v are sets, then one cannot have both $u \in v$ and $v \in u$.

Proof as problem.

9. The axiom of infinity. If one deals only with finite sets, the naive concept of a set would suffice; on the other hand, any construction of mathematics from set theory will surely require infinite sets. Hence one of our axioms must assert the existence of an infinite set. We take a specific such axiom, asserting the existence of the set ω of non-negative integers. This depends

upon the following explicit construction of a model of the set of integers.

Regard the null set 0 as the integer 0, and define each positive integer n as the set whose elements are all preceding non-negative integers:

$$1 = \{0\}, \quad 2 = \{0, 1\}, \quad 3 = \{0, 1, 2\}.$$

For any given integer n, its successor $s(n) = n+1$ is then to be the set whose only elements are n and the elements of n. This is expressible directly in terms of our α-function as

(9.1) $$s(n) = \alpha(n, n)$$

or in ordinary notation as $s(n) = n \cup \{n\}$.

More generally we set $s(y) = \alpha(y, y)$ for any set y.

<u>LEMMA 9.2</u>. For any set y, $y \subset s(y)$ and $y \neq s(y)$.

PROOF: Since $s(y) = \alpha(y, y)$ contains all elements of y, by A3, $y \subset s(y)$ is immediate. If $y = s(y)$, then $y \in \alpha(y, y) = s(y) = y$, hence $y \in y$, contrary to Corollary 8.1.

In this fashion our present axioms allow for the construction of an infinite number of (different) sets n. We want more: a set with an infinite number of elements.

Call a set u <u>s-closed</u> if both

(9.3) $$0 \in u \;\&\; [(\forall y)(y \in u \longrightarrow s(y) \in u)].$$

By axiom B, we may form the class C of all s-closed sets u, and hence the intersection ω of all the s-closed sets u, thus

(9.4) $x \in \omega$ if and only if x belongs to every s-closed set u.

Thus ω is (intuitively) the class of all non-negative integers n, defined as above. The axiom of infinity is

C4. The class ω is a set.

THEOREM 9.5. The set ω is s-closed; i.e., $0 \in \omega$ and $y \in \omega$ implies $s(y) \in \omega$.

PROOF: Since for any s-closed set u, $0 \in u$, we have $0 \in \omega$ by (9.4). If $y \in \omega$, then by (9.4) $y \in u$ for every s-closed set u. Hence, by (9.3); $s(y) \in u$, for every such u. Thus by (9.4) again, $s(y) \in \omega$, and ω is s-closed.

THEOREM 9.6. There exists a set v, not the null set, such that for every $x \in v$ there is a set $y \in v$ with $x \subset y$ and $x \neq y$.

PROOF: The set ω has the properties here required of v. Indeed, $0 \in \omega$, hence ω has elements, and thus is not the null set. Secondly, if $x \in \omega$, then $s(x) \in \omega$ by Theorem 9.5. By Lemma 9.2, $x \subset s(x)$ and $x \neq s(x)$, so s(x) has the properties required of y in the Theorem.

Theorem 9.6 may be used as an alternative formulation of the axiom of infinity.

10. The Peano postulates. The set of natural numbers (non-negative integers) may be defined in many fashions; but for any definition it may be characterized by the usual Peano postulates.

THEOREM 10.1. The set ω and the function s on ω to ω satisfy the Peano postulates, to wit:

(10.2) $0 \in \omega$

(10.3) $n \in \omega$ implies $s(n) \in \omega$

(10.4) n and $m \in \omega$ and $s(n) = s(m)$ imply $n = m$

(10.5) For all $n \in \omega$, $s(n) \neq 0$.

(10.6) (Mathematical Induction). If u is a subset of ω, such that $0 \in u$ and $(\forall n)[n \in u \longrightarrow s(n) \in u]$, then $u = \omega$.

PROOF: The properties (10.2) and (10.3) are immediate by Theorem 9.5. Since $s(n)$ is a set containing n, for every n, it is not the null set, hence (10.5). If $s(n) = s(m)$, then $m \in s(m) = s(n)$. Hence, by the definition of $s(n) = n \cup \{n\}$, we have $m = n$, or $m \in n$. By the reverse argument we have $n = m$ or $n \in m$. Hence, if $m \neq n$, then $m \in n$ and $n \in m$, a contradiction to Corollary 8.2.

Finally, to prove the induction principle, let u be a subset of ω with the cited properties. Then u is s-closed, hence $w \subset u$ by the definition of ω as the intersections of all s-closed sets.

Alternative forms of the induction principle may also be obtained. For example, let $W(x)$ be a formula involving x as a free variable, and let $W(n)$ be the formula obtained by replacing x throughout W by n. Then the induction principle takes the form

$$W(0) \& [\forall(n)\ (n \in \omega \ \& \ W(n)) \longrightarrow W(s(n))]$$
$$\longrightarrow (\forall m)(m \in \omega \longrightarrow W(m)).$$

The usual development of the addition and multiplication of integers, etc., can be based on these postulates.

11. The Axiom of Choice. We adopt the axiom of choice in the following strong form.

E. There is a function C such that for each non-void set u, there exists an x with $< x, u > \in C$ and $x \in u$.

In other words, C is defined for every non-empty argument u, and the value C(u) of the function is an element x in the set u. Thus the function C simultaneously chooses an element for each non-empty set.

On the basis of the preceding axioms, one may state as precise theorems the equivalence of this form of the axiom of choice to the other familiar forms (e.g., Zorn's Lemma).

12. Summary of the axioms. The complete system of axioms for sets and classes is as follows.

A1. Axiom of extensionality ($A = B$ and $A \in C$ imply $B \in C$).

A2. Existence of the empty set 0.

A3. For any sets u, x, existence of the set $\lambda(u, x) = u \cup \{ x \}$.

B. Existence of the set of all n-tuples $< x_1, \ldots, x_n >$ satisfying any formula W in which $x_1, \ldots, x_n$ are free.

C1. Existence of the set P(u) of all subsets of the set u.

C2. Existence of the union of all elements x of a set u.

C3. The image of a set by a function is a set.

C4. The class ω of all integers is a set (axiom of infinity).

D. (Restrictive axiom) Every non-empty set u has an element x with $u \cap x = 0$.

E. The axiom of choice.

The axioms without C4 (axiom of infinity) are consistent, for they are satisfied by finite sets and classes.

Gödel (loc. cit.) has proved that if the axioms with D and E omitted are consistent, then the whole set of axioms is consistent. The real force of the axioms thus consists in the construction of a (denumerably) infinite set, by axiom C4, and axioms C1 and C2, which then allow the construction of still larger infinite sets.

13. Ordinal numbers. The axioms (especially the restrictive axiom D) 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

(13.1) $\omega, \omega + 1, \omega + 2, \ldots$.

The first ordinal beyond these, 2ω, may be defined as the set of all finite ordinals and all ordinals (13.1). The general principle will be that of defining each ordinal number as the set of all preceding ordinals.

Specifically, call a set u complete if $x \in u$ implies $x \subset u$; i.e., if every element of u is also a subset of u. An equivalent formulation is: u is complete if $t \in x$ and $x \in u$ imply $t \in u$.

DEFINITION: Any ordinal number α is a complete set such that for any two elements x and y of α, either

(13.2) $x \in y$, $y \in x$, or $x = y$.

By Corollaries 8.1 and 8.2, it follows at once 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 $x \in z$. Hence the set α is (linearly) ordered by the $\in$-relation E. We may thus say that an ordinal number is a complete set ordered by the $\in$-relation. One proves readily

<u>THEOREM 13.3</u>. Every integer $n \in \omega$ is an ordinal, and ω is an ordinal.

By axiom B, we may construct the <u>class</u> Ω of all ordinals. This class is not a set.

<u>THEOREM 13.4</u>. The elements of an ordinal α are the complete subsets of α different from α.

PROOF: Let $x \in \alpha$. Then, since α is complete, $x \subset \alpha$. Also x cannot equal α, for then $\alpha = x \in \alpha$, in contradiction to 8.1. To prove x complete, take any $y \in x$. Then $y \in x \subset \alpha$, hence $y \in \alpha$ and $y \subset \alpha$, by the completeness of α. Now take any element $t \in y$. Then $t \in \alpha$. By (13.2) we have for the two elements x, t of α one of the alternatives $t \in x$, $t = x$, or $x \in t$. In the third case $x \in t$, $t \in y$, and $y \in x$, a situation which violates axiom D. In the second case $t = x$, we have $x = t \in y$ and $y \in x$, a contradiction to Corollary 8.2. Hence we have only the first case: for every $t \in y$, we get $t \in x$. This means that $y \subset x$, so that any element of x is a subset of x.

Conversely, let x be any complete subset of α with $x \neq \alpha$:

(13.5) $\qquad x \subset \alpha, \quad x \neq \alpha, \quad x$ complete.

Form the complement of x in α

(13.6) $$z = \alpha \cap x' \quad .$$

By axiom D, there exists in z an element d with

(13.7) $$d \cap z = 0 \qquad d \in z \quad ,$$

and hence $d \notin x$.

We intend to prove d = x. Let t be any 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. Therefore $t \in \alpha$. But $t \in d$, hence by (13.7) $t \notin z$. Hence t is in the complement of z in α; i.e., $t \in x$. Since this holds for every $t \in d$, we have $d \subset x$.

Conversely, let r be any element of x. Since r and d are elements of α, we have one of the three cases

$$r \in d, \qquad r = d, \qquad d \in r.$$

If r = d, then $d \in x$, hence $d \notin z$ by (13.6), in contradiction to (13.7). If $d \in r$, then $d \in r$, $r \in x$; since x is complete this gives $d \in r$, $r \subset x$, hence $d \in x$, leading to the same contradiction. There remains only the case $r \in d$, which also shows that every element r of x is an element of d, hence $x \subset d$. We therefore have x = d.

Since $d \in \alpha$, Axiom A1 yields $x \in \alpha$; thus x is an element of α, as asserted by the Theorem.

COROLLARY 13.8. If α is an ordinal, every element x of α is an ordinal.

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

THEOREM 13.9. If A is any non-empty class of ordinals, then

$$\beta = \bigcap_{\alpha \in A} \alpha$$

is an ordinal, and is contained in the class A.

PROOF: Clearly β is a set. To prove it complete, let $x \in \beta$ and $t \in x$. Then for every α in A, $t \in x \in \beta \subset \alpha$, hence $t \in x \in \alpha$, hence $t \in \alpha$. Thus $x \subset \alpha$ for every α in A, hence $x \subset \beta$ by the definition of the intersection. This makes β complete. Since β is contained in some α_0 in A, and α_0 is ordered by membership, so is β. Hence β is an ordinal.

If β is not in the class A, then $\beta \neq \alpha$ for every $\alpha \in A$. Since $\beta \subset \alpha$, by its definition, Theorem 13.4 proves $\beta \in \alpha$. Because this holds for every α, $\beta \in \beta$, by the definition of β. This contradicts 8.1.

COROLLARY 13.9.1. If α and γ are ordinals, then one of the relations

$$\alpha \subset \gamma, \quad \alpha = \gamma, \quad \gamma \subset \alpha$$

holds.

PROOF: Apply the Theorem to the Class $A = \{\alpha, \gamma\}$. It follows that intersection $\alpha \cap \gamma$ is one of α or γ, q.e.d.

COROLLARY 13.9.2. If α and γ are ordinals, then $\alpha \in \gamma$ if and only if $\alpha \subset \gamma$ and $\alpha \neq \gamma$.

PROOF: If $\alpha \subset \gamma$ and $\alpha \neq \gamma$, then $\alpha \in \gamma$ by Theorem 13.4. Conversely, if $\alpha \in \gamma$, then $\alpha \subset \gamma$, and $\alpha = \gamma$ is impossible.

COROLLARY 13.9.3. The class of all ordinals is linearly ordered by either of the relations of Corollary 13.9.2. We write this relation $\alpha < \gamma$.

14. <u>Construction of ordinals</u>. A class A is well ordered by a relation R if A is linearly ordered by R and if every subclass of A has a first element in this order.

<u>THEOREM 14.1</u>. The class Ω of all ordinal numbers is well ordered by the relation $\alpha < \gamma$.

PROOF: Let A be any subclass of Ω, and apply Theorem 13.9 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 14.2. If α is an ordinal, α is well ordered by the relation $\beta < \gamma$ for its elements.

PROOF: Any subclass of α is a subclass of the class of all ordinals, hence has a first element.

<u>THEOREM 14.3</u>. The elements of an ordinal α are all ordinals smaller than α.

PROOF: If $x \in \alpha$, then x is an ordinal (Corollary 13.8) and $x \in \alpha$, hence $x < \alpha$. Conversely, if β is an ordinal and $\beta < \alpha$, then $\beta \in \alpha$.

<u>THEOREM 14.4</u>. Every complete set of ordinals is an ordinal.

PROOF: 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 14.5. If α is an ordinal, so is $s(\alpha)$, and $s(\alpha)$ is the first ordinal larger than α.

PROOF: $s(\alpha)$ has as elements α and the elements of α, hence $s(\alpha)$ is a set of ordinals. To show $s(\alpha)$ complete, let $x \in s(\alpha)$. Then either $x \in \alpha$, in which case $x \subset \alpha \subset s(\alpha)$ by the completeness of α, or $x = \alpha$, in which case $x \subset s(\alpha)$ by the definition of $s(\alpha)$. Hence $s(\alpha)$ is a complete set of ordinals, thus an ordinal by the Theorem.

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

This corollary allows us to produce an ordinal larger than any given ordinal α. Also, since 0 is trivially an ordinal, we may prove by induction that any positive integer $n = s(n-1)$ is an ordinal.

COROLLARY 14.6. 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 $\alpha \in u$.

PROOF: By definition, β is a set whose elements (the elements of the α's) are ordinals. Also β is 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 complete. As a complete set of ordinals, β is 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 $\beta = \alpha < \alpha_1$, a contradiction to the definition of β. Thus β is larger than every α.

This corollary also allows us to construct larger ordinals from given ones. In particular, it may be used to prove ω an ordinal.