Basic Set Theory

John L. Bell

I. Sets and Classes.

We distinguish between *objects* and *classes.* Any collection of objects is deemed to form a *class* which is uniquely determined by its *elements.* We write

$$a \in A$$

to indicate that the object a is an element or *member* of the class A. *We assume that every member of a class is an object.* Lower-case letters *a, b, c, x, y, z, …* will always denote objects, and later, *sets.*

Equality between classes is governed by the *Axiom of Extensionality:*

Axiom $$A = B \;\Leftrightarrow\; \forall x[x \in A \Leftrightarrow x \in B].$$

One class is said to be a *subclass* of another if every element of the first is an element of the second. This relation between classes is denoted by the symbol $\subseteq$. Thus we make the

Definition $$A \subseteq B \;\Leftrightarrow\; \forall x[x \in A \Rightarrow x \in B].$$

A subclass B of a class A such that $B \neq A$ is called a *proper* subclass of A.

Every property of objects determines a class. Suppose $\varphi(x)$ is the given property of objects x; the class *determined* by this property is denoted by

$$\{x\colon \varphi(x)\},$$

which we read as *the class of all x such that* $\varphi(x)$. *Church's scheme* is an axiom guaranteeing that the class named in this way behaves in the manner expected:

Axiom $$\forall y[y \in \{x\colon \varphi(x)\} \;\Leftrightarrow\; \varphi(y)].$$

Among classes we single out the *universal* class V comprising *all* objects and the *empty* class $\varnothing$ which has *no* members. Thus we make the

Definition $$V = \{x\colon x = x\} \qquad \varnothing = \{x\colon x \neq x\}.$$

We shall sometimes write 0 for $\varnothing$.

A *set* is a class which is also an object. The purpose of a *theory of sets* is to formulate existence principles which ensure the presence of sufficiently many sets to enable mathematics to be done.

Russell's Paradox shows that *not every class can be a set.* For consider the class

$$\{x\colon x \notin x\} = R$$

(here $x \notin x$ stands for "not $x \in x$"). Suppose this class were a set r. Then it follows from Church's scheme that

$$r \in r \Leftrightarrow r \notin r,$$

a contradiction. Therefore R is not a set.

In the present formulation of the theory of sets we quantify *only over objects*, and *not* over classes in general. We do, on the other hand, *name* many classes and state *principles* which apply to all classes.

Definitions

$$\{a\} =_{\text{df}} \{x\colon x = a\}$$
$$\{a_1, \ldots, a_n\} =_{\text{df}} \{x\colon x = a_1 \vee \ldots \vee x = a_n\}$$
$$\forall x{\in}A\ \varphi(x) \Leftrightarrow_{\text{df}} \forall x[x \in A \Rightarrow \varphi(x)]$$
$$\exists x{\in}A\ \varphi(x) \Leftrightarrow_{\text{df}} \exists x[x \in A \wedge \varphi(x)]$$
$$\{x{\in}A\colon \varphi(x)\} =_{\text{df}} \{x\colon x \in A \wedge \varphi(x)\}$$
$$x, y, \ldots, z \in A \Leftrightarrow_{\text{df}} x \in A \wedge y \in A \wedge \ldots \wedge z \in A$$
$$a_1 \in a_2 \in \ldots \in a_n \Leftrightarrow_{\text{df}} a_1 \in a_2 \wedge \ldots \wedge a_{n-1} \in a_n$$
$$A \cup B =_{\text{df}} \{x\colon x \in A \vee x \in B\}$$
$$A \cap B =_{\text{df}} \{x\colon x \in A \wedge x \in B\}$$
$$-A =_{\text{df}} \{x\colon x \notin A\}$$
$$A - B =_{\text{df}} \{x\colon x \in A \vee x \in B\}$$

Notice that $\cup$, $\cap$ and $-$ satisfy the following laws of *Boolean algebra*

$$A \cup B = B \cup A, \qquad A \cap B = B \cap A;$$
$$A \cup A = A, \qquad A \cap A = A;$$
$$(A \cup B) \cap B = B, \qquad (A \cap B) \cup B = B$$
$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \qquad A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$$
$$A \cup -A = V, \qquad A \cap -A = \varnothing$$

At this point we make the simplifying assumption that *everything is a class.* This is of course equivalent to asserting that every *object* is a class, i.e., a set. It follows that the universal class V is also *the class of all sets,* that is,

$$X \in V \Leftrightarrow X \text{ is a set.}$$

Moreover, *any* class A is now a class—or *family*—of sets, and the sets in A can be unioned or intersected together to form a new class. Formally, we make the

Definitions $\bigcup A =_{df} \{y: \exists x \in A\ (y \in x)\}$ $\bigcap A =_{df} \{y: \forall x \in A\ (y \in x)\}$

$\bigcup A$ and $\bigcap A$ are called the *union* and *intersection* of A, respectively. One now easily proves the

Theorem $\bigcup \varnothing = \varnothing$ $\bigcap \varnothing = V.$ ■

For any class B, the sets $x \subseteq B$—the *subsets* of B—form another class called the *power class* of B:

Definition $\mathbf{P}B =_{df} \{x: x \subseteq B\}.$

We now introduce *Zermelo's axioms* for set theory.

The *axiom* (or scheme) *of separation* tells us that whenever we separate the elements of a *set* into a subclass, the result is again a set. That is, *every subclass of a set is a set.* Thus we have the

Axiom $A \subseteq a \in V \Rightarrow A \in V,$

or, equivalently, for any property $\varphi(x)$,

$$a \in V \Rightarrow \{x \in a: \varphi(x)\} \in V.$$

One easily derives from this the

Theorem $V \notin V.$ ■

Since the axiom of separation does not give us any sets to start with, we need to assume that there is at least one set. But then it follows from the axiom of separation that $\varnothing$, which is clearly a subclass of every class, will also have to be a set. In that case we might as well assume the *axiom of the empty set:*

Axiom $\varnothing \in V.$

We also assume the axioms of *pairing, union, and power set:*

Axioms

$$\forall x \forall y \, \{x, y\} \in V$$
$$A \in V \Rightarrow \bigcup A \in V$$

$$A \in V \Rightarrow \mathbf{P}A \in V.$$

II. Relations and Functions.

As we have seen, every property of objects determines a unique class. Properties of *pairs* of objects are called relations; does every relation determine a class? This will be the case provided that each pair of objects can itself be regarded as an object. Zermelo's axioms in fact guarantee this.

Since not all relations are symmetric, what we require is not the *unordered* pair $\{a, b\}$ but some notion of *ordered* pair $\langle a, b\rangle$. An adequate definition of ordered pair can be introduced by iterating the operation of unordered pair in an asymmetric manner. Thus we make the

Definition $\langle a, b\rangle =_{df} \{\{a\}, \{a, b\}\}$.

One then proves without much trouble the

Theorem $\langle a, b\rangle = \langle x, y\rangle \Leftrightarrow a = x \wedge b = y.$ ■

To dispel the impression that this definition has fallen out of a clear blue sky, notice that $\langle a, b\rangle$ as defined above is the set of *initial segments* of the ordered set $\{a, b\}$ in which a precedes b.

The *Cartesian product* of two sets A and B is introduced by means of the

Definition $A \times B =_{df} \{\langle x, y\rangle : x \in A \wedge y \in B\}$.

Here the r.h.s. is an abbreviation for $\{z: \exists x \in A \exists y \in B\, [z = \langle x, y\rangle\}$: we shall use similar abbreviations in the sequel. Notice that $A \times B \subseteq \mathbf{PP}(A \cup B)$, whence

$$A, B \in V \Rightarrow A \times B \in V.$$

A (binary) relation may now be regarded as a class of ordered pairs. For if $\varphi(x, y)$ is a property of pairs x, y of objects, membership in the class $\{\langle x, y\rangle : \varphi(x, y)\}$ completely determines the pairs having the property φ. Since $V \times V$ is evidently the largest class of ordered pairs, we make the

Definitions

$\mathrm{Rel}[R] \Leftrightarrow_{df} R \subseteq V \times V$ (R is a *relation*)

R is a *relation on* $A \Leftrightarrow_{df} R \subseteq A \times A$

$xRy \Leftrightarrow_{df} \langle x, y\rangle \in R$

$\mathrm{dom}\ R =_{df} \{x: \exists y.\ xRy\}$ (the *domain* of R)

$\mathrm{ran}\ R =_{df} \{y: \exists x.\ xRy\}$ (the *range* of R)

A single-valued relation is called a *function.* Thus we make the further

Definitions $\exists!x\varphi(x) \Leftrightarrow_{df} \exists y\forall x[\varphi(x) \Leftrightarrow x = y\}$
$\text{Fun}[F] \Leftrightarrow_{df} \text{Rel}[F] \wedge \forall x{\in}\text{dom } F\, \forall y\forall z[xFy \wedge xFz \Leftrightarrow y = z]$ (F is a *function.*)

If F is a function, we define the function value $F(x)$ or Fx so that $F(x) = V$ if $x \notin \text{dom } F$. This is convenient because $V \notin V$, so to say that F is defined at x is just to say that $F(x) \neq V$. Formally we make the

Definition $F(x) =_{df} \bigcap y\text{: Fun}[F] \wedge xFy\}$.

One now proves

Cantor's Theorem. There is no function F for which there is a set a with dom $F = a$ and ran $F = \mathbf{P}a$.
Proof. Suppose $\text{Fun}[F]$, dom $F = a \in V$ and ran $F \subseteq \mathbf{P}a$. Define $b = \{x{\in}a\text{: } x \notin F(x)\}$. Then $b \in \mathbf{P}a$. If $b \in \text{ran } F$, then $b = F(c)$ for some $c \in a$. But then

$$c \in b \Leftrightarrow c \notin F(c) = b,$$

a contradiction. Therefore $c \notin \text{ran } F$, so that ran $F \neq \mathbf{P}a$. ■

It will also be convenient to make the

Definitions $R^{-1} =_{df} \{\langle x, y\rangle\text{: } xRy\}$ (the *inverse* of R)
$R[A] =_{df} \{y\text{: } \exists x{\in}A.\ xRy\}$
$R \circ S =_{df} \{\langle x, z\rangle\text{: } \exists y[xSy \wedge yRz]\}$[1]
$R\lceil A =_{df} \{\langle x, y\rangle\text{: } x \in A \wedge xRy\}$
$F\text{: } A \to B \Leftrightarrow_{df} \text{Fun}[F] \wedge \text{dom } F = A \wedge \text{ran } F = B$ (F is a function *from* A *to* B)
F is *one-one* or *injective* or an *injection*
$\Leftrightarrow_{df} \forall x{\in}\text{dom } F\, \forall y{\in}\text{dom } F\, [F(x) = F(y) \Rightarrow x = y]$
$F\text{: } A \to B$ is *onto* B or *surjective* or a *surjection* $\Leftrightarrow_{df} \text{ran } F = B$
$F\text{: } A \to B$ is *bijective* or a *bijection* $\Leftrightarrow_{df} F$ is one-one and onto B
$B^A =_{df} \{f\text{: } [f\text{: } A \to B]\}$.

Since $B^A \subseteq \mathbf{P}(A \times B)$, it follows that, if $A, B \in V$, then $B^A \in V$. Also, for any set a, and any class B, every element $f \in B^a$ is a function $f\text{: } a \to B$. But what about the converse? Is every function $F\text{: } B \to a$ an element of B^a? To ensure that this is the case we add the *Axiom of Replacement:*

[1] Note here that, if F and G are functions, then $(F \circ G)(x) = F(G(x))$.

Axiom $\quad \text{Fun}[F] \wedge \text{dom } F \in V \Rightarrow \text{ran } F \in V.$

From this one easily derives the required

Theorem $\quad \text{Fun}[F] \wedge \text{dom } F \in V \Rightarrow F \in V.$ ■

The notion of ordered pair can be iterated to yield that of *ordered n-tuple*: thus we make the

Definition $(n \geq 3) \quad \langle x_1, \ldots, x_n \rangle =_{\text{df}} \langle x_1, \ldots, x_{n-1} \rangle, x_n \rangle$

$$A^n =_{\text{df}} \{\langle x_1, \ldots, x_n \rangle : x_1, \ldots, x_n \in A\}.$$

A subset of A^n is called an *n-ary relation on A.*

Theorem. $\quad \langle x_1, \ldots, x_n \rangle = \langle y_1, \ldots, y_n \rangle \Leftrightarrow x_1 = y_1 \wedge \ldots \wedge x_n = y_n.$ ■

Indexed sets. An *indexing* of a set A is a function $f: I \to A$ from a set I — called the *index set—onto A.* One then usually writes:

$$a_i \text{ for } f(i), \quad \{a_i : i \in I\} \text{ for } A, \quad \bigcup_{i \in I} a_i \text{ for } \bigcup A, \quad \bigcap_{i \in I} a_i \text{ for } \bigcap A.$$

III. Well-orderings and Ordinals.

Definitions. A (*partial*) *ordering* on a class A is a relation $\leq$ on A satisfying, for all $x, y, z \in A$,

(i) $x \leq x$ (*reflexivity*)

(ii) $x \leq y \wedge y \leq z \Rightarrow x \leq z$ (*transitivity*)

(iii) $x \leq y \wedge y \leq x \Rightarrow x = y$ (*antisymmetry*).

If in addition $\leq$ satisfies, for all $x, y \in A$,

(iv) $x \leq y \vee x \leq y$ (*dichotomy*),

then $\leq$ is called a *total* or *linear* ordering of A.

A *strict linear ordering of* A is a relation $<$ on A satisfying, for all $x, y, z \in A$,

(a) $x \not< x$ (*irreflexivity*)

(b) $x < y \wedge y < z \Rightarrow x < z$ (*transitivity*)

(c) $x < y \vee y < x \vee x = y$ (*trichotomy*).

Linear orderings and strict linear orderings are interchangeable in view of the fact that

$\leq$ is a linear ordering $\Leftrightarrow$ $<$ ($=_{df} \leq \wedge \neq$) is a strict linear ordering,
$<$ is a strict linear ordering $\Leftrightarrow$ $\leq$ ($=_{df} < \vee =$) is a linear ordering.

If $\leq$ is a linear ordering, etc., of a set A, we shall call the pair $\mathfrak{A} = \langle A, \leq \rangle$ a *linearly ordered set*, etc. The set A is called the *underlying set* of $\mathfrak{A}$. We occasionally identify $\mathfrak{A}$ with its underlying set A.

Definition. A *well-ordering* of a class A is a linear ordering $\leq$ of A such that

(a) each nonempty subset $X \subseteq A$ has a (necessarily unique) *least* element w.r.t. $\leq$, i.e., an element $a \in X$ such that $a \leq x$ for all $x \in X$ [2];

(b) for any $a \in A$, $\{x \in A: x \leq a\}$ is a set.

A *strict well-ordering* of A is a strict linear ordering $<$ of A such that

(i) each nonempty subset $X \subseteq A$ has a (necessarily unique) *minimal* element w.r.t. $<$, i.e., an element $a \in X$ such that $x \not< a$ for all $x \in X$;

(ii) for any $a \in A$, $\{x \in A: x < a\}$ is a set.

Observe the following two facts:

[2] In view of this condition the stipulation that $\leq$ be a linear ordering is redundant: consider ant two-element subset of A.

- If $<$ is a strict well-ordering of A and $X \subseteq A$, then a $<$-minimal element of X is also the $\leq$-least element of X.
- $<$ is a strict well-ordering $\Leftrightarrow$ $\leq$ is a well-ordering.

Principle of Induction for Strict Well-Orderings. If $<$ is a strict well-ordering of a class A, then for any property $\varphi(x)$, the following is true:

$$\forall x \in A[\forall y[y < x \Rightarrow \varphi(y) \Rightarrow \varphi(x)] \Rightarrow \forall x \in A \varphi(x).$$

And conversely, a strict linear ordering of a class A is a strict well-ordering iff it satisfies the principle of induction and condition (ii) above.

Definitions An *equivalence relation* on a class A is a relation R on A which is reflexive, transitive, and *symmetric,* i.e. satisfies $R = R^{-1}$. If A is a set, each equivalence relation R on A gives rise to *equivalence classes* (which are, in fact, sets): for each $a \in A$, the R-equivalence class determined by a is $a/R =_{df} \{ x \in A: aRx\}$.

Given two partially ordered sets $\mathfrak{X} = \langle X, \leq \rangle$ and $\mathfrak{Y} = \langle Y, \sqsubseteq \rangle$, a function $f: X \rightarrow Y$ is said to be *order-preserving* if, for any $x, y \in X$,

$$x \leq y \Rightarrow f(x) \sqsubseteq f(y).$$

An order-preserving function which is bijective and whose inverse is also order-preserving is called an *order-isomorphism.* Thus a bijection f between partially ordered sets $\mathfrak{X}$ and $\mathfrak{Y}$ is an order-isomorphism iff, for any $x, y \in X$,

$$x \leq y \Leftrightarrow f(x) \sqsubseteq f(y).$$

If there is an order-isomorphism between $\mathfrak{X}$ and $\mathfrak{Y}$, we write $\mathfrak{X} \cong \mathfrak{Y}$, and say that they are (*order-*) *isomorphic.*

Now let $\mathcal{C}$ be a class of partially ordered sets. The isomorphism relation $\cong$ is clearly an equivalence relation on $\mathcal{C}$. A function $F: \mathcal{C} \rightarrow \mathcal{C}$ is called an *order-type operation* on $\mathcal{C}$ if it satisfies

- $\forall \mathfrak{X} \in \mathcal{C} \forall \mathfrak{Y} \in \mathcal{C}[F\mathfrak{X} = F\mathfrak{Y} \Leftrightarrow \mathfrak{X} \cong \mathfrak{Y}]$
- $\forall \mathfrak{X} \in \mathcal{C}[F\mathfrak{X} \cong \mathfrak{X}]$.

This amounts to saying that F picks one member from each isomorphism class $\{\mathfrak{X} \in \mathcal{C}: \mathfrak{X} \cong \mathfrak{A}\}$ for $\mathfrak{A} \in \mathcal{C}$.

If $\mathcal{C}$ is the class of *all* partially ordered sets—or even just the subclass of linearly ordered sets—it is not possible to define an order-type operation on $\mathcal{C}$ explicitly. However, we show that when $\mathcal{C}$ is the class $\mathcal{W}$ of *well-ordered* sets, the corresponding order-type operation can actually be defined through the concept of *ordinal number.*

Thus we seek a class $\mathcal{O} \subseteq \mathcal{W}$ and a surjection $F: \mathcal{W} \to \mathcal{O}$ satisfying

$$\forall \mathfrak{x} \in \mathcal{W} \forall \mathfrak{y} \in \mathcal{W}[F\mathfrak{x} = F\mathfrak{y} \Leftrightarrow \mathfrak{x} \cong \mathfrak{y}]$$
$$\forall \mathfrak{x} \in \mathcal{W}[F\mathfrak{x} \cong \mathfrak{x}].$$

The underlying sets of the members of $\mathcal{O}$ will be called *ordinals.* We now define ordinals explicitly.

Definitions X is *transitive: Trans*$(X) \Leftrightarrow_{\text{df}} \forall x \in X[x \subseteq X]$

X is an *ordinal: Ord*$(X) \Leftrightarrow_{\text{df}}$ X is a transitive set and the relation $\in | X =_{\text{df}} \{\langle x, y\rangle \in X \times X: x \in y\}$ is a strict well-ordering of X.

We use letters α, β, γ, ξ, η, ζ to denote ordinals. Note that $0 = \varnothing$ is an ordinal.

Theorem. Each element of an ordinal is an ordinal.
Proof. Let α be an ordinal and $x \in \alpha$. Then $x \subseteq \alpha$, so, since $\in | \alpha$ is a well-ordering of α, $\in | x$ is a well-ordering of x. If $z \in y \in x$, then $y \in \alpha$ because $x \subseteq \alpha$. Since $\in | \alpha$ is a transitive relation on α, it follows that $z \in x$. So x is transitive and hence an ordinal. ■

Definitions . For ordinals α, β we write

$$\alpha < \beta \text{ for } \alpha \in \beta; \quad \alpha \leq \beta \text{ for } (\alpha < \beta \vee \alpha = \beta).$$

Let $\langle P, \leq \rangle$ be a partially ordered set. An *initial segment* of P is a subset X of P such that

$$\forall x \in X \forall y \in P[y \leq x \Rightarrow y \in X];$$

it is *proper* if $X \neq P$. Thus by an initial segment of an ordinal α we mean a subset X of α such that $\forall \xi \in X \forall \eta \in \alpha[\eta \in \xi \Rightarrow \eta \in X]$;

Theorem. The only initial segments of an ordinal are the ordinal itself and its initial segments.
Proof. Let X be an initial segment of an ordinal α. If $X \neq \alpha$, then $\alpha - X$ has a $<$-least element ξ. If $\eta < \xi$, then $\eta \in X$; if $\eta \geq^3 \xi$, then $\eta \notin X$, for

[3] Here we write $x \geq y$ for $y \leq x$, and similarly for ">".

otherwise $\xi \in X$ because X is an initial segment of α. Hence $X = \{\eta \in \alpha: \eta < \xi\} = \xi \cap \alpha = \xi$ since $\xi \subseteq \alpha$. ■

Corollary. For ordinals α, β,

$$\alpha \leq \beta \Leftrightarrow \alpha \subseteq \beta.$$

Hence $0 \leq \alpha$ for any α. ■

Theorem. For each ordinal α, $\alpha \notin \alpha$.
Proof. If $\xi \in \alpha$, then $\xi \notin \xi$ since $\in | \alpha$ is a strict ordering on α. In particular, if $\alpha \in \alpha$ then $\alpha \notin \alpha$, and the contention follows. ■

Theorem. For ordinals α, β, exactly one of $\alpha = \beta$, $\alpha < \beta$, $\beta < \alpha$ holds.
Proof. Put $\xi = \alpha \cap \beta$. Then ξ is an initial segment of α, for if $y \in x \in \xi$ then $x \in \alpha$ and $x \in \beta$, whence $y \in \xi$. Similarly, ξ is an initial segment of β. Hence

$$[\xi = \alpha \text{ or } \xi \in \alpha] \text{ and } [\xi = \beta \text{ or } \xi \in \beta].$$

Accordingly there are four (actually three) possibilities:

$\xi = \alpha$ and $\xi = \beta$, in which case $\alpha = \beta$;
$\xi = \alpha$ and $\xi \in \beta$, in which case $\alpha \in \beta$;
$\xi = \beta$ and $\xi \in \alpha$, in which case $\beta \in \alpha$;
$\xi \in \alpha$ and $\xi \in \beta$, in which case $\xi \in \alpha \cap \beta = \xi$, contradicting $\xi \notin \xi$.

One can't have $\alpha = \beta$ and $\alpha \in \beta$ (or $\beta \in \alpha$) since $\beta \notin \beta$ (and $\alpha \notin \alpha$); nor can one have $\alpha \in \beta$ and $\beta \in \alpha$, for then $\beta \in \alpha \subseteq \beta$, contradicting $\beta \notin \beta$. The result follows. ■

Definition $ORD = \{x: Ord(x)\}$.

Theorem. The relation < (i. e., $\in$) is a (strict) well-ordering of *ORD.*
Proof. Most of this has been established above. It only remains to show that any nonempty subset X of *ORD* has a <-minimal element. Choose any $\alpha \in X$; if $\alpha \cap X = \varnothing$, then α is minimal in X; if $\alpha \cap X \neq \varnothing$, then, as a nonempty subset of α, $\alpha \cap X$ has a minimal element which is easily seen to be minimal in X also. ■

Theorem. *ORD* is not a set.
Proof. Suppose *ORD* were a set a. Then $\in | a$ is a well-ordering of a and each element of a is an ordinal, hence a subset of a. Thus a is transitive,

hence an ordinal, and so $a \in a$, contradicting the fact that no ordinal can be a member of itself. ■

Theorem. **(i)** For any ordinal α, the least ordinal $> \alpha$ is $\alpha + 1 =_{df} \alpha \cup \{\alpha\}$.

(ii) For each set X of ordinals, $\bigcup X$ is an ordinal which is the least upper bound of X w.r.t. $<$.

Proof. **(i)** It is easy to verify that $\alpha + 1$ is an ordinal, and clearly $\alpha < \alpha + 1$.Finally, if $\gamma > \alpha$, then $\alpha \subseteq \gamma$ and $\alpha \in \gamma$, whence $\alpha \cup \{\alpha\} \subseteq \gamma$, i.e., $\alpha + 1 \leq \gamma$.

(ii) Put $\beta = \bigcup X$. To begin with, β is an ordinal. For if $\varnothing \neq Z \subseteq \beta$, then $Z \cap \alpha \neq \varnothing$ for some $\alpha \in X$, and $Z \cap \alpha$ has a least element which is clearly also the least element of Z. Thus $\in | \beta$ is a well-ordering of β. Also, β is transitive: for if $y \in x \in \beta$, then $x \in \alpha$ for some $\alpha \in X$, whence $y \in \alpha$, and so $y \in \beta$. Finally, we have $\beta \geq \alpha$ for any $\alpha \in X$; if $\alpha \leq \gamma$ for all $\alpha \in X$, then $\alpha \subseteq \gamma$ for all $\alpha \in X$; so $\beta \subseteq \gamma$, whence $\beta \leq \gamma$. Thus β is the least upper bound for X as claimed. ■

Theorem. Let α and β be ordinals and let f be an order-isomorphism of $\langle \alpha, < \rangle$ with $\langle \beta, < \rangle$. Then $\alpha = \beta$ and f is the identity.

Proof. Suppose f is not the identity, and let ξ be the least element of α for which $f(\xi) \neq \xi$. Then $f(\eta) = \eta$ for all $\eta \in \xi$, whence $\xi \subseteq \beta$, and so ξ, as an ordinal, is an initial segment of β. It follows that $\xi \in \beta$ or $\xi = \beta$. In the latter case f carries the proper initial segment ξ of α onto β, contradicting the assumption that it is an isomorphism. So $\xi \in \beta$. Since f is an isomorphism, for each $\eta < \xi$ we have $f(\xi) > f(\eta) = \eta$. Therefore $f(\xi) \geq \xi$; since $f(\xi) \neq \xi$, it follows that $f(\xi) > \xi$. So if $\gamma \in \alpha$, then

$$\gamma < \xi \Rightarrow f(\gamma) = \gamma < \xi, \quad \gamma \geq \xi \Rightarrow f(\gamma) \geq f(\xi) > \xi.$$

Hence $\xi \notin \text{ran } f$, contradicting the fact that $\xi \in \beta$ and, again, the assumption that f is an isomorphism. ■

From the fact that $<$ is a strict well-ordering of *ORD* we immediately obtain the

- **Least ordinal principle.** For any property $\varphi(x)$,

$$\exists \alpha \varphi(\alpha) \rightarrow \exists! \alpha [\varphi(\alpha) \wedge \forall \beta [\varphi(\beta) \Rightarrow \alpha \leq \beta]],$$

as well as the

- **Principle of transfinite induction.** For any property $\varphi(x)$,

$$\forall\alpha[\forall\beta(\beta < \alpha \Rightarrow \varphi(\beta)) \Rightarrow \varphi(\alpha)] \Rightarrow \forall\alpha\varphi(\alpha).$$

Here the implication $\forall\beta(\beta < \alpha \Rightarrow \varphi(\beta)) \Rightarrow \varphi(\alpha)$ is called the *induction step.*

Theorem. Principle of transfinite recursion. Suppose we are given $F: V \to V$. Then

(i) $\forall\alpha(\exists!g: \alpha \to V)\forall\beta<\alpha\ [g(\beta) = F(g\restriction\beta)]$.

(ii) Defining G: $ORD \to V$ by $G(\alpha) = F(g_\alpha)$, where g_α is the unique function g given in **(i)**, we have

$$\forall\alpha[G(\alpha) = F(G\restriction\alpha)].$$

Proof. (i) Let α be any ordinal. We want to show that there is a unique $g: \alpha \to V$ for which

(*) $$\forall\beta<\alpha\ [g(\beta) = F(g\restriction\beta)].$$

Uniqueness of g. Suppose that g and h both satisfy (*), and that $g \neq h$. Let β be the least ordinal for which $g(\beta) \neq h(\beta)$. Then $g\restriction\beta = h\restriction\beta$, whence

$$g(\beta) = F(g\restriction\beta) = F(h\restriction\beta)] = h(\beta),$$

a contradiction.

Existence of g. We use the principle of transfinite induction. Suppose that for each $\beta < \alpha$ there is a unique $g_\beta: \beta \to V$ such that

$$\forall\gamma<\beta\ [g_\beta(\gamma) = F(g_\beta\restriction\gamma)].$$

In view of uniqueness, for $\gamma < \beta < \alpha$ we have $g_\gamma = g_\beta\restriction\gamma$. Now define $g: \alpha \to V$ by

$$g(\beta) = F(g_\beta) \text{ for } \beta < \alpha.^4$$

If $\gamma < \beta < \alpha$ we have

$$g(\gamma) = F(g_\gamma) = F(g_\beta\restriction\gamma) = g_\beta(\gamma),$$

so that $g\restriction\beta = g_\beta$. Therefore

$$\forall\beta<\alpha[F(g\restriction\beta) = F(g_\beta) = g(\beta)],$$

[4] To be precise, g is defined by $g = \{<\beta, F(g_\beta)>: \beta < \alpha\}$, which the axiom of replacement guarantees is a set.

completing the induction step and the proof of **(i).**

(ii) By uniqueness we have $g_\alpha \lceil \beta = g_\beta$ for $\beta < \alpha$, so that

$$G(\beta) = F(g_\beta) = F(g_\alpha \lceil \beta) = g_\alpha(\beta).$$

Hence $G \lceil \alpha = g_\alpha$, so that

$$G(\alpha) = F(g_\alpha) = F(G \lceil \alpha). \qquad \blacksquare$$

Definition An ordinal λ is a *limit* ordinal, written $Lim(\lambda)$, if it is neither 0 nor of the form $\alpha + 1$.

Another form of transfinite induction. For any property $\varphi(x)$,

$$[\varphi(0) \wedge \forall\alpha[\varphi(\alpha) \Rightarrow \varphi(\alpha + 1)] \wedge \forall\lambda[[Lim(\lambda) \wedge \forall\alpha < \lambda\varphi(\alpha)] \Rightarrow \varphi(\lambda)]] \Rightarrow \forall\alpha\varphi(\alpha).$$

This form of transfinite induction is easily derived from the original one.

Another form of transfinite recursion. Suppose given $a_0 \in V$, $F_1, F_2: V \to V$. Then there is a unique $G: ORD \to V$ such that

$$G(0) = a_0, \quad G(\alpha + 1) = F_1(G(\alpha)), \quad G(\lambda) = F(G \lceil \lambda) \text{ for limit } \lambda.$$

To obtain G, apply the original principle of transfinite recursion to the function $F: V \to V$ defined by

$$F(x) = \begin{cases} \varnothing & \text{if Fun[x]} \wedge \text{dom } x \in ORD \\ a_0 & \text{if } x = \varnothing \\ F_1(x(\alpha)) & \text{if Fun}[x] \wedge \text{dom } x = \alpha + 1 \\ F_2(x) & \text{if Fun}[x] \wedge \exists\lambda[\, Lim\,(\lambda) \wedge \text{dom } x = \lambda] \end{cases}$$

We use this form of transfinite recursion to define the operations of *addition* and *multiplication* on ordinals as follows:

$$\alpha + 0 = \alpha, \quad \alpha + (\beta + 1) = (\alpha + \beta) + 1, \quad \alpha + \lambda = \bigcup\{\alpha + \xi: \xi < \lambda\} \text{ for limit } \lambda;$$

$$\alpha.0 = 0, \quad \alpha.(\beta + 1) = \alpha.\beta + \alpha, \quad \alpha.\lambda = \bigcup\{\alpha.\xi: \xi < \lambda\} \text{ for limit } \lambda.$$

Now we can prove the

Theorem. Each well-ordered set $\langle A,<\rangle$ is order-isomorphic to a unique ordinal.

Proof. Uniqueness follows from the theorem on p. 12. To establish existence, define $G: ORD \to V$ by

$$G(\alpha) = \begin{cases} \text{least element of } A \text{ not in } \{G(\beta): \beta < \alpha\} \text{ if } \{G(\beta): \beta < \alpha\} \neq A \\ \text{some fixed } u \notin A \text{ if } \{G(\beta): \beta < \alpha\} = A. \end{cases}$$

That is, apply transfinite recursion to the function $F: V \to V$ defined by

$$F(x) = \begin{cases} \text{least element of } A - \text{ran } x \text{ if } A - \text{ran } x \neq \varnothing \\ u \text{ if } A \subseteq \text{ran } x \end{cases}$$

Observe that if $A \neq \{G(\beta): \beta < \alpha\}$ then $G(\alpha) \neq G(\beta)$ for $\beta < \alpha$. Therefore, if $A \neq \{G(\beta): \beta < \alpha\}$ for all α, then G would define an injection of ORD into A. Thus G^{-1} would define a function from a subset of A onto ORD, so by the axiom of replacement ORD would be a set, a contradiction. Accordingly there is a least ordinal α_0 for which $A = \{G(\beta): \beta < \alpha_0\}$. Since $A \neq \{G(\gamma): \gamma < \beta\}$, G is injective on α_0. Moreover, it is order-preserving. For if $\gamma < \beta < \alpha_0$, then since G is injective on α_0, we have

$$G(\beta) \in A - \{G(\delta): \delta < \gamma\}.$$

But $G(\gamma)$ has been chosen to be the least member of $A - \{G(\delta): \delta < \gamma\}$, whence $G(\gamma) < G(\beta)$.

Therefore $G \restriction \alpha_0$ is an order-isomorphism between α_o and $\langle A,<\rangle$. ■

This theorem shows that we have achieved the objective stated on page 10.

Corollary. Suppose $X \subseteq \alpha$. Then there is a unique $\beta \leq \alpha$ such that $\langle X,<\rangle \cong \langle \beta,<\rangle$.

Proof. There is a unique β and an order-isomorphism f: $\langle \beta,<\rangle \to \langle X,<\rangle$. We show that $\beta \leq \alpha$. If $\alpha \leq \beta$, then f is an order-preserving function from β to β. We claim that $f(\xi) \geq \xi$ for all $\xi < \beta$. If not, let η be the least ordinal such that $f(\eta) < \eta$. Then $f(f(\eta)) < f(\eta) < \eta$, contradicting the choice of η, and proving the claim. Since $f(\xi) < \alpha$ for all $\xi < \beta$, it follows that $\xi < \alpha$ for

all $\xi < \beta$. Hence $\beta \leq \alpha$, so that $\alpha = \beta$. Therefore $\alpha \leq \beta \Rightarrow \alpha = \beta$, whence $\beta \leq \alpha$. ■

Next, we introduce the natural numbers.

Definition x is a *natural number*: $N(x) \Leftrightarrow_{\text{df}} Ord(x) \wedge \forall \alpha \leq x \neg Lim(\alpha)$.

Natural numbers are also called *finite ordinals.* An ordinal which is not finite is called *infinite.*

The following is now easily proved:

Theorem (i) For any class A,

$$0 \in A \wedge \forall x[x \in A \Rightarrow x + 1 \in A] \Rightarrow \forall x[N(x) \Rightarrow x \in A].$$

(ii) If α, β are finite ordinals, so are $\alpha + \beta$, $\alpha.\beta$. ■

Notice that the natural numbers in order are 0, $1 = \{0\}$, $2 = \{0, 1\}$, $3 = \{0, 1, 2\}$, …, $n + 1 = \{0, 1, 2, \ldots, n\}$, … .

Our penultimate axiom is the *Axiom of Infinity:*

Axiom *There is an infinite ordinal.*

Let us denote the *least* infinite ordinal by ω. Then we have the

Theorem. (i) ω is the least limit ordinal
(ii) ω is the set of all finite ordinals.
Proof. (i) Evidently $\omega \neq 0$; if $\alpha < \omega$, then α is finite, so $\alpha + 1$ is also finite; hence $\alpha + 1 \neq \omega$. Thus ω is a limit ordinal. If β is any limit ordinal, clearly β must be infinite, so $\omega \leq \beta$. This proves **(i).**
(ii) If α is finite, we cannot have $\omega \leq \alpha$ because ω is a limit ordinal by **(i).** Therefore $\alpha < \omega$, i.e., $\alpha \in \omega$. Conversely, if $\gamma \in \omega$, then $\gamma < \omega$, so γ is finite by definition of ω. ■

The axiom of infinity may be stated in the following three equivalent forms:

- The class of finite ordinals is a set;
- There is a limit ordinal;
- $\exists x[0 \in x \wedge \forall y[y \in x \Rightarrow y \cup \{y\} \in x]]$.

Zermelo-Fraenkel set theory **ZF** is the axiomatic theory based on the axioms of separation, empty set, pairing, union, power set, replacement, and infinity.

IV. Cardinal Numbers and the Axiom of Choice

We begin with the

Definitions $X \approx Y \Leftrightarrow$ there is a bijection between X and Y.

$X \preccurlyeq Y \Leftrightarrow$ there is an injection of X into Y

$$X \prec Y \Leftrightarrow X \preccurlyeq Y \wedge X \not\approx Y.$$

It is easily seen that $\approx$ is an equivalence relation on V. "$X \approx Y$" is read "X and Y are *equipollent*".

For the present we shall assume that we have a function $|.|: V \to V$ called the *cardinality function* satisfying the following *axiom of cardinalities:*

Axiom $$\forall X \in V \forall Y \in V[|X| = |Y| \Leftrightarrow X \approx Y]$$

$|X|$ will be called the *cardinality* of X. Later on we will show (once we have the *axiom of choice*) that the function $|.|$ can actually be defined, and the axiom of cardinalities derived from the other axioms.

Definition. $$CARD =_{\text{df}} \{|X|: X \in V\}.$$

Members of *CARD* are called *cardinals.* We shall use letters $\mathfrak{m}$, $\mathfrak{n}$, $\mathfrak{p}$, etc. for cardinals.

Definitions $$X + Y =_{\text{df}} (X \times \{0\}) \cup (Y \times \{1\})$$
$$\coprod_{i \in I} A_i =_{\text{df}} \bigcup_{i \in I} A_i \times \{i\}$$

If $\mathfrak{m} = |X|$, $\mathfrak{n} = |Y|$, we define

$$\mathfrak{m} + \mathfrak{n} =_{\text{df}} |X + Y|, \quad \mathfrak{m}.\mathfrak{n} =_{\text{df}} |X \times Y|, \quad |, \quad \mathfrak{m}^{\mathfrak{n}} =_{\text{df}} |X^Y|,$$
$$\mathfrak{m} \leq_c \mathfrak{n} =_{\text{df}} X \preccurlyeq Y, \quad \mathfrak{m} <_c \mathfrak{n} =_{\text{df}} X \prec Y.$$

For an indexed set of cardinals $\{\mathfrak{m}_i: i \in I\}$ with $\mathfrak{m}_i = |A_i|$,

$$\sum_{i \in I} \mathfrak{m}_i =_{\text{df}} \left| \coprod_{i \in I} A_i \right|.^5$$

[5] The legitimacy of this definition requires the axiom of choice.

For a natural number k,

$$k.\mathfrak{m} =_{\text{df}} \mathfrak{m} + \mathfrak{m} + \ldots + \mathfrak{m}, \quad \mathfrak{m}^k = \mathfrak{m}.\mathfrak{m} \ldots \mathfrak{m} \text{ (both } k \text{ times)}.$$

Theorem.
$$\mathfrak{m} + \mathfrak{n} = \mathfrak{n} + \mathfrak{m}, \quad \mathfrak{m}.\mathfrak{n} = \mathfrak{n}.\mathfrak{m}$$
$$(\mathfrak{m} + \mathfrak{n}) + \mathfrak{p} = \mathfrak{m} + (\mathfrak{n} + \mathfrak{p}), \quad (\mathfrak{m}.\mathfrak{n}).\mathfrak{p} = \mathfrak{m}.(\mathfrak{n}.\mathfrak{p})$$
$$(\mathfrak{m} + \mathfrak{n}).\mathfrak{p} = \mathfrak{m}.\mathfrak{p} + \mathfrak{n}.\mathfrak{p}$$
$$\mathfrak{m} \leq_{\mathfrak{c}} \mathfrak{m}' \wedge \mathfrak{n} \leq_{\mathfrak{c}} \mathfrak{n}' \Rightarrow \mathfrak{m} + \mathfrak{n} \leq_{\mathfrak{c}} \mathfrak{m}' + \mathfrak{n}'$$
$$\text{if } \mathfrak{m}_i \leq \mathfrak{m} \text{ for all } i \in I, \text{ then } \sum_{i \in I} \mathfrak{m}_i \leq_{\mathfrak{c}} |I|.\mathfrak{m}$$

Theorem (Cantor). $|A| \leq |\mathbf{P}A|$. ■

Axiom of Choice. $\forall X \in V(\exists f\colon X \to \bigcup X)\forall x \in X[x \neq \varnothing \Rightarrow f(x) \in x]$.
Such a function f is called a *choice function* on X.

Definition. A nonempty partially ordered set P is said to be *inductive* if each nonempty linearly ordered subset (also called a *chain*) has an upper bound in P. An element of P is *maximal* if it is not strictly smaller than any element of P.

Using the axiom of choice we now prove

Zorn's Lemma. Each inductive set has a maximal element.
Proof. Let $\langle A, \leq \rangle$ be a partially ordered set, let h be a choice function for $\mathbf{P}A$ and let

$$\mathfrak{B} = \{X \subseteq A\colon X \text{ has a strict upper bound}\}.$$

(Here by a *strict* upper bound for a subset X of P we mean an upper bound not in X.) Define $f\colon \mathfrak{B} \to A$ by

$$f(X) = h(\{x\colon x \text{ is a strict upper bound for } X\}).$$

Now let e be any set which is not a member of A. By transfinite recursion define $F\colon ORD \to V$ by

$$F(\alpha) = \begin{cases} f(\{F(\beta)\colon \beta < \alpha\}) \text{ if } \{F(\beta)\colon \beta < \alpha\} \in \mathfrak{B} \\ e \text{ if } \{F(\beta)\colon \beta < \alpha\} \notin \mathfrak{B} \end{cases}$$

Observe that ran $F \subseteq A \cup \{e\}$. If ran $F \subseteq A$, then $\{F(\beta): \beta < \alpha\} \in \mathfrak{B}$ for all α. Hence, if $\beta < \alpha$, then $F(\alpha) \neq F(\beta)$. In that case F^{-1} is a function from A onto *ORD*, so that *ORD* would be a set by the axiom of replacement, a contradiction.

It follows that F takes the value e somewhere. Let α_0 be the first ordinal at which it does. If $\alpha < \alpha_0$, then $F(\alpha) \in A$, so $\{F(\beta): \beta < \alpha\} \in \mathfrak{B}$ and thus $F(\alpha)$ is a strict upper bound for $\{F(\beta): \beta < \alpha\}$. Therefore

$$\beta < \alpha < \alpha_0 \Rightarrow F(\beta) < F(\alpha).$$

Hence $F \lceil \alpha_0$ is an order-preserving injection of α_0 into $\langle A, \leq \rangle$. In that case $\{F(\beta): \beta < \alpha_0\}$ is a linearly ordered subset of A and so has an upper bound a. If $a < b$, then b would be a *strict* upper bound for $\{F(\beta): \beta < \alpha_0\}$, so that $\{F(\beta): \beta < \alpha_0\} \in \mathfrak{B}$ and $F(\alpha_0) \neq e$, a contradiction. Therefore a is a maximal element of P. ■

Corollary. The Well-Ordering Theorem. Every set can be well-ordered, and is therefore equipollent to an ordinal.

Proof. Let $\mathfrak{B}$ be the set of all pairs $\langle B, \leq \rangle$ with $B \subseteq A$ and $\leq$ a well-ordering of B. Then $\mathfrak{B} \neq \varnothing$; partially order $\mathfrak{B}$ by

$$\langle B, \leq \rangle \sqsubseteq \langle B', \leq' \rangle \Leftrightarrow B \subseteq B', \leq = \leq' \lceil B \text{ and } B \text{ is an initial segment of } B'.$$

We claim that $\langle \mathfrak{B}, \sqsubseteq \rangle$ is inductive. For let $\mathcal{C} = \{\{\langle B_i, \leq_i \rangle : i \in I\}$ be any chain in $\mathfrak{B}$. Then $\leq = \bigcup_{i \in I} \leq_i$ is easily seen to be a linear ordering on $B = \bigcup_{i \in I} B_i$; we show that it is a well-ordering. Let X be a nonempty subset of B. Then $X \cap B_i \neq \varnothing$ for some $i \in I$ and so $X \cap B_i \neq \varnothing$ has a least element, a, say, w.r.t $\leq_i$. If $x \in X$, then, since $\mathcal{C}$ is a chain, there is $j \in I$ such that $B_i \cup \{x\} \subseteq B_j$. If $x < a$, then $x \in X \cap B_i$ since B_i is an initial segment of B_j. This contradicts the choice of a, so $a \leq x$ and therefore a is the least element of X. Accordingly $\leq$ well-orders B.

Accordingly $\langle B, \leq \rangle$ is an upper bound for $\mathcal{C}$ and therefore $\langle \mathfrak{B}, \sqsubseteq \rangle$ is inductive. Consequently, Zorn's lemma applies that it has a maximal element $\langle D, \leq \rangle$. We claim that $D = A$, thereby proving the result. If, on the contrary, $b \in A - D$, then we define $\leq'$ on $D' = D \cup \{b\}$ by

$$\leq' \lceil D = D = \leq ; \quad x <' b \text{ for } x \in D.$$

Clearly $\langle D', \leq' \rangle \in \mathfrak{B}$ is strictly greater than $\langle D, \leq \rangle$ w.r.t. $\sqsubseteq$, contradicting the maximality of the latter. ■

Corollary. Let $\mathfrak{m}$ and $\mathfrak{n}$ be cardinals. Then

(i) Trichotomy: $\mathfrak{m} <_c \mathfrak{n}$ or $\mathfrak{n} <_c \mathfrak{m}$ or $\mathfrak{m} = \mathfrak{n}$

(ii) Antisymmetry: $\mathfrak{m} \leq_c \mathfrak{n}$ & $\mathfrak{n} \leq_c \mathfrak{m} \Rightarrow \mathfrak{m} = \mathfrak{n}$

Proof. **(i)** follows immediately from the fact that each cardinal is equipollent to an ordinal. As for **(ii),** it suffices to show that, for any ordinals α, β, if $\alpha \leq_c \beta$ and $\beta \leq_c \alpha$, then $\alpha \approx \beta$. Let α^* and β^* be the least ordinals such that $\alpha \approx \alpha^*$ and $\beta \approx \beta^*$. Suppose that $\alpha \leq_c \beta$ and $\beta \leq_c \alpha$. Then $\alpha^* \leq_c \beta^*$ and $\beta^* \leq_c \alpha^*$. Let f: $\alpha^* \to \beta^*$ be an injection, and write X for $f[\alpha^*]$. Then $\alpha^* \approx X \subseteq \beta^*$. By the Corollary on p. 18, there is $\gamma \leq \beta^*$ such that $\gamma \approx X$. But then $\gamma \approx \alpha^* \approx \alpha$, so that $\alpha^* \leq \gamma$ by definition of α^*. It follows that $\alpha^* \leq \beta^*$. Similarly $\beta^* \leq \alpha^*$, so that $\alpha^* = \beta^*$, whence $\alpha \approx \beta$. ■

Definition. Let A be a set, and $\mathfrak{m}$ a cardinal.

A is *finite* if $A \approx \alpha$ for some finite ordinal α.

A is *infinite* if A is not finite.

$\mathfrak{m}$ is *finite* if $\mathfrak{m} = |A|$ for some finite set A.

$\mathfrak{m}$ is *infinite* if $\mathfrak{m}$ is not finite.

Theorem. If $\mathfrak{m}$ is an infinite cardinal, then $\mathfrak{m}^k = \mathfrak{m}$ for any natural number k.

To prove this theorem we need some preliminary lemmas.

Lemma 1. For any infinite cardinal $\mathfrak{m}$ and any natural number k, we have $k.\mathfrak{m} \leq_c \mathfrak{m}^2$.

Proof. Let $\mathfrak{m} = |A|$. Since A is infinite, we can choose k distinct elements $a_1, \ldots, a_k$ of A. Define an injection $A + A + \ldots + A$ (k times) $\to A \times A$ by sending an element x of the i^{th} copy of A in $A + A + \ldots + A$ to the pair $\langle a_i, x \rangle$ in $A \times A$. ■

Lemma 2. Each infinite set A has a subset which is equipollent to ω.

Proof. We know that $A \approx \alpha$ for some ordinal α. Since A is infinite, we must have $\alpha \geq \omega$, whence $\omega \preccurlyeq A$. ■

Lemma 3. $\omega \times \omega \approx \omega$.

Proof. Like the usual proof that the rational numbers are countable. ■

Proof of the Theorem. Let $\mathfrak{m}$ be an infinite cardinal; we show that $\mathfrak{m}^2 = \mathfrak{m}$; the theorem then follows by induction on k.

Let $\mathfrak{m} = |A|$ and (Lemma 2) let $B \subseteq A$ satisfy $B \approx \omega$. Then (Lemma 3) there is a bijection f_0: $B \to B \times B$. Let $\mathscr{F}$ be the set of pairs $\langle X, f \rangle$ where $B \subseteq X \subseteq A$ and f is a bijection between X and $X \times X$ such that $f_0 \subseteq f$. Partially order $\mathscr{F}$ by stipulating that

$$\langle X, f \rangle \sqsubseteq \langle X', f' \rangle \Leftrightarrow X \subseteq X' \text{ and } f \subseteq f'.$$

Then $\langle \mathscr{F}, \sqsubseteq \rangle$ is clearly inductive and hence by Zorn's Lemma has a maximal element $\langle C, g \rangle$. We show that $|C| = \mathfrak{m}$, thereby proving the theorem.

Suppose on the contrary that $|C| <_c \mathfrak{m}$. Then since $\mathfrak{n} = |C|$ is infinite and $\mathfrak{n}^2 = \mathfrak{n}$ (recall that C is equipollent with $C \times C$), we have, using Lemma 1,

$$\mathfrak{n} \leq_c 2.\mathfrak{n} \leq_c 3.\mathfrak{n} \leq_c \mathfrak{n}^2 = \mathfrak{n}.$$

It follows that $3.\mathfrak{n} = 2.\mathfrak{n} = \mathfrak{n}$. From $\mathfrak{n} <_c \mathfrak{m}$ we infer that $|A - C| >_c \mathfrak{n}$; for if not then

$$|A| \leq_c \mathfrak{n} + \mathfrak{n} = 2.\mathfrak{n} = \mathfrak{n},$$

contradicting $|A| >_c |C| = \mathfrak{n}$.

Accordingly there is a subset $Y \subseteq A - C$ such that $|Y| = \mathfrak{n}$; put $Z = C \cup Y$. We show that there is a bijection h: $Z \to Z \times Z$ such that $g \subseteq h$. For we have

$$Z \times Z = (C \times C) \cup (C \times Y) \cup (Y \times C) \cup (Y \times Y),$$

and the sets on the r.h.s. are disjoint. Since $C \approx Y$, we have

$$|C \times Y| = |Y \times C| = |Y \times Y| = \mathfrak{n}^2 = \mathfrak{n},$$

so that

$$|(C \times Y) \cup (Y \times C) \cup (Y \times Y)| = 3.\mathfrak{n} = \mathfrak{n}.$$

Thus there is a bijection g' of Y onto $(C \times Y) \cup (Y \times C) \cup (Y \times Y)$. So the map h of Z into $Z \times Z$ defined by $h \restriction C = g$ and $h \restriction Y = g'$ is a bijection and $g \subseteq h$.

But this contradicts the maximality of $\langle C, g \rangle$. therefore $|C| <_c \mathfrak{m}$ is impossible, so $|C| = \mathfrak{m}$ and the result is proved. ■

Corollary. If $\mathfrak{m}$, $\mathfrak{n}$ are cardinals $\neq 0$, of at least one is infinite, then

$$\mathfrak{m}.\mathfrak{n} = \mathfrak{m} + \mathfrak{n} = \text{larger of } \mathfrak{m}, \mathfrak{n}.$$

Proof. Suppose $\mathfrak{m} \leq_c \mathfrak{n}$. Then

$$\mathfrak{n} \leq_c \mathfrak{m}.\mathfrak{n} \leq_c \mathfrak{n}^2 = \mathfrak{n}; \quad \mathfrak{n} \leq_c \mathfrak{m} + \mathfrak{n} \leq_c 2.\mathfrak{n} \leq_c \mathfrak{n}^2 = \mathfrak{n}. \quad \blacksquare$$

Now we can make the

Definition. For any set A,

$$|A| = \textit{least } \alpha \textit{ such that } A \approx \alpha.$$

Then we have the

Theorem. **(i)** $\forall A \in V \;\; A \approx |A|$

(ii) $\text{Card}(x) \Leftrightarrow \text{Ord}(x) \wedge |x| = x$

$\Leftrightarrow \text{Ord}(x) \wedge \forall \beta < x \;\; \beta \prec x.$

(iii) $\mathfrak{m} \leq_c \mathfrak{n} \;\; \Leftrightarrow \mathfrak{m} \leq \mathfrak{n}$ (as ordinals)

(iv) $\mathfrak{m}$ is infinite $\Leftrightarrow \omega \leq \mathfrak{m}$. $\blacksquare$

Theorem. Each finite ordinal is a cardinal.

Proof. We argue by induction. Clearly Card(0). If α is a finite ordinal which is also a cardinal, but $\alpha + 1 = \alpha \cup \{\alpha\}$ is not, then there is an ordinal $\gamma < \alpha + 1$ and a bijection f: $\alpha + 1 \to \gamma$. Then $\gamma \neq 0$ since $\alpha + 1 \neq 0$. Hence $\gamma = \beta + 1 = \beta \cup \{\beta\}$ for some β, and since $\gamma < \alpha + 1$, it follows that $\beta < \alpha$. If $f(\alpha) = \beta$, then $f \lceil \alpha$ is a bijection between α and β which is impossible since α is a cardinal. Hence $f(\alpha) = \xi \neq \beta$. Because f is bijective, there is $\eta \in \alpha \cup \{\alpha\}$ such that $f(\eta) = \beta$. Since $\eta \neq \xi$, we have $\eta \in \alpha$. The map g: $\alpha \to \beta$ defined by $g(x) = f(x)$ for $x \neq \eta$ and $g(\eta) = \xi$ is accordingly a bijection, contradicting the assumption that α is a cardinal. Therefore we have shown that, if α is a finite ordinal which is also a cardinal, so is $\alpha + 1$, and the result follows by induction. $\blacksquare$

Theorem. ω is the least infinite cardinal.

Proof. If this were not the case, then $|\omega| = \alpha$ for some finite α. Take any finite ordinal $\beta > \alpha$ (e.g. $\beta = \alpha + 1$). Then $|\beta| \leq |\omega| = \alpha$, contradicting the fact that β, as a finite ordinal, must be a cardinal. So ω is an infinite cardinal. It must be the least such because every smaller cardinal is finite. $\blacksquare$

Theorem. For each set A of cardinals there is a cardinal $\mathfrak{m}$ which exceeds every member of A.

Proof. Take $\mathfrak{m} = |\mathbf{P}(\cup A)|$. $\blacksquare$

By transfinite recursion we define the cardinal $\aleph_\alpha$ ("aleph-α") for each ordinal α as follows:

$$\aleph_\alpha = \text{least cardinal } \mathfrak{m} \text{ such that } \omega \leq \mathfrak{m} \text{ and } \mathfrak{m} \notin \{\aleph_\gamma : \gamma < \alpha\}.$$

Theorem. **(i)** $\text{Card}(\aleph_\alpha) \wedge \omega \leq \aleph_\alpha \wedge \aleph_0 = \omega \wedge (\alpha < \beta \Rightarrow \aleph_\alpha < \aleph_\beta)$.
(ii) $(\text{Card}(\mathfrak{m}) \wedge \omega < \mathfrak{m}) \Rightarrow \exists\alpha \ \mathfrak{m} = \aleph_\alpha$.

Proof. Most of **(i)** follows immediately from the definition of $\aleph_\alpha$. Only the last part requires proof. Thus suppose $\alpha < \beta$. Then $\aleph_\beta \notin \{\aleph_\gamma : \gamma < \beta\}$, so *a fortiori* $\aleph_\beta \notin \{\aleph_\gamma : \gamma < \alpha\}$. But $\aleph_\alpha$ is the *least* infinite cardinal not in $\{\aleph_\gamma : \gamma < \alpha\}$, so $\aleph_\alpha < \aleph_\beta$.

To prove **(ii)**, we first observe that, since the map $\alpha \mapsto \aleph_\alpha$ is injective, a straightforward application of the axiom of replacement shows that $\{\aleph_\alpha : \alpha \in \text{ORD}\}$ is not a set. So, given an infinite cardinal $\mathfrak{m}$, there must be an α for which $\aleph_\alpha \notin \mathfrak{m}$. Thus $\mathfrak{m} \leq \aleph_\alpha$. If $\mathfrak{m} = \aleph_\alpha$ we are through. If $\mathfrak{m} < \aleph_\alpha$ then since $\aleph_\alpha$ is the least infinite cardinal not in $\{\aleph_\gamma : \gamma < \alpha\}$, we must have $\mathfrak{m} \in \{\aleph_\gamma : \gamma < \alpha\}$. So $\mathfrak{m} = \aleph_\gamma$ for some $\gamma < \alpha$. ■

Accordingly the alephs form an enumeration of all the infinite cardinals.

Problems

1. Is $V \times V = V$?

2. **(i)** Find a bijection between $A^{B \times C}$ and $(A^B)^C$.

(ii) What are $A^{\varnothing}$, $A^{\{\varnothing\}}$, $\varnothing^A$?

(iii) Find a necessary and sufficient condition on A, B for $A^B \cap B^A \neq \varnothing$.

(iv) Show from first principles that $A^A \neq A$ for any set A with at least two elements.

(v) Show that $\{x: x \notin \bigcup x\} \notin V$, and generalize.

3. For each set A write 1_A for the identity map on A.

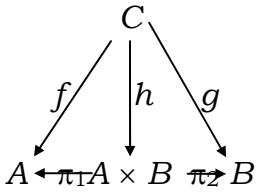
(i) Let $f: A \to B$ with $A \neq \varnothing$. Show that f is injective iff there is a map $r: B \to A$ such that $r \circ f = 1_A$. When is r *unique*?

(ii) Let $f: A \to B$ and suppose that there is a map $s: B \to A$ such that $f \circ s = 1_B$: s is called a *section* of f. Show that, in that case, f is onto. Does *every* onto map have a section?

4. Let $f: A \to B$. Define $\varphi: \mathbf{P}B \to \mathbf{P}A$ by $\varphi(X) = f^{-1}[X]$ for $X \subseteq B$. Show that f is injective (surjective) iff φ is surjective (injective).

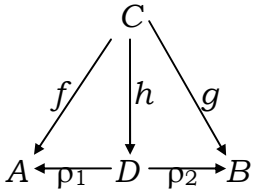
5. Define the *projection* maps $\pi_1: A \times B \to A$, $\pi_2: A \times B \to B$ by $\pi_1(\langle a, b\rangle) = a$, $\pi_2(\langle a, b\rangle) = b$ for $a \in A$, $b \in B$.

(i) Given maps $f: C \to A$, $g: C \to B$, find the *unique* map $h: C \to A \times B$ for which the diagram

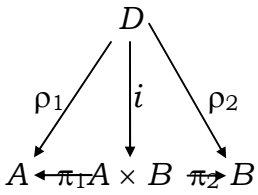


commutes (i.e. $\pi_1 \circ h = f$, $\pi_2 \circ h = g$).

(ii) Suppose that $\rho_1: D \to A$, $\rho_2: D \to B$, are such that, for any $f: C \to A$, $g: C \to B$, there is a *unique* $h: C \to D$ such that the diagram



commutes. Show that there is a *unique* bijection $i: D \to A \times B$ such that the diagram



commutes. What does this tell us about $A \times B$?

6. Let $\{X_i\colon i \in I\}$ be a family of subsets of a set B and let $f\colon A \to B$. Show that $f^{-1}[\bigcap_{i\in I} X_i] = \bigcap_{i\in I} f^{-1}[X_i]$ and $f^{-1}[\bigcup_{i\in I} X_i] = \bigcup_{i\in I} f^{-1}[X_i]$. If $\{Y_i\colon i \in I\}$ is a family of subsets of A, show that $f[\bigcup_{i\in I} X_i] = \bigcup_{i\in I} f[X_i]$. Does $f[\bigcap_{i\in I} X_i] = \bigcap_{i\in I} f[X_i]$ always?

7. Let $\{A_{ij}\colon \langle i, j\rangle \in I \times J\}$ be a family of sets. Using the axiom of choice, prove the *distributive law*

$$\bigcap_{i\in I}\bigcup_{j\in J} A_{ij} = \bigcup_{f\in J^I}\bigcap_{i\in I} A_{if(i)}$$

Deduce the dual formula

$$\bigcup_{i\in I}\bigcap_{j\in J} A_{ij} = \bigcap_{f\in J^I}\bigcup_{i\in I} A_{if(i)}\,.$$

Show, conversely, that the distributive law implies the axiom of choice.

8. Let $f\colon X \to E$ be an injection of a set X into a subset $E \subseteq X$. Define $D = \{y\in X\colon \exists n\in\omega\ \exists x\in X\text{–}E.\ y = f^n x\}$, where $f^n x$ is defined recursively by $f^0 x = x$, $f^{n+1}x = f(f^n x)$. Let $h\colon X \to E$ be the map defined by $h\restriction D = f\restriction D$, $h\restriction X\text{–}D = 1_{X\text{–}D}$. Show that h is bijective, and deduce the *Schröder-Bernstein theorem:* if $A \preccurlyeq B$ and $B \preccurlyeq A$, then $A \approx B$.

9. For a natural number $n > 0$, an *n-family* is a family $\mathfrak{A}$ of sets with $|A| = n$ for each $A \in \mathfrak{A}$. *Without* using the axiom of choice, prove that **(i)** if for some $k > 0$ every kn-family has a choice function, then so does every n-family; **(ii)** if every 2-family has a choice function, then so does every 4-family.

10. A set A is *transfinite* if $\omega \preccurlyeq A$ and *Dedekind infinite* if there is a proper subset B of A for which $B \approx A$. *Without* using the axiom of choice, show that, if A is infinite, then $\mathbf{PP}A$ is transfinite. Again without using the axiom of choice, show that transfiniteness is equivalent to Dedekind infiniteness, and that the former implies infiniteness. Assuming the axiom of choice, show that infiniteness implies transfiniteness.

11. A set A is said to be *countable* if $|A| \le \aleph_0$. Show that: **(i)** the union of two countable sets is countable; **(ii)** assuming the axiom of choice, the union of a countable family of countable sets is countable; **(iii)** the set of all finite subsets and of all finite sequences of members of a countable set is countable; **(iv)** a set is countable iff it is the union of a chain of

finite sets; **(v)** there is an uncountable set which is the union of a chain of countable sets.

12. The *transitive closure* of a set A is the least transitive set which includes A. Show that any set has a transitive closure.

13. A partially ordered set $\langle A,\leq\rangle$ is said to be *complete* if every subset of A has a least upper bound and a greatest lower bound. Let f be an order preserving map of a complete partially ordered set A into itself. Show that f has a *fixed point*, i.e. there is $a \in A$ for which $f(a) = a$.

14. Let **C** be the set of all continuous maps of the set $\mathbb{R}$ of real numbers into itself. What is $|\mathbf{C}|$? What is $|\mathbb{R}^{\mathbb{R}}|$? Are they the same?

15. A family $\mathscr{F}$ of subsets of a set A is said to be *local* if, for any subset X of A, $X \in \mathscr{F}$ iff every finite subset of X is in $\mathscr{F}$. Using Zorn's lemma, show that every local family has a maximal member w.r.t. $\subseteq$. Conversely, show that this principle implies Zorn's lemma.

16. Derive the axiom of choice directly from Zorn's lemma.

17. Use Zorn's lemma to give a direct proof of the comparability principle: for any sets A, B, either $A \preccurlyeq B$ or $B \preccurlyeq A$.

18. A family $\mathfrak{U}$ of subsets of a set A is called an *ultrafilter* on A if (a) $\varnothing \notin \mathfrak{U}$, (b) $X, Y \in \mathfrak{U} \Rightarrow X \cap Y \in \mathfrak{U}$, (c) for all $X \subseteq A$, either $X \in \mathfrak{U}$ or $A - X \in \mathfrak{U}$.

- **(i)** Let $\mathfrak{U}$ be an ultrafilter on A and suppose that $X \in \mathfrak{U}$ and $X \subseteq Y \subseteq A$. Show that $Y \in \mathfrak{U}$.
- **(ii)** Show that for any $a \in A$ the family $\{X \subseteq A: a \in X\}$ is an ultrafilter on A. An ultrafilter of this form is called *principal*.
- **(iii)** A family $\mathfrak{F}$ of subsets of A is said to be *neighbourly* if for finite subfamily of $\mathfrak{F}$ has a nonempty intersection. A neighbourly family $\mathfrak{F}$ is said to be *maximal* if the only neighbourly family which includes $\mathfrak{F}$ is $\mathfrak{F}$ itself. Show that ultrafilters and maximal neighbourly families coincide.
- **(iv)** Let $\mathfrak{F}$ be a neighbourly family of subsets of A. Use Zorn's lemma to show that $\mathfrak{F}$ is included a maximal neighbourly family, and hence an ultrafilter on A.
- **(v)** Suppose that A is infinite. Show that the family of *cofinite* (i.e., complement of finite) subsets of A is neighbourly, and deduce the existence of nonprincipal ultrafilters on A.

19. The *axiom of foundation* is the assertion $\forall x \neq \varnothing\ \exists y \in x\ [x \cap y \neq \varnothing\}$. Show that the axiom of foundation implies that

(*) there does not exist an infinite descending sequence of sets

$x_0 \ni x_1 \ni x_2 \ni \ldots x_n \ni \ldots$

Deduce that, if the axiom of foundation holds, there can be no nonempty class A such that $A \times A = A$. Show also that, assuming the axiom of choice, (*) implies the axiom of foundation.

20. Define the sets R_α recursively by $R_\alpha = \{x\colon \exists\xi{<}\alpha\ x \subseteq R_\xi\}$. Show that **(i)** $R_0 = \varnothing$, $R_{\alpha+1} = \mathbf{P}R_\alpha$, $\mathrm{Lim}(\lambda) \Rightarrow R_\lambda = \bigcup_{\xi<\lambda} R_\xi$; **(ii)** each R_α is transitive; **(iii)** $\beta < \alpha \Rightarrow R_\beta \subseteq R_\alpha \wedge R_\beta \in R_\alpha$; **(iv)** $\forall\alpha\ \alpha \subseteq R_\alpha$; **(v)** the assertion $\forall x\, \exists\alpha\ x \in R_\alpha$ is equivalent to the axiom of foundation.