B. Background in Set Theory

Intuitively, a "set" is a collection of objects called "members" of the set. In mathematics the notion of a "set" is taken as undefined, as is the relation of membership, and axioms are put down for these to follow. We will not do so in this "naive" treatment. Of course, it is well known that an undisciplined approach easily leads to logical difficulties such as the "set of all sets that do not contain themselves." These problems are handled in set theory by careful treatment of the axioms. But that is not the purpose of this appendix. We merely intend to set down terminology and notation that the reader must already have a feeling for, or he would not be studying this book. We will briefly discuss some "obvious" concepts and results, and will then prove some things that are not so obvious.

The terms "collection" or "family" are synonyms of "set," although the term "family" is usually used only for somewhat complicated sets such as a family of subsets of a set or a family of functions. The term "class" is often used as a synonym for "set," but in axiomatic set theory, it is used for a more encompassing concept: a "set" is a class that is a member of another class. A "proper class" is a class that is not a "set." The phrase "the class of all sets that do not contain themselves" is meaningful, but "the class of all classes that do not contain themselves" is not. We will not worry about such things, but we will avoid the use of the term "class" when we mean a "set." (An exception to this is the use of "class" in the phrase "equivalence class" which is traditional.)

We shall use the logical symbols $\exists$ to mean "there exists", $\exists 1$ to mean "there exists a unique," $\forall$ to mean "for all," $\ni$ to mean "such that," $\Rightarrow$ to mean "implies," $\Leftarrow$ to mean "is implied by," and $\Leftrightarrow$ to mean "if and only if."

If an object x is a member of a set S then we write $x \in S$. If not then we write $x \notin S$. If $P(x)$ is a statement about objects x which can be true or false for a given object x, then $\{x \mid P(x)\}$ stands for the set of all objects for which $P(x)$ is true, provided this does in fact define a set. If S is a set then $\{x \in S \mid P(x)\}$ is the same as $\{x \mid x \in S \text{ and } P(x)\}$.

If S and T are sets then we say S is contained in T, or S is a "subset" of T if $x \in S \Rightarrow x \in T$. This is denoted by $S \subset T$ or $T \supset S$. The statement $S \subset S$ is true for all sets S. If $S \subset T$ is false then we write $S \not\subset T$.

The "empty set" $\varnothing$ is the unique set with no objects, i.e., $x \in \varnothing$ is false for all objects x. The statement $\varnothing \subset S$ is true for all sets S.

The "union" of two sets S and T is $S \cup T = \{x \mid x \in S \text{ or } x \in T\}$. The "or" here is always inclusive, i.e., in the previous sentence it means $x \in S$ or $x \in T$ or both $x \in S$ and $x \in T$. The "intersection" of two sets S and T is $S \cap T = \{x \mid x \in S$ and $x \in T\}$. The "difference" of two sets is $S - T = \{s \in S \mid s \notin T\}$.

If A is a collection of sets then $\bigcup\{S \mid S \in A\} = \{x \mid \exists S \in A \ni x \in S\}$ and $\bigcap\{S \mid S \in A\} = \{x \mid \forall S \in A, x \in S\}$. If $\{S_\alpha \mid \alpha \in A\}$ is an "indexed" family of sets, we also use the notation $\bigcup S_\alpha = \{x \mid \exists \alpha \in A \ni x \in S_\alpha\}$ and $\bigcap S_\alpha = \{x \mid \forall \alpha \in A, x \in S_\alpha\}$.

Unions, intersections and differences follow these laws:

$$S \cup T = T \cup S, \qquad S \cap T = T \cap S,$$
$$R \cup (S \cup T) = (R \cup S) \cup T, \qquad R \cap (S \cap T) = (R \cap S) \cap T,$$
$$R \cup (S \cap T) = (R \cup S) \cap (R \cup T), \qquad R \cap (S \cup T) = (R \cap S) \cup (R \cap T),$$
$$R \cup \bigcap S_\alpha = \bigcap (R \cup S_\alpha), \qquad R \cap \bigcup S_\alpha = \bigcup (R \cap S_\alpha),$$
$$(X - S) \cup (X - T) = X - (S \cap T), \qquad (X - S) \cap (X - T) = X - (S \cup T),$$
$$\bigcup (X - S_\alpha) = X - \bigcap S_\alpha, \qquad \bigcap (X - S_\alpha) = X - \bigcup S_\alpha,$$
$$(\bigcup S_\alpha) \cap (\bigcup T_\beta) = \bigcup (S_\alpha \cap T_\beta), \qquad (\bigcap S_\alpha) \cup (\bigcap T_\beta) = \bigcap (S_\alpha \cup T_\beta).$$

The "cartesian product," or simply the "product" of two sets S and T is the set of ordered pairs $S \times T = \{\langle s,t \rangle | s \in S, t \in T\}$. We sometimes use (s,t) instead of $\langle s,t \rangle$ to denote an ordered pair.

A "relation" R between two sets S and T is a set of ordered pairs $R \subset S \times T$. We usually write $s\,R\,t$ to mean $\langle s,t \rangle \in R$. For example, $\in$ is a relation between a set of objects and a collection of sets. Another example is the relation $x \leq y$ between the set $\mathbf{R}$ of real numbers and itself.

The "domain" of a relation $R \subset S \times T$ is $\{t | \exists s \in S \ni s\,R\,t\}$ and the "range" of R is $\{s | \exists t \in T \ni s\,R\,t\}$.

A "function" f from the set X to the set Y is a relation $f \subset Y \times X$ with domain X such that $(x \in X, y \in Y, y' \in Y, y f x$ and $y' f x) \Rightarrow y = y'$. One writes $y = f(x)$ to mean yfx. We also use $f: X \to Y$, and variants of this to mean that f is a function from X to Y. The notation $x \mapsto y$ is also used for $y = f(x)$.

A function $f: X \to Y$ is said to be "injective" or "one–one into" if $f(a) = f(b) \Rightarrow a = b$. It is said to be "surjective" or "onto" if $y \in Y \Rightarrow \exists x \in X \ni y = f(x)$. It is said to be "bijective" or a "one–one correspondence" if it is both injective and surjective.

The identity function on X taking every member of X to itself is denoted by 1_X, or simply by 1 when that is not ambiguous.

If R and S are relations (in particular, if they are functions) then we define the "composition" of R and S to be

$$R \circ S = \{\langle a,c \rangle | \exists b \ni a\,R\,b \text{ and } b\,S\,c\},$$

and the "inverse" of R to be

$$R^{-1} = \{\langle a,b \rangle | b\,R\,a\}.$$

It is easy to see that $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$. It is also elementary that $g \circ f$ is a function when f and g are both functions.

If $R \subset Y \times X$ is a relation and $A \subset X$ then we put $R(A) = \{y \in Y | \exists a \in A \ni y\,R\,a\}$. Note that, for a function $f: X \to Y, f(A) \subset Y$ is defined for $A \subset X$ and $f^{-1}(B) \subset X$ is defined for $B \subset Y$.

If $f: X \to Y$ and $A \subset X$ then let $f|_A = f \cap (Y \times A)$, the "restriction" of f to A.

B.1. Definition. A relation $R \subset X \times X$ is an *equivalence relation* on X if:

(1) (reflexive) $\qquad x\,R\,x \quad$ for all $x \in X$,

(2) (symmetric) $\quad x\,R\,y \;\Rightarrow\; y\,R\,x,$
(3) (transitive) $\quad x\,R\,y$ and $y\,R\,z \;\Rightarrow\; x\,R\,z.$

B.2. Definition. If R is an equivalence relation on X then we put

$$[x] = \{y \in X \mid x\,R\,y\}.$$

This is called the *equivalence class* of x.

B.3. Proposition. *If R is an equivalence relation, then $[x]=[y] \Leftrightarrow x\,R\,y$. Also $[x] \cap [y] \neq \varnothing \Leftrightarrow [x]=[y]$.* □

In other words the equivalence classes $[x]$ partition X into disjoint subsets whose union is X.

B.4. Definition. If R is an equivalence relation on X then the set of equivalence classes $\{[x] \mid x \in X\}$ is denoted by X/R. There is the canonical surjection $\phi: X \to X/R$ given by $\phi(x) = [x]$.

B.5. Definition. If X is a set then its *power set* is $\mathscr{P}(X) = \{A \mid A \subset X\}$. Also let $\mathscr{P}_0(X) = \mathscr{P}(X) - \{\varnothing\}$.

B.6. Definition. If X and Y are sets, put $Y^X = \{f \mid f: X \to Y\}$.

B.7. Proposition. *If 2 denotes the set $\{0,1\}$ of two elements then the correspondence $A \leftrightarrow \chi_A$ between $\mathscr{P}(X)$ and 2^X given by*

$$\chi_A(x) = \begin{cases} 0 & \text{if } x \notin A, \\ 1 & \text{if } x \in A, \end{cases}$$

is a bijection. □

B.8. Definition. A *partial ordering* on a set X is a relation $\leq$ on X such that:

(1) (reflexive) $\quad a \leq a$ for all $a \in X,$
(2) (antisymmetric) $\quad a \leq b$ and $b \leq a \;\Rightarrow\; a = b,$
(3) (transitive) $\quad a \leq b$ and $b \leq c \;\Rightarrow\; a \leq c.$

A set together with a partial ordering is called a *partially ordered set* or a *poset*.

B.9. Definition. A poset X is said to be *totally ordered* (or *simply ordered* or *linearly ordered* or a *chain*) if $a, b \in X \Rightarrow$ either $a \leq b$ or $b \leq a$.

B.10. Definition. A function $f: X \to X$ on a poset is called *isotone* if $x \leq y \Rightarrow f(x) \leq f(y)$.

B.11. Definition. If $(X, \leq)$ is a poset and $A \subset X$ then $x \in X$ is an *upper bound* for A if $a \in A \Rightarrow a \leq x$. The element x is a *least upper bound* or *lub* for A if it is an upper bound and x' an upper bound for $A \Rightarrow x \leq x'$. Similarly for *lower* bound and *greatest lower bound* or *glb*. Also, *supremum* = *least upper bound* and *infimum* = *greatest lower bound*, and *sup* and *inf* are abbreviations of these.

B.12. Definition. A *lattice* is a poset such that every two element subset has an lub and a glb. It is a *complete lattice* if every subset has an lub and a glb.

B.13. Proposition. *If S is a set then $\mathscr{P}(S)$ is partially ordered by inclusion (i.e., by $\subset$) and is a complete lattice.* $\square$

B.14. Proposition. *If X is a complete lattice and $f: X \to X$ is isotone, then f has a fixed point, i.e., $\exists x \in X \ni f(x) = x$.*

Proof. Let $Y = \{x \in X \mid f(x) \geq x\}$ and put $y_0 = \sup(Y)$. Note that $y \in Y \Rightarrow f(y) \geq y \Rightarrow f(f(y)) \geq f(y) \Rightarrow f(y) \in Y$. Also $y \in Y \Rightarrow y \leq y_0 \Rightarrow y \leq f(y) \leq f(y_0) \Rightarrow f(y_0)$ is an upper bound for $Y \Rightarrow f(y_0) \geq y_0 \Rightarrow y_0 \in Y \Rightarrow f(y_0) \in Y \Rightarrow f(y_0) \leq y_0$. Since we had the opposite inequality, we conclude that $f(y_0) = y_0$. $\square$

B.15. Proposition. *Let $f: X \to Y$ and $g: Y \to X$ be functions. Then there are sets $A \subset X$ and $B \subset Y$ such that $f(A) = B$ and $g(Y - B) = X - A$.*

Proof. Consider the power set $\mathscr{P}(X)$ ordered by inclusion. It is a complete lattice by Proposition B.13. If $S \subset X$ then let $h(S) \in \mathscr{P}(X)$ be $h(S) = X - g(Y - f(S))$.

If $S \subset T$ then it is easy to see that $h(S) \subset h(T)$, so that h is isotone. By Proposition B.14 there is a subset $A \subset X$ such that $h(A) = A$. Let $B = f(A)$. Then $g(Y - B) = g(Y - f(A)) = X - h(A) = X - A$. $\square$

B.16. Definition. A totally ordered set X is said to be *well ordered* if every non-empty subset has a least element. That is, $\varnothing \neq A \subset X \Rightarrow \exists a \in A \ni (b \in A \Rightarrow a \leq b)$. (Of course, the least element of A is $\mathrm{glb}(A)$.) If $x \in X$ then its *initial segment* is

$$\mathrm{IS}(x) = \{y \in X \mid y < x\},$$

and its *weak initial segment* is

$$\mathrm{WIS}(x) = \{y \in X \mid y \leq x\}.$$

Also, if $x \in X$ and is not the least upper bound of X (which may not exist) then we put

$$\mathrm{succ}(x) = \mathrm{glb}\{y \in X \mid y > x\},$$

the *successor* of x.

Note that every subset of a well ordered set is well ordered.

B.17. Lemma. *Let X be a poset such that every well ordered subset has an lub in X. If $f: X \to X$ is such that $f(x) \geq x$ for all $x \in X$, then f has a fixed point.*

PROOF. Pick an element $x_0 \in X$. Let S be the collection of subsets $Y \subset X$ such that:

(1) Y is well ordered with least element x_0 and successor function $f|_{Y - \{\mathrm{lub}\, Y\}}$.
(2) $x_0 \neq y \in Y \Rightarrow \mathrm{lub}_X(\mathrm{IS}_Y(y)) \in Y$.

For example, $\{x_0\} \in \mathrm{S}$, $\{x_0, f(x_0)\} \in \mathrm{S}$, etc. We need the following sublemmas (A) and (B):

(A) If $Y \in \mathrm{S}$ and $Y' \in \mathrm{S}$, then Y is an initial segment of Y' or vice versa.

To prove (A) let $V = \{x \in Y \cap Y' \mid \mathrm{WIS}_Y(x) = \mathrm{WIS}_{Y'}(x)\}$. Suppose first that V has a last element v. If v is not the last element of Y then $\mathrm{succ}_Y(v) = f(v)$. If v is not the last element of Y' then $\mathrm{succ}_{Y'}(v) = f(v)$. Hence if neither of Y, Y' is an initial segment of the other then $f(v) \in V$, whence $f(v) = v$ and we are done.

If, on the contrary, V has no last element, let $z = \mathrm{lub}_X(V)$. If $Y \neq V \neq Y'$ then it follows from (2) that $z \in Y \cap Y'$ (because if $y = \inf(Y - V)$ then $V = \mathrm{IS}_Y(y)$ and therefore $z = \mathrm{lub}_X(\mathrm{IS}_Y(y)) \in Y$ by (2)). Therefore, $z \in V$, a contradiction, proving (A).

(B) The set $Y_0 = \bigcup \{Y \mid Y \in \mathrm{S}\}$ is in S.

To prove (B) note that if $y_0 \in Y \in \mathrm{S}$ then it follows from (A) that $\{y \in Y_0 \mid y < y_0\} = \mathrm{IS}_Y(y_0)$ and so this subset is well ordered with successor function f. This implies immediately that Y_0 is well ordered and satisfies (1). Also $\mathrm{lub}_X(\mathrm{IS}(y_0)) \in Y \subset Y_0$ which gives condition (2) for Y_0. Thus (B) is proved.

Now we complete the proof of Lemma B.17. Let $y_0 = \mathrm{lub}_X(Y_0)$. If $y_0 \notin Y_0$ then $Y_0 \cup \{y_0\} \in \mathrm{S}$ and so $y_0 \in Y_0$ after all. If $f(y_0) > y_0$ then $Y_0 \cup \{f(y_0)\} \in \mathrm{S}$ contrary to the definition of Y_0. Thus $f(y_0) = y_0$ as desired. □

B.18. Theorem. *The following statements are equivalent:*

(A) *For each set X, there is a function $f: \mathscr{P}_0(X) \to X$ such that $f(S) \in S$ for all $\varnothing \neq S \subset X$.*
(B) *If X is a poset such that every well ordered subset has an* lub *in X then X contains a maximal element, i.e., an element $a \in X \ni a' \geq a \Rightarrow a' = a$.*
(C) *(Maximal Chain Theorem.) If X is a poset then X contains a maximal chain, i.e., a chain not properly contained in any other chain in X.*
(D) *(Maximality Principle.) If X is a poset such that every chain in X has an upper bound, then X has a maximal element.*
(E) *(Zermelo, Well-Ordering Theorem.) Every set can be well ordered.*
(F) *If $f: X \to Y$ is surjective then there is a section $g: Y \to X$ of f, i.e., an injection $g: Y \to X$ such that $f \circ g = 1_Y$.*
(G) *(Axiom of Choice.) If $\{S_\alpha \mid \alpha \in A\}$ is an indexed family of nonempty sets S_α then there exists a function $f: A \to \bigcup S_\alpha$ such that $f(\alpha) \in S_\alpha$ for all $\alpha \in A$.*

PROOF. (A)⇒(B): Assume (B) is false. Then let $X_a = \{x \in X \mid x > a\}$. By assumption $X_a \neq \varnothing$ for all $a \in X$. Let $g: \mathscr{P}_0(X) \to X$ be a choice function. Define

$f: X \to X$ by $f(a) = g(X_a) > a$. Then $f(x) > x$ for all $x \in X$ contrary to Lemma B.17.

(B)$\Rightarrow$(C): Let $\mathbf{S}$ be the collection of all chains in X ordered by inclusion. If $\mathbf{C} \subset \mathbf{S}$ is a chain of chains (i.e., $Y_1, Y_2 \in \mathbf{C} \Rightarrow Y_1 \subset Y_2$ or $Y_2 \subset Y_1$) then $\bigcup\{Y \mid Y \in \mathbf{C}\}$ is a chain. Therefore *every* chain in $\mathbf{S}$ has a lub. By (B) there is a maximal element of $\mathbf{S}$, i.e., a maximal chain.

(C)$\Rightarrow$(D): Pick a maximal chain C and note that if x is an upper bound of C, then x is maximal.

(D)$\Rightarrow$(E): Consider the collection W of elements of the form $(U, \ll)$ where $U \subset X$ and $\ll$ is a well ordering on U. Order these by $(U, \ll) \leq (V, \ll') \Leftrightarrow$ they are equal or $(U, \ll)$ is an initial segment of $(V, \ll')$ and $\ll$ is the restriction of $\ll'$ to $U \times U$.

As in the proof of Lemma B.17 we see that every chain in W has a (least) upper bound, namely, the union of its elements. Thus (D) implies that there exists a maximal (with respect to $\leq$) well ordering, say $(U, \ll)$.

We claim that $U = X$. If not, let $x \in X - U$ and define $(U \cup \{x\}, \ll')$ where $\ll' = \ll \cup (U \times \{x\})$, i.e., make x larger than anything in U. This contradicts maximality of $(U, \ll)$.

(E)$\Rightarrow$(F): Well order X and let $g(y)$ be the first element of $f^{-1}(y)$. Then $f \circ g(y) = y$.

(F)$\Rightarrow$(G): Let $S = \bigcup S_\alpha$ and $X = \{\langle s, \alpha \rangle \in S \times A \mid x \in S_\alpha\}$. Let $p_s: X \to S$ and $p_A: X \to A$ be the projections $p_S\langle s, \alpha \rangle = s$ and $p_A\langle s, \alpha \rangle = \alpha$. Then p_A is onto since each $S_\alpha \neq \varnothing$. Thus there is a section $g: A \to X$ for p_A; i.e., $g(\alpha) = \langle s, \alpha \rangle$ for some $s \in S_\alpha$. Let $f = p_S \circ g: A \to S$. Then f is a choice function since $f(\alpha) = p_S g(\alpha) = p_S\langle s, \alpha \rangle = s$ for some $s \in S_\alpha$, all $\alpha \in A$.

(G)$\Rightarrow$(A): For $T \in \mathscr{P}_0(X)$ define $S_T = T$. Then $\mathscr{P}_0(X) = \{S_T \mid T \in \mathscr{P}_0(X)\}\}$ is an indexed collection of nonempty sets. Note that $\bigcup S_T = X$ since, for any $x \in X, x \in \{x\} = S_{\{x\}}$. By (G) there is a function $f: \mathscr{P}_0(X) \to \bigcup S_T = X$ such that $f(T) \in S_T = T$ for any $\varnothing \neq T \subset X$. $\square$

The Maximality Principle (D) is often inappropriately referred to as "Zorn's Lemma." It is actually due, independently, to R.L. Moore and Kuratowski, a dozen years before Zorn.

The only numbered results in this appendix that depend on the Axiom of Choice are Theorems B.28 and B.26(d). (The latter requires only a countable number of arbitrary choices, and so is relatively innocuous.)

B.19. Definition. Two sets X and Y are said to have the *same cardinal number* if there exists a one–one correspondence between them.

Given a set $\mathbf{S}$ of sets, this relation is an equivalence relation on $\mathbf{S}$. If $X \in \mathbf{S}$ we denote the equivalence class of X by $\operatorname{card}(X)$. We also write $\operatorname{card}(X) \leq \operatorname{card}(Y)$ if there exists an injection $f: X \to Y$.

B.20. Theorem (Schroeder–Bernstein). *If* $\operatorname{card}(X) \leq \operatorname{card}(Y)$ *and* $\operatorname{card}(Y) \leq \operatorname{card}(X)$ *then* $\operatorname{card}(X) = \operatorname{card}(Y)$.

PROOF. (Note that this proof does not use the Axiom of Choice.) By hypothesis there exist injections $f: X \to Y$ and $g: Y \to X$. By Proposition B.15 there exist subsets $A \subset X$ and $B \subset Y$ such that $f(A) = B$ and $g(Y - B) = X - A$. Therefore $f|_A: A \leftrightarrow B$ and $g|_{Y-B}: Y - B \leftrightarrow X - A$ are one–one correspondences. Put them together. □

B.21. Corollary. *The ordering $\leq$ on the cardinals is a partial ordering.* □

It is not hard to see that, assuming the Axiom of Choice in the guise of the Well-Ordering Theorem, the cardinals are well ordered by $\leq$. This is, in fact, equivalent to the Axiom of Choice.

B.22. Theorem. *For any X,* $\operatorname{card}(X) < \operatorname{card}(\mathscr{P}(X))$.

PROOF. The relation $\operatorname{card}(X) \leq \operatorname{card}(\mathscr{P}(X))$ holds because of the injection $x \mapsto \{x\}$. Let $f: X \to \mathscr{P}(X)$ be any function. Put $A = \{x \in X \mid x \notin f(x)\}$. We claim that there can be no $y \in X$ with $A = f(y)$. If there is such a y then

$$y \in A \quad \Rightarrow \quad y \notin f(y) = A$$

and

$$y \notin A \quad \Rightarrow \quad y \in f(y) = A,$$

so neither possibility is tenable. Thus there never exists a surjection $f: X \to \mathscr{P}(X)$. □

The symbol ω is used to denote the set of nonnegative integers with the usual ordering. Let $\omega' = \omega \cup \{\omega\}$, tacking on a last element. Note that $\operatorname{card}(\omega)$ is the least infinite cardinal.

B.23. Definition. *A* set X is said to be *countable* if there exists an injection $f: X \to \omega$.

B.24. Lemma. *The product $\omega \times \omega$ is countable.*

PROOF. The function $f: \omega \times \omega \to \omega$ given by $f(n, k) = (2n + 1)2^k - 1$ is a bijection. □

B.25. Lemma. *If $f: X \to Y$ is an injection with $X \neq \varnothing$ then there exists a surjection $g: Y \to X$ such that $g \circ f = 1_X$.*

PROOF. For some $x_0 \in X$ let $g(y)$ be x_0 for $y \notin f(X)$ and $g(y) = f^{-1}(y)$ for $y \in f(X)$. □

B.26. Theorem.

(a) *If X is countable and $f: X \to Y$ is onto then Y is countable.*
(b) *A subset of a countable set is countable.*

(c) *X, Y countable $\Rightarrow X \times Y$ countable.*
(d) *A countable union of countable sets is countable.*

Proof. For (a) let $g: X \to \omega$ be injective and define $h(y) = \inf g(f^{-1}(y))$. Then $h: Y \to \omega$ is an injection.

Part (b) is trivial.

For (c), if $f: X \to \omega$ and $g: Y \to \omega$ are injections then the composition of $f \times g: X \times Y \to \omega \times \omega$ with the injection $\omega \times \omega \to \omega$, given by Lemma B.24, gives an injection $X \times Y \to \omega$.

For (d), suppose that X_α is a countable set defined for $\alpha \in A \neq \varnothing$ which is countable. Then let $f_\alpha: \omega \to X_\alpha$ be a surjection and $g: \omega \to A$ a surjection. Let $h: \omega \times \omega \to \bigcup\{X_\alpha | \alpha \in A\}$ be given by $h(n, k) = f_{g(n)}(k)$. Then h is surjective and so $\bigcup X_\alpha$ is countable by (a). □

In general, it can be shown that if X, Y are nonempty, and not both finite, then $\operatorname{card}(X \cup Y) = \max(\operatorname{card}(X), \operatorname{card}(Y)) = \operatorname{card}(X \times Y)$. The consequence that $\operatorname{card}(X \times X) = \operatorname{card}(X)$, whenever X is infinite, is equivalent to the Axiom of Choice.

B.27. Theorem. *If* $\mathbf{R}$ *is the set of reals then* $\operatorname{card}(\mathbf{R} - \{0\}) = \operatorname{card}(\mathscr{P}(\omega)) > \operatorname{card}(\omega)$.

Proof. Let $\mathbf{R}_+$ denote the nonnegative reals. The injection $\mathbf{R}_+ \hookrightarrow \mathbf{R}$ and the injection $\mathbf{R} \to \mathbf{R}_+$ given by $x \mapsto e^x$ show that $\operatorname{card}(\mathbf{R}) = \operatorname{card}(\mathbf{R}_+)$. Similarly, $\operatorname{card}(\mathbf{R}) = \operatorname{card}(\mathbf{R} - \{0\})$. We shall exhibit a bijection $\mathbf{R}_+ \leftrightarrow \mathscr{P}(\omega)$.

First write each positive real number r in its continued fraction expansion

$$r = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

where the a_i are integers with $a_0 \geq 0$ and $a_i > 0$ for $i > 0$. We shall denote this expansion by $r = [a_0, a_1, a_2, \ldots]$. A terminating (rational) continued fraction will be written in the form

$$r = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots \; a_{s-1} + \cfrac{1}{a_s + 1}}}$$

and condensed to $[a_0, a_1, \ldots, a_s]$. In particular, an integer $n > 0$ is written as $n = (n-1) + 1 = [n-1]$. With this understanding, a continued fraction representing r is uniquely determined by r. Thus this determines a one–one correspondence of the positive reals r with the sequences, infinite or finite, $[a_0, a_1, \ldots]$. Also let the real 0 correspond to the empty sequence. Finally, let a sequence $[a_0, a_1, \ldots]$ correspond to the subset $\{a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots\}$ of ω. This is easily seen to be a one–one correspondence between the non-negative reals and subsets of ω, as claimed. The real number 0 corresponds to the empty subset of ω. $\square$

It should be noted that, in the above correspondence, the rationals correspond to the finite subsets of ω. Thus the set of all finite subsets of ω is countable. (I believe the foregoing proof is due to A. Gleason.)

B.28. Theorem. *There exists an uncountable well-ordered set Ω' with last element Ω such that $x < \Omega \Rightarrow \mathrm{IS}(x)$ is countable.*

PROOF. Well order the reals and put on an extra element x_0 at the end. Then let Ω be the least element in the ordering such that $\mathrm{IS}(\Omega)$ is uncountable. This exists since $\mathrm{IS}(x_0)$ has cardinality that of $\mathbf{R}$ which is greater than that of ω. Then $\Omega' = \mathrm{WIS}(\Omega)$ is the desired set. Note that by an equivalence, one can regard Ω' as $\Omega \cup \{\Omega\}$. $\square$

We shall refer to Ω, as in Theorem B.28, as "the least uncountable ordinal" and to other elements of Ω' as "countable ordinal numbers."

B.29. Theorem. *If* $\mathrm{card}(X) = \mathrm{card}(X \times Y)$ *then* $\mathrm{card}(\mathscr{P}_0(X)) = \mathrm{card}(\mathscr{P}_0(X) \times \mathscr{P}_0(Y))$.

PROOF. By assumption there is a one–one correspondence $f: X \times Y \to X$. This induces a one–one correspondence $F: \mathscr{P}(X \times Y) \to \mathscr{P}(X)$ by $F(S) = \{f(x, y) \mid \langle x, y \rangle \in S\}$; i.e., $F(S) = f(S)$. But $g: \mathscr{P}_0(X) \times \mathscr{P}_0(Y) \to \mathscr{P}_0(X \times Y)$ given by $g(\langle S, T \rangle) = S \times T$ is an injection, and so $F \circ g: \mathscr{P}_0(X) \times \mathscr{P}_0(Y) \to \mathscr{P}_0(X)$ is also an injection. There is also an injection $\mathscr{P}_0(X) \to \mathscr{P}_0(X) \times \mathscr{P}_0(Y)$ (unless $Y = \varnothing$, in which case the result is trivial) and so the contention follows from the Schroeder–Bernstein Theorem (Theorem B.20). $\square$

B.30. Corollary. *For any positive integer n we have* $\mathrm{card}(\mathbf{R}^n) = \mathrm{card}(\mathbf{R})$.

PROOF. By Lemma B.24, $\mathrm{card}(\omega \times \omega) = \mathrm{card}(\omega)$. By Theorem B.27, $\mathrm{card}(\mathbf{R}) = \mathrm{card}(\mathbf{R} - \{0\}) = \mathrm{card}(\mathscr{P}_0(\omega))$, so Theorem B.29 implies that $\mathrm{card}(\mathbf{R}^2) = \mathrm{card}(\mathbf{R} \times \mathbf{R}) = \mathrm{card}(\mathbf{R})$. If we know that $\mathrm{card}(\mathbf{R}^n) = \mathrm{card}(\mathbf{R})$ then $\mathrm{card}(\mathbf{R}^{n+1}) = \mathrm{card}(\mathbf{R} \times \mathbf{R}^n) = \mathrm{card}(\mathbf{R} \times \mathbf{R}) = \mathrm{card}(\mathbf{R})$ and so an induction finishes the proof. $\square$

As mentioned before, the Axiom of Choice implies similar facts for arbitrary infinite cardinals, but Theorem B.29 and Corollary B.30 do not depend on the Axiom of Choice.

In this book, we shall often make use of the Axiom of Choice without explicit mention. In cases where use of the axiom is known to be crucial, we