Notes on Set Theory

Robert H. C. Moir
The University of Western Ontario
Dept. of Philosophy
robert@moir.net

Contents

1 Fundamentals of Sets and Classes

1.1 Classes and Sets

Sets are definite collections of objects considered themselves as a single object. We will consider a *class* to be any collection of distinct objects. Thus, a *set* is any collection of objects that can be a member of a class. For any given class A, we write $a \in A$ to state that a is an *element* or *member* of A, or that A *contains* a.

1.2 Principle of Extensionality

If P is some property such that it can be considered to be a property of an arbitrary object x, then the *extension* of P is the class of objects that has the property P. The extension of P, the class of all objects x with property P is denoted by

$$\{x \mid Px\},$$

where Px means that x has P.

A class is taken to be identified by its elements. So if A and B are any classes such that

$$x \in A \Leftrightarrow x \in B,$$

then $A = B$. This is called the *principle of extensionality.*

1.3 Axiom of Pairing

One of the basic principles of set theory is that for any two objects a and b, the class $\{a, b\}$ is a set. This is called the *axiom of pairing.* This entails that for any object a the class $\{a\}$ is a set. This is called *the singleton of* a.

1.4 Axiom of Subsets

There is a special set that has no elements called the *empty set*, which is denoted by $\varnothing$. It can be expressed as the class

$$\varnothing = \{x \mid x \neq x\}.$$

Let A and B be two classes. If each element B is an element of A, then B is called a *subclass* of A. This is denoted as $B \subseteq A$. If B is a subclass of A, but $B \neq A$, then B is a *proper subclass* of A, which we denote by $B \subset A$. Now, if $B \subseteq A$ and A is a set, then B is also a set. This is called the *axiom of subsets.* Thus, if A is a set, it follows that the class

$$\{x \in A \mid Px\}$$

is also a set.

1.5 Proper Classes

Some classes are so large that they cannot be comprehended as single objects, *i.e.* they cannot be sets. Such classes are called *proper classes*. An example of such a class is the *Russell class* defined as the class

$$R = \{x \mid x \text{ is a set such that } x \notin x\}.$$

To see that this is a proper class, suppose that R is a set. It must be that $R \in R$ or $R \notin R$. Well, if $R \in R$, then by definition $R \notin R$. On the other hand if $R \notin R$, then $R \in R$ by definition of R. These two contradictions establish that R cannot be a set, and hence it is a proper class. Another proper class is the *class of all sets* V,

$$V = \{x \mid x \text{ is a set}\}.$$

To see that this is a proper class, assume that V is a set. Since each member of the Russell class is in V, it follows that $R \subseteq V$. Then, by the axiom of subsets, R must be a set, but this is a contradiction. Thus, V is a proper class.

1.6 Axiom of Union

If A is any class, then

$$\bigcup A =_{def} \{x \mid x \in y \text{ for some } y \in A\}.$$

$\bigcup A$ is called the *union class of* A. If A is a set, then so is $\bigcup A$. This is called the *axiom of union*. Notice that the members of $\bigcup A$ are the members of the members of A. The *union* (or *join*) of two classes A and B is written as

$$A \cup B =_{def} \{x \mid x \in A \text{ or } x \in B\}.$$

If A and B are sets, then their union $A \cup B$ is a set by the axiom of union.

IF A is any class, then

$$\bigcap A =_{def} \{x \mid x \in y \text{ for every } y \in A\}.$$

This is called the *intersection class of* A. This is the class of members of members of A that are members of every member of A. If A and B are classes, then

$$A \cap B =_{def} \{x \mid x \in A \text{ and } x \in B\}.$$

This is called the *intersection* (or *meet*) of A and B. If A and B are sets, then $\bigcap A$ and $A \cap B$ are sets. This follows from the axiom of union and the axiom of subsets.

If A is any class, then

$$A^c =_{def} \{x \mid x \notin A\}$$

This is called the *complement* of A. If A and B are classes, then

$$B \backslash A =_{def} B \cap A^c.$$

This is called the *difference between* B *and* A. If A and B are sets, $B \backslash A$ is a set, but A^c could be a proper class. $B \backslash A$ is also called the *relative complement of* A *in* B. Without potentially referring to proper classes we have

$$B \backslash A = \{x \mid x \in B \text{ and } x \notin A\}.$$

1.7 Axiom of Power Set

If A is any class, then

$$\mathcal{P}A =_{def} \{x \mid x \text{ is a set such that } x \subseteq A\}.$$

This is called the *power class of A*. If A is a set, then $\mathcal{P}A$ is a set. This is called the *axiom of power set*. From the point of view of the class/set distinction, what is different between this axiom and the previous ones?

1.8 Axiom of Infinity

The axioms presented so far do not entail that any infinite sets exist. This requires the *axiom of infinity*, which states that there is a set Z such that $\varnothing \in Z$ and such that for every set $x \in Z$ also $x \cup \{x\} \in Z$. The natural numbers are the set

$$\mathbb{N} = \{1, 2, 3, \ldots\}.$$

A special infinite class is the *ordered* class of all natural numbers, ordered in terms of magnitude, which is denoted by ω. The axiom of infinity is equivalent to the statement that ω is a set.

2 Constructions I (Products and Coproducts)

Many of the axioms presented provide ways of generating new sets from old ones. This section considers other important set theoretic constructions.

2.1 Finite Products

An *ordered pair* of two objects a and b, not necessarily distinct, is an object $\langle a, b\rangle$ such that

$$\langle a, b\rangle = \langle c, d\rangle \Leftrightarrow a = c \text{ and } b = d.$$

Note that $\langle a, b\rangle \neq \{a, b\}$, which is not ordered. We can define an ordered pair to be

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

Why does this succeed as a definition?

An important class that can be constructed from two classes A and B is the class

$$A \times B =_{def} \{\langle a, b\rangle \mid a \in A \text{ and } b \in B\}.$$

This is called the (*cartesian*) *product* of A and B. If A and B are sets, then so is their product. An ordered *n-tuple* of n objects $a_1, a_2, \ldots, a_n$ is the natural generalization of an ordered pair of two objects. An n-tuple of these objects is denoted as

$$\langle a_1, a_2, \ldots, a_n\rangle.$$

The (*cartesian*) *product* of any classes $A_1, A_2, \ldots, A_n$ is the class

$$A_1 \times A_2 \times \cdots \times A_n =_{def} \{\langle x_1, x_2, \ldots, x_n\rangle \mid x_1 \in A_1, x_2 \in A_2, \ldots, x_n \in A_n\}.$$

We may sometimes write

$$\prod_{i=1}^{n} A_i =_{def} A_1 \times A_2 \times \cdots \times A_n.$$

This can also be written less explicitly as $\prod_i A_i$ or simply just as $\prod A_i$ where there is no confusion. The *n-th cartesian power* of a class A is the product of A with itself n times:

$$A^n = \underbrace{A \times A \times \cdots \times A}_{n \text{ times}}.$$

2.2 Finite Coproducts

Another important class that can be constructed from two classes A and B is the class

$$A + B =_{def} \{\langle 1, a\rangle \mid a \in A\} \cup \{\langle 2, b\rangle \mid b \in B\}.$$

This is called the *coproduct* or *disjoint union* of A and B. Similarly to the case of products, we can construct the coproduct of n classes A_i, $i \in I = \{1, 2, \ldots, n\}$:

$$A_1 + A_2 + \cdots + A_n =_{def} \{\langle i, x_i\rangle \mid i \in I \text{ and } x_i \in A_i\}.$$

Similarly to the case of products we may write

$$\coprod_{i=1}^{n} A_i =_{def} A_1 + A_2 + \cdots + A_n,$$

which can be written elliptically as either $\coprod_i A_i$ or $\coprod A_i$ where there is no confusion.

2.3 Finite Unions and Intersections

Taking our attention back to unions and intersections, we have a special notation for the union and intersection of n classes A_i, $i \in \{1, 2, \ldots, n\}$. We have that

$$\bigcup_{i=1}^{n} A_i =_{def} A_1 \cup A_2 \cup \cdots \cup A_n$$

and that

$$\bigcap_{i=1}^{n} A_i =_{def} A_1 \cap A_2 \cap \cdots \cap A_n.$$

Similarly to above we may write $\bigcup_i A_i$ for unions and $\bigcap_i A_i$ for intersections if no confusion will arise. Notice that there is an analogy between intersections and products and an analogy between unions and co-products.

As we would expect, all the statements of this section apply *mutatis mutandis* replacing 'class' with 'set.'

3 Functions

3.1 Functions

A function f associates to some object x another object $f(x)$. The function f maps x to $f(x)$. The class of all objects x that are assigned to some value $f(x)$ is called the *domain* of f, which is denoted $dom(f)$. The *graph* of f is the following class of ordered pairs:

$$\{\langle x, f(x)\rangle \mid x \in dom(f)\}.$$

A function associates only one object with each element of its domain. This can be stated by saying that a function is not a one-many map, *i.e.* it does not map any element of its domain to more than one object. A function can be many-one, which is to say it sends many members of its domain to the same object, but not one-many. The class of all objects that the elements of the domain of a function f are mapped to is called the *image* of f, and it denoted by $im(f)$.

A *function* is defined to be a class f of ordered pairs that have the property that whenever both $\langle x, y\rangle \in f$ and $\langle x, z\rangle \in f$ then $y = z$. The *domain* of f is the class

$$dom(f) =_{def} \{x \mid \langle x, y\rangle \in f \text{ for some } y\}.$$

If $x \in dom(f)$, then the *value of* f *at* x, denoted $f(x)$, is the unique y such that $\langle x, y\rangle \in f$. The *image* (or *range*) of f is then the class

$$im(f) =_{def} \{f(x) \mid x \in dom(f)\}.$$

It follows that f is identical to its graph.

A function f is a map *from* A *to* B (or that f maps *into* B) if $dom(f) = A$ and $im(f) \subseteq B$. In such a case the class B is called the *codomain* of f, which is denoted as $cod(f)$.

3.2 Surjections, Injections and Bijections

A function f is *surjective* (or *onto*) if it is a function from A to B with the property that $im(f) = cod(f) = B$. Equivalently f is surjective if for any $z \in B$ there is an $x \in A$ such that $f(x) = z$. A surjective function is called a *surjection*.

A function f is *injective* (or *one-to-one* or *1-1*) if it is a function from A to B with the property that whenever x and y are distinct members of $dom(f)$ then $f(x)$ and $f(y)$ are distinct. Equivalently, f is injective if for any $x, y \in A$, if $f(x) = f(y)$ then $x = y$. Notice the symmetry with the definition of a function, since a characteristic property of a function is that for any $x, y \in A$, if $x = y$ then $f(x) = f(y)$. An injective function is called an *injection*.

A function f is *bijective* if it is a map from A to B that is both injective and surjective. A bijective function is called a *bijection*.

3.3 Axiom of Replacement

If f is a function and $dom(f)$ is a set, then $im(f)$ is also a set. This is called the *axiom of replacement*. If f is an function from A to B, which we may denote as $f\colon A \longrightarrow B$, and A and B are both sets, then both $im(f)$ and f are sets. This does not actually require the axiom of replacement to prove, however, since it follows from the previously stated axioms. It follows from the axiom of replacement, however, that for any function from any set into any class, the image of f is a set.

Let f be a function $f\colon A \longrightarrow B$. Then the *restriction function* to a subset $C \subseteq A$, denoted $f|_C$ is the set $f \cap C \times B$. This is function generated by the restriction of the domain of f to C.

3.4 Image and Preimage

Let f be a function $f\colon A \longrightarrow B$. The *image* of an element $x \in A$ is the element $f(x) \in B$. The image of a subclass $X \subseteq A$ is the class

$$f(X) =_{def} \{y \in B \mid x \in X \text{ and } f(x) = y\}.$$

This makes the image a function from $\mathcal{P}A$ to B.

The *preimage* or *inverse image* of a subclass $Y \subseteq B$, is the class

$$f^{-1}(Y) =_{def} \{x \in A \mid y \in Y \text{ and } f(x) = y\}.$$

If $b \in B$, then the inverse image $f^{-1}(\{b\})$ of the singleton $\{b\}$ is called a *fibre*. We also call $f^{-1}(\{b\})$ the *inverse image* or *preimage of* b and write it as $f^{-1}(b)$ since it is essentially equivalent to $f^{-1}(\{b\})$.

In the case that $f\colon A \longrightarrow B$ is injective, the inverse image of each $b \in im(f) \subseteq B$ is a unique element $a \in A$. In such a case the map g that takes each $b \in im(f)$ to its preimage $a \in A$ is a function. This function is called the *inverse function of* f, and is written as f^{-1}. If we define the identity function on a class A to be the class

$$1_A = \{\langle x, x\rangle \mid x \in A\},$$

then it follows that $f^{-1} \circ f = 1_A$. In the case that f is also bijective, then the domain of f is all of B. We may see that in this case it is also true that $f \circ f^{-1} = 1_B$. If for any function $f\colon A \longrightarrow B$ there is a g such that $g \circ f = 1_A$ and $f \circ g = 1_B$, then f is a bijection and $g = f^{-1}$.

3.5 Composition

Let f be a function $f\colon A \longrightarrow B$ and let g be a function $g\colon B \longrightarrow C$. The *composite of* g *with* f, written $g \circ f$, is the function defined such that each $a \in A$ is mapped to $g(f(a)) \in C$. This is the function that takes each element of A to its image in B under f and then takes this element to its image in C under g. It is common to read $g \circ f$ as 'g follows f.'

3.6 Exponentials

Given two classes A and B, the class of all functions from A to B is the class

$$B^A =_{def} \{f \mid f \text{ is a function and } dom(f) = A \text{ and } cod(f) = B\}$$

The class B^A is called a *function class* or *exponential class*. In case that A and B are sets, so is B^A. This provides another way of constructing new sets from old. You may notice that if M is a set with m elements, and N is a set with n elements, $m, n \in \mathbb{N}$, then the exponential M^N has m^n elements.

3.7 Arbitrary Union, Intersection, Product and Coproduct

With the notion of a function defined, we can now look at the generalization of the union, intersection, product and coproduct constructions. This requires two new concepts. A *family of sets* $\mathcal{C}$ indexed by the *index set* I is a function $f\colon I \longrightarrow \mathcal{C}$. The sets in the family are the sets $X_i = f(i)$ for $i \in I$. We may think of a family of sets as a kind of generalization of an ordered n-tuple. Accordingly, we may write $\mathcal{C} = \langle X_i \mid i \in I\rangle$ to stress the analogy to ordered tuples.

The *union* and *intersection* of a family of sets are defined as

$$\begin{aligned} \bigcup_{i\in I} X_i &= \bigcup\{X_i \mid i \in I\} \\ \bigcap_{i\in I} X_i &= \bigcap\{X_i \mid i \in I\}. \end{aligned}$$

Note that the union is always defined and the the intersection is defined provided that $I \neq \varnothing$.

Using families of sets we can also generalize the product and coproduct. Let $\mathcal{C} = \langle X_i \mid i \in I\rangle$ be a family of sets defined by a function $f\colon I \longrightarrow \mathcal{C}$. Then a *choice function* for the family is a function $g\colon I \longrightarrow \bigcup\mathcal{C}$ with the property that $g(i) \in f(i)$ for all $i \in I$. Essentially, the function g chooses a representative from each member of the family. The (*cartesian*) *product*

$$\prod_{i\in I} X_i$$

is defined to be the set of all choice functions for the family $\mathcal{C} = \langle X_i \mid i \in I\rangle$.

It is not at all clear that such a set exists in general. Its existence cannot be established from any of the set theoretic axioms we have considered, since it relies essentially on the existence of choice functions, which we cannot know exist in general. Their existence in for general families of sets requires another axiom called the *axiom of choice*. This axiom has a special kind of status in set theory, which is discussed in section 9. Products of families in which one of the members of the family is empty are empty and, hence, do not exist. The converse, however, is equivalent to the axiom of choice.

Let $\mathcal{C} = \langle X_i \mid i \in I\rangle$ be a family of sets. The *coproduct*

$$\coprod_{i\in I} X_i$$

is defined to be the set $\{\langle i, x_i\rangle \mid i \in I \text{ and } x_i \in X_i\}$. The existence of such a set in general also relies on the axiom of choice.

4 Relations

4.1 Relations and Properties

A *binary relation on* A is a class

$$R = \{\langle x, y\rangle \mid x, y \in A\}.$$

Thus, a function is a special kind of binary relation. An *n-ary relation* is a subclass R of A^n. A *property on* A is a unary relation on A, *i.e.* a subclass of A. A 0-ary relation on A is defined to be a designated element of A.

4.2 Properties of Relations

There are several important properties that a binary relation on A can have. Let R be a binary relation on A. We will write xRy if $\langle x, y\rangle \in R$. The properties we consider here are the following, where we assume that $x, y \in A$ are arbitrary:

- R is *reflexive* if xRx.
- R is *symmetric* if whenever xRy also yRx.
- R is *antisymmetric* if whenever xRy then yRx does not hold.
- R is *weakly antisymmetric* if whenever xRy and yRx then $x = y$.
- R is *transitive* if whenever xRy and yRz then xRz.
- R is *connected* if xRy or yRx
- *trichotomy* holds for R if exactly one of xRy, yRx and $x = y$ holds.

4.3 Equivalence Relations and Equivalence Classes

Particular combinations of these properties yield important kinds of relations. An *equivalence relation* on A is a relation R that is reflexive, symmetric and transitive. A simple, but important example of an equivalence relation on A is the identity function 1_A. This is also called the *diagonal* function (or relation) on A. Notice that a relation on a class that relates pairs its elements that can be mapped onto one another bijectively is an equivalence relation.

Let R be an equivalence relation on A. For each $a \in A$ we write

$$[a]_R = \{x \mid xRa\}.$$

The class $[a]_R$ is called the *R-class* of a, or the *equivalence class of a modulo R*. We may also, when there is no confusion, drop the subscript R and write $[a]$. It follows that for any $a, b \in A$, $[a] = [b]$ iff aRb, and that each $a \in A$ is a member of exactly one equivalence class, *viz.*, a member of $[a]$. This entails that for any $a, b \in A$, if $[a] \neq [b]$, then $[a] \cap [b] = \varnothing$.

4.4 Partitions

Any class $S \subseteq \mathcal{P}A$ of subclasses of A that has the property that for any $X, Y \in S$, either $X = Y$ or $X \cap Y = \varnothing$, and $\bigcup S = A$ is called a *partition* of A. We may notice that the class of equivalence classes generated by an equivalence relation R on A, is a partition of A.

4.5 Equivalence Relation Theorem

Now we may state the following theorem:

Theorem 4.1 (Equivalence Relation Theorem)

1. If R is an equivalence relation on A, then the class
$$P = \{\, [a]_R \mid a \in A\}$$
of R-classes is a partition of A.
2. If P is a partition of A, then
$$R = \{\langle x, y\rangle \mid x, y \in p \text{ for some } p \in P\}$$
is an equivalence relation on A.
3. The constructions in 1. and 2. are mutually inverse.

4.6 Partial Orders

A *sharp partial order* on A is a binary relation R on A that is both antisymmetric and transitive. A *blunt partial order* on A is a binary relation on R on A that is reflexive, weakly antisymmetric and transitive. You may notice that given any set A, the inclusion relation $\subseteq$ generates a blunt partial order on $\mathcal{P}A$, and the strict inclusion relation $\subset$ generates a sharp partial order on A.

4.7 Total Orders

A *sharp total order* on A is a binary relation R on A that is transitive and for which trichotomy holds. A *blunt total order* on A is a binary relation R on A that is connected, weakly antisymmetric and transitive. Notice that the natural ordering relations $\leq$ and $<$ on the natural numbers are blunt and sharp total order relations on $\mathbb{N}$ respectively.

Note that we may use sharp order and *strict order* interchangeably, and blunt order and *non-strict order* interchangeably.

4.8 Order Theorem

The following theorem establishes a close relation between strict and non-strict orders:

Theorem 4.2 (Order Theorem)

1. If R is a strict order on A, then
$$S = R \cup \{\langle x, x\rangle \mid x \in A\}$$
is a non-strict order on X.
2. If S is a non-strict order on A, then
$$R = S \cap \{\langle x, y\rangle \in S \mid x \neq y\}$$
is a strict order on X.
3. The constructions in 1. and 2. are mutually inverse.

5 Constructions II (Quotients)*

5.1 Classes of Fibres

Another important construction involves the construction of an injective and a bijective function from any function f. Let f be any function f: $A \longrightarrow B$. Consider the class Q of all fibres of f, *i.e.* the class

$$Q = \{f^{-1}(b) \mid b \in im(f)\}.$$

It follows from definitions that for any $b_1, b_2 \in B$, $f^{-1}(b_1) \cap f^{-1}(b_2) = \varnothing$ and that $\bigcup\{f^{-1}(b) \mid b \in B\} = A$. Thus, Q is a partition of A. By definition each member of the fibre $f^{-1}(b)$, where $b \in B$, is associated under f with a unique member of B, namely b. Likewise the fibre itself is naturally associated with a unique member $b \in B$. Also, each $b \in B$ is associated with a unique member of Q, the related fibre. Thus the function g: $Q \longrightarrow im(f)$ defined as

$$g = \{\langle f^{-1}(b), b\rangle \mid b \in im(f)\}$$

is a bijection from Q to $im(f)$. Considered as a function from Q to B, which is set-theoretically identical, g is an injection, but not in general a bijection since we do not know that it is onto. In case that f is a surjection, however, the function g from Q to B will be a bijection. Thus, we may see that we can construct an injective and a bijective function given any function f, and that for surjective functions we just get a bijection.

5.2 Quotients

We may see here a connection with the consideration of partitions and equivalence classes in the previous section. Using f we can define an equivalence relation $\sim$ defined such that for any $x_1, x_2 \in A$,

$$x_1 \sim x_2 \Leftrightarrow f(x_1) = f(x_2).$$

The equivalence classes under $\sim$ are just the fibres of f. Thus, we may see that the class of equivalence classes under $\sim$ is just the class Q defined above. This class is called the *quotient class* and is denoted as $A/\sim$. The function g: $A/\sim \longrightarrow B$ defined above is called the *quotient map*.

5.3 First Isomorphism Theorem

Let f be a function from A to B. The *kernel* of a function f, denoted $ker(f)$, is defined to be the relation on A defined as

$$ker(f) =_{def} \{\langle x_1, x_2\rangle \in A \times A \mid f(x_1) = f(x_2)\}.$$

Theorem 5.1 (First Isomorphism Theorem) Let f be a function from A to B. Then

1. $im(f)$ is a subclass of B;
2. $ker(f)$ is an equivalence relation on A;
3. f induces a bijection from $A/ker(f)$ to $im(f)$.

6 Finite and Infinite Sets

6.1 Cardinality

In the remainder of these notes, unless otherwise specified, we will restrict our attention to sets. For any two sets X and Y that have finitely many elements, they have the same number of elements iff there is a bijection from X to Y. Thus, if we know that there is a bijection from X to Y then we know that X and Y have the same number of elements even if we do not know how many elements each of the sets has. The cardinality of a set is taken to be a measure of the size of the set. In set theory, two sets X and Y are defined to have the same *cardinality* if there is a bijection from X to Y. We write $|X| = |Y|$ if the two sets have the same cardinality. Note that at this point, the notion of the cardinality of a set, which would be denoted $|X|$, is not defined.

The set X *has smaller cardinality than* the set Y, written as $|X| \leq |Y|$, if there is an injection from X into Y. If $|X| \leq |Y|$ and X and Y do not have the same cardinality, then X *has strictly smaller cardinality than* Y, which we write as $|X| < |Y|$.

There are (at least) two important theorems in this context.

6.2 Schröder-Bernstein Theorem

Theorem 6.1 (Schröder-Bernstein Theorem) If there is an injective function from X to Y and an injective function from Y to X, then there is a bijection between X and Y. This is to say that

$$(|X| \leq |Y| \wedge |Y| \leq |X|) \Rightarrow |X| = |Y|.$$

6.3 Cantor's Theorem

Theorem 6.2 (Cantor's Theorem) For any set X, there is an injection from X into $\mathcal{P}X$ but not bijection between them. This is to say that

$$|X| < |\mathcal{P}X|.$$

6.4 Finite and Infinite Sets

For convenience we will redefine the set of natural numbers in the following way

$$\mathbb{N} =_{re\text{-}def} \mathbb{N} \cup \{0\}.$$

Then we will define the number n to be the set $\{0, 1, \ldots, n-1\} \subset \mathbb{N}$. A set X, then, has n elements if there is a bijection from X to n. It follows that if there is a bijection between m and n, then $m = n$. From this it follows that for any set X there is at most one natural number n such that there is a bijection from X to n. A set is defined to be *finite* if such an n exists, and *infinite* otherwise.

An alternative definition of finite and infinite sets is due to Dedekind. A set X is *Dedekind infinite* if there is a bijection from X to a proper subset of X, and is *Dedekind finite* otherwise. It turns out that if we have the axiom of choice, the definitions of finiteness and infiniteness are equivalent, but without choice they are not.

6.5 Countable Sets

A set is defined to be *countable* if it is finite of it has the same cardinality as the set of natural numbers $\mathbb{N}$. A set is *countably infinite* if it has the same cardinality as $\mathbb{N}$. We have that there is a surjection from $\mathbb{N}$ to a set X iff X is countable. Many different constructions involving only countable sets yield countable sets. It can be shown that the union of countably many countable sets is itself a countable set. Also, the product or coproduct of two countable sets is countable. Hence, finite products and coproducts of countable sets are countable.

Notice that Cantor's theorem entails that there are sets that are not countable, *i.e.* infinite sets that are strictly larger than the set of natural numbers $\mathbb{N}$. Such a set, *viz.* a set X such that there exists no surjection from $\mathbb{N}$ onto X, is an *uncountable* set. Before we consider any such sets, we have the following theorem:

Theorem 6.3 The sets $\mathbb{Z}$ of integers and $\mathbb{Q}$ of rational numbers are both countable.

Interestingly, however, we have the following theorem:

Theorem 6.4 The set $\mathbb{R}$ of real numbers is uncountable.

The first two proofs of theorem 6.4 were given by Georg Cantor in 1874 and 1891.

7 Well-order and Induction

This section is devoted to a consideration of the notion of a well-ordering and its relation to mathematical induction.

7.1 Well-Orders and the Least Number Principle

A *well-order* on a set X is a total order $<$ on X with the property that any non-empty subset of X has a least element. The natural numbers $\mathbb{N}$ with their natural ordering form a *well-ordered set*, which is to say the order relation is a well-order. This set $\langle \mathbb{N}, < \rangle$ is the simplest infinite well-ordered set. Any finite totally ordered set is well-ordered. That $\langle \mathbb{N}, < \rangle$ is a well-ordered set is called the *least number principle*. The least number principle (LNP) has an important relationship to the *principle of mathematical induction* (PMI).

7.2 The Weak Principle of Mathematical Induction

The PMI has a strong and a weak form. The weak form of PMI states that in order to prove that $\forall n Pn$ (*i.e.* that all numbers have the property P), it is sufficient to prove that (1) $P0$ and (2) for any n, if Pn, then $P(n+1)$ or, in symbols, $\forall n[Pn \Rightarrow P(n+1)]$. Giving an induction proof of a proposition, then, involves two steps: the *base case*, which establishes that $P0$; and the *induction step* which establishes that $\forall n[Pn \Rightarrow P(n+1)]$. This latter step requires showing that on the assumption that Pn, called the *induction hypothesis*, it follows that $P(n+1)$, where n is an arbitrary number. Notice that the induction step, on its own, does not establish that Pn for any n.

7.3 Induction Proof Example

As an illustration of weak induction, consider the following proposition:

Proposition 7.1 For any natural number n,

$$0 + 1 + \cdots + n = n(n+1)/2. \tag{1}$$

Proof. We define a property P such that Pn iff (1) holds for all n. Then, we seek to show that $\forall n Pn$ by weak induction.

Base case: For $n = 0$ both sides of (1) are 0. Thus we have that $P0$.

Induction step: Let n be any number such that Pn. The induction hypothesis is that equation (1) is true for n. Then we have that

$$\begin{aligned} 0 + 1 + \cdots + n + (n+1) &= n(n+1)/2 + (n+1) \\ &= (n+1)(n/2+1) \\ &= (n+1)(n+2)/2, \end{aligned}$$

where the first line follows from the induction hypothesis. This entails that $P(n+1)$. So we have shown that $Pn \Rightarrow P(n+1)$. This completes the proof. □

7.4 The Strong Principle of Mathematical Induction

The PMI also has a strong version, which states that if it is the case that for any n, if for any $m < n$ we have Pm, then Pn, then it follows that $\forall nPn$. In symbols we have that if $\forall n[\forall m < nPm \Rightarrow Pn]$, then $\forall nPn$. A proof by strong induction requires that it is proved that if for any n such that $\forall m < nPm$ holds, then Pn holds. In this case there is no base case, it is just necessary to show that Pn follows from from the assumption that $\forall m < nPm$. This assumption is called the *induction hypothesis*.

7.5 An Equivalence Theorem

We are now in a position to state the following theorem:

Theorem 7.2 The following three statements are equivalent:

1. The least number principle;
2. The weak principle of mathematical induction;
3. The strong principle of mathematical induction.

For a proof see Machover (1996).

7.6 The Principle of Induction

There is an extension of the principle of mathematical induction into general (*i.e.* possibly transfinite) well-ordered sets. This is the *principle of induction*:

Theorem 7.3 (Principle of Induction) Let $\langle X, < \rangle$ be a well-ordered set. Let P be a property that may hold for elements of X. Suppose that, for all $x \in X$, if for every element $y < x$ has property P, then x has property P. Then it follows that every element of X has property P.

We may see that the principle of induction is a generalization of the strong PMI discussed above.

8 Ordinals and Cardinals*

The extension of the notions of order and size that come from the natural numbers $\mathbb{N}$ into the domain of infinite sets involves the definition of ordinal and cardinal numbers, respectively.

8.1 Ordinals

Given a totally ordered set $\langle X, < \rangle$, and an element $a \in X$, we define the *section* X_a, to be the set of all elements of X that are less than a:

$$X_a = \{x \in X \mid x < a\}.$$

An *ordinal* is defined to be a well-ordered set $\langle X, < \rangle$ with the property that $X_a = a$ for all $a \in X$. Thus, each element of X is the set of all its predecessors.

The initial elements of the class of ordinals are the following:

$$\begin{aligned} 0 &= \varnothing \\ 1 &= \{\varnothing\} = \{0\} \\ 2 &= \{\varnothing, \{\varnothing\}\} = \{0, 1\} \\ 3 &= \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} = \{0, 1, 2\} \\ &\vdots \end{aligned}$$

In general we have that $n = \{0, 1, \ldots, n-1\}$, as we had defined n above. Thus, each natural number is an ordinal.

The axiom of infinity establishes that the class of all finite ordinals is a set, which is denoted by ω. The subsequent ordinal is $\omega \cup \{\omega\}$, and so on.

8.2 Supremum and Infimum

Let $\prec$ be a partial order on a class A and let $B \subseteq A$. If $u \in A$ and $x \prec u$ for all $x \in B$, then u is called an *upper bound of* (or *for*) B *with respect to* (*wrt*) $\prec$. If u is the least element of the class of upper bounds for B wrt $\prec$, *i.e.* if u is an upper bound for B and if $u \prec x$ for any upper bound x for B wrt $\prec$, then u is called the *least upper bound* (written lub) or *supremum* (written sup) for B wrt $\prec$. The notions of lower bound and greatest lower bound (glb), or infimum (inf), are defined similarly.

From now on we will use lower-case Greek letters, mainly 'α,' 'β,' 'γ,' 'λ,' 'ξ,' and 'η,' as variables ranging over the ordinals. We may now state the following theorem:

Theorem 8.1

1. If α is an ordinal, then so is $\alpha \cup \{\alpha\}$ (with $\beta < \alpha$ for all $\beta \in \alpha$). Moreover, $\alpha \cup \{\alpha\}$ is the supremum of α;
2. If X is a set of ordinals then its union set $\bigcup X$ is an ordinal. Moreover, $\bigcup X$ is the supremum of X.

8.3 Zero, Successor and Limit Ordinals

With this we may define the three main types of ordinals. The first is just zero, the least ordinal. The second type are successor ordinals. The *successor ordinal* of an ordinal α is $\alpha \cup \{\alpha\}$. We will write $\alpha + 1$ to denote the successor of α. The positive natural numbers are all successor ordinals. The third type consists of limit ordinals. A non-zero ordinal λ is a *limit ordinal* if it is the union of all its predecessors:

$$\lambda = \bigcup_{\alpha<\lambda} \alpha.$$

You may notice that a successor ordinal cannot be a limit ordinal. ω is the least limit ordinal.

8.4 Order-Isomorphism

Consider two ordered sets $\langle X, \prec\rangle$ and $\langle Y, \prec\rangle$. An *isomorphism of ordered sets* is a function f from X to Y that is both bijective and it preserves order, *i.e.* it has the property that for any $x_1, x_2 \in X$, if $x_1 \prec x_2$ then $f(x_1) \prec f(x_2)$. If there is an isomorphism between two ordered sets, then those sets are said to be *isomorphic*. If α and β are isomorphic ordinals, it can be seen that $\alpha = \beta$. An isomorphism of ordered sets is also called an *order-isomorphism*.

8.5 Representation Theorem for Well-ordered Sets

Theorem 8.2 (Representation Theorem for Well-ordered Sets) Any well-ordered set is isomorphic to a unique ordinal.

8.6 Least Ordinal Principle and Transfinite Induction

It is an important result that the class W of all ordinals is not a set. Nevertheless, if A is any non-empty class of ordinals, then A has a least member. This is called the *least ordinal principle*. Analogous to the case of PMI, the least ordinal principle (LOP) can be used to develop other forms of the principle of (transfinite) induction (PTI) defined in terms of ordinals.

Let P be a property of ordinals. Then we have the following theorems:

Theorem 8.3 (Strong Principle of Transfinite Induction) If A is a class of ordinals such that for every ordinal ξ

$$\eta \in A \text{ for every } \eta < \xi \Rightarrow \xi \in A,$$

then $A = W$.

Theorem 8.4 (Weak Principle of Transfinite Induction) If A is a class of ordinals satisfying the following three conditions

1. $\varnothing \in A$,
2. for every ordinal ξ, if $\xi \in A$ then so is its successor,
3. for every limit ordinal λ, if $\lambda \subseteq A$ then $\lambda \in A$,

then $A = W$.

8.7 Cardinals

The definition of cardinal numbers is given in terms of ordinals. Two sets X and Y are *equipollent* if there is a bijection from X to Y, and we write $X \approx Y$. The *cardinality* $|X|$ *of a set* X is defined to be the least ordinal α such that X is equipollent to α. Thus the cardinal numbers of finite sets are just the natural numbers.

The smallest infinite cardinal is ω. Considered as a cardinal number, the notation $\aleph_0$ is used to denote the smallest infinite cardinal. It is possible to define a (necessarily unique) function F such that $dom(f) = W$ and for every ordinal α,

$$F(\alpha) = \text{the least infinite cardinal not belonging to } im(F|_\alpha).$$

This enables us to define the infinite cardinals precisely to be

$$\aleph_\alpha =_{def} F(\alpha).$$

It follows that for any ordinals α and β, if $\alpha < \beta$ then $\aleph_\alpha < \aleph_\beta$.

8.8 Ordinal and Cardinal Arithmetic

The ordinal numbers and cardinal numbers each have their own arithmetic with addition, multiplication and exponentiation operations. The two arithmetics are quite different. We will not consider ordinal and cardinal arithmetic here, but we will note some if their properties or consequences.

The arithmetic of finite ordinals, is equivalent to the ordinary arithmetic of the natural numbers. Interestingly, iterations of the arithmetic operations involving countable ordinals never takes you out of the countable ordinal numbers. For instance ω, ω^ω, ω^{ω^ω}, *etc.* are all countable ordinals. Even the limit ordinal ϵ_0 of the series (*i.e.* the infinite sum) of these ordinals a countable ordinal. The set ω_1 of countable ordinals is the least uncountable ordinal. Hence $\aleph_1 = \omega_1$. This ordinal cannot be constructed from countable ordinals by the operations of ordinal arithmetic.

8.9 The (Generalized) Continuum Hypothesis

An important conjecture in this context is the continuum hypothesis. Note that in cardinal arithmetic, the cardinality of the exponential set Y^X of functions from X to Y has cardinality $|Y|^{|X|}$. Since the set-theoretic continuum can be defined in terms of the set of functions from $\mathbb{N}$ to $2 = \{0, 1\}$, *i.e.* sequences of zeros and ones, it follows from the construction that the cardinality $\mathfrak{c}$ of the continuum is $2^{\aleph_0}$. The *continuum hypothesis* (CH) then states that

$$\mathfrak{c} = \aleph_1$$

or that

$$\aleph_1 = 2^{\aleph_0}.$$

It is the statement that the cardinality of the continuum is the cardinality of the least uncountable cardinal.

This generalizes to the *generalized continuum hypothesis* (GCH), which states that

$$\aleph_{\alpha+1} = 2^{\aleph_\alpha}$$

for any ordinal α. This is essentially the statement that the power set operation yields the minimal increase in cardinality. It was shown by Gödel that GCH is consistent with ZFC set theory (see the end of this section for a brief mention of what ZFC is). Cohen subsequently showed that it was also independent of ZFC. Thus, ZFC offers no way to decide on the truth value of GCH. The same applies in the case of the axiom of choice (AC) (see section 9). Again, Gödel proved consistency and Cohen proved independence.

Both ordinal and cardinal arithmetic can establish propositions of the sort that state that countable unions of countable sets are countable. The basic properties of cardinal arithmetic establish that products and coproducts of infinite sets with the same cardinality preserve cardinality. Also products and coproducts of infinite sets with different cardinality have the cardinality of the set with the greatest cardinality. Note, however, that many of these proofs assume the axiom of choice (see section 9).

8.10 The Hierarchy of Sets

The class of all sets can be generated in a sequence of stages of construction, which can be indexed by ordinals. This is so since any well-ordered set is order-isomorphic to a unique ordinal. Let V_α be the set of all sets constructed at stage α. Then we define the V_α inductively in the following way:

$$\begin{aligned} V_0 &= \varnothing; \\ V_{\alpha+1} &= \mathcal{P}V_\alpha; \\ V_\lambda &= \bigcup_{\alpha<\lambda} V_\alpha \text{ for limit ordinals } \lambda. \end{aligned}$$

It follows from this that we have that

$$\begin{aligned} V_1 &= \{\varnothing\}; \\ V_2 &= \{\varnothing, \{\varnothing\}\}; \\ V_3 &= \{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}; \end{aligned}$$

V_4 is a set of 16 elements; *etc.*

It can be shown that this process generates all of the sets. This is to say that every set is contained in V_α for some ordinal α. Defining V to be the class of all sets, we have that

$$V = \bigcup_{\alpha \in W} V_\alpha.$$

Using transfinite induction it can be shown that $V_\alpha \subseteq V_\beta$ for $\alpha < \beta$. Also using induction, it can be shown that for any ordinal $\alpha \subseteq V_\alpha$, and hence that $\alpha \in V_{\alpha+1}$.

Intuitively, the hierarchy of sets can be pictured as a letter 'V', with V_0 at the bottom and each successive, and larger, V_α expanding above it. This picture motivates the notation V for the class of all sets.

8.11 A Mention of ZFC and NBGC

Naturally, the quasi-axiomatic presentation up to now is a rough sketch of the theory of sets and classes. This presentation of set theory should, however, be sufficient for doing mathematics on an informal level. Although, some care needs to be taken with respect to whether or not one is assuming the axiom of choice. We will consider issues concerned with choice in the following section. No attempt has been made to have anything anywhere approaching a rigorous presentation of (some small part of) set theory, there being many, many, many, of these in existence, and I do not endeavour to add to them.

What has been presented here is something like a hybrid of *Zermelo-Fraenkel set theory* (ZF), which is a theory of sets, *von-Neumann-Bernays-Gödel set theory* (NBG), which is a theory of sets and classes, and both of these systems with the axiom of choice, ZFC and NBGC respectively. Since ZFC and NBGC are (proof-theoretically) equivalent theories, there is nothing particularly misleading in this presentation. For ZFC, all of the definitions and statements that involve classes that are not proper classes will be well-defined or true in case we replace 'class' with 'set.' The notes also highlight some definitions and statements that are also true of proper classes and, consequently, give some indication of what is available descriptively in NBGC.

9 The Axiom of Choice*

9.1 Choice and Choice Functions

A *choice function* on a class $\mathcal{S}$ of sets is a function g with $dom(g) = \mathcal{S}$, such that $g(X) \in X$ for every $X \in \mathcal{S}$. The *axiom of choice* states that if $\mathcal{S}$ is a set of non-empty sets then there is a choice function on $\mathcal{S}$. This axiom is of a great degree of interest philosophically since it is a *prima facie* natural principle, yet it has some rather unusual and quite unexpected consequences.

9.2 Russell's Shoes and Socks

As an introduction to the subtleties introduced by the axiom of choice, we consider a riddle due to Bertrand Russell. Consider the following problem:

1. Suppose that a drawer contains infinitely many pairs of shoes. Construct a set containing one shoe from each pair.

2. Same question for pairs of socks.

The first part of this problem has an easy solutions, since we could simply take the left shoe from each pair. The situation in the second part, however, is more problematic since there is no general rule to select one sock from each pair as there is in part 1.

Turning back to consideration of choice functions, the question posed is asking if there is a choice function on the set of sets of pairs of socks that picks one sock from each pair, or whether the image of this function is a set. Put another way, the question is asking whether there is a set S containing one element from each of an infinite set of two-element sets. The problem is that there is no apparent way to construct such a set S. There is nothing contradictory about considering such a set to exist yet, interestingly, there is also nothing contradictory about considering it not to exist.

The existence of a choice function for any set of sets $\mathcal{S}$ is equivalent to the statement that the cartesian product of any set of non-empty sets is non-empty. In this connection, Russell called the axiom of choice the *multiplicative axiom*.

9.3 Consequences of the Axiom of Choice

Assuming that the axiom of choice holds in general has a variety of strange and sometimes paradoxical consequences. A favourite example of this among mathematicians is the *Banach-Tarski paradox*. If the axiom of choice is assumed, then it can be shown that it is possible to partition a unit sphere into a finite number of pieces, which can be reassembled (by rigid motions) to form two unit spheres, or a sphere of radius 2. Although this may seem to make the axiom of choice unacceptable, the proof of this result relies on the sphere being decomposed into *non-measurable sets*, which, intuitively, and restricting attention to subsets of $\mathbb{R}^3$, are sets with no definable length, area or volume. Thus, the sets involved in the construction are very peculiar indeed. But that such sets exist is guaranteed by the axiom of choice.

Working in the axiom of choice's favour is its ability to generate very nice general results in mathematics. One such result is that every vector space has a basis. Another is that the compactness theorem holds for propositional logic based on any set of propositional variables. That this follows from choice is, perhaps, connected with the 'spookiness' (or dubiousness?) of some of the arguments generated using compactness. Another important theorem is Tychonoff's theorem, which states that the product of a family of compact spaces is compact in the product topology.

Choice also enables the proof of following theorems:

1. Every set is equipollent to a unique cardinal number;
2. Every partial ordering of a set X can be extended to a linear ordering of X;
3. Every lattice with unit and at least one other element has a maximal ideal (Maximal Ideal Theorem for Lattices);
4. Every Boolean algebra has a prime ideal;
5. Every Boolean algebra has an ultrafilter;
6. Every Boolean algebra is isomorphic to a set algebra (Stone Representation Theorem);
7. Every subgroup of a free group is a free group (Nielsen-Schreier Theorem);
8. For any field F, and algebraic closure of F exists and is unique (up to isomorphism);
9. Every subspace of a separable metric space is separable.

9.4 Equivalents to the Axiom of Choice

Because choice is such a natural principle, it is quite easy to assume it without realizing it. The proofs of some of the propositions and theorems stated earlier in these notes require choice. But, because choice, is to some extent, a controversial axiom, one reason for which being its highly non-constructive character, it is wise to be aware of when it is being assumed. Accordingly, it is very useful to be familiar with other principles that are logically equivalent to it. Some statements that are equivalent to the axiom of choice are the following:

1. Every set is equipollent to an ordinal;
2. The class of cardinal numbers is well-ordered;
3. For every set X there is a well-ordering on X (Well-Ordering Theorem);
4. If $\mathcal{S}$ is a set of finite character, then for every $S \in \mathcal{S}$ there is an $M \in \mathcal{S}$ such that $S \subseteq M$ and M is maximal in $\mathcal{S}$ wrt $\subset_{\mathcal{S}}$ (Tukey-Teichmüller Lemma);
5. Let $\langle X, \prec \rangle$ be a partially ordered set and let $\mathcal{C}$ be the set of all chains in $\langle X, \prec \rangle$. Then every member of $\mathcal{C}$ is included in some member of $\mathcal{C}$ that is maximal wrt $\subset_{\mathcal{C}}$ (Hausdorff Maximality Principle);
6. Let $\langle X, \prec \rangle$ be a partially ordered set such that every chain in it has an upper bound in X. Then for each $x \in X$ there is some $u \in X$ such that u is maximal in X wrt $\prec$ and such that $x \preceq u$ (Zorn's Lemma);

7. The cartesian product of any family of non-empty sets is non-empty (Multiplicative Axiom);

8. Every lattice with a unit element and at least one other element contains a maximal ideal.

10 The Number Systems†

10.1 The Natural Numbers

With an appreciation of the fundamentals of set theory, we may now turn to consider the number systems fundamental to mathematics from a set theoretic point of view. We have already defined the set of natural numbers $\mathbb{N}$. In the context of number theory, the mathematical structure we are interested in is not simply $\mathbb{N}$, but it together with an order relation $<$ and two binary operations $+$ and $\cdot$ such that the Peano axioms are satisfied. We will consider the Peano axioms informally.

10.1.1 Peano Axioms

The first five axioms of Peano, which could be considered to be the axioms of order, are:

1. 0 is a natural number;
2. Every natural number m has a unique successor $s(m)$;
3. If m and n are distinct natural numbers, then $s(m) \neq s(n)$;
4. For every every natural number m, $s(m) \neq 0$;
5. The Principle of Mathematical Induction.

Clearly, the first five axioms are satisfied by establishing that ω is a set.

The remaining four axioms fix the properties of the $+$ and $\cdot$ operations. A binary operation on a set X is essentially a function from X^2 to X that is endowed with certain properties. The addition operation $+$ and the multiplication operation $\cdot$ are required to satisfy the four axioms for any m and n:

1. $m + 0 = m$;
2. $m + s(n) = s(m + n)$;
3. $m \cdot 0 = 0$;
4. $m \cdot s(n) = m \cdot n + m$.

10.1.2 A Mention of Models and Structures

Defining two binary operations $+$ and $\cdot$ on ω that satisfy the above and the natural ordering we can construct the 4-tuple $\langle \omega, <, +, \cdot \rangle$, which is a mathematical structure that satisfies all of the Peano axioms. This structure is a *model* of Peano arithmetic. The Peano axioms determine a unique algebra of propositions, *viz.*, the truths of arithmetic. Our structure $\langle \omega, <, +, \cdot \rangle$ is a concrete realization of this abstract algebra of propositions, in the sense that each member of the abstract algebra of Peano propositions is a true statement in our structure. This sort of concrete realization of an abstract algebra of propositions is called a *model* of that algebra of propositions.

Note that strictly speaking the claim that $\langle \omega, <, +, \cdot \rangle$ is a model of Peano arithmetic is false. In order for a mathematical structure to be a model of an algebra of propositions it must contain not only the relevant mathematical operations (which are functions in the abstract language), but also relations corresponding to the predicates of the language of the algebra of propositions. In general, though not universally, this includes the identity relation $=$, which is just the identity relation on the universe, which in this case is $\mathbb{N}$. In some cases this also includes certain designated individuals, considered to be 0-ary relations of the abstract language. In this case it is necessary to identify 0 with $\varnothing$ and 1 with $\{\varnothing\}$. In this informal presentation I am just overlooking these details and taking the equality and required relations to be implicit.

Note also that a model of Peano arithmetic is a 4-tuple $\langle S, <, +, \cdot \rangle$ consisting of a set S, an order relation on S satisfying the first five axioms, and two binary operations on S satisfying the last four axioms. Thus, our structure, $\langle \omega, <, +, \cdot \rangle$ is a (very) special case of a model of Peano arithmetic.

When there is no confusion, we will simply abbreviate $\langle \omega, <, +, \cdot \rangle$ and represent it by $\mathbb{N}$. It is important, however, to distinguish between, the bare set $\mathbb{N}$, the (explicitly) ordered version of it $\langle \omega, < \rangle$, which contains information about order relations between numbers, and the 4-tuple $\langle \omega, <, +, \cdot \rangle$, which has more structure associated with which propositions of arithmetic come out true. Intuitively ω has a strictly greater degree of structure than the bare set corresponding to it (take an equinumerous set with no natural ordering), $\langle \omega, < \rangle$ has a strictly greater degree of structure than ω, and $\langle \omega, <, +, \cdot \rangle$ has a strictly greater degree of structure than $\langle \omega, < \rangle$.

With this intuitive idea of the degree of structure possessed by $\mathbb{N}$, we can discuss an important property of $\mathbb{N}$ as a model of Peano arithmetic. Every model of Peano arithmetic is *structurally identical* or *isomorphic* to $\mathbb{N}$. This means that every collection of objects and relations that satisfies Peano's axioms will be isomorphic to $\mathbb{N}$. This property of the Peano axioms is called *categoricity*. The categoricity of the Peano axioms justifies identifying the finite ordinals with the natural numbers. This amounts to taking $\mathbb{N}$ to represent the equivalence class of models.

10.2 The Integers

The approach that we will take with the remaining number systems will be much less formal. Two important laws of the natural numbers is the *cancellation laws* for addition and multiplication (except by zero):

$$a + c = b + c \Rightarrow a = b;$$

$$ac = bc, c \neq 0 \Rightarrow a = b.$$

The difference $a - b$ between two numbers is only defined when $a + x = b$ has a solution for $x \in \mathbb{N}$, which occurs iff $a \leq b$. Thus subtraction can be defined only for certain (decomposable) numbers. In order to have subtraction defined everywhere, it is necessary to extend the natural numbers to include negative numbers, which involves the construction of the integers.

We may now see that we must define an operation $-$ that takes every pair of numbers $\langle a, b\rangle$ to their difference $x = a - b$. Each such ordered pair must correspond to a solution of the equation $a + x = b$. Different ordered pairs should determine the same difference, which means that the numbers should be represented as equivalence classes of ordered pairs. The equivalence relation should identify two ordered pairs $\langle a, b\rangle$ and $\langle c, d\rangle$ when their differences are equal, *i.e.* when $a - b = c - d$. This is equivalent to the proposition $a + d = b + c$, which is well-defined in $\mathbb{N}$.

Let $\sim$ be a relation on $\mathbb{N} \times \mathbb{N}$ that defined such that $\langle a, b\rangle = \langle c, d\rangle$ iff $a + d = b + c$. It is easy to see that this is an equivalence relation. We may then define the set of *integers* to be

$$\mathbb{Z} =_{def} (\mathbb{N} \times \mathbb{N})/\sim .$$

Let $[a, b]$ denote the equivalence class containing $\langle a, b\rangle$. Then we may define addition, multiplication and order on $\mathbb{Z}$ in the following way:

- $[a, b] + [c, d] = [a + c, b + c]$;
- $[a, b] \cdot [c, d] = [ac + bd, ad + bc]$;
- $[a, b] \leq [c, d]$ iff $a + d \leq b + c$.

To see what motivates these definitions, notice that $[a, b]$ is intended to represent the integer $a - b$. It is an exercise to show that these operations preserve equivalence, and that the usual arithmetic properties of the integers hold in $\mathbb{Z}$ (It is required to show that $\mathbb{Z}$ is a *commutative ring with identity* and has no *zero divisors*).

Notice that the map that takes a to $[a, 0]$ is an injection from $\mathbb{N}$ into $\mathbb{Z}$ that preserves order, addition and multiplication. This enables us to redefine $\mathbb{N}$ so that $\mathbb{N} \subset \mathbb{Z}$. This amounts to an *embedding* of $\mathbb{N}$ into $\mathbb{Z}$.

10.3 The Rational Numbers

To generate the rational numbers from the integers we must add solutions to the equations $ax = b$ with $a \neq 0$; and the equations $ax = b$ and $cx = d$ should have the same solution if $ad = bc$. Let $\sim$ be a relation on the set

$$\{\langle a, b\rangle \in \mathbb{Z}^2 \mid a \neq 0\}$$

by the requirement that $\langle a, b\rangle \sim \langle c, d\rangle$ iff $ad = bc$. We define the set of *rational numbers* to be

$$\mathbb{Q} =_{def} \mathbb{Z}^2/\sim .$$

Let $[a, b]$ denote the equivalence class containing $\langle a, b\rangle$. Then we may define addition, multiplication and order on $\mathbb{Z}$ in the following way:

- $[a, b] + [c, d] = [ad + bc, bd]$;
- $[a, b] \cdot [c, d] = [ac, bd]$;
- $[a, b] \leq [c, d]$ iff $abd^2 \leq b^2cd$.

It is an exercise to show that these operations preserve equivalence, and that the usual arithmetic properties of the rational numbers hold in $\mathbb{Q}$, which entails showing that $\mathbb{Q}$ is an *ordered field*. Notice that the map from $\mathbb{Z}$ to $\mathbb{Q}$ taking a to $[a, 1]$ is an embedding of $\mathbb{Z}$ into $\mathbb{Q}$.

10.4 The Real Numbers

The set of real numbers contains many gaps that can be approximated as closely as one likes but not reached by rational numbers. These, irrational, numbers are numbers like $\sqrt{2}$, π and e. The construction of the real numbers completes the rationals to include all of the irrational numbers and fill in the gaps.

The two standard ways of constructing the real numbers are by *equivalence classes of Cauchy sequences* or by *Dedekind cuts*. It is possible to sketch the latter construction. These constructions are necessarily more involved as a result of the fact that the move to the real numbers increases the cardinality.

10.4.1 Dedekind Cuts

The idea of Dedekind cuts is that each real number r partitions the set of rationals into two sets, $\{x \in \mathbb{Q} \mid x \leq r\}$ and $\{x \in \mathbb{Q} \mid x > r\}$. Each real number determines a unique cut, since between each two rationals there is a real number. The construction, then, identifies each real number with the corresponding cut.

A *Dedekind cut* is a partition of $\mathbb{Q}$ into two subsets L and R with the properties:

- Every element of L is smaller than every element of R;
- R has no least element.

Then we define the set of real numbers $\mathbb{R}$ to be the set of all Dedekind cuts, which we write as ordered pairs $\langle L, R\rangle$. The task is then to define order, addition and multiplication on $\mathbb{R}$. We will not examine this here.

10.4.2 Axioms for the Real Numbers

It can be shown from the resulting construction that $\mathbb{R}$ is a totally ordered field and satisfies the *principle of the supremum*:

> Every non-empty set of real numbers which has an upper bound has a supremum.

This is also called the *order-completeness* property. The axiomatic presentation of the theory of the real numbers characterizes the real numbers as an order-complete totally ordered field. It can be shown that these axioms are categorical. Thus, $\mathbb{R}$ is *the* order-complete totally ordered field.

10.5 The Complex Numbers

We conclude this presentation with a brief consideration of the complex numbers. The motivation for the extension of the reals to the complex numbers comes from the desire to solve equations like $x^2 + 1 = 0$. The solution to this equation is the imaginary unit $i = \sqrt{-1}$.

A general complex number $z = x + iy$, where $x, y \in \mathbb{R}$. x is called the *real part* of z and y the *imaginary part*. We can construct the set of complex numbers in such a way that $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ by identifying a complex number z with its real and imaginary parts. The addition and multiplication operations are then defined as

- $\langle a, b\rangle + \langle c, d\rangle = \langle a + c, b + d\rangle$;
- $\langle a, b\rangle \cdot \langle c, d\rangle = \langle ac - bd, ad + bc\rangle$.

It can then be shown that $\mathbb{C}$ is a field and that it is *algebraically closed*.

References

[1] Cameron, Peter J. *Sets, Logic and Categories.* New York: Springer, 2002.

[2] Jech, Thomas J. *The Axiom of Choice.* New York: American Elseivier Publishing Company, Inc., 1973.

[3] Lawvere, William F. and Robert Rosenbrugh. *Sets for Mathematics.* New York: CUP, 2003.

[4] Machover, Moché. *Set Theory, Logic and Their Limitations.* Cambridge: CUP, 1996.