# CATEGORY THEORY
# TOPIC 30 - FOUNDATIONS

PAUL L. BAILEY

Abstract. We now wish to begin constructing new categories, in the context of category theory itself, and this give an opportunity to review the limitations of set theory. In particular, we wish to make precise the concept of "the class of all sets". This document describes the problems this notion entails, lists and discusses the **ZFC** axioms of set theory, and then briefly lists the axioms of **NBG** which may be used to form a valid class theory.

## 1. Set Theory Paradoxes

As the usage of the vocabulary of sets became more frequent during the latter half of the 19th century, various problems occurred when claims were made about sets which were constructed without proper justification. These paradoxes led to vigorous attempts to lay a firmer foundation for set theory. We give two of the more famous paradoxes, both involving the impossibility of a set of all sets.

1.1. **Russell's Paradox.** Suppose that $V$ is a set which contains all sets. We note that $V$ is an element of itself, that is, $V \in V$; but this is not where we find the paradox.

Construct the set
$$W = \{A \in V \mid A \notin A\}.$$
Since $W$ is a set, we see that $W \in V$. Now either $W \in W$ or $W \notin W$.

But if $W \in W$, then the condition for being in $W$ is not met by $W$, so $W \notin W$, a contradiction. On the other hand, if $W \notin W$, then the condition for being in $W$ is met by $W$, so $W \in W$, also a contradiction. So something is wrong; either our method of forming the subset $W$ is not allowable, or $V$ itself does not exist.

1.2. **Power Sets.** Let $X$ be any set. We also wish to be able to form the set of all subsets of $X$, which we label $\mathcal{P}(X)$. The power set is bigger than the original set, in the sense that there is no surjective function from $X$ to $\mathcal{P}(X)$. To see this, let $f : X \to \mathcal{P}(X)$ be any function; we show that $f$ is not surjective.

Define the subset $A \subset X$ by
$$A = \{x \in X \mid x \notin f(x)\}.$$
Suppose that $f(a) = A$ for some $a \in X$. Then either $a \in A$ or $a \notin A$. If $a \in A$, then $a \notin f(a) = A$, a contradiction. If $a \notin A$, then $a \in f(a) = A$, also a contradiction. Thus it cannot be the case that $f(a) = A$ for any $a \in X$. Therefore, $f$ is not surjective.

However, if $X \subset Y$ is nonempty, then there exists a surjective function $Y \to X$ by mapping $y \to y$ if $y \in X$ and $y \to x_0$ otherwise, for some fixed $x_0 \in X$.

Now suppose that $V$ is a set which contains all sets. Then $\mathcal{P}(V) \subset V$, so we can build a surjective function $V \to \mathcal{P}(V)$. Herein lies the paradox of $V$.

*Date*: February 3, 2018.

## 2. ZFC Axioms

The *Zermelo-Fraenkel* axioms are intended to place set theory on a solid logical foundation. Together with the Axiom of Choice, these form the **ZFC** axioms of set theory, upon which the bulk of modern mathematics is based. It should be noted that there is some variation in the literature as to exactly which axioms to list; the list here is logically equivalent to other versions.

The primitive entity of **ZFC** is a *set.* If $a$ and $s$ are sets, the sentence $a \in s$ is defined, and is either true or false.

**Axiom 1** (Axiom of Extension)**.** *Two sets are equal if and only if they have the same elements.*

$$\forall A, \forall B \; : \; A = B \iff (\forall C \; : \; C \in A \Leftrightarrow C \in B)$$

**Axiom 2** (Axiom of the Existence)**.** *There is a set with no elements.*

$$\exists \varnothing, \forall x \; : \; \neg(x \in \varnothing)$$

**Axiom 3** (Axiom of Pairing)**.** *If $A$ and $B$ are sets, then there is a set containing $A$ and $B$ as its only elements.*

$$\forall A, \forall B, \exists C, \forall D \; : \; D \in C \iff (D = A \vee D = B)$$

**Axiom 4** (Axiom of Union)**.** *If $A$ is a set, there is a set whose elements are precisely the elements of the elements of $A$.*

$$\forall A, \exists B, \forall C \; : \; C \in B \iff (\exists D \; : \; C \in D \wedge D \in A)$$

**Axiom 5** (Axiom of Infinity)**.** *There is a set $S$ such that $\varnothing$ is in $S$ and whenever $A$ is in $S$, so is $A \cup \{A\}$.*

$$\exists S \; : \; \varnothing \in S \wedge (\forall A \; : \; A \in S \Rightarrow A \cup \{A\} \in S)$$

**Axiom 6** (Axiom of Powers)**.** *If $A$ is a set, there is a set whose elements are precisely the subsets of $A$.*

$$\forall A, \exists \mathcal{P}(A), \forall B \; : \; B \in \mathcal{P}(A) \iff (\forall C \; : \; C \in B \Rightarrow C \in A)$$

**Axiom 7** (Axiom of Regularity)**.** *If $A$ is a nonempty set, there is an element of $A$ which is disjoint from $A$.*

$$\forall A \; : \; \neg(A = \varnothing) \Rightarrow (\exists B \; : \; B \in A \wedge \neg(\exists C \; : \; C \in A \wedge C \in B))$$

**Axiom 8** (Axiom of Specification)**.** *Given any set $A$ and any proposition $p(x)$, there is a subset of $A$ containing precisely those $x$ for which $p(x)$ is true.*

$$\forall A, \exists B, \forall C \; : \; C \in B \iff C \in A \wedge p(C).$$

**Axiom 9** (Axiom of Replacement)**.** *Given any set $A$ and any proposition $p(x, y)$ where $p(x, y_1)$ and $p(x, y_2)$ implies $y_1 = y_2$, there is a set containing precisely those $y$ for which $p(x, y)$ is true for some $x$ in $A$.*

**Axiom 10** (Axiom of Choice)**.** *Given any set of nonempty sets, there is a set the contains exactly one element in each of the nonempty sets.*

## 3. Remarks regarding ZFC

The axioms dictate how new sets can be formed from existing sets. Let us take a moment to see how various well known examples play out in this context.

Axiom 1 tells us when two sets are equal. Axiom 2 declares the existence of the empty set. Axiom 5 declares the existence of an infinite set. Axiom 7 declares the nonexistence of certain sets. Axioms 3, 4, 6, 8, 9, and 10 supply the means of constructing new sets from existing ones.

3.1. **The Axiom of Extension.** Axiom 1 is the Axiom of Extension, which says that a set is completely determined by its elements. Thus, even though we may sometimes think of a set as a list of elements, we need to keep in mind that the order in which the elements are listed is not information included in the set; moreover, listing an element more than once has no effect on the set.

3.2. **The Axiom of Existence.** Axiom 2 is the Axiom of the Existence, also know as the Axiom of the Empty Set, which asserts that the empty set exists. Since a set is completely determined by the elements it contains, the empty set is unique.

The existence of the empty set can be deduced from Axiom 5 (Infinity) and Axiom 8 (Specification); however, since Axiom 5 appears to presuppose the existence of the empty set, we feel it is more readable to first include Axiom 2 (Existence).

We call this the Axiom of Existence because it is the only axiom that states that a set exists. Strictly speaking, all sets that are known to exist within **ZFC** are build from the empty set and the remaining axioms.

Axiom 5 (the Axiom of Infinity) also supplies an existing set, which happens to contain $\varnothing$. We note that the existence of $\varnothing$ actually follows from Axiom 8 (the Axiom of Specification) together with the existence of any set $X$, because $\varnothing = \{x \in X \mid x \neq x\}$.

3.3. **The Axiom of Specification.** Axiom 8 is the Axiom of Specification, also known as the Axiom of Separation, the Axiom of Comprehension, or the Axiom of Subsets. This tells us how we may build sets based on a logical proposition. It avoids Russell's Paradox by demanding that the set we build is a subset of an existing set.

3.4. **The Axioms of Union and Pairing.** A collection of sets is a set containing other sets; in formal set theory, all elements are sets, so a collection is just a set, but the word collection is still psychologically useful.

Given a collection $\mathcal{C}$ of sets, Axiom 4 (Union) states that the union of the sets in the collection exists. Combine this with Axiom 8 (Specification) to deduce that the intersection of the sets in the collection exists. We denote the union by $\cup\mathcal{C}$ and the intersection by $\cap\mathcal{C}$.

Now if we are given two sets, we can use Axiom 3 (Pairing) to form a collection whose elements are these two sets, and then we can take the union or intersection. So the union and intersection of two sets exist. We can repeat this to see that the union and intersection of finitely many sets exist.

3.5. **The Axiom of Regularity.** Axiom 7, the Axiom of Regularity, rules out certain constructions from being sets. In particular, one can conclude that there is no set $A$ such that $A = \{A\}$, relieving us from trying to understand the phrase "It's turtles all the way down".

3.6. **The Axiom of Replacement.** Axiom 9 is the Axiom of Replacement; it asserts that the image of a set under any definable function is also a set. The axiom of replacement was not part of Ernst Zermelo's 1908 axiomatization of set theory (**Z**). Its introduction by Adolf Fraenkel in 1922 is what makes modern set theory Zermelo-Fraenkel set theory (**ZF**).

3.7. **The Axiom of Choice.** Axiom 10 is the Axiom of Choice; it is the addition of this axiom that turns **ZF** into **ZFC**. In the presence of **ZF**, the Axiom of Choice is equivalent to each the following statements.

- The Cartesian product of any family of nonempty sets is nonempty.
- Every surjective function has a right inverse.
- Well-Ordering Theorem: Every set can be well-ordered.
- Tarkski's Thorem: For every infinite set $A$, there is a bijective map between $A$ and $A \times A$.
- Hausdorff Maximality Principle: In any partially ordered set, every totally ordered subset is contained in a maximal totally ordered subset.
- Zorn's Lemma: Any partially ordered set in which every chain (totally ordered subset) has an upper bound contains a maximal element.

## 4. Set Based Structures

We will use all of the set based structures which you have studied earlier. Any doubt as to the existence of such structures can be put to rest using the axioms of set theory. We list some of what will be used:

- Notions of subsets, the empty set, and power sets
- Set operations, including union, intersection, and complement
- Collections of sets and partitions
- Cartesian products of sets
- Functions and their properties
  - Injective, surjective, and bijective
  - Image and preimage
  - Identity map, inverse functions
- Relations and their properties
  - Reflexive, Symmetric, Antisymmetric, Transitive, and Definite
  - Partial orders and total orders
  - Equivalence relations
- Binary operations
  - identities and inverses
  - commutativity, associativity, and distributivity

## 5. Numbers

5.1. **The Natural Numbers.** The axioms are sufficient to declare the existence of the natural numbers. The natural numbers may be defined by setting 0 equal to the empty set given by Axiom 2 (Existence), and for each natural number $n$ previously defined, we define the *successor* of $n^+ = n \cup \{n\}$. Now Axiom 5 (Infinity) states that there exists a set containing 0 and all of its successors; call this set $S$. By Axioms 8 and 6 (Specification and Powers), the following set exists:

$$\mathcal{T} = \{T \in \mathcal{P}(S) \mid T \text{ contains } 0 \text{ and all of its successors}\}.$$

By the previous comment, the following set exists:

$$N = \cap\mathcal{T}.$$

This $N$ is a minimal set containing 0 and all of its successors; it is the set of natural numbers.

Let $\mathbb{N}$ denote the set of natural number. We define addition and multiplication on $\mathbb{N}$ using recursion, via a set-theoretic theorem known as the *recursion theorem.* We define order on $\mathbb{N}$ by inclusion: $m \leq n$ if an only if $m \subset n$.

5.2. **The Integers.** The integers are formed from the natural numbers using an equivalence relation on ordered pairs of integers. Specifically, we think of an ordered pair $(a, b)$ as the integer $a - b$, but since we don't have subtraction yet, we use addition to set up the integers.

Define a relation on the set $\mathbb{N} \times \mathbb{N}$ by

$$(a, b) \sim (c, d) \quad \Leftrightarrow \quad a + d = b + c.$$

This is an equivalence relation. Let $[a, b]$ denote the equivalence class of $(a, b)$. Define additional and multiplication by

$$[a, b] + [c, d] = [a + c, b + d] \quad \text{and} \quad [a, b] \cdot [c, d] = [ac + bd, ad + bc].$$

Define order on the set of equivalence classes by

$$[a, b] \leq [c, d] \quad \Leftrightarrow \quad a + d \leq b + c.$$

Let $\mathbb{Z}$ denote the set of integers, defined to be this set of equivalence classes. Finally, embed $\mathbb{N}$ in $\mathbb{Z}$ via $n \mapsto [n, 0]$, and identify $\mathbb{N}$ with its image under this map. It is convenient to set $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$.

5.3. **The Rational Numbers.** The rational numbers are formed similarly; we define an equivalence relation on $\mathbb{Z} \times \mathbb{Z}^*$ by

$$(a, b) \sim (c, d) \quad \Leftrightarrow \quad ad = bc.$$

Here $(a, b)$ represents the rational number $a/b$. Let $[a, b]$ denote the equivalence class of $(a, b)$. Define additional and multiplication by

$$[a, b] + [c, d] = [ad + bc, bd] \quad \text{and} \quad [a, b] \cdot [c, d] = [ac, bd].$$

Define order on the set of equivalence classes by

$$[a, b] \leq [c, d] \quad \Leftrightarrow \quad ad \leq bc.$$

Let $\mathbb{Q}$ denote the set of rational numbers, defined to be this set of equivalence classes. Embed $\mathbb{Z}$ into $\mathbb{Q}$ via $n \mapsto [n, 1]$.

5.4. **The Real Numbers.** The real numbers are the geometric completion of the rational numbers. They are designed to correspond to the set of all possible distances (forward and backward) on a line.

The need to formally define the real numbers was explored by Cantor and Dedekind at the end of the $19^{\text{th}}$ century. There are two common methods to make the leap from rationals to reals; we may use the *Dedekind completeness axiom*, which say that every subset of the real number which is bounded above has a least upper bound, or we may use the *Cauchy completeness axiom*, which says that a Cauchy sequence of real numbers converges inside the reals. We outline the first approach.

A *Dedekind cut* is a subset of the rational numbers, $C \subset \mathbb{Q}$, which is bounded above but has no maximum element, with the property that if $a, b \in \mathbb{Q}$ with $a < b$ and $b \in C$, then $a \in C$. These in effect are all the rational numbers less than a given (intended) real number. This allows us to find the "gaps" in the rational number line, where the irrational numbers may be found. Thus a cut corresponds to a rational number if its complement in $\mathbb{Q}$ contains a minimal element; otherwise, it corresponds to an irrational number. Addition and multiplication are defined setwise, with appropriate modifications to deal with negative numbers. Order is inclusion.

We let $\mathbb{R}$ equal the set of all Dedekind cuts.

5.5. **The Complex Numbers.** The complex numbers are the algebraic completion of the real numbers; this means that it is possible to solve any polynomial equation over the complex numbers. Complex number are appropriately viewed as ordered pairs of $\mathbb{R}$; that is, $\mathbb{C} = \mathbb{R}^2$, endowed with the additional structure of complex multiplication.

5.6. **The Cardinal Numbers.** Let $X$ and $Y$ be sets. We say that $X$ and $Y$ have the same *cardinality*, and write $|X| = |Y|$, if there exists a bijective function $X \to Y$. Moreover, we say that the cardinality of $X$ is less than or equal to that of $Y$, and write $|X| \leq |Y|$, if there exists an injective function $X \to Y$.

The famous *Schöeder Bernstein Theorem* states that if there exists an injective function $X \to Y$ and an injective function $Y \to X$, then there exists a bijective function $X \to Y$. Thus, if $|X| \leq |Y|$, and $|Y| \leq |X|$, we know that $|X| = |Y|$. This can be shown without the use of the Axiom of Choice.

However, suppose we do not assume the Axiom of Choice. Then it is impossible to show that there is either an injective function $X \to Y$ or an injective function $Y \to X$. Similarly, it is impossible to prove that there is either an injective function $X \to Y$ or a surjective function $X \to Y$. So, without the Axiom of Choice, we cannot say that either $|X| \leq |Y|$, or $|Y| \leq |X|$. The relative sizes of $X$ and $Y$ may be incomparable. However, with the axiom of choice, either $|X| \leq |Y|$ or $|Y| \leq |X|$.

Within **ZFC**, given a "universal" set $U$, it is possible to define the set of *cardinal numbers* in $U$ as follows. Let $\mathcal{U} = \mathcal{P}(U)$ be the power set of $U$. Consider the relation $\leftrightarrow$ on $\mathcal{U}$ given by

$$X \leftrightarrow Y \quad \Leftrightarrow \quad \exists \text{ bijective function } X \to Y.$$

This is clearly an equivalence relation on $\mathcal{U}$, so it forms a partition of $\mathcal{U}$. Let $\overline{X}$ denote the equivalence class of $X$, and let $\overline{\mathcal{U}}$ denote the set of equivalence classes. Then $X \leftrightarrow Y \quad \Leftrightarrow \quad \overline{X} = \overline{Y} \quad \Leftrightarrow \quad |X| = |Y|$. Now we *define* $|X|$ to be equal to $\overline{X}$, in which case we call $|X|$ the *cardinal number* of $X$. We may define a relation on $\overline{\mathcal{U}}$ by

$$\overline{X} \leq \overline{Y} \quad \Leftrightarrow \quad \exists \text{ injective function } X \to Y.$$

This relation is well-defined; moreover, it satisfies

- $\overline{X} \leq \overline{Y}$ or $\overline{Y} \leq \overline{X}$ (Definiteness)
- $\overline{X} \leq \overline{Y}$ and $\overline{Y} \leq \overline{X}$ implies $\overline{X} = \overline{Y}$ (Antisymmetry)
- $\overline{X} \leq \overline{Y}$ and $\overline{Y} \leq \overline{Z}$ implies $\overline{X} \leq \overline{Z}$ (Transitivity)

These properties say that $\leq$ is a *total order* relation on $\overline{\mathcal{U}}$. The totality property depends on the Axiom of Choice; without it, we have only a partial order.

## 6. Classes

It quickly became apparent that the lack of a universal set is a drawback of Zermelo's set theory. Various re-axiomatizations were proposed to remedy this. In articles published in 1925 and 1928, John von Neumann restated axioms adequate to develop set theory. This evolved to become what is now called *von Neumann-Bernays-Gödel set theory* (**NBG**), which is a conservative extension of **ZFC**.

The primitive entity of **NBG** is the *class*. A *set* is a class that may be an element of another class. Other classes are called *proper classes*.

Let $a$ and $s$ be two individuals. Then the atomic sentence $a \in s$ is defined if $a$ is a set and $s$ is a class. In other words, $a \in s$ is defined unless $a$ is a proper class.

We list the axioms of **NBG** which apply specifically to classes; the remaining axioms are equivalent to **ZFC** for the classes which are sets.

**Axiom 11. (Axiom of Class Extensionality)**
*Two classes are equal if and only if they have the same elements.*

**Axiom 12. (Axiom of Class Foundation)**
*Each nonempty class is disjoint from at least one of its elements.*

**Axiom 13. (Axiom of Class Comprehension)**
*Given any proposition $p(x)$, which is true or false for each set $x$, there is a class whose elements are precisely those sets $x$ for which $p(x)$ is true.*

**Axiom 14. (Axiom of Limitation of Size)**
*A class $C$ is a set if and only if there is no bijection between $C$ and the class $V$ of all sets.*

Note that the class $V$ of all sets exists by the Axiom of Class Comprehension, where $p(x) =$ "$x$ is a set".

Department of Mathematics, BASIS Scottsdale
*E-mail address*: pbailey@basised.com