Bare minimum: set theory

Abstract

A condensed overview of set theory. Bare minima are meant to give a rough overview, by no means complete, of the subject so that one at least knows what there is to know. It is mainly intended as background for those interested in participating in the Assumptions of Physics (https://assumptionsofphysics.org) project.

1   Introduction

Set theory is important for those working on Assumptions of Physics for at least three reasons. First, it is a foundational framework in mathematics. Second, it showcases a successful attempt at concept generalization. Third, it shows how a formal system is bootstrapped.

There are roughly two branches of set theory: naive and axiomatic. Naive set theory is built on intuitive concepts, and as such is not fully formalized and is open to potential paradoxes. Axiomatic set theory provides a fully formalized axiomatic system that aims to close those problems. We will start with naive set theory, which is the best setting for defining all the basic notions and getting a sense of how the framework works. Then we will turn our attention to axiomatic set theory to get a sense of what problems it solves and how, and whether those solutions mesh with foundational goals in physics.

For more details on set theory, see for example [1, 2].

2   Naive set theory

We take sets and elements as primitive objects, with no formal definition. This is the only difference between naive and axiomatic set theory, therefore all the formal definitions that follow apply to both. A **set** A is a collection of elements. If an element e belongs to the set we say e **is in** A, noted $e \in A$. A set A can be defined by listing the elements, noted $A = \{e_1, e_2, e_3\}$. A set A can be defined by specifying how to build the elements through a rule (i.e. predicate) $P(e)$, and is noted $A = \{e \mid P(e)\}$. The rule may include symbols like for all $\forall$, exists $\exists$, logical connectors AND $\wedge$, OR $\vee$ and NOT $\neg$.

Common sets: $\varnothing$, $\{a\}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$

Definition 1 (Common sets)**.** *The **empty set** $\varnothing$ is the set with no elements. A **singleton** is the set of a single element $\{a\}$. The **set of natural numbers** $\mathbb{N} = \{0, 1, 2, ...\}$. The **set of positive natural numbers** $\mathbb{N}^+ = \{n \in \mathbb{N} \mid n > 0\}$. The **set of integer numbers** $\mathbb{Z} = \{..., -2, -1, 0, +1, +2, ...\}$. The **set of rational numbers** $\mathbb{Q}$. The **set of real numbers** $\mathbb{R}$. The **set of complex numbers** $\mathbb{C}$. A **universe** U is the set that, within a specific context, contains all entities under study.*

2.1   Basic definitions

Subset, superset: $\subseteq$, $\supseteq$

Definition 2 (Set relationships)**.** *Let A and B be two sets. If all elements in A are also contained in B, we say A is a **subset** of B, noted $A \subseteq B$. In this case, we also say B is a **superset** of A, noted $B \supseteq A$. Additionally, if B contains elements that A does not contain, we say A is a **strict subset** of B, noted $A \subset B$, and B*

$A \subset B$
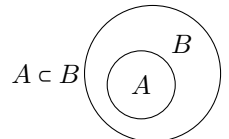
is a **strict superset** of A, noted $B \supset A$. If the two sets contain exactly the same elements, then we say the sets are **equal**, noted $A = B$. If the two sets have no element in common they are said to be **disjoint**.

Power set:
$\mathcal{P}(A)$, 2^A

Definition 3 (Power set). *Given a set A, its **power set**, noted $\mathcal{P}(A)$ or 2^A, is the set of all subsets of A. That is, $\mathcal{P}(A) = \{B \mid B \subseteq A\}$.*

Set operations:
$\cup$, $\cap$, $\smallsetminus$, $^{\complement}$

Definition 4 (Set operations). *We define the following operations between two sets A and B:*

Union. *Noted $A \cup B$, the union of A and B is the set of all elements that are contained by A, B or both.*
Intersection. *Noted $A \cap B$, the intersection of A and B is the set of all elements that are contained by both.*
Set difference (or relative complement). *Noted $A \smallsetminus B$, A minus B is the set of all elements in A that are not contained by B.*
Complement (or absolute complement). *Noted $A^{\complement}$, represents all elements that are not contained in A. Note that this operation is context specific: we need to know that, within a certain context, U is the set of all elements under study; then the complement of $A \subseteq U$ is $A^{\complement} = U \smallsetminus A$.*

Let $\mathcal{A} \subseteq \mathcal{P}(A)$ be a set of subsets of A. The operations of union $\bigcup\limits_{A \in \mathcal{A}} A$ and intersection $\bigcap\limits_{A \in \mathcal{A}} A$ can be extended to span all elements.

Ordered pair:
(a, b)

Definition 5 (Ordered pair). *Given two elements a and b, the **ordered pair** (a, b) specifies two objects in that order. That is $(a, b) \neq (b, a)$. This can be constructed by setting $(a, b) = \{\{a\}, \{a, b\}\}$.*

Cartesian
product: ×

Definition 6 (Cartesian product). *Given two sets A and B, the **Cartesian product** $A \times B = \{(a, b) \mid \forall a \in A, b \in B\}$ is the set of all possible ordered pairs between the elements of A and B.*

Definition 7 (Disjoint union). *Let A and B be two sets. The disjoint union $A \sqcup B$ is the union where the elements of the sets are always treated as distinct. That is, given $A \sqcup B = (\{0\} \times A) \cup (\{1\} \times B)$.*


$A \times B$

Remark. Ordered pairs and disjoint unions exemplify the effort in set theory to construct *all objects* starting from just two primitives: set and set inclusion. The downside is that the constructions do not map exactly to our intuition. For example, we still think as an ordered pair as a set, so we want to write $b \in (a, b)$, though, looking at the definition b itself is not directly a member of (a, b).

2.2 Relations and functions

Remark. Relations provide a common foundation for functions (i.e. the relation is the set of pairs $(x, f(x))$), equivalences and order (i.e. the relation is the set of pairs (a, b) such that $a = b$ or $a \leq b$ respectively). This exemplifies the effort in set theory to find general structures.

Binary
relation: aRb

Definition 8 (Binary relation). *Given a set A, called **domain**, and a set B, called **codomain**, a **binary relation** is a set $R \subseteq A \times B$ of ordered pairs. We say a is R-**related** to b, noted aRb if $(a, b) \in R$. If $A = B$ (i.e. domain and codomain coincide) the relation is said to be **homogeneous**, **heterogeneous** if not.*

R

Image and
preimage

Definition 9 (Image and preimage). *If $X \subseteq A$, the **image of** X **under** R is the set of all elements in B that are related to at least one element in X. The **image of** R is the set of all elements in B that are related to at least one element in A (i.e. the image of the full set A).*

R^{-1}

2
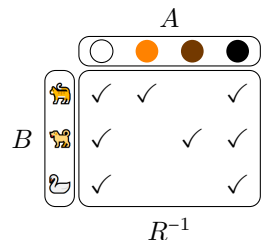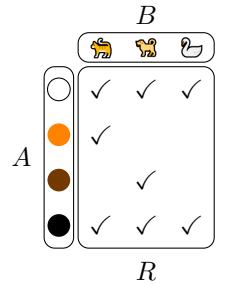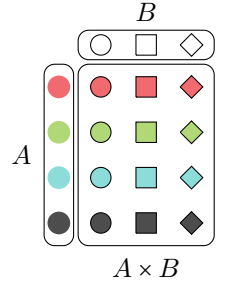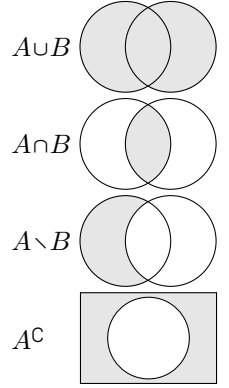
*If $Y \subseteq B$, the **preimage of** Y **under** R is the set of all elements in A that are related to at least one element in Y. The **preimage of** R is the set of all elements in A that are related to at least one element in B (i.e. the preimage of the full set B).*

Converse: R^{-1}

Definition 10 (Converse relationship). *Given a relationship R between A and B, the **converse** or **inverse** relationship $R^{-1} \subseteq B \times A$ is the relationship where the order of the pairs (i.e. domain and codomain) is inverted. That is, $R^{-1} = \{(b, a) \mid (a, b) \in R\}$.*

Relation properties

Definition 11 (Relation properties). *Let R be a binary relation between A and B. We define the following properties:*

Functional or right-unique. *An element $a \in A$ is related to at most one element $b \in B$. That is, if aRb and aRc then $b = c$.*
Injective or left-unique. *An element $b \in B$ is related to at most one element $a \in A$. That is, if aRb and cRb then $a = c$.*
Serial or left-total. *Every element $a \in A$ is related to at least one element $b \in B$. That is, for each $a \in A$ there exists at least one $b \in B$ such that aRb. In this case, the preimage of R is the whole A.*
Surjective or right-total. *Every element $b \in B$ is related to at least one element $a \in A$. That is, for each $b \in B$ there exists at least one $a \in A$ such that aRb. In this case, the image of R is the whole B.*

From those basic properties, we define the following:

One-to-one. *Injective and functional (i.e. left-unique and right-unique).*
One-to-many. *Injective and not functional (i.e. left-unique and not right-unique).*
Many-to-one. *Not injective and functional (i.e. not left-unique and right-unique).*
Many-to-many. *Not injective nor functional (i.e. not left-unique nor right-unique).*

(Partial) functions: $f : A \to B$

Definition 12 (Function). *A **partial function** $f : A \to B$ is a binary relationship between A and B that is functional (i.e. right-unique). The **graph** of the function $G \subset A \times B$ is the function expressed as ordered pairs. A **total function**, or simply a **function**, is a partial function that is also serial (i.e. left-total, defined on the whole domain). A function is **bijective** if it is injective and surjective.*

When the functional form $f(a)$ is known, this can be expressed as $a \mapsto f(a)$ (read "a maps to f of a").

Remark. Note that the domain and the codomain can be any set, including scalar products. Therefore $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for which $(x, y) \mapsto \sqrt{x^2 + y^2}$ gives us the Euclidean norm of a two dimensional vector.

Identity: Id_A

Definition 13 (Identity function). *Given a set A, the identity function $\mathrm{Id}_A : A \to A$ is the function such that $\mathrm{Id}_A(a) = a$ for all $a \in A$.*

Inverse: f^{-1}

Proposition 14 (Inverse function). *Let $f : A \to B$ be a function. The corresponding converse relationship f^{-1} is a function if and only if f is bijective. In this case $f^{-1} : B \to A$ is also bijective and is called the **inverse** of f.*

$f(X), f^{-1}(Y)$

Definition 15 (Image/preimage functions). *Let $f : A \to B$ be a function. This induces a relation $f : \mathcal{P}(A) \to \mathcal{P}(B)$ that associates each subset of A to its image and a relation $f^{-1} : \mathcal{P}(B) \to \mathcal{P}(A)$ that associates each subset of B with its preimage.*

Proposition 16 (Properties of image/preimage functions). *The induced relations $f : \mathcal{P}(A) \to \mathcal{P}(B)$ and $f^{-1} : \mathcal{P}(B) \to \mathcal{P}(A)$ are functions and satisfy the following properties for all $\mathcal{A} \subseteq \mathcal{P}(A)$ and $\mathcal{B} \subseteq \mathcal{P}(B)$:*

1. $f\left(\bigcup_{U \in \mathcal{A}} U \right) = \bigcup_{U \in \mathcal{A}} f(U)$

$2.\ f^{-1}(\bigcup_{V \in \mathcal{B}} V) = \bigcup_{V \in \mathcal{B}} f^{-1}(V)$

$3.\ f^{-1}(\bigcap_{V \in \mathcal{B}} V) = \bigcap_{V \in \mathcal{B}} f^{-1}(V)$

Remark. With abuse of notation, f and f^{-1} indicate both the function between elements and sets. Note that $f^{-1} : \mathcal{P}(B) \to \mathcal{P}(A)$ always exists even if f is not bijective. Also note that the image of the intersection is not in general the intersection of the image, and how maps that preserve set theoretic structures (e.g. topologies and σ-algebras) are defined based on preimages.

Fiber

Definition 17 (Fiber)**.** *Given a function $f : A \to B$, the **fiber of** $b \in B$ **under** f is the subset of A of where the function takes the value b. That is, the preimage $f^{-1}(\{b\})$ of the singleton. Fibers over real functions are also called **level sets**.*

Restriction: $f|_C$

Definition 18 (Restriction)**.** *Let $f : A \to B$ be a function and let $C \subseteq A$. Then the **restriction of** f **over** C is the function $f|_C : C \to B$ such that $f|_C(c) = f(c)$ for all $c \in C$.*

Function and relation composition: $R \circ S$, $f \circ g$

Definition 19 (Composition)**.** *Let A, B and C be three sets. Let R be a binary relationship between A and B and S be a binary relationship between B and C. Then the **composition of** R **and** S is the binary relationship $S \circ R \subseteq A \times C$ between A and C such that $(a, c) \in S \circ R$ if we can find $b \in B$ such that aRb and bSc.*

Remark. The order of composition follows those of functions: if $f : A \to B$ and $g : B \to C$, composition leads to $c = g(f(a)) = (g \circ f)(a)$.

Composition properties

Proposition 20 (Composition properties)**.** *Relation composition follows the following properties:*

- *Composition is associative: $R \circ (T \circ S) = (R \circ T) \circ S$*
- *Converse of composition: $(R \circ S)^{-1} = S^{-1} \circ R^{-1}$*
- *Composition of functional/injective/serial/surjective relations is functional/injective/serial/surjective.*
- *If $f : A \to B$ and $g : B \to C$ are (partial) functions, then $g \circ f$ is a (partial) function.*
- *Let $f : A \to B$ be a bijective function, then $f \circ f^{-1} = \mathrm{Id}_B$ and $f^{-1} \circ f = \mathrm{Id}_A$.*

Sets of functions: B^A

Definition 21 (Sets of functions)**.** *The set of all possible functions $f : A \to B$ is noted B^A.*

Proposition 22. *Let $2 = \{0, 1\}$ denote a set with two elements. Then $\mathcal{P}(A)$ and 2^A are in one-to-one correspondence.*

Remark. The notation for function sets is due to the cardinality. If A has n elements and B has m elements, each function is given by choosing an element of B for each element of A, that is m choices n times, that is m^n. The power set is noted 2^A because each subset $U \subseteq A$ can be identified by a unique function $f : A \to \{0, 1\}$, where for all $a \in A$ then $f(a) = 0$ corresponds to $a \notin U$ and $f(a) = 1$ corresponds to $a \in U$.

Definition 23 (Set closure under operations)**.** *Let $f : A_1 \times A_2 \times ... \times A_n \to B$ be an n-ary function. We say a set A **is closed under** f if whenever the function is applied to elements of A it returns an element of A. That is, $A \subseteq A_i$ for all $1 \le i \le n$, $A \subseteq B$ and $f(A, A, ..., A) \subseteq A$.*

Remark. For example, integers are closed under addition, subtraction and multiplication but are not closed under division. Most mathematical structures (groups, topologies, σ-algebras, vector spaces, ...) define sets closed under some operations.

2.3   Homogeneous relations, equivalences and orders

Remark. We will cover the notions of orders that are necessary for the definition of ordinal and cardinal numbers. Orders will be treated more in detail in a Bare Minimum dedicated to order theory.

Reflexivity, symmetry, transitivity, ...

Definition 24 (Homogeneous relation properties). *Let $R \subseteq A \times A$ be a homogeneous binary relation between A and itself. We define the following properties:*

Reflexive. *Every element $a \in A$ is related to itself. That is, aRa for all $a \in A$.*
Irreflexive. *No element $a \in A$ is related to itself. That is, aRa for no $a \in A$.*
Symmetric. *If a is related to b, then b is related to a. That is, if aRb then bRa.*
Antisymmetric. *If a is related to b and b is related to a, then a and b are the same element. That is, if aRb and bRa then $a = b$.*
Transitive. *If a is related to b and b is related to c, then a is related to c. That is, if aRb and bRc then aRc.*
Connected. *Given two distinct elements, one is related to the other. That is, given $a \neq b$, either aRb or bRa.*

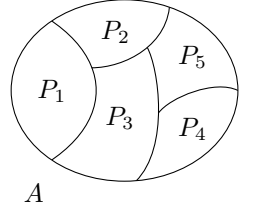Order, equivalence: $\leq$, $<$, $\equiv$

Definition 25 (Orders and equivalence). *A **partial order** (noted for example $\leq$, $\lesssim$, $\preceq$) is a homogeneous binary relationship that is reflexive, antisymmetric and transitive. A **linear order** or **total order** is an order that is connected. An order is **strict** (noted for example $<$, $\prec$) if it is irreflexive instead of reflexive. A **strict linear order** or **strict total order** is an order that is strict and connected.*

*An **equivalence relation** (noted for example $\equiv$, $\sim$) is a homogeneous binary relationship that is reflexive, symmetric and transitive.*

Partitions

Definition 26 (Partitions). *A **partition** of a set A is a collection $P \subset \mathcal{P}(A)$ of disjoint subsets that cover all of A. That is:*

1. $\bigcup_{B \in P} B = A$
2. $B \cap C = \varnothing$ *for all $B, C \in P$ where $B \neq C$*

Eq. classes and quotient set: $[a]_\sim, a_{/\sim}, A_{/\sim}$

Definition 27 (Equivalence classes and quotient set). *Let A be a set and $\sim$ an equivalence relation on A. Let $a \in A$. The **equivalence class of** a **by** $\sim$ is the set $[a]_\sim = a_{/\sim} = \{b \in A \mid b \sim a\}$ of all the elements of A equivalent to a. The **quotient set of** A **by** $\sim$ $A_{/\sim} = \{a_{/\sim} \mid a \in A\}$ is the set of all of the equivalence classes by $\sim$.*

Proposition 28 (Quotient sets are partitions). *Every quotient set is a partition, and every partition is a quotient set. That is, let A be a set and $\sim$ an equivalence class, then $A_{/\sim}$ is a partition. Conversely, let P be a partition of A, then $\sim = \{(a, b) \mid a, b \in X$ for some $X \in P\}$ is an equivalence relation and $A_{/\sim} = P$.*

Definition 29. *Given a function $f : A \to B$ the **equivalence relation determined by** f is the relation $R_f = \{(a_1, a_2) \mid f(a_1) = f(a_2)\}$.*

Proposition 30 (Function canonical decomposition). *Every function $f : A \to B$ can be written as the composition $f = t \circ s \circ r$ of the following three functions:*

$r : A \to A_{/R_f}$ *defined as $a \mapsto [a]_{R_f}$; this function is surjective*
$s : A_{/R_f} \to f(A)$ *defined as $[a]_{R_f} \mapsto f(a)$; this function is bijective*
$t : f(A) \to B$ *defined as $t(b) = b$; this function is injective.*

Remark. All the above concepts are of general use in all areas of mathematics. Note that they are all constructed on two primitives: the notion of sets and of set inclusion.

2.4 Indexed families and sequences

Families, sequences:
$\{x_i\}_{i \in I}$,
$\{x_i\}_{i=1}^{\infty}$

Definition 31 (Families). *An **indexed family**, or simply a **family**, $\{x_i\}_{i \in I}$ is a collection of elements identified by an **index set** I. More formally, an indexed family is a tuple (X, I, x) where X and I are sets and $x : I \to X$ is a function such that $i \mapsto x_i$. A **sequence** is a family for which the index set is a contiguous interval of natural numbers. A **finite sequence** of n elements is noted $\{x_i\}_{i=1}^{n}$ while an **infinite sequence** is noted $\{x_i\}_{i=1}^{\infty}$.*

$\bigcap\limits_{i \in I} A_i$, $\bigcup\limits_{i \in I} A_i$

Definition 32 (Operations on families of sets). *Let $\{A_i\}_{i \in I}$ be a family of sets, the operations of intersection $\bigcap\limits_{i \in I} A_i$ and union $\bigcup\limits_{i \in I} A_i$ can be extended to span all members.*

Set limit:
$\lim\limits_{i \to \infty} A_i = A$

Definition 33 (Set-theoretic limit). *Let $\{A_i\}_{i=1}^{\infty}$ be a sequence of sets. We define the **limit infimum** as $\liminf\limits_{i \to \infty} A_i = \bigcup\limits_{i \geq 1} \bigcap\limits_{j \geq i} A_j$ and the **limit supremum** as $\limsup\limits_{i \to \infty} A_i = \bigcap\limits_{i \geq 1} \bigcup\limits_{j \geq i} A_j$. If the two are equal to the same set then we say the sequence **converges**, the **set-theoretic limit** $\lim\limits_{i \to \infty} A_i = A$ exists and it is equal to that set.*

Monotone sequence:
$A_1 \subseteq A_2 \subseteq ...$
$A_1 \supseteq A_2 \supseteq ...$

Definition 34 (Monotone sequence). *A **monotone sequence** is a sequence of sets $\{A_i\}_{i=1}^{\infty}$ where each set is a subset (or superset) of the preceding. More specifically, we distinguish four cases:*

Increasing *when $A_i \subseteq A_{i+1}$ for all $i \geq 1$*
Decreasing *when $A_i \supseteq A_{i+1}$ for all $i \geq 1$*
Strictly increasing *when $A_i \subset A_{i+1}$ for all $i \geq 1$*
Strictly decreasing *when $A_i \supset A_{i+1}$ for all $i \geq 1$*

$A_i \supset A_{i+1}$

Monotone convergence

Proposition 35 (Monotone convergence). *All monotone sequences converge. For increasing sequences the limit is given by $\lim\limits_{i \to \infty} A_i = \bigcup\limits_{i \geq 1} A_i$, while for decreasing sequences the limit is given by $\lim\limits_{i \to \infty} A_i = \bigcap\limits_{i \geq 1} A_i$.*
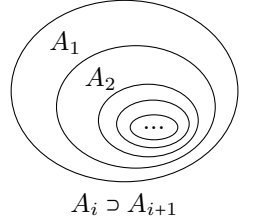
3 Axiomatic set theory

Axiomatic set theory was introduced to eliminate some technical problems that would make naive set theory inconsistent. There is no single approach and there is no single set of axioms. For example, constructive set theory, as all constructive mathematics, rejects "nonconstructive proofs" in which the existence of an object is proved with no explicit algorithm as to how to construct it. This will lead to different choices than the ones outlined here. Much of the work on the foundations of mathematics in set theory is to clarify the relationship between the axioms: which ones are required to get certain results, which ones are independent and which ones are inconsistent.

This is an important insight for those working on Assumptions of Physics: the foundations of mathematics are not absolute. We have to be conscious of the consequences of those choices and understand whether they are suitable for the foundations of physics.

3.1 Paradoxes

While they are generally called paradoxes, the following are really contradictions and that is why axiomatic set theory needs to avoid them. The first type of paradoxes for naive set theory arises from the use of natural language and the meaning of the words, typically when this is self-referential.

Berry's paradox. Consider the expression "the smallest positive integer not definable in under sixty letters." Given that there are only twenty-six letters in the

English language, there are finitely many phrases under sixty letters and therefore only finitely many integers can be defined in under sixty letters. Therefore there are integers not definable under sixty letters and, since the integers are ordered, there will be a least one. However, the above expression contains fifty-seven letters. We have a contradiction.

This type of paradoxes is solved by restricting to symbolic expressions of a formal language, such as predicate logic. The consequence is that, at the foundational level, mathematics does not keep track of what the symbols mean.

The second type of paradoxes is purely logical, so they can be constructed in the formal system.

Russell's paradox. Let R be the set of all sets that do not contain themselves. That is, $R = \{x \mid x \notin x\}$. Suppose $R \notin R$ then, by definition, $R \in R$. Suppose $R \in R$ then, by definition, $R \notin R$. We have a contradiction.

These are solved by carefully specifying the rules that are allowed for the creation of sets, by picking the "right" axioms. Part of the solution revolves around the following concepts.

3.2   Classes, sets and elements

A **class** is any collection, therefore **sets are classes**. Some classes are not sets. These are called **proper classes** and are exactly the collections that cannot be elements of a class. Therefore, sets are classes that can be contained by other classes, while proper classes cannot. The difference between sets and proper classes is not in what they contain but in whether they can be contained. Proper classes are "too large" to be sets in the following sense: given a class A, if we can find a set B that contains all elements of A, then A will be a set as well. Therefore proper classes can never be subclasses of sets, they are bigger than all sets. The class of all sets, for example, is a proper class. So is the class of all cardinal numbers.

Note that axiomatic set theory does not distinguish between objects that can or cannot contain other objects. All elements are classes, all elements can, in principle, contain something. In fact, since equality of two objects coincides on whether they have the same elements, two objects must contain something different to be different. Numbers will be constructed as sets starting from the empty set. As a consequence, the proposition $3 \in 4$ is valid, and happens to be true if 3 and 4 are naturals, though it is not true if 3 and 4 are reals. From the perspective of Assumptions of Physics, this is not great as the framework allows propositions that are not physically meaningful. Radically different approaches to the foundations of mathematics (e.g. through category theory) do not exhibit this particular problem, though they exhibit others (e.g. identity of an object can be defined only up to its relationships with other objects).
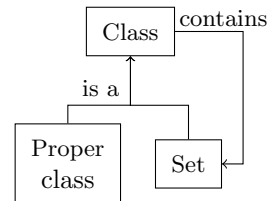
3.3   Axiomatization

One strategy for the axiomatization of set theory, refered to as ZermeloFraenkel (ZFC), only defines sets. The C in ZFC stands for the addition of the axiom of choice. It avoids paradoxes by restricting set construction to subsets only. Proper classes are not objects that live in the framework.

Another strategy, refered to as von NeumannBernaysGdel (NBG), starts by defining classes and then defines sets as those classes that can be elements of classes. It avoids paradoxes by restricting set construction on predicates over sets. Since a proper class cannot be an element, if the definition creates a contradiction for an item x (i.e. we have both $x \in S$ and $x \notin S$) then it must be that the element x is a proper class.

The two approaches are proved to be equivalent.

Set theory axioms   **Axioms of set theory**. Here are the axioms of a particular formulation of axiomatic set theory taken from [1] that follows NBG:

1. Axiom of extent. If two classes have the same elements, then they are identical.
2. Axiom of class construction. Let $P(x)$ be a predicate expressed in terms of the symbols $\in, \vee, \wedge, \neg, \Rightarrow, \exists, \forall$, brackets, variables $x, y, z, A, B, ...$ Then there exists a class C that consists of all the elements (i.e. sets) x which satisfy $P(x)$.
3. Every subclass of a set is a set.
4. $\varnothing$ is a set.
5. If a and b are sets, then $\{a, b\}$ is a set.
6. If $\mathcal{A}$ is a set of sets, then, $\bigcup\limits_{A \in \mathcal{A}} A$ is a set.
7. If A is a set, the power set $\mathcal{P}(A)$ is a set.
8. Axiom of foundation. If $A \neq \varnothing$ is a set, there exists an element $a \in A$ such that $a \cap A = \varnothing$.
9. Axiom of replacement. If A is a set and $f : A \to B$ is a surjective function, then B is a set.
10. Axiom of choice. Every set has a choice function. That is, given a set A, there exists a function $f : \mathcal{P}(A) \smallsetminus \{\varnothing\} \to A$ such that for all non-empty $B \subseteq A$ we have $f(B) \in B$.
11. Axiom of infinity. There exists a successor set (as defined in Definition 37).

Remark. Axiom 1 serves to define class equality. Axiom 2 defines how to define classes based on predicates. Axioms 3-7 tell us that some classes, constructable by axiom 2, are sets. Axiom 8 guarantees that we cannot have infinite sequences $\{A_i\}_{i=1}^{\infty}$ such that $A_{i+1} \in A_i$. In particular, a set cannot be an element of itself. Axiom 9 tells us that any set that is "smaller" than a set (i.e. we can map another set onto it) is a set. Axiom 10, the axiom of choice, guarantees that we can always pick an element from a set, though it does not tell us how. Axiom 11 tells us that there are sets with infinite elements. Once the natural numbers are assumed to exist, higher cardinality of infinity can be constructed through, for example, the power set.

The axiom of choice has been the focus of particular investigation as it has been regarded as controversial for some of its consequences. While most mainstream mathematicians accept it, constructive mathematics rejects the axiom of choice as it does not lead to constructive proofs: it claims some objects exist even though they may not even be expressible or computable. It also implies the law of excluded middle (i.e. a proposition is either true or not), which constructive mathematics rejects. What does or does not require the axiom of choice is therefore ample subject of study.

Proposition 36. *The following statements are equivalent:*

- *Axiom of choice: every set has a choice function*
- *Well-ordering theorem: every set can be well ordered (defined below)*
- *Trichotomy: given two sets, either they have equal cardinality, or one has lower cardinality than the other (i.e. $|A| < |B|$, $|A| = |B|$ or $|A| > |B|$)*
- *Zorn's lemma: every non-empty poset (i.e. partially ordered set) for which every chain (i.e. linearly ordered subset) has an upper bound contains at least one maximal element*
- *Every vector space has a basis*

4   Numeric constructions

In set theory, numbers are defined to be sets with particular properties. These constructions are instructive to gain insights into the foundations of mathematics. Note that, for Assumptions of Physics, we found it more productive to understand numbers as an order-theoretic concept (i.e. the integer numbers can be thought

of as a linearly ordered set where each element has an immediate successor and an immediate predecessor).

4.1 Standard numbers

Definition 37. *Let A be a set. We define the **successor** $A^+ = A \cup \{A\}$. A set S is called a **successor set** if it contains the empty set (i.e. $\varnothing \in S$ and if contains the successor of every element (i.e. if $X \in S$ then $X^+ \in S$). The set of **natural numbers** $\mathbb{N}$ is the intersection of all successor sets. The set of **positive naturals** is $\mathbb{N}^+ = \mathbb{N} \smallsetminus \{\varnothing\}$.*

Remark. The set of the natural numbers exists only because the axiom of infinity tells us it does. The negation of the axiom of infinity can be consistently chosen, which would lead to a different version of set theory. The natural numbers so constructed are known as von Neumann ordinals. They are as follows:

- $0 = \{\} = \varnothing$
- $1 = \{0\} = \{\varnothing\}$
- $2 = \{0, 1\} = \{\varnothing, \{\varnothing\}\}$
- $3 = \{0, 1, 2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$
- $4 = \{0, 1, 2, 3\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}\}$
- ...

It can be shown that they obey Peano's axioms, another common way to axiomatize the natural numbers and their arithmetic.

Definition 38. *The **integers** can be defined from the natural numbers as an ordered pair of sign and value. That is, $\mathbb{Z} = (\{0\} \times \mathbb{N}) \cup (\{1\} \times \mathbb{N}^+)$.*

Definition 39. *The **rationals** can be defined from the integer numbers as the equivalence class that leads to the same simplification. That is, let $Q = \mathbb{Z} \times \mathbb{Z}$ be all pairs of integers. Let $\sim\, \subset Q \times Q$ be an equivalence relation such that $(a_1, b_1) \sim (a_2, b_2)$ if and only if $a_1 b_2 = a_2 b_1$ (i.e. the pairs are proportional). The set of rationals is $\mathbb{Q} = Q_{/\sim}$. Given two integers $a, b \in \mathbb{Z}$, the rational number $\frac{a}{b}$ is the equivalence class $[(a, b)]_\sim$.*

Definition 40. *The **reals** can be defined from the rationals as "cuts" that divide the rationals in two contiguous sets. That is, let a real number be a partition (A, B) such that $A \cup B = \mathbb{Q}$, $A \neq \varnothing$, $B \neq \varnothing$ and $a < b$ for all $a \in A$ and $b \in B$. The set of all reals $\mathbb{R}$ is the set of all such partitions.*

Remark. Operations such as additions, multiplications, ... can be defined appropriately as functions of the respective sets.

4.2 Ordinal numbers

To extend natural numbers beyond finite numbers we need to make a distinction. Cardinal numbers (i.e. one, two, three, ...) identify the number of elements in a set. Ordinal numbers (i.e. first, second, third) identify a position within a sequence. Cardinals and ordinals are in one to one correspondence only for finite values. For infinite values, cardinals are a subset of the ordinals.

Remark. To define ordinals, we extend the notion of sequences to arbitrary levels of infinity through well-ordered sets. Note that we introduced the non-standard symbols $\preceq$, $\sim$ and $\langle A \rangle$ to make the notation more compact.

Poset: $(A, \leq_A)$ **Definition 41** (Poset)**.** *A **partially ordered set**, or **poset**, is a set with an associated partial order. That is, a tuple $(A, \leq_A)$ where A is a set and $\leq_A$ is a partial order over A. A **linearly ordered set** is a poset whose order is linear.*

Initial segment   **Definition 42** (Initial segment)**.** *The **initial segment** of a poset $(A, \leq_A)$ determined by $a \in A$ is the set of all elements before a. That is, $S_a = \{x \in A \,|\, x < a\}$.*

Order homo/ isomorphism	**Definition 43** (Order morphisms)**.** *Let $(A, \leq_A)$ and $(B, \leq_B)$ be two posets. An **order homomorphism** is a function $f : A \to B$ that preserves the ordering. That is, $a_1 \leq_A a_2$ implies $f(a_1) \leq_B f(a_2)$. An **order isomorphism** is a bijective order homomorphism whose inverse is also an order homomorphism (i.e. the posets have identical order). That is, $b_1 \leq_B b_2$ implies $f^{-1}(b_1) \leq_A f^{-1}(b_2)$.*

Order homo/
isomorphism

Definition 43 (Order morphisms)**.** *Let $(A, \leq_A)$ and $(B, \leq_B)$ be two posets. An **order homomorphism** is a function $f : A \to B$ that preserves the ordering. That is, $a_1 \leq_A a_2$ implies $f(a_1) \leq_B f(a_2)$. An **order isomorphism** is a bijective order homomorphism whose inverse is also an order homomorphism (i.e. the posets have identical order). That is, $b_1 \leq_B b_2$ implies $f^{-1}(b_1) \leq_A f^{-1}(b_2)$.*

Well-order

Definition 44 (Well-order)**.** *A **well-ordered** set is a linearly ordered set $(A, \leq_A)$ where each subset has a least (i.e. first) element. That is, for every $B \subseteq A$, there exists an $a \in B$ such that $a \leq_A b$ for all $b \in B$. Let A and B be two well-ordered sets. We say A **has lower or equal ordinality than** B, noted $A \preceq B$, if A is orderwise identical to the beginning of B. That is, A is order isomorphic to an initial segment of B. They have **the same ordinality**, noted $A \sim B$, if they have the same order. That is, they are order isomorphic.*

Proposition 45 (Ordinality)**.** *If A and B are two well-ordered sets, one of them has lower or equal ordinality than the other. That is, either $A \preceq B$ or $B \preceq A$. If $\mathcal{A}$ is a collection of well-ordered sets, then one element has least ordinality.*

Remark. Simple examples of well-ordered sets:

- a finite sequence: $(a_1, a_2, ..., a_n)$
- an infinite sequence: $(a_1, a_2, ...)$
- an infinite sequence with elements afterwards: $(a_1, a_2, ..., b_1, b_2, ..., b_6)$
- three infinite sequences, one after the other: $(a_1, a_2, ..., b_1, b_2, ..., c_1, c_2, ...)$

In a well-ordered set, every element always has an immediate successor, but not necessarily an immediate predecessor (e.g. b_1 in the third example).

Ordinals:
ORD

Definition 46 (Ordinals)**.** *An **ordinal** is a set A that includes the previous and only the previous ordinals (this makes $a \in A$ mean $a \prec A$). That is, for all elements $a, b, c \in A$ (i.e. previous ordinals of A):*

1. *$a \notin a$ (each previous ordinal is not before itself)*
2. *$a \in b$ implies $b \notin a$ (if a previous ordinal is before another, then the latter is not before the former)*
3. *$a \in b$ and $b \in c$ implies $a \in c$ (previous ordinals are transitive)*
4. *$x \in a$ implies $x \in A$ (an ordinal before a previous ordinal is a previous ordinal)*

*We note **ORD** as the class of all ordinals.*

Proposition 47 (Natural numbers are ordinals)**.** *The natural numbers are ordinals. That is, $\mathbb{N} \subset \mathbf{ORD}$.*

Remark. Ordinal numbers so defined generalize the construction of von Neumann ordinals (i.e. a natural number is a set that includes all previous natural numbers) to all possible well-orders.

Proposition 48 (Ordinals are well-ordered)**.** *All ordinals are well-ordered sets. That is, for each $A \in \mathbf{ORD}$, $(A, \leq_A)$ is a well-ordered set where $a \leq_A b$ if either $a \in b$ or $a = b$. Moreover, the class **ORD** is well ordered. That is, given a set of ordinals, one has lower or equal ordinality than all others.*

$\langle A \rangle, \omega$

Proposition 49 (Well-orders are ordinals)**.** *For every well-ordered set A there is a unique ordinal that fully characterizes its order. That is, there is a unique $\langle A \rangle \in \mathbf{ORD}$ such that $\langle A \rangle \sim A$. We say $\langle A \rangle$ is the **ordinality of** A. The ordinality of an infinite sequence is denoted by ω. That is, $\omega = (0, 1, 2, ...) = \langle \{x_i\}_{i=1}^{\infty} \rangle$ for all $\{x_i\}_{i=1}^{\infty}$.*

Remark. As an example, consider the von Neumann ordinal $3 = (0, 1, 2)$ as a well-ordered set. Similarly, the ordinality of the well-ordered set (a_1, a_2, a_3) is $3 = (0, 1, 2)$ as they share the same ordering. The definitions are such that these ideas extend to infinite sets.

Definition 50 (Ordinal addition)**.** *The addition of two ordinals is defined by concatenating the first order with the second. Let $(A, \leq_A)$ and $(B, \leq_B)$ be two well-ordered sets. Let $(A \sqcup B, \leq)$ be the poset where:*

- $(0, a_1) \leq (0, a_2)$ *if $a_1 \leq_A a_2$*
- $(1, b_1) \leq (1, b_2)$ *if $b_1 \leq_B b_2$*
- $(0, a) \leq (1, b)$ *for all $a \in A$ and $b \in B$.*

*Then $(A \sqcup B, \leq)$ is well-ordered and the ordinal $\langle A \sqcup B \rangle$ is fully determined by $\langle A \rangle$ and $\langle B \rangle$. Therefore we can define the **ordinal addition** $\langle A \rangle + \langle B \rangle = \langle A \sqcup B \rangle$.*

Remark. Addition is non-commutative (i.e. $\langle A \rangle + \langle B \rangle \neq \langle B \rangle + \langle A \rangle$) for infinite ordinals. For example $(a_1, a_2, ...) \sqcup (b_1, b_2, b_3) = (a_1, a_2, ..., b_1, b_2, b_3)$ is not the same as $(b_1, b_2, b_3) \sqcup (a_1, a_2, ...) = (b_1, b_2, b_3, a_1, a_2, ...) \sim (c_1, c_2, ...)$. Therefore $\omega + 3 \neq 3 + \omega = \omega$.

Definition 51 (Ordinal multiplication)**.** *The multiplication of two ordinals is defined by ordering their Cartesian product. Let $(A, \leq_A)$ and $(B, \leq_B)$ be two well-ordered sets. Let $(A \times B, \leq)$ be the poset where:*

- $(a_1, b_1) \leq (a_2, b_2)$ *if $b_1 \leq_B b_2$*
- $(a_1, b) \leq (a_2, b)$ *if $a_1 \leq_A a_2$*

*Then $(A \times B, \leq)$ is well-ordered and the ordinal $\langle A \times B \rangle$ is fully determined by $\langle A \rangle$ and $\langle B \rangle$. Therefore we can define the **ordinal multiplication** $\langle A \rangle \langle B \rangle = \langle A \times B \rangle$.*

Remark. Muliplication is non-commutative (i.e. $\langle A \rangle \langle B \rangle \neq \langle B \rangle \langle A \rangle$) for infinite ordinals. For example $(a_1, a_2, ...) \times (b_1, b_2) = ((a_1, b_1), (a_2, b_1), ..., (a_1, b_2), (a_2, b_2), ...)$ is not the same as $(b_1, b_2) \times (a_1, a_2, ..., a_n, ...) = ((b_1, a_1), (b_2, a_1), (b_1, a_2), (b_2, a_2), ...) \sim (c_1, c_2, ...)$. Therefore $\omega 2 \neq 2\omega = \omega$.

With addition and multiplication, we can get a sense of how the ordinals work. ω represents a countable sequence (i.e. $\omega = (0, 1, 2, ...) = (\mathbb{N}, \leq)$). $\omega + 3$ is a countable sequence followed by three elements. $\omega + \omega = \omega 2$ represents a countable sequence followed by a countable sequence. $\omega 6$ represents a sequence of 6 countable sequences. $\omega\omega = \omega^2$ represents a countable sequence of countable sequences. ω^3 represents a countable sequence of countable sequences of countable sequences. ω^ω extends the recursion infinitely many times and so on. There is no upper limit to the ordinals.

Adding an ordinal of a lower power to the left does not change the ordinal. Adding finitely many elements at the beginning of an infinite sequence still gives us an infinite sequence (i.e. $n + \omega = \omega$). In the same way, adding an infinite sequence before an infinite sequence of infinite sequences still gives us an infinite sequence of infinite sequences (i.e. $\omega + \omega^2 = \omega^2$).

Proposition 52 (Cantor normal form)**.** *Every ordinal number $\langle A \rangle$ can be written as a sum of powers of ω. That is, $\langle A \rangle = \omega^{\alpha_1} c_1 + \omega^{\alpha_2} c_2 + ... + \omega^{\alpha_n} c_n$ where $n \in \mathbb{N}$, $c_i \in \mathbb{N}^+$ for all $1 \leq i \leq n$ and $\alpha_i \in \mathbf{ORD}$ such that $\alpha_i > \alpha_j$ for all $i > j$.*

4.3 Cardinal numbers

Remark. The cardinality of a set (i.e. how many elements it contains) is defined in terms of injection/bijection: a set is smaller than another if it can be injected into it, and it is as big as another if it can be put into a one-to-one correspondence. Note that we introduced the non-standard symbols $\hookrightarrow$ and $\leftrightarrow$ to make the notation more compact.

Definition 53 (Cardinal comparison)**.** *Let A and B be two sets. If there exists a injective function $f : A \to B$ then we say A **has lower or equal cardinality than** B, noted $A \hookrightarrow B$. If there exists a bijective function we say they have **the same cardinality**, noted $A \leftrightarrow B$.*

Proposition 54 (Cardinal comparability of sets)**.** *Given two sets A and B, one must have lower or equal cardinality than the other. That is, either $A \hookrightarrow B$ or $B \hookrightarrow A$. Additionally, there exists a bijective function between A and B if and only if there exists an injective function in both directions. That is, $A \leftrightarrow B$ if and only if $A \hookrightarrow B$ and $B \hookrightarrow A$.*

Proposition 55 (Well-ordering theorem)**.** *Every set can be well-ordered.*

Remark. The well-ordering theorem allows us to reuse ordinal numbers to quantify cardinality. As two different well-ordered sets may have the same cardinality, we'll take the lowest ordinal between them.

Cardinals: **CRD**

Definition 56 (Cardinal number)**.** *A **cardinal** is an ordinal $a \in \textbf{ORD}$ that has lower or equal ordinality than any ordinal that has the same cardinality. That is, $a \leq b$ for all $b \in \textbf{ORD}$ such that $a \leftrightarrow b$. The class of all cardinals is noted **CRD** and, by definition $\textbf{CRD} \subset \textbf{ORD}$.*

$|A|$

Proposition 57 (Cardinality of a set)**.** *For every set A there is a unique cardinal that identifies the cardinality of A. That is, there exists a unique $|A| \in \textbf{CRD}$ such that $|A| \leftrightarrow A$. We say $|A|$ is the **cardinality of** A.*

Proposition 58 (Cantor's theorem)**.** *A set has strictly lower cardinality than its power set. That is, $|A| < |\mathcal{P}(A)|$. If $\alpha = |A|$ then we note $|\mathcal{P}(A)| = 2^{\alpha}$*

Aleph numbers: $\aleph_{\alpha}$

Proposition 59 (Aleph numbers)**.** *The infinite cardinals are well-ordered. They are noted $\aleph_{\alpha}$ where $\alpha \in \textbf{ORD}$.*

$\aleph_0$

Definition 60 (Aleph-nought)**.** *The smallest infinite cardinal $\aleph_0$ denotes countable infinity.*

Continuum: $2^{\aleph_0}$

Definition 61 (Cardinality of continuum)**.** *The cardinality of the continuum is the cardinality of the powerset of a countable set $2^{\aleph_0}$.*

Remark. Sets that have countable cardinality:

- natural $\mathbb{N}$, integer $\mathbb{Z}$ and rational $\mathbb{Q}$ numbers
- ordinals such as ω, $\omega + 3$, $\omega 2$, ω^{ω}

Sets that have cardinality of the continuum:

- powerset of natural, integers and rationals: $\mathcal{P}(\mathbb{N})$, $\mathcal{P}(\mathbb{Z})$, $\mathcal{P}(\mathbb{Q})$
- real $\mathbb{R}$ and complex $\mathbb{C}$ numbers
- real and complex Euclidean spaces: $\mathbb{R}^n$, $\mathbb{C}^n$
- continuous functions over Euclidean spaces: $C(\mathbb{R}^n, \mathbb{R}^m)$
- standard topology and Borel algebra on $\mathbb{R}^n$ (i.e. the set of open sets and Borel set respectively)
- countable sequences of numbers: $\mathbb{N}^{\mathbb{N}}$, $\mathbb{Z}^{\mathbb{N}}$, $\mathbb{Q}^{\mathbb{N}}$, $\mathbb{R}^{\mathbb{N}}$, $\mathbb{C}^{\mathbb{N}}$

Sets that have cardinality greater than the continuum:

- powerset of the real and complex numbers: $\mathcal{P}(\mathbb{R})$, $\mathcal{P}(\mathbb{C})$
- set of all functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

We should note that all physically meaningful sets have at most the cardinality of the continuum. This is not a coincidence. In science we are interested in objects that can be experimentally identified. Every experimental test can acquire finite information in finite time and, assuming arbitrarily long time, we can run at most countably many tests. Countable sequences of numbers have the cardinality of the continuum, which is therefore the best we can distinguish. Even if nature allows for wider variety, it would be merged into equivalence classes of objects that are experimentally indistinguishable. Many of the issues of higher cardinals can therefore be safely set aside by those working on Assumptions of Physics.

$\aleph_1$

Definition 62 (Aleph-one)**.** *Let Ω be the set of all countable ordinal numbers. Its cardinality $|\Omega| = \aleph_1$ is the second aleph number.*

Remark. While it can be proven that the cardinality of the continuum $2^{\aleph_0}$ is an aleph number, which one is undetermined in standard set theory.

Continuum hypothesis

Continuum hypothesis There is no infinite set with cardinality in between that of the continuum and countable infinity. That is, $2^{\aleph_0} = \aleph_1$.

Remark. The continuum hypothesis is proven to be independent of the axioms of set theory. It can be taken to be true or to be false without inconsistency. As the axiom deals with sets of cardinality lower than the continuum, this aspect of set theory may be interesting for those working on Assumptions of Physics. It could be, in principle, that the distinguishable configurations in space-time have a cardinality that is neither countable nor that of the continuum.

References

[1] Charles C Pinter. *A book of set theory.* Dover Publications Corporation, 2014.

[2] Alex Svirin. https://www.math24.net/topics-set-theory/, 2021.