2. Set theory

Every since about mid 1800s mathematicians realized that sets provide a useful tool to express mathematical statements and proofs. We will now review some elementary facts from this theory. Throughout we will mostly use the "naive" version of the theory where all "sensible" operations on sets are allowed. We will nonetheless point out deficiencies of the naive theory and give a quick overview of Zermelo's axiomatic.

2.1 Naive set theory.

We start with the basic setting of naive set theory:

- The basic building blocks of naive set theory are *sets*, to be denoted by capital letters A, B, etc. Sets are basically "containers" collecting other objects which, at least in pure set theory, are themselves sets. For any two sets A, B, we thus assume the existence of a (logical) proposition $A \in B$, whose TRUE value designates that A *belongs to* or *is an element of* B. We will write $A \notin B$ for $\neg(A \in B)$.
- In order to be able to form sets from other sets, we put forward a basic assumption, termed the *Comprehension Principle*, which states that, for any predicate $P(X)$ whose parameter is a set,
$$\{X\colon P(X)\} \text{ is a set} \tag{2.1}$$

Using the Comprehension Principle we can construct many objects we are used to from prior experience with set theory. For instance, we can define the *empty set* by
$$\emptyset := \{X\colon \text{FALSE}\} \tag{2.2}$$
where "FALSE" stands for a logical proposition that takes only FALSE value. For a set A, we can define its *complement* by
$$A^{\mathrm{c}} := \{X\colon X \notin A\}. \tag{2.3}$$
There are also a number of familiar operations on pairs of sets. Indeed, given sets A, B,
$$\begin{aligned} A \cup B &:= \{X\colon X \in A \vee X \in Y\} \\ A \cap B &:= \{X\colon X \in A \wedge X \in Y\} \\ A \smallsetminus B &:= \{X\colon X \in A \wedge X \notin B\} \\ A \triangle B &:= (A \smallsetminus B) \cup (B \smallsetminus A), \end{aligned} \tag{2.4}$$
define their *union*, *intersection*, *set difference* and *symmetric difference*, respectively.

The Comprehension Principle is rather strong to give us far more than the above. For instance, for each set A we can define the *singleton* $\{A\}$ containing just A by
$$\{A\} := \{X\colon X = A\} \tag{2.5}$$
Here we used the equality sign "=" in the meaning of *sameness* or *identity*. (This symbol is sometimes introduced as part of the setup of the theory; if not, then we define it as $A = B := \forall X\colon X \in A \Leftrightarrow X \in B$.) Similarly, we can *pair* two sets A and B into
$$\{A, B\} := \{X\colon X = A \vee X = B\}. \tag{2.6}$$
Another useful object is the set of all subsets of A, termed the *power set*, defined by
$$\mathcal{P}(A) := \{X\colon X \subseteq A\}, \tag{2.7}$$

where we made use of the binary relation A *is a subset of* B with the definition

$$A \subseteq B := (\forall X \in A\colon X \in B). \tag{2.8}$$

(This is just a shorthand induced by the relation $\in$.)

A rather important consequence of the Comprehension Principle is that it implies existence of infinite sets. Indeed, first generalize the first line in (2.4) to

$$\bigcup A := \bigl\{X\colon (\exists B \in A\colon X \in B)\bigr\} \tag{2.9}$$

for the union of all sets contained in A. (Note that $A \cup B = \bigcup\{A, B\}$.) The set

$$I := \Bigl\{X\colon \bigcup X \subseteq X\Bigr\} \tag{2.10}$$

is then closed under the operation $X \mapsto X \cup \{X\}$ (meaning: $\forall X \in I\colon X \cup \{X\} \in I$) because

$$\bigcup\bigl(X \cup \{X\}\bigr) = X \cup \bigcup X \underset{X\in I}{=} X \subseteq X \cup \{X\}, \tag{2.11}$$

where we used various general facts $\cup$ and $\subseteq$ whose proof we leave to the reader.

Noting that $\varnothing \in I$ because $\bigcup\varnothing = \varnothing \subseteq \varnothing$, we now recursively check that I is *infinite* according to the following definition: A set B is said to be infinite if there is an injective map $f\colon B \to B$ such that $\mathrm{Ran}(f) := \{f(X)\colon X \in B\}$ is a proper subset of B (meaning: $\mathrm{Ran}(f) \subseteq B$ yet $\mathrm{Ran}(f) \neq B$). For I in (2.11), this is witnessed by the map

$$f(X) := X \cup \{X\} \tag{2.12}$$

which is injective because $X \cup \{X\} = Y \cup \{Y\}$ is reduced to $X = Y$ (for $X, Y \in I$) by taking the union (see the second equality in (2.11)), but is not onto because f does not have $\varnothing$ in its range. (The concepts of "injective", "onto", etc have yet to be introduced so this explanation is mainly for those already in command of these terms.) The intuition of the above construction is that I contains all elements in the set

$$\Bigl\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}, \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}, \dots\Bigr\} \tag{2.13}$$

which is "obviously" infinite.

Unfortunately, in the late 1900s it became gradually apparent that the naive set theory contains inconsistencies. These started at large infinite sets but, ultimately, surfaced in the following elementary mind-boggling paradox discovered by B. Russell in 1901:

Theorem 2.1 (Russell's antinomy) *In naive set theory,*

$$\{X\colon X \notin X\} \tag{2.14}$$

is not a set. In particular, the comprehension principle is inconsistent.

Proof. Suppose, by way to contradiction, $A := \{X\colon X \notin X\}$ is a set. Then either $A \in A$ is TRUE or $A \notin A$ is TRUE. But if $A \in A$ is TRUE then the predicate defining A forces $A \notin A$ while if $A \notin A$ is TRUE then $A \in A$ because otherwise A would otherwise be included in (2.14). Either possibility leads to a contradiction and so A is not a set. $\square$

2.2 Zermelo's axiomatic.

It is fair to say that Russell's observation rattled the foundations of mathematics of that day. A number of solutions was gradually proposed some of which permeate various treatments of set theory till today. Here we will follow the solution proposed by E. Zermelo in 1908 which, ultimately, leads to the so called *ZFC set theory* used prevalently throughout analysis.

An important novelty of Zermelo's approach was that one should rely on *axiomatic* formulation of set theory modeled, in some way, on a similar approach to classical Euclidean geometry. To dispense with Russell's paradox, Zermelo proposed to restrict the Comprehension Principle by requiring only:

- **Separation axiom:**
$$\forall B \text{ set} : \{X \in B \colon P(X)\} \text{ is a set} \tag{2.15}$$

Here we purposefully deviate from our earlier convention that tells us write the set (2.15) as $\{X \colon X \in B \wedge P(X)\}$. Since the point of writing $X \in B$ is to enforce the "separation," we will adhere to this practice throughout the course. Russell's argument then gives:

Lemma 2.2 *Let B be a set and let $A := \{X \in B \colon X \notin X\}$. Then $B \notin A \ \wedge\ A \notin A \ \wedge\ A \notin B$.*

Proof. Assuming $B \in A$ we get $B \in B$ AND $B \notin B$, a contradiction. So we must have $B \notin A$ as claimed. Similarly, $A \in A$ implies $A \notin A$, a contradiction. So $A \notin A$ holds as well. But $A \notin A$ then forces $A \notin B$ because otherwise we would have $A \in A$, a contradiction. □

Unfortunately, with the Comprehension Principle gone as stated, we lose the ability to define many of the above sets — specifically, $\varnothing$ (as there could no sets at all), $A \cup B$ (as there could be no sets subsuming both A and B), $\{A\}$ (as there could be no set containing A) and, for similar reasons, $\mathcal{P}(A)$ and I. Further axioms are thus needed. We will now state Zermelo's axioms in bullet-point format:

- **Axiom of Extensionality:**
$$\forall A, B \ : A = B \Leftrightarrow \Big(\forall X \colon\ X \in A \Leftrightarrow X \in B\Big). \tag{2.16}$$
This axiom ensures that a set is uniquely determined by its elements.
- **Empty set axiom:**
$$\exists \varnothing\, \forall X \colon X \notin \varnothing \tag{2.17}$$
Thanks to Axiom of Extensionality, $\varnothing$ is the unique set with this property.
- **Pairset axiom:**
$$\forall X\, \forall Y\, \exists A\, \forall Z \colon Z \in A \ \Leftrightarrow\ \big(X = Z \vee Y = Z\big). \tag{2.18}$$
Again, the resulting set A is determined uniquely so we henceforth denote it $\{X, Y\}$ when X and Y are different and $\{X\}$ when $X = Y$.
- **Axiom of Union:**
$$\forall A\, \exists B\, \forall X \ : \ X \in B \Leftrightarrow \Big(\exists C \in A \colon X \in C\Big) \tag{2.19}$$
This is a bit hard to parse at first sight. Here A is a set of sets and B is the union of all elements in the elements of A. We use the notation $\bigcup A$ for B.
- **Powerset axiom:**
$$\forall A\, \exists B\, \forall X \ : \ X \subseteq A \Leftrightarrow X \in B \tag{2.20}$$
Here B is thus the powerset $\mathcal{P}(A)$, i.e., the set of all subsets, of A. Again, the powerset is unique by the Axiom of Extensionality.

- **Axiom of infinity:**

$$\exists I : \varnothing \in I \wedge \Big(\forall X\colon\; X \in I \Rightarrow \{X\} \in I\Big) \tag{2.21}$$

The notation $\{X\}$ is meaningful by the Pairset Axiom.

Here are some remarks on the above. First, as noted earlier, (2.16) requires the notion of identity — represented by the equality sign — to be part of the setup of the theory. Otherwise (2.16) can be read as a definition of "=" sign:

$$\forall A, B\colon \quad A = B \;:=\; \Big(\forall X\colon\; X \in A \Leftrightarrow X \in B\Big) \tag{2.22}$$

meaning that two sets are said to be equal when they have exactly the same elements. Yet another way to characterize equality is via:

Lemma 2.3 *We have*

$$\forall A, B\colon \quad A = B \;\Leftrightarrow\; A \subseteq B \wedge B \subseteq A \tag{2.23}$$

Proof. We first note that, given two predicates $P(x)$ and $Q(x)$ depending on x, we have

$$\big(\forall x\colon P(x)\big) \wedge \big(\forall x\colon Q(x)\big) \;\Leftrightarrow\; \forall x\colon P(x) \wedge Q(x) \tag{2.24}$$

To verify this, note that $\forall x\colon P(x)$ means that $\mathrm{Ran}(P)$ is $\{\mathrm{TRUE}\}$, and similarly $\forall x\colon Q(x)$ means $\mathrm{Ran}(Q) = \{\mathrm{TRUE}\}$. The proposition on the left thus equivalent to both ranges being equal $\{\mathrm{TRUE}\}$, which is equivalent to that on the right. To get (2.23) from this, we set $P(X) := X \in A$ and $Q(X) := X \in B$ and apply (2.8) and (2.22) or the Axiom of Extensionality. □

Second, the set-theoretical notation $\bigcup A$ for the "union of all sets in A" is often substituted by other, more intuitive, expressions such as

$$\bigcup\{X\colon X \in A\} \quad \text{or even} \quad \bigcup_{X \in A} X \tag{2.25}$$

With the union postulated to exist, we can define general *intersection* by

$$\bigcap A := \Big\{X \in \bigcup A\colon (\forall C \in A\colon X \in C)\Big\}. \tag{2.26}$$

with similar alternative notations as in (2.25). Similarly, while we will keep writing $\mathcal{P}(A)$ for the power set of A in this course, in practice we at times write $\{X\colon X \subseteq A\}$ in blatant violation of the Separation Axiom.

While the Emptyset Axiom ensures existence of a set, the Axiom of Infinity ensures existence of an infinite set. These may not exist otherwise in our universe (for instance, if our universe of sets is the collection of finite subsets of the naturals). The Axiom of Infinity in particular ensures existence of a set that contains

$$\Big\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\{\{\varnothing\}\}\}, \dots\Big\} \tag{2.27}$$

as a subset. This axiom is phrased already with its important applications — namely, the construction of the naturals, which will arise from the set in (2.27) — in mind. (We deviate from what stated initially in class for exactly this reason.)

Zermelo's system includes one additional axiom, namely, the *Axiom of Choice* which we will get to in the next section. Another axiom — called *Axiom of Replacement* — was added later to the system by A. Fraenkel and, independently, T. Skolem leading to (what is now called) ZFC axiomatic — where the acronym stands for "Zermelo-Fraenkel with Axiom of Choice." The latter axiom is somewhat intricate to explain and we will not state it explicitly. As we shall see, we can actually get a good way into real analysis without needing either of them.

The Comprehension Principle can be retained in the axiomatic system by introducing the notion of a *class*. This is just another way to describe a "collection of sets with a given property" except that it is generally too large to be automatically called a set. By definition, every set is a class but the class is a set only if it is contained in a set. Russell's paradox then shows only that there are classes that are not sets.

Further details (and inspiration for the above presentation) can be found in Yannis Moschovakis' *Notes on Set Theory*, which is a wonderful advanced-undergraduate introduction to set theory.