Qiuzhen Lectures on Analysis

BIN GUI

September 6, 2023

1 Basic set theory and numbers

In this chapter, we discuss informally some of the basic notions in set theory and basic properties about numbers. A more thorough treatment can be found in [Mun, Ch. 1] (for set theory) and [Rud-P, Ch. 1] (for numbers).

Let me first list the notations and conventions that will be used throughout the notes. We use frequently the abbreviations:

iff=if and only if

LHS=left hand side RHS=reft hand side

$\exists$=there exists $\forall$=for all

i.e.=id est=that is=namely

c.f.=compare/check/see/you are referred to

resp.=respectively WLOG=without loss of generality

The topics marked with $\star\star$ are technical and/or their methods are rarely used in later studies. You can skim or skip them on first reading. The topics marked with $\star$ are interesting, but not necessarily technical. They are not essential for understanding the later part of the notes.

1.1 Basic operations and axioms

Intuitively, a set denotes a collection of elements. For instance:

$$\mathbb{Z} = \{\text{all integers}\} \qquad \mathbb{N} = \{n \in \mathbb{Z} : n \geqslant 0\} \qquad \mathbb{Z}_+ = \{n \in \mathbb{Z} : n > 0\}$$

have infinitely many elements. (In this course, we will not be concerned with the rigorous construction of natural numbers and integers from Peano axioms.) We also let

$$\mathbb{Q} = \{\text{all rational numbers}\} \qquad \mathbb{R} = \{\text{all real numbers}\}$$

assuming that rational and real numbers exist and satisfy the properties we are familiar with in high school mathematics.

Set theory is the foundation of modern mathematics. It consists of several Axioms telling us what we can do about the sets. For example, the following way of describing sets

$$\{x : x \text{ satisfies property...}\} \tag{1.1}$$

is illegal, since it gives **Russell's paradox**: Consider

$$S = \{A : A \text{ is a set and } A \notin A\} \tag{1.2}$$

If S were a set, then $S \in S \Rightarrow S \notin S$ and $S \notin S \Rightarrow S \in S$. This is something every mathematician doesn't want to happen.

Instead, the following way of defining sets is legitimate:

$$\{x \in X : x \text{ satisfies property } \dots\} \tag{1.3}$$

where X *is a given set*. For instance, we can define the **difference** of two sets:

$$A \backslash B = A - B = \{x \in A : x \notin B\}$$

So let us figure out the legal way of defining unions and intersections of sets. The crucial point is that we assume the following axiom:

Axiom. If $\mathscr{A}$ is a set of sets, then there exists a set X such that $A \subset X$ for all $A \in \mathscr{A}$.

Thus, if $\mathscr{A}$ is a set of sets, let X be a set larger than (i.e. "$\supset$") every member of $\mathscr{A}$, then we can define the **union** and the **intersection**

$$\bigcup_{A \in \mathscr{A}} A = \{x \in X : \text{there exists } A \in \mathscr{A} \text{ such that } x \in A\} \tag{1.4a}$$

$$\bigcap_{A \in \mathscr{A}} A = \{x \in X : \text{for all } A \in \mathscr{A} \text{ we have } x \in A\} \tag{1.4b}$$

It is clear that this definition does not rely on the particular choice of X.

Remark 1.1. In many textbooks, it is not uncommon that sets are defined as in (1.1). You should interpret such definition as (1.3), where the set X is omitted because it is clear from the context. For instance, if the context is clear, the set $\{x \in \mathbb{R} : x \geqslant 0\}$ could be simply written as $\{x : x \geqslant 0\}$ or even $\{x \geqslant 0\}$. By the same token, the phrase "$\in X$" in (1.4) could be omitted. So we can also write

$$A \cup B = \{x : x \in A \text{ or } x \in B\} \qquad A \cap B = \{x : x \in A \text{ and } x \in B\}$$

which are special cases of (1.4).

Remark 1.2. In the same spirit, when discussing subsets of a given "large" set X, and if X is clear from the context, we shall write $X \backslash A$ (where $A \subset X$) as A^c and call it the **complement** of A.

Example 1.3. We have

$$\bigcup_{x \in (1,+\infty)} [0, x) = [0, +\infty) \qquad \bigcap_{n \in \mathbb{Z}_+} (0, 1 + 1/n) = (0, 1] \qquad \bigcup_{n \in \mathbb{N}} (0, 1 - 1/n] = (0, 1)$$

The readers may notice that these examples are not exactly in the form (1.4). They are actually unions and intersections of indexed families of sets. (See Def. 1.9.) We need some preparation before discussing this notion.

Axiom. If $A_1, \dots, A_n$ are sets, their **Cartesian product** exists:

$$A_1 \times \cdots \times A_n = \{(a_1, \dots, a_n) : a_i \in A_i \text{ for all } 1 \leqslant i \leqslant n\}$$

where two elements $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$ of the Cartesian product are viewed as equal iff $a_1 = b_1, \dots, a_n = b_n$. We also write

$$(a_1, \dots, a_n) = a_1 \times \cdots \times a_n$$

especially when a, b are real numbers and (a, b) can mean an open interval.

If $A_1 = \cdots = A_n = A$, we write the Cartesian product as A^n. □

Example 1.4. Assume that the set of real numbers $\mathbb{R}$ exist. Then the set of complex numbers $\mathbb{C}$ is defined to be $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ as a set. We write $(a, b) \in \mathbb{C}$ as $a + b\mathbf{i}$ where $a, b \in \mathbb{R}$. Define

$$(a + b\mathbf{i}) + (c + d\mathbf{i}) = (a + c) + (b + d)\mathbf{i}$$
$$(a + b\mathbf{i}) \cdot (c + d\mathbf{i}) = (ac - bd) + (ad + bc)\mathbf{i}$$

Define the zero element 0 of $\mathbb{C}$ to be $0 + 0\mathbf{i}$. More generally, we consider $\mathbb{R}$ as a subset of $\mathbb{C}$ by viewing $a \in \mathbb{R}$ as $a + 0\mathbf{i} \in \mathbb{C}$. This defines the usual arithmetic of complex numbers.

If $z = a + b\mathbf{i}$, we define its **absolute value** $|z| = \sqrt{a^2 + b^2}$. Then $z = 0$ iff $|z| = 0$. We define the **(complex) conjugate** of z to be $\overline{z} = a - b\mathbf{i}$. Then $|z|^2 = z\overline{z}$.

If $z \neq 0$, then there clearly exists a unique $z^{-1} \in \mathbb{C}$ such that $zz^{-1} = z^{-1}z = 1$: $z^{-1} = |z|^{-2} \cdot \overline{z}$. Thus, using the language of modern algebra, $\mathbb{C}$ is a **field**.[1] □

The axiom of Cartesian product allows us to define relations and functions:

Definition 1.5. If A, B are sets, a subset R of $A \times B$ is called a **relation**. For $(a, b) \in A \times B$, we write aRb iff $(x, y) \in R$. We understand "aRb" as "a is related to b through the relation R".

Definition 1.6. A relation f of A, B is called a **function** or a **map** (or a **mapping**), if for every $a \in A$ there is a unique $b \in B$ such that afb. In this case, we write $b = f(a)$.

When we write $f : A \to B$, we always mean that A, B are sets and f is a function from A to B. A and B are called respectively the **domain** and the **codomain** of f. (Sometimes people also use the words "source" and "target" to denote A and B.)

[1] The readers can easily find the definition of fields online or through textbooks (e.g. [Rud-P, Def. 1.12]). We will not present the full definition of fields in the notes. Just keep in mind some typical (counter)examples: $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields. $\mathbb{Z}$ is not a field, because not every non-zero element of $\mathbb{Z}$ has an inverse. The set of quaternions $\{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} : a, b, c, d \in \mathbb{R}\}$ is not a field because it is not commutative ($\mathbf{ij} = -\mathbf{ji} = \mathbf{k}$). The set of rational functions $P(x)/Q(x)$, where P, Q are polynomials with coefficients in $\mathbb{R}$ and $Q \neq 0$, is a field.

If $E \subset A$ and $F \subset B$, we define the **image under** f of E and the **preimage under** f of F to be

$$f(E) = \{b \in B : \exists a \in E \text{ such that } b = f(a)\}$$
$$f^{-1}(F) = \{a \in A : f(a) \in F\}.$$

$f(A)$ is simply called the **image** of f, or the **range** of f. If $b \in B$, $f^{-1}(\{b\})$ is often abbreviated to $f^{-1}(b)$. The function

$$f|_E : E \to B \qquad x \mapsto f(x)$$

is called the **restriction** of f to E. □

The intuition behind the definition of functions is clear: we understand functions as the same as their graphs. So a subset f of the "coordinate plane" $A \times B$ is the graph of a function iff it "intersects every vertical line precisely once".

Definition 1.7. A function $x : \mathbb{Z}_+ \to A$ is called a **sequence in** A. We write $x(n)$ as x_n, and write this sequence as $(x_n)_{n \in \mathbb{Z}_+}$.

Many people write such a sequence as $\{x_n\}_{n \in \mathbb{Z}_+}$. We do not use this notation, since it can be confused with $\{x_n : n \in \mathbb{Z}_+\}$ (the range of the function x).

Axiom. If X is a set, then the **power set** 2^X exists, where

$$2^X = \{\text{Subsets of } X\}$$

Example 1.8. The set $2^{\{1,2,3\}}$ has 8 elements: $\varnothing$, $\{1\}$, $\{2\}$, $\{3\}$, $\{1,2\}$, $\{1,3\}$, $\{2,3\}$, $\{1,2,3\}$. Surprisingly, $8 = 2^3$. As we shall see in Exp. 1.50 and Cor. 1.53, this relationship holds more generally, which explains the terminology 2^X.

Now we are ready to define indexed families of sets.

Definition 1.9. An **indexed family of sets** $(S_i)_{i \in I}$ is defined to be a function $S : I \to 2^X$ for some sets I, X. We write $S(i)$ as S_i. (So S_i is a subset of X.) I is called the **index set**. Define

$$\bigcup_{i \in I} S_i = \bigcup_{T \in S(I)} T \qquad \bigcap_{i \in I} S_i = \bigcap_{T \in S(I)} T$$

Note that $S(I)$ is the image of the function S.

Example 1.10. In the union $\bigcup_{x \in (1,+\infty)} [0, x)$, the index set is $I = (1, +\infty)$, and X can be the set of real numbers $\mathbb{R}$. Then $S : I \to 2^X$ is defined to be $S_i = S(i) = [0, i)$.

Exercise 1.11. Let $f : A \to B$ be a function. We say that f is **injective** if for all $a_1, a_2 \in A$ satisfying $a_1 \neq a_2$ we have $f(a_1) \neq f(a_2)$. We say that f is **surjective** if for each $b \in B$ we have $f^{-1}(b) \neq \varnothing$. f is called **bijective** if it is both surjective and bijective. Define the **identity maps** $\mathrm{id}_A : A \to A, a \mapsto a$ and id_B in a similar way. Prove that

$$f \text{ is injective} \iff \text{there is } g : B \to A \text{ such that } g \circ f = \mathrm{id}_A \tag{1.5a}$$

$$f \text{ is surjective} \iff \text{there is } g : B \to A \text{ such that } f \circ g = \mathrm{id}_B \tag{1.5b}$$

$$f \text{ is bijective} \iff \text{there is } g : B \to A \text{ such that } g \circ f = \mathrm{id}_A \text{ and } f \circ g = \mathrm{id}_B \tag{1.5c}$$

Show that the g in (1.5a) (resp. (1.5b), (1.5c)) is surjective (resp. injective, bijective).

The equivalence (1.5b) is subtler, since its proof requires Axiom of Choice.

Axiom. Let $(S_i)_{i \in I}$ be an indexed family of sets. The **Axiom of Choice** asserts that if $S_i \neq \varnothing$ for all $i \in I$, then there exists a function (the **choice function**)

$$f : I \to \bigcup_{i \in I} S_i$$

such that $f(i) \in S_i$ for each $i \in I$.

Intuitively, the axiom of choice says that for each $i \in I$ we can choose an element $f(i)$ of S_i. And such choice gives a function f.

Example 1.12. Let $f : A \to B$ be surjective. Then each member of the family $(f^{-1}(b))_{b \in B}$ is nonempty. Thus, by axiom of choice, there is a choice function g defined on the index set B such that $g(b) \in f^{-1}(b)$ for each b. Clearly $f \circ g = \mathrm{id}_B$.

Remark 1.13. Suppose that each member S_i of the family $(S_i)_{i \in I}$ has exactly one element. Then the existence of a choice function does not require Axiom of Choice: Let $X = \bigcup_{i \in I} S_i$ and define relation

$$f = \{(i, x) \in I \times X : x \in S_i\}$$

Then one checks easily that this relation between I and X is a function, and that it is the (necessarily unique) choice function of $(S_i)_{i \in I}$.

According to the above remark, one does not need Axiom of Choice to prove (1.5a) and (1.5c). Can you see why?

1.2 Partial and total orders, equivalence relations

Definition 1.14. Let A be a set. A **partial order** (or simply an **order**) $\leqslant$ on A is a relation on $A \times A$ satisfying for all $a, b, c \in A$ that:

- (Reflexivity) $a \leqslant a$.
- (Antisymmetry) If $a \leqslant b$ and $b \leqslant a$ then $a = b$.
- (Transitivity) If $a \leqslant b$ and $b \leqslant c$ then $a \leqslant c$.

We write $b \geqslant a$ iff $a \leqslant b$. Write $a > b$ iff $a \geqslant b$ and $a \neq b$. Write $a < b$ iff $b > a$. So $\geqslant$ is also an order on A. The pair $(A, \leqslant)$ is called a **partially ordered set**. A partial order $\leqslant$ on A is called a **total order**, if for every $a, b \in A$ we have either $a \leqslant b$ or $b \leqslant a$.

Example 1.15. The following are examples of orders.

- Assume that $\mathbb{R}$ exists. Then $\mathbb{R}$ has the canonical total order, which restricts to the total order of $\mathbb{Z}$. This is the total order that everyone is familiar with.
- Let X be a set. Then $(2^X, \subset)$ is a partially ordered set.
- $\mathbb{R}^2$ is a partially ordered set, if we define $(a, b) \leqslant (c, d)$ to be $a \leqslant c$ and $b \leqslant d$.

Definition 1.16. A relation $\sim$ on a set A is called an **equivalence relation**, if for all $a, b, c \in A$ we have

- (Reflexivity) $a \sim a$.
- (Symmetry) $a \sim b$ iff $b \sim a$.
- (Transitivity) If $a \sim b$ and $b \sim c$ then $a \sim c$.

Later, we will use the notions of partial orders and equivalence relation not just for a set, but for a collection of objects "larger" than a set. See Sec. 1.4.

Definition 1.17. Let A be a set, together with an equivalence relation $\sim$. Define a new set

$$A/\sim = \{[a] : a \in A\}$$

where the notion $[a]$ can be understood in the following two equivalent ways (we leave it to the readers to check the equivalence):

(1) $[a]$ is a new symbol. We understand $[a]$ and $[b]$ as equal iff $a \sim b$.

(2) $[a] = \{x \in A : x \sim a\}$

We call $[a]$ the **equivalence class** (or the **residue class**) of a, and call $A/\sim$ the **quotient set** of A under $\sim$. The surjective map $\pi : a \in A \mapsto [a] \in A/\sim$ is called the **quotient map**.

Exercise 1.18. Prove that every surjective map is equivalent to a quotient map. More precisely, prove that for every surjection $f : A \to B$, there is an equivalence relation $\sim$ on A and a bijective map $\Phi : A/\sim \to B$ such that the following diagram commutes:

$$\begin{array}{ccc} & A & \\ {\scriptstyle\pi}\swarrow & & \searrow{\scriptstyle f} \\ A/\sim & \xrightarrow{\Phi} & B \end{array} \tag{1.6}$$

This is the first time we see commutative diagrams. Commutative diagrams are very useful for indicating that certain maps between sets are "equivalent" or are satisfying some more general relations. For example, (1.6) shows that the maps f and π are equivalent, and that this equivalence is implemented by the bijection Φ. The formal definition of commutative diagrams is the following:

Definition 1.19. A diagram (i.e. some sets denoted by symbols, and some maps denoted by arrows) is called a **commutative diagram**, if all directed paths in the diagram with the same start and endpoints lead to the same result.

Here is an example of commutative diagram in linear algebra. This example assumes some familiarity with the basic properties of vector spaces and linear maps.[2]

Example 1.20. Let V, W be vector spaces over a field $\mathbb{F}$ with finite dimensions m, n respectively. Let $e_1, \dots, e_m$ be a basis of V, and let $\varepsilon_1, \dots, \varepsilon_n$ be a basis of W. We know that there are unique linear isomorphisms $\Phi : \mathbb{F}^m \xrightarrow{\simeq} V$ and $\Psi : \mathbb{F}^n \xrightarrow{\simeq} W$ such that

$$\Phi(a_1, \dots, a_m) = a_1 e_1 + \cdots + a_m e_m \qquad \Psi(b_1, \dots, b_n) = b_1 \varepsilon_1 + \cdots + b_n \varepsilon_n$$

Let $T : V \to W$ be a **linear map**, i.e., a map satisfying $T(a\xi + b\eta) = aT\xi + bT\eta$ for all $a, b \in \mathbb{F}, \xi, \eta \in V$. Then there is a unique $n \times m$ matrix $A \in \mathbb{F}^{n\times m}$ such that the following diagram commutes:

$$\begin{array}{ccc} \mathbb{F}^m & \xrightarrow[\simeq]{\Phi} & V \\ {\scriptstyle A}\downarrow & & \downarrow{\scriptstyle T} \\ \mathbb{F}^n & \xrightarrow[\simeq]{\Psi} & W \end{array} \tag{1.7}$$

namely, $T\Phi = \Psi A$. This commutative diagram tells us that T is equivalent to its **matrix representation** A under the bases $e_\bullet, \varepsilon_\star$, and that this equivalence is implemented by the linear isomorphisms Φ (on the sources) and Ψ (on the targets).

Commutative diagrams are ubiquitous in mathematics. You should learn how to read and understand the intuitive meanings of commutative diagrams. We will see more examples in the future of this course.

[2]Again, we refer the readers to Internet or any Linear Algebra textbook (e.g. [Axl]) for the definition of vector spaces and linear maps.

1.3 $\mathbb{Q}$, $\mathbb{R}$, and $\overline{\mathbb{R}} = [-\infty, +\infty]$

Using equivalence classes, one can construct rational numbers from integers, and real numbers from rationals. We leave the latter construction to the future, and discuss the construction of rationals here.

Example 1.21 (Construction of $\mathbb{Q}$ from $\mathbb{Z}$)**.** Define a relation on $\mathbb{Z}\times\mathbb{Z}^\times$ (where $\mathbb{Z}^\times = \mathbb{Z}\backslash\{0\}$) as follows. If $(a,b),(a',b') \in \mathbb{Z}\times\mathbb{Z}^\times$, we say $(a,b) \sim (a',b')$ iff $ab' = a'b$. It is a routine check that $\sim$ is an equivalence relation. Let $\mathbb{Q} = (\mathbb{Z}\times\mathbb{Z}^\times)/\sim$, and write the equivalence class of (a,b) as a/b or $\frac{a}{b}$. Define additions and multiplications in $\mathbb{Q}$ to be

$$\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd} \qquad \frac{a}{b}\cdot\frac{c}{d} = \frac{ac}{bd}$$

We leave it to the readers to check that this definition is **well-defined**: If $(a,b) \sim (a',b')$ and $(c,d) \sim (c',d')$ then $(ad+bc,bd) \sim (a'd'+b'c',b'd')$ and $(ac,bd) \sim (a'c',b'd')$.

We regard $\mathbb{Z}$ as a subset of $\mathbb{Q}$ by identifying $n \in \mathbb{Z}$ with $\frac{n}{1}$. (This is possible since the map $n \in \mathbb{Z} \mapsto \frac{n}{1} \in \mathbb{Q}$ is injective.) Each $a/b \in \mathbb{Q}$ has additive inverse $\frac{-a}{b}$. If $a/b \in \mathbb{Q}$ is not zero (i.e. $(a,b) \not\sim (0,1)$), then a/b has multiplicative inverse b/a. This makes $\mathbb{Q}$ a field: the field of **rational numbers**.

If $a/b \in \mathbb{Q}$, we say $a/b \geqslant 0$ if $ab \geqslant 0$. Check that this is well-defined (i.e., if $(a,b) \sim (a',b')$, then $ab \geqslant 0$ iff $a'b' \geqslant 0$). More generally, if $a/b, c/d \in \mathbb{Q}$, we say $\frac{a}{b} \geqslant \frac{c}{d}$ if $\frac{a}{b} - \frac{c}{d} \geqslant 0$. Check that $\geqslant$ is a total order on $\mathbb{Q}$. Check that $\mathbb{Q}$ is an Archimedean ordered field, defined below. □

Definition 1.22. A field $\mathbb{F}$, together with a total order $\leqslant$, is called an **ordered field**, if for every $a,b,c \in \mathbb{F}$ we have

- (Addition preserves $\leqslant$) If $a \leqslant b$ then $a + c \leqslant b + c$.
- (Multiplication by $\mathbb{F}_{\geqslant 0}$ preserves $\leqslant$) If $a, b \geqslant 0$ then $ab \geqslant 0$.

These two properties relate $\leqslant$ to $+$ and $\cdot$ respectively.

Remark 1.23. Many familiar properties about inequalities in $\mathbb{Q}$ hold for an ordered field. For instance:

$$\begin{gathered}
a \geqslant b \ \wedge\ c \geqslant d \qquad \Longrightarrow \qquad a + c \geqslant b + d \\
a \geqslant 0 \qquad \Longleftrightarrow \qquad -a \leqslant 0 \\
a \geqslant 0 \ \wedge\ b \geqslant c \qquad \Longleftrightarrow \qquad ab \geqslant ac \\
a \leqslant 0 \ \wedge\ b \geqslant c \qquad \Longleftrightarrow \qquad ab \leqslant a \\
a^2 \geqslant 0 \\
0 < a \leqslant b \qquad \Longrightarrow \qquad 0 < b^{-1} \leqslant a^{-1}
\end{gathered}$$

Check them yourself, or see [Rud-P, Prop. 1.18].

Definition 1.24. We say that an ordered field $\mathbb{F}$ satisfies **Archimedean property** if for each $a, b \in \mathbb{F}$ we have

$$a > 0 \qquad \Longrightarrow \qquad \exists n \in \mathbb{N} \text{ such that } na > b$$

where na means $\underbrace{a + \cdots + a}_{n}$.

Prop. 1.27 gives an important application of Archimedian property. We will use this in the construction of $\mathbb{R}$ from $\mathbb{Q}$, and in the proof that $\mathbb{Q}$ is dense in $\mathbb{R}$.

Definition 1.25. Let $\mathbb{F}$ be a field. A subset $\mathbb{E} \subset \mathbb{F}$ is called a **subfield** of $\mathbb{F}$, if $\mathbb{E}$ contains the 1 of $\mathbb{F}$, and if $\mathbb{E}$ is closed under the operations of addition, multiplication, taking negative, and taking inverse in $\mathbb{F}$ (i.e. if $a, b \in \mathbb{E}$ then $a + b, ab, -a \in \mathbb{E}$, and $a^{-1} \in \mathbb{E}$ whenever $a \neq 0$). We also call $\mathbb{F}$ a **field extension** of $\mathbb{E}$, since $\mathbb{E}$ is clearly a field.

Note that if $\mathbb{E}$ is a subfield of $\mathbb{F}$, the 0 of $\mathbb{F}$ is in $\mathbb{E}$ since $0 = 1 + (-1) \in \mathbb{E}$.

Definition 1.26. Let $\mathbb{E}$ be an ordered field. A field extension $\mathbb{F}$ of $\mathbb{E}$ is called an **ordered field extension**, if $\mathbb{F}$ is equipped with a total order $\leqslant$ such that $\mathbb{F}$ is an ordered field, and if the order $\leqslant$ of $\mathbb{F}$ restricts to that of $\mathbb{E}$. We also call $\mathbb{E}$ an **ordered subfield** of $\mathbb{F}$.

Our typical example of ordered field extension will be $\mathbb{Q} \subset \mathbb{R}$.

Proposition 1.27. *Let $\mathbb{F}$ be an ordered field extension of $\mathbb{Q}$. Assume that $\mathbb{F}$ is Archimedean. Then for every $x, y \in \mathbb{F}$ satisfying $x < y$, there exists $p \in \mathbb{Q}$ such that $x < p < y$.*

Proof. Assume $x, y \in \mathbb{F}$ and $x < y$. Then $y - x > 0$ (since $y - x \neq 0$ and $-x + x \leqslant -x + y$). By Archimedean property, there exists $n \in \mathbb{Z}_+$ such that $n(y - x) > 1$. So $y - x > \dfrac{1}{n}$ and hence $x + \dfrac{1}{n} < y$.

Let us prove that the subset

$$A = \left\{k \in \mathbb{Z} : \frac{k}{n} \leqslant x\right\}$$

is nonempty and bounded from above in $\mathbb{Z}$. By Archimedean property, there is $m \in \mathbb{Z}_+$ such that $m > nx$, i.e. $\dfrac{m}{n} > x$. So for each $k \in \mathbb{Z}_+$ satisfying $k \geqslant m$, we have $\dfrac{k}{n} = \dfrac{m}{n} + \dfrac{k - m}{n} > x$. Therefore, for each $k \in A$ we have $k < m$. So A is bounded above. By Archimedean property again, there is $l \in \mathbb{Z}_+$ such that $\dfrac{l}{n} > -x$. So $-\dfrac{l}{n} < x$, and hence A is nonempty.

We now use the fact that *every nonempty subset of $\mathbb{Z}$ bounded above has a maximal element*. Let $k = \max A$. Since $k+1 \notin A$, we have $x < \dfrac{k+1}{n}$. Since $\dfrac{k}{n} \leqslant x$, we have

$$\frac{k+1}{n} = \frac{k}{n} + \frac{1}{n} \leqslant x + \frac{1}{n} < y$$

This proves $x < p < y$ with $p = \dfrac{k+1}{n}$. □

To introduce $\mathbb{R}$ formally, we need more definitions:

Definition 1.28. Let $(X, \leqslant)$ be a partially ordered set and $E \subset X$. An **upper bound of E in X** is an element $x \in X$ satisfying $e \leqslant x$ for all $e \in E$. An upper bound $x \in X$ of E is called a **least upper bound** or a **supremum** if $x \leqslant y$ for every upper bound $y \in Y$ of E. In this case, we write the supremum as $\sup E$. It is not hard to check that supremums are unique if they exist.

We leave it to the readers to define **lower bounds** and the **greatest lower bound** (if exists) of E, also called the **infinimum** and is denoted by $\inf E$. □

Definition 1.29. Let $(X, \leqslant)$ be a partially ordered set. We say that X satisfies the **least-upper-bound property**, if every every nonempty subset $E \subset X$ which is bounded above (i.e. E has an upper bound) has a supremum in X. The **greatest-lower-bound property** is defined in the opposite way.

Example 1.30. $\mathbb{Z}$ satisfies the least-upper-bound and the greatest-lower-bound property: Let $A \subset \mathbb{Z}$. If A is bounded above (resp. below), then the maximum $\max A$ (resp. minimum $\min A$) exists and is the supremum (resp. infinimum) of A.

Example 1.31. Let X be a set. Then $(2^X, \subset)$ satisfies the least-upper-bound and the greatest-lower-bound property: Let $\mathscr{A} \subset 2^X$, i.e., $\mathscr{A}$ is a set of subsets of X. Then $\mathscr{A}$ is bounded from above by X, and is bounded from below by $\varnothing$. Moreover,

$$\sup \mathscr{A} = \bigcup_{A \in \mathscr{A}} A \qquad \inf \mathscr{A} = \bigcap_{A \in \mathscr{A}} A$$

Theorem 1.32. *There is an ordered field extension of $\mathbb{Q}$ which is Archimedian and satisfies the least-upper-bound property. This field is denoted by $\mathbb{R}$. Its elements are called **real numbers**.*

By taking negative, we see that $\mathbb{R}$ also satisfies the greatest-lower-bound property.

Remark 1.33. The ordered field extensions satisfying the conditions in Thm. 1.32 are unique "up to isomorphisms". (The words "**isomorphism**" and "equivalence" are often interchangeable, though "isomorphism" is more often used in the

algebraic setting, whereas "equivalence" can be used in almost every context. For example, in point-set topology, "equivalence" means "homeomorphism".) We leave it to the readers to give the precise statement. We will not use this uniqueness in this course.

Note that to compare two extensions $\mathbb{F}, \mathbb{R}$ of $\mathbb{Q}$, it is very confusing to regard $\mathbb{Q}$ as a subset of both $\mathbb{F}$ and $\mathbb{R}$. You'd better consider two different injective maps $\tau : \mathbb{Q} \to \mathbb{F}$ and $\iota : \mathbb{Q} \to \mathbb{R}$ preserving the algebraic operations and the order of $\mathbb{Q}$, and use a commutative diagram to indicate that τ and ι are equivalent. (Thus, what's happening here is that we have an equivalence of maps, not just an equivalence of the fields $\mathbb{F}$ and $\mathbb{R}$.) □

Definition 1.34. Let $-\infty, +\infty$ be two different symbols, and extend the total order $\leqslant$ of $\mathbb{R}$ to the **extended real line**

$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$$

by letting $-\infty < x < +\infty$ for all $x \in \mathbb{R}$. Define for each $x \in \mathbb{R}$ that

$$x \pm \infty = \pm\infty + x = \pm\infty \qquad +\infty - (-\infty) = +\infty$$
$$x \cdot (\pm\infty) = \pm\infty \cdot x = \begin{cases} \pm\infty & \text{if } x > 0 \\ \mp\infty & \text{if } x < 0 \end{cases}$$
$$\frac{x}{\pm\infty} = 0$$
$$\frac{\pm\infty}{x} = x^{-1} \cdot (\pm\infty) \qquad \text{if } x \neq 0$$

If $a, b \in \overline{\mathbb{R}}$ and $a \leqslant b$, we define **intervals** with endpoints a, b:

$$\begin{aligned} [a,b] &= \{x \in \overline{\mathbb{R}} : a \leqslant x \leqslant b\} \qquad (a,b) = \{x \in \overline{\mathbb{R}} : a < x < b\} \\ (a,b] &= \{x \in \overline{\mathbb{R}} : a < x \leqslant b\} \qquad [a,b) = \{x \in \overline{\mathbb{R}} : a \leqslant x < b\} \end{aligned} \tag{1.8}$$

So $\mathbb{R} = (-\infty, +\infty)$ and $\overline{\mathbb{R}} = [-\infty, +\infty]$.

In this course, unless otherwise stated, an interval always means one of the four sets in (1.8). The first two intervals are called respectively a **closed interval** and an **open interval**.

Remark 1.35. Clearly, every subset E of $\overline{\mathbb{R}}$ is bounded and has a supremum and an infinimum. We have that $\sup E = +\infty$ iff E is not bounded above in $\mathbb{R}$, and that $\inf E = -\infty$ iff E is not bounded below in $\mathbb{R}$.

1.4 Cardinalities, countable sets, and product spaces Y^X

Definition 1.36. Let A and B be sets. We say that A and B have the same **cardinality** and write $\operatorname{card}(A) = \operatorname{card}(B)$ (or simply $A \approx B$), if there is a bijection $f : A \to B$. We write $\operatorname{card}(A) \leqslant \operatorname{card}(B)$ (or simply $A \lesssim B$) if A and a subset of B have the same cardinality.

Exercise 1.37. Show that $\mathrm{card}(A) \leqslant \mathrm{card}(B)$ iff there is an injection $f : A \to B$, iff there is a surjection $g : B \to A$. (You need either Axiom of Choice or its consequence (1.5b) to prove the last equivalence.)

It is clear that $\approx$ is an equivalence relation on the collection of sets. It is also true that $\lesssim$ is a partial order: Reflexivity and transitivity are easy to show. The proof of antisymmetry is more involved:

Theorem 1.38 (Schröder-Bernstein)**.** *Let A, B be two sets. If $A \lesssim B$ and $B \lesssim A$, then $A \approx B$.*

Proof $\star\star$. Assume WLOG that $A \subset B$. Let $f : B \to A$ be an injection. Let $A_n = f^n(A)$ defined inductively by $f^0(A) = A$, $f^n(A) = f(f^{n-1}(A))$. Let $B_n = f^n(B)$. Then

$$B_0 \supset A_0 \supset \cdots \supset B_n \supset A_n \supset B_{n+1} \supset \cdots$$

In particular, $C = \bigcap_{n\in\mathbb{N}} A_n$ equals $\bigcap_{n\in\mathbb{N}} B_n$. Note that f gives a bijection $B_n \backslash A_n \to B_{n+1}\backslash A_{n+1}$ (since f gives bijections $B_n \to B_{n+1}$ and $A_n \to A_{n+1}$). Therefore, we have a bijection $g : B \to A$ defined by

$$g(x) = \begin{cases} f(x) & \text{if } x \in B_n\backslash A_n \text{ for some } n \in \mathbb{N} \\ x & \text{otherwise} \end{cases}$$

where "otherwise" means either $x \in C$ or $x \in A_n\backslash B_{n+1}$ for some n. □

Intuition about the above proof: View B as an onion. The layers of B are $B_n\backslash A_n$ (the odd layers) and $A_n\backslash B_{n+1}$ (the even layers). The bijection g maps each odder layer to the subsequent odd one, and fixes the even layers and the core C.

Example 1.39. If $-\infty < a < b < +\infty$, then $(0, 1) \approx (a, b)$.

Proof. $f : (0, 1) \to (a, b)$ sending x to $(b - a)x + a$ is a bijection. □

Example 1.40. If $-\infty < a < b < +\infty$, then $\mathbb{R} \approx (a, b)$

Proof. By the previous example, it suffices to prove $\mathbb{R} \approx (-2, 2)$. Define $f : \mathbb{R} \to \mathbb{R}$ as follows. We let $f(0) = 0$. Let $n \in \mathbb{N}$ and suppose that $f(n)$ has been defined. Then $f|_{(n,n+1]}$ is $f(x) = 2^{-n} \cdot (x - n) + f(n)$. This defines f on $[0, +\infty)$. On $(-\infty, 0)$, f is defined by $f(x) = -f(-x)$. Then f gives a bijection $\mathbb{R} \to (-2, 2)$, finishing the proof. (Checking the surjectivity of f is tedious. But it suffices to check that f is injective, and conclude the proof using Schröder-Bernstein.) □

Alternatively, one may use the tangent function to give a bijection between $(-\pi/2, \pi/2)$ and $\mathbb{R}$. I have avoided this method, since piecewisely linear functions are more elementary than trigonometric functions. The mathematically rigorous definition of trigonometric functions and the verification of their well-known properties are far from easy tasks.

Proposition 1.41. *Let I be an interval with endpoints $a < b$ in $\overline{\mathbb{R}}$. Then $I \approx \mathbb{R}$.*

Proof. Let $A = (0,1) \cup \{-\infty, +\infty\}$. By Exp. 1.40, we have

$$(a,b) \subset I \lesssim \mathbb{R} \lesssim \overline{\mathbb{R}} \approx A \approx [0,1] \subset (-2,2) \approx (a,b)$$

The proof is finished by Schröder-Bernstein Thm. 1.38. □

Definition 1.42. A set A is called **finite** if $A \lesssim \{1, \dots, n\}$ for some $n \in \mathbb{Z}_+$. A is called **countable** if $A \lesssim \mathbb{N}$.

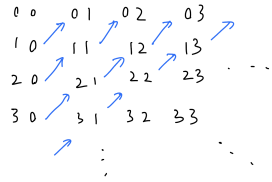
Clearly, a set A is finite iff either $A \approx \varnothing$ or $A \approx \{1, \dots, n\}$ for some $n \in \mathbb{Z}_+$.

Remark 1.43. Let $A \subset \mathbb{N}$. If A is bounded above, then $A \subset \{0, \dots, n\}$ and hence A is finite. If A is not bounded above, then we can construct a strictly increasing sequence $(x_n)_{n\in\mathbb{N}}$ in A. (Pick any $x_0 \in A$. Suppose we have $x_n \in A$. Since x_n is not an upper bounded of A, there is $x_{n+1} \in A$ larger than x_n. So $(x_n)_{n\in\mathbb{N}}$ can be constructed inductively.) This gives an injection $\mathbb{N} \to A$. Therefore $A \gtrsim \mathbb{N}$, and hence $A \approx \mathbb{N}$ by Schröder-Bernstein.

It follows that if $B \lesssim \mathbb{N}$, then either B is a finite set, or $B \approx \mathbb{N}$. Therefore, "a set B is countable and infinite" means the same as "$B \approx \mathbb{N}$". □

Theorem 1.44. *A countable union of countable sets is countable. In particular, $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$.*

Proof. Recall Exe. 1.37. Let $A_1, A_2, \dots$ be countable sets. Since each A_i is countable, there is a surjection $f_i : \mathbb{N} \to A_i$. Thus, the map $f : \mathbb{N} \times \mathbb{N} \to \bigcup_i A_i$ defined by $f(i,j) = f_i(j)$ is surjective. Therefore, it suffices to show that there is a surjection $\mathbb{N} \to \mathbb{N} \times \mathbb{N}$. This is true, since we have a bijection $g : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ where $g(0), g(1), g(2), \dots$ are $(0,0)$, $(1,0)$, $(0,1)$, $(2,0)$, $(1,1)$, $(0,2)$, $(3,0)$, $(2,1)$, $(1,2)$, $(0,3)$, etc., as shown by the figure



□

Later, when we have learned Zorn's Lemma (an equivalent form of Axiom of Choice), we will be able to prove the following generalization of $\mathbb{N} \times \mathbb{N} \approx \mathbb{N}$. So we defer the proof of the following theorem to a later section.

Theorem 1.45. *Let X be a infinite set. Then $X \times \mathbb{N} \approx X$.*

Our next goal is to prove an exponential law $a^{b+c} = a^b \cdot a^c$ for cardinalities. For that purpose, we first need to define the set-theoretic operations that correspond to the summation $b + c$ and the exponential a^b.

Definition 1.46. We write $X = \bigsqcup_{\alpha\in\mathscr{A}} A_\alpha$ and call X the **disjoint union** of $(A_\alpha)_{\alpha\in\mathscr{A}}$, if $X = \bigcup_{\alpha\in\mathscr{A}} A_\alpha$ and $(A_\alpha)_{\alpha\in\mathscr{A}}$ is a family of pairwise disjoint sets (i.e. $A_\alpha \cap A_\beta = \varnothing$ if $\alpha \neq \beta$). If moreover $\mathscr{A} = \{1, \dots, n\}$, we write $X = A_1 \sqcup \cdots \sqcup A_n$.

Definition 1.47. Let X, Y be sets. Then

$$Y^X = \{\text{functions } f : X \to Y\} \tag{1.9}$$

A more precise definition of Y^X (in the spirit of (1.3)) is $\{f \in X \times Y \mid f : X \to Y \text{ is a function}\}$. In the special case that $X = \varnothing$, we set

$$Y^\varnothing = \{\varnothing\} \tag{1.10}$$

Namely, there is precisely one function $\varnothing \to Y$, which is $\varnothing$ as a subset of $\varnothing \times Y$.

This new notation is compatible with the old one $Y^n = Y \times \cdots \times Y$:

Example 1.48. Let $n \in \mathbb{Z}_+$. We have $Y^{\{1,\dots,n\}} \approx Y^n$ due to the bijection

$$Y^{\{1,\dots,n\}} \to Y^n \qquad f \mapsto (f(1), \dots, f(n))$$

Remark 1.49. The above example suggests that in the general case that X is not necessarily finite, we can view each function $f : X \to Y$ as $(f(x))_{x\in X}$, an **indexed family of elements** of Y with index set X. Thus, intuitively and hence not quite rigorously,

$$Y^X = \underbrace{Y \times Y \times \cdots}_{\text{card}(X) \text{ pieces}} \tag{1.11}$$

This generalizes the intuition in Def. 1.7 that a function $f : \mathbb{Z}_+ \to Y$ is equivalently a sequence $(f(1), f(2), f(3), \dots)$.

The viewpoint that Y^X is a **product space** with index set X is very important and will be adopted frequently in this course. □

Example 1.50. Let X be a set. For each $A \subset X$, define the **characteristic function** $\chi_A : X \to \{0, 1\}$ to be

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

Then we have $2^X \approx \{0, 1\}^X$ since the following map is bijective:

$$2^X \to \{0, 1\}^X \qquad A \mapsto \chi_X$$

Its inverse is $f \in \{0, 1\}^X \mapsto f^{-1}(1) \in 2^X$.

Proposition 1.51 (Exponential Law). *Suppose that $X = A_1 \sqcup \cdots \sqcup A_n$. Then*

$$Y^X \approx Y^{A_1} \times \cdots \times Y^{A_n}$$

Proof. We have a bijection

$$\begin{aligned} \Phi : Y^X &\to Y^{A_1} \times \cdots \times Y^{A_n} \\ f &\mapsto (f|_{A_1}, \ldots, f|_{A_n}) \end{aligned} \tag{1.12}$$

where we recall that $f|_{A_i}$ is the restriction of f to A_i. □

Exercise 1.52. Assume that $A_1, \ldots, A_n$ are subsets of X. Define Φ by (1.12). Prove that Φ is injective iff $X = A_1 \cup \cdots \cup A_n$. Prove that Φ is surjective iff $A_1, \ldots, A_n$ are pairwise disjoint.

Corollary 1.53. *Let X, Y be finite sets with cardinalities $m, n \in \mathbb{N}$ respectively. Assume that $Y \neq \varnothing$. Then Y^X is a finite set with cardinality n^m.*

Proof. The special case that $m = 0$ (i.e. $X = \varnothing$, cf. (1.10)) and $m = 1$ is clear. When $m > 1$, assume WLOG that $X = \{1, \ldots, m\}$. Then $X = \{1\} \sqcup \cdots \sqcup \{m\}$. Apply Prop. 1.51 to this disjoint union. We see that $Y^X \simeq Y \times \cdots \times Y \simeq \{1, \ldots, n\}^m$ has n^m elements. □

We end this section with some (in)equalities about the cardinalities of product spaces. To begin with, we write $X \precnsim Y$ (or $\mathrm{card}(X) < \mathrm{card}(Y)$) if $X \precsim Y$ and $X \not\approx Y$.

Proposition 1.54. *Let X, Y be sets with $\mathrm{card}(Y) \geqslant 2$ (i.e. Y has at least two elements). Then $X \precnsim Y^X$. In particular, $X \precnsim 2^X$.*

Proof. The case $X = \varnothing$ is obvious since $0 < 1$. So we assume $Y \neq \varnothing$. Clearly $2^X \simeq \{0, 1\}^X$ is $\precsim Y^X$. So it suffices to prove $X \precnsim 2^X$. Since the map $X \to 2^X$ sending x to $\{x\}$ is injective, $X \precsim 2^X$. Let us prove $X \not\approx 2^X$.

Assume that $X \approx 2^X$. So there is a bijection $\Phi : X \to 2^X$ sending each $x \in X$ to a subset $\Phi(x)$ of X. Motivated by Russell's Paradox (1.2), we define

$$S = \{x \in X : x \notin \Phi(x)\}$$

Since Φ is surjective, there exists $y \in X$ such that $S = \Phi(y)$. If $y \in \Phi(y)$, then $y \in S$, and hence $y \notin \Phi(y)$ by the definition of S. If $y \notin \Phi(y)$, then $y \notin S$, and hence $y \in \Phi(y)$ by the definition of S. This gives a contradiction. □

Remark 1.55. Write $\{1, \ldots, n\}^X$ as n^X for short. Assuming that real numbers have decimal, binary, or (more generally) base-n presentations where $n \in \mathbb{Z}_{\geqslant 2}$, then $\mathbb{R} \approx n^{\mathbb{N}}$. So by Prop. 1.54, $\mathbb{N} \precnsim \mathbb{R}$, i.e. $\mathbb{R}$ *is uncountable*. The base-n presentations of real numbers suggest that $\mathrm{card}(n^{\mathbb{N}})$ is independent of n. This fact can be proved by elementary methods without resorting to the analysis of real numbers:

Theorem 1.56. *Let X be an infinite set. Then*

$$2^X \approx 3^X \approx 4^X \approx \cdots \approx \mathbb{N}^X$$

Proof. First, we assume that $X = \mathbb{N}$. Clearly, for each $n \in \mathbb{Z}_{\geqslant 2}$ we have $2^X \lesssim n^X \lesssim \mathbb{N}^X$. Since elements of $\mathbb{N}^X$ are subsets of $X \times \mathbb{N}$ (i.e. elements of $2^{X\times\mathbb{N}}$), we have

$$\mathbb{N}^X \subset 2^{X\times\mathbb{N}} \simeq 2^X$$

since $X \times \mathbb{N} \approx X$ by Thm. 1.44. So $2^X \approx n^X \approx \mathbb{N}^X$ by Schröder-Bernstein.

As pointed out earlier (cf. Thm. 1.45), it can be proved by Zorn's Lemma that $X \times \mathbb{N} \approx X$ for every infinite set X. So the same conclusion holds for such X. □

Index

References

[Axl] Axler, S. (2015). Linear algebra done right. 3rd ed.
[Mun] Munkres, J. (2000). Topology. Second Edition.
[Rud-P] Rudin, W. (1976). Principles of Mathematical Analysis. 3rd ed.

YAU MATHEMATICAL SCIENCES CENTER, TSINGHUA UNIVERSITY, BEIJING, CHINA.
E-mail: binguimath@gmail.com bingui@tsinghua.edu.cn