

# PMATH 351: Real Analysis

Instructor Stephen New  
 $\text{\LaTeX}$ er Iris Jiang

Spring 2020

## Contents

1	Cardinality .....	2
2	Metric Spaces .....	12

# 1. Cardinality

## Definition 1.1 — domain, range, image, inverse image.

Let  $X$  and  $Y$  be sets and let  $f : X \rightarrow Y$ . Recall the **domain** of  $f$  and the **range** of  $f$  are the sets

$$\text{Domain}(f) = X, \text{Range}(f) = f(X) = \{f(x) | x \in X\}$$

for  $A \subseteq X$ , the **image** of  $A$  under  $f$  is the set

$$f(A) = \{f(x) | x \in A\}$$

For  $B \subseteq Y$ , the **inverse image** of  $B$  under  $f$  is the set

$$f^{-1}(B) = \{x \in X | f(x) \in B\}$$

## Definition 1.2 — Composite.

Let  $X, Y$  and  $Z$  be sets, let  $f : X \rightarrow Y$  and let  $g : Y \rightarrow Z$ . We define the **composite** function  $(g \circ f)(x) = g(f(x))$  for all  $x \in X$

## Definition 1.3 — injective, surjective, bijective.

We say that  $f$  is **injective** (or **one-to-one**) when for every  $y \in Y$  there exists **at most** one  $x \in X$  such that  $f(x) = y$ . Equivalently,  $f$  is injective when for all  $x_1, x_2 \in X$ , if  $f(x_1) = f(x_2)$  then  $x_1 = x_2$ .

We say that  $f$  is **surjective** (or **onto**) when for every  $y \in Y$  there exists **at least** one  $x \in X$  such that  $f(x) = y$ . Equivalently,  $f$  is surjective when  $\text{Range}(f) = Y$

We say that  $f$  is **bijective** (or **invertible**) when  $f$  is both injective and surjective, that is when for every  $y \in Y$  there exists exactly one  $x \in X$  such that  $f(x) = y$ . When  $f$  is both injective and surjective, that is when for every  $y \in Y$  there exists exactly one  $x \in X$  such that  $f^{-1} : Y \rightarrow X$  such that for all  $y \in Y$ ,  $f^{-1}(y)$  is equal to the unique element  $x \in X$  such that  $f(x) = y$ . Note that when  $f$  is bijective so is  $f^{-1}$ , and in this case we have  $(f^{-1})^{-1} = f$

**Theorem 1.1** Let  $f : X \rightarrow Y$  and let  $g : Y \rightarrow Z$ . Then

- (1) If  $f$  and  $g$  are both injective then so is  $g \circ f$
- (2) If  $f$  and  $g$  are both surjective then so is  $g \circ f$
- (3) If  $f$  and  $g$  are both invertible then so is  $g \circ f$ , and in this case  $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$

*Proof.*

- (1) Suppose that  $f$  and  $g$  are both injective. Let  $x_1, x_2 \in X$ . If  $g(f(x_1)) = g(f(x_2))$  then since  $g$  is injective we have  $f(x_1) = f(x_2)$ , and then since  $f$  is injective we have  $x_1 = x_2$ . Thus  $g \circ f$  is injective.
- (2) Suppose that  $f$  and  $g$  are both surjective. Given  $z \in Z$ , since  $g$  is surjective we can choose  $y \in Y$  so that  $g(y) = z$ , then since  $f$  is surjective we can choose  $x \in X$  so that  $f(x) = y$ , and then we have  $g(f(x)) = g(y) = z$ . Thus  $g \circ f$  is surjective.
- (3) Follows (1) and (2). ■

**Definition 1.4 — identity function.**

For a set  $X$ , we define the **identity function** on  $X$  to be the function  $I_X : X \rightarrow X$  given by  $I_X(x) = x$  for all  $x \in X$ . Note that for  $f : X \rightarrow Y$  we have  $f \circ I_X = f$  and  $I_Y \circ f = f$ .

**Definition 1.5 — inverse.**

Let  $X$  and  $Y$  be sets and let  $f : X \rightarrow Y$ . A **left inverse** of  $f$  is a function  $g : Y \rightarrow X$  given by  $g \circ f = I_X$ . Equivalently, a function  $g : Y \rightarrow X$  is a left inverse of  $f$  when  $g(f(x)) = x$  for all  $x \in X$ .

A **right inverse** of  $f$  is a function  $h : Y \rightarrow X$  such that  $f \circ h = I_Y$ . Equivalently, a function  $h : Y \rightarrow X$  is a right inverse of  $f$  when  $f(h(y)) = y$  for all  $y \in Y$ .

**Theorem 1.2** Let  $X$  and  $Y$  be nonempty sets and let  $f : X \rightarrow Y$ . Then

- (1)  $f$  is injective  $\iff f$  has a left inverse.
- (2)  $f$  is surjective  $\iff f$  has a right inverse.
- (3)  $f$  is bijective  $\iff f$  has a left inverse  $g$  and a right inverse  $h$ , and in this case we have  $g = h = f^{-1}$ .

*Proof.*

- (1) Suppose first that  $f$  is injective. Since  $X \neq \emptyset$  we can choose  $a \in X$  and then define  $g : Y \rightarrow X$  as follows: if  $y \in \text{Range}(f)$  then (using the fact the  $f$  is injective) we define  $g(y)$  to be the unique element  $x_y \in X$  with  $f(x_y) = y$ , and if  $y \notin \text{Range}(f)$ , then we define  $g(y) = a$ . Then for every  $x \in X$  we have  $y = f(x) \in \text{Range}(f)$ , so  $g(y) = x_y = x$ , that is  $g(f(x)) = x$ . Conversely, if  $f$  has a left inverse, say  $g$ , then  $f$  is injective since for all  $x_1, x_2 \in X$ , if  $f(x_1) = f(x_2)$  then  $x = g(f(x_1)) = g(f(x_2)) = x_2$ .
- (2) Suppose first that  $f$  is onto. For each  $y \in Y$ , choose  $x_y \in X$  with  $f(x_y) = y$ , then define  $g : X \rightarrow Y$  by  $g(y) = x_y$  (We need the Axiom of Choice for this). Then  $g$  is a right inverse of  $f$  since for every  $y \in Y$  we have  $f(g(y)) = f(x_y) = y$ . Conversely, if  $f$  has a right inverse, say  $g$ , then  $f$  is onto since given any  $y \in Y$  we can choose  $x = g(y)$  and then we have  $f(x) = f(g(y)) = y$ .
- (3) Suppose first that  $f$  is bijective. The inverse function  $f^{-1} : Y \rightarrow X$  is a left inverse for  $f$  because given  $x \in X$  we can let  $y = f(x)$  and then  $f^{-1}(y) = x$  so that  $f^{-1}(f(x)) = f^{-1}(y) = x$ . Similarly,  $f^{-1}$  is a right inverse for  $f$  because given  $y \in Y$  we can let  $x$  be the unique element in  $X$  with  $y = f(x)$  and then we have  $x = f^{-1}(y)$  so that  $f(f^{-1}(y)) = f(x) = y$ . Conversely, suppose that  $g$  is a left inverse for  $f$  and  $h$

is a right inverse for  $f$ . Since  $f$  has a left inverse, it is injective by (1). Since  $f$  has a right inverse, it is surjective by (2). Since  $f$  is injective and surjective, it is bijective. As shown above, the inverse function  $f^{-1}$  is both a left inverse and a right inverse. Finally, note that  $g = f^{-1} = h$  because for all  $y \in Y$  we have

$$g(y) = g(f(f^{-1}(y))) = f^{-1}(y) = f^{-1}(f(h(y))) = h(y)$$

■

### Corollary 1.3

Let  $X$  and  $Y$  be sets. Then there exists an injective map  $f : X \rightarrow Y$  if and only if there exists a surjective map  $g : Y \rightarrow X$ .

*Proof.* Suppose  $f : X \rightarrow Y$  is an injective map. Then  $f$  has a left inverse. Let  $g$  be a left inverse of  $f$ . Since  $g \circ f = I_X$ , we see that  $f$  is a right inverse of  $g$ . Since  $g$  has a right inverse,  $g$  is surjective. Thus, there is a surjective map  $g : Y \rightarrow X$ . Similarly, if  $g : Y \rightarrow X$  is surjective, then it has a right inverse  $f : X \rightarrow Y$  which is injective. ■

### Definition 1.6 — same cardinality, less than or equal to, less than.

Let  $A$  and  $B$  be sets. We say that  $A$  and  $B$  have the **same cardinality**, and write  $|A| = |B|$ , when there exists a bijective map:  $f : A \rightarrow B$  (or equivalently when there exists a bijective map  $g : B \rightarrow A$ ).

We say that the cardinality of  $A$  is **less than or equal to** the cardinality of  $B$ , and write  $|A| \leq |B|$ , when there exists an injective map  $f : A \rightarrow B$  (or equivalently a surjective map  $g : B \rightarrow A$ ).

We say that the cardinality of  $A$  is **less than** the cardinality of  $B$ , and write  $|A| < |B|$ , when  $|A| \leq |B|$  and  $|A| \neq |B|$ , (that is when there exists an injective map  $f : A \rightarrow B$  but there does not exist a bijective map  $g : A \rightarrow B$ ).

We also write  $|A| \geq |B|$  when  $|B| \leq |A|$ ; and  $|A| > |B|$  when  $|B| < |A|$ .

■ **Example 1.1** Let  $\mathbb{N} = \{n \in \mathbb{Z} | n \geq 0\} = \{0, 1, 2, \dots\}$ .

- (1) The map  $f : \mathbb{N} \rightarrow 2\mathbb{N}$  given by  $f(k) = 2k$  is bijective, so  $|2\mathbb{N}| = |\mathbb{N}|$ .
- (2) The map  $g : \mathbb{N} \rightarrow \mathbb{Z}$  given by  $g(2k) = k$  and  $g(2k+1) = -k-1$  for  $k \in \mathbb{N}$  is bijective, so we have  $|\mathbb{Z}| = |\mathbb{N}|$ .
- (3) The map  $h : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$  given by  $h(k, l) = 2^k(2l+1) - 1$  is bijective, so we have  $|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$ .

### Theorem 1.4 For all sets $A$ , $B$ and $C$

- (1)  $|A| = |A|$
- (2) If  $|A| = |B|$  then  $|B| = |A|$
- (3) If  $|A| = |B|$  and  $|B| = |C|$ , then  $|A| = |C|$
- (4)  $|A| \leq |B| \iff (|A| = |B| \text{ or } |A| < |B|)$
- (5) If  $|A| \leq |B|$  and  $|B| \leq |C|$ , then  $|A| \leq |C|$

*Proof.*

- (1) holds because the identity function  $I_A : A \rightarrow A$  is bijective.
- (2) holds because if  $f : A \rightarrow B$  is bijective then so is  $f^{-1} : B \rightarrow A$ .
- (3) holds because if  $f : A \rightarrow B$  and  $g : B \rightarrow C$  are bijective then so is the composite  $g \circ f : A \rightarrow C$

■

**Definition 1.7 — finite, infinite, countable.**

Let  $A$  be a set. For each  $n \in \mathbb{N}$ , let  $S_n = \{0, 1, 2, \dots, n-1\}$ . For  $n \in \mathbb{N}$ , we say that the cardinality of  $A$  is equal to  $n$ , or that  $A$  has  $n$  **elements**, and we write  $|A| = n$ , when  $|A| = |S_n|$ .

We say that  $A$  is **finite** when  $|A| = n$  for some  $n \in \mathbb{N}$ . We say  $A$  is **infinite** when  $A$  is not finite. We say that  $A$  is **countable** when  $|A| = |\mathbb{N}|$ .

**Note 1.1** When a set  $A$  is finite with  $|A| = n$ , and when  $f : A \rightarrow S_n$  is a bijection, if we let  $a_k = f^{-1}(k)$  for each  $k \in S_n$  then we have  $A = \{a_0, a_1, \dots, a_{n-1}\}$  with the elements  $a_k$  distinct. Conversely, if  $A = \{a_0, a_1, \dots, a_{n-1}\}$  with the elements  $a_k$  all distinct, then we define a bijection  $f : A \rightarrow S_n$  by  $f(a_k) = k$ . Thus we see that  $A$  is finite with  $|A| = n$  if and only if  $A$  is of the form  $A = \{a_0, a_1, \dots, a_{n-1}\}$  with the elements  $a_k$  all distinct. Similarly, a set  $A$  is countable if and only if  $A$  is of the form  $A = \{a_0, a_1, a_2, \dots\}$  with the elements  $a_k$  all distinct.

**Note 1.2** For  $n \in \mathbb{N}$ , if  $A$  is a finite set with  $|A| = n + 1$  and  $a \in A \setminus \{a\} = n$ . Indeed, if  $A = \{a_0, a_1, \dots, a_n\}$  with the elements  $a_i$  distinct, and if  $a = a_k$  so that we have  $A \setminus \{a\} = \{a_0, a_1, \dots, a_{k-1}, a_{k+1}, \dots, a_n\}$ , then we can define a bijection  $f : S_n \rightarrow A \setminus \{a\}$  by  $f(i) = a_i$  for  $0 \leq i < k$  and  $f(i) = a_{i+1}$  for  $k \leq i < n$ .

**Theorem 1.5** Let  $A$  be a set. Then the following are equivalent:

- (1)  $A$  is infinite
- (2)  $A$  contains a countable subset
- (3)  $|\mathbb{N}| \leq |A|$
- (4) There exists a map  $f : A \rightarrow A$  which is injective but not surjective

*Proof.*

- (1)  $\implies$  (2) Suppose  $A$  is infinite. Since  $A \neq \emptyset$  we can choose an element  $a_0 \in A$ . Since  $A \neq \{a_0\}$  we can choose an element  $a_1 \in A \setminus \{a_0\}$ . Since  $A \neq \{a_0, a_1\}$  we can choose  $a_2 \in A \setminus \{a_0, a_1\}$ . Continue this procedure: having chosen distinct elements  $a_0, a_1, \dots, a_{n-1} \in A$ , since  $A \neq \{a_0, a_1, \dots, a_{n-1}\}$  we can choose  $a_n \in A \setminus \{a_0, a_1, \dots, a_{n-1}\}$ . In this way we obtain  $\{a_0, a_1, a_2, \dots\} \subseteq A$ .
- (2)  $\iff$  (3) Suppose that  $A$  contains a countable subset, say  $\{a_0, a_1, a_2, \dots\} \subseteq A$  with the element  $a_i$  distinct. Since  $a_i$  are distinct, the map  $f : \mathbb{N} \rightarrow A$  given by  $f(k) = a_k$  is injective, and so we have  $|\mathbb{N}| \leq |A|$ . Conversely as a map from  $\mathbb{N} \rightarrow f(\mathbb{N})$  where  $f$  is bijective, so we have  $|\mathbb{N}| = |f(\mathbb{N})|$  hence  $f(\mathbb{N})$  is a countable subset of  $A$ .
- (2)  $\implies$  (4) Suppose that  $A$  has a countable subset, say  $\{a_0, a_1, a_2, \dots\} \subseteq A$  with the element  $a_i$  distinct. Define  $f : A \rightarrow A$  by  $f(a_k) = a_{k+1}$  for all  $k \in \mathbb{N}$  and by  $f(b) = b$  for all  $b \in A \setminus \{a_0, a_1, a_2, \dots\}$ . Then  $f$  is injective but not surjective (the element  $a_0$  is not in the range of  $f$ ).
- (4)  $\implies$  (1) To prove this we shall prove that if  $A$  is finite then every injective map  $f : A \rightarrow A$  is surjective. We prove this by induction on the cardinality of  $A$ .  
The only set  $A$  with  $|A| = 0$  is the set  $A \neq \emptyset$ , and then the only function  $f : A \rightarrow A$  is the empty function, which is surjective.  
Since that base case may appear too trivial, let us consider the next case. Let  $n = 1$  and let  $A$  be a set with  $|A| = 1$ , say  $A = \{a\}$ . The only function  $f : A \rightarrow A$  is the function given by  $f(a) = a$ , which is surjective.  
Let  $n \geq 1$  and suppose, inductively, that for every set  $A$  with  $|A| = n$ , every injective

map  $f : A \rightarrow A$  is surjective. Let  $B$  be a set with  $|B| = n + 1$  and let  $g : B \rightarrow B$  be injective.

Suppose, for a contradiction, that  $g$  is not surjective. Choose an element  $b \in B$  which is not in the range of  $g$  so that we have  $g : B \rightarrow B \setminus \{b\}$ . Let  $A = B \setminus \{b\}$  and let  $f : A \rightarrow A$  be given by  $f(x) = g(x)$  for all  $x \in A$ . Since  $g : B \rightarrow A$  is injective and  $f(x) = g(x)$  for all  $x \in A$ ,  $f$  is also injective. Again since  $g$  is injective, there is no element  $x \in B \setminus \{b\}$  with  $g(x) = g(b)$ , so there is no element  $x \in A$  with  $f(x) = g(b)$ , and so  $f$  is not surjective. Since  $|A| = n$ , this contradicts the induction hypothesis. Thus  $g$  must be surjective.

By the Principle of Induction, for every  $n \in \mathbb{N}$  and for every set  $A$  with  $|A| = n$ , every injective function  $f : A \rightarrow A$  is surjective. ■

### Corollary 1.6

Let  $A$  and  $B$  be sets.

- (1) If  $A$  is countable then  $A$  is infinite
- (2) When  $|A| \leq |B|$ , if  $B$  is finite so is  $A$  (equivalently if  $A$  is infinite then so is  $B$ )
- (3) If  $|A| = n$  and  $|B| = m$  then  $|A| = |B|$  if and only if  $n = m$
- (4) If  $|A| = n$  and  $|B| = m$  then  $|A| \leq |B|$  if and only if  $n \leq m$
- (5) When one of the two sets  $A$  and  $B$  is finite, if  $|A| \leq |B|$  and  $|B| \leq |A|$  then  $|A| = |B|$

*Proof.*

- (1) If  $A$  is countable then  $A$  contains a countable subset (itself), so  $A$  is infinite by Theorem 1.5.
- (2) Suppose that  $|A| \leq |B|$  and that  $|A|$  is infinite. Since  $A$  is infinite, we have  $|\mathbb{N}| \leq |A|$  (by Theorem 1.5). Since  $|\mathbb{N}| \leq |A|$  and  $|A| \leq |B|$  we have  $|\mathbb{N}| \leq |B|$  (by Theorem 1.4). Since  $|\mathbb{N}| \leq |B|$ ,  $B$  is infinite (by Theorem 1.5).
- (3) Suppose that  $|A| = n$  and  $|B| = m$ . If  $n = m$  then we have  $S_n = S_m$  and so  $|A| = |S_n| = |S_m| = |B|$ . Conversely, suppose that  $|A| = |B|$ . Suppose, for a contradiction, that  $n \neq m$ , say  $n > m$ , and note that  $S_m \subsetneq S_n$ . Since  $|A| = |B|$  we have  $|S_n| = |A| = |B| = |S_m|$  so we must have  $n = m$ .
- (4) Suppose  $|A| = n$  and  $|B| = m$ . If  $n \leq m$  then  $S_n \subseteq S_m$  so the inclusion map  $I : S_n \rightarrow S_m$  is injective and we have  $|A| = |S_n| \leq |S_m| = |B|$ . Conversely, suppose that  $|A| \leq |B|$  and suppose, for a contradiction, that  $n > m$ . Since  $|A| \leq |B|$  we have  $|S_n| = |A| \leq |B| = |S_m|$  so we can choose an injective map  $f : S_n \rightarrow S_m$ . Since  $n > m$  we have  $S_m \subsetneq S_n$  so we can consider  $f$  as a map  $f : S_n \rightarrow S_m$ , and this map is injective but not surjective. This contradicts Theorem 1.5, and so  $n \leq m$ .
- (5) Suppose that one of the two sets  $A$  and  $B$  is finite, and that  $|A| \leq |B|$  and  $|B| \leq |A|$ . If  $A$  is finite then, since  $|B| \leq |A|$ , (2) implies that  $B$  is finite. If  $B$  is finite then, since  $|A| \leq |B|$ , (2) implies that  $A$  is finite. Thus, in either case, we see that  $A$  and  $B$  are both finite. Since  $A$  and  $B$  are both finite with  $|A| \leq |B|$  and  $|B| \leq |A|$ , we must have  $|A| = |B|$  by (3) and (4). ■

**Theorem 1.7** Let  $A$  be a set. Then  $|A| \leq |\mathbb{N}| \iff A$  is finite or countable.

*Proof.* First we claim that every subset of  $\mathbb{N}$  is either finite or countable. Let  $A \subseteq \mathbb{N}$  and suppose that  $A$  is not finite.

Since  $A \neq \emptyset$ , we can set  $a_0 = \min\{A\}$  (using the Well-Ordering Property of  $\mathbb{N}$ ). Note that

$\{0, 10, \dots, a_0\} \cap A = \{a_0\}$ .

Since  $A \neq \{a_0\}$  (so the set  $A \setminus \{a_0\}$  is nonempty), we can set  $a_1 = \min\{A \setminus \{a_0\}\}$ . Then we have  $a_0 < a_1$  and  $\{0, 1, \dots, a_1\} \cap A = \{a_0, a_1\}$ .

Since  $A \neq \{a_0, a_1\}$  we can set  $a_2 = \min\{A \setminus \{a_0, a_1\}\}$ . Then we have  $a_0 < a_1 < a_2$  and  $\{0, 1, 2, \dots, a_2\} \cap A = \{a_0, a_1, a_2\}$ .

We continue the procedure: having chosen  $a_0, a_1, \dots, a_{n-1} \in A$  with  $a_0 < a_1 < \dots < a_{n-1}$  such that  $\{0, 1, \dots, a_{n-1}\} \cap A = \{a_0, a_1, \dots, a_{n-1}\}$ . Since  $A \neq \{a_0, a_1, \dots, a_{n-1}\}$ , we can set  $a_n = \min\{A \setminus \{a_0, a_1, \dots, a_{n-1}\}\}$  and then we have  $a_0 < a_1 < \dots < a_{n-1} < a_n$  and  $\{0, 1, \dots, a_n\} \cap A = \{a_0, a_1, \dots, a_n\}$ .

In this way, we obtain a countable set  $\{a_0, a_1, a_2, \dots\} \subseteq A$  with  $a_0 < a_1 < a_2 < \dots$  with the property that for all  $m \in \mathbb{N}$ ,  $\{0, 1, 2, \dots, a_m\} \cap A = \{a_0, a_1, \dots, a_m\}$ .

Since  $0 \leq a_0 < a_1 < a_2 < \dots$ , it follows (by induction) that  $a_k \geq k$  for all  $k \in \mathbb{N}$ . It follows in turn that  $A \subseteq \{a_0, a_1, a_2, \dots\}$  because given  $m \in A$ , since  $m \leq a_m$  we have

$$m \in \{0, 1, 2, \dots, m\} \cap A \subseteq \{0, 1, 2, \dots, a_m\} \cap A = \{a_0, a_1, \dots, a_m\}.$$

Thus  $A = \{a_0, a_1, a_2, \dots\}$  and the elements  $a_i$  are distinct, so  $A$  is countable. This proves our claim that every subset of  $\mathbb{N}$  is either finite or countable.

Suppose that  $|A| \leq |\mathbb{N}|$  and choose an injective map  $f : A \rightarrow \mathbb{N}$ . Since  $f$  is injective, when we consider it as a map  $f : A \rightarrow f(A)$ , it is bijective, and so  $|A| = |f(A)|$ . Since  $f(A) \subseteq \mathbb{N}$ , the previous paragraph shows that  $f(A)$  is either finite or countable. If  $f(A)$  is finite with  $|f(A)| = n$  then  $|A| = |f(A)| = |S_n|$ , and if  $f(A)$  is countable then we have  $|A| = |f(A)| = |\mathbb{N}|$ . Thus  $A$  is finite or countable. ■

**Theorem 1.8** Let  $A$  be a set. Then

- (1)  $|A| < |\mathbb{N}| \iff A$  is finite
- (2)  $|\mathbb{N}| < |A| \iff A$  is neither finite nor countable
- (3) if  $|A| \leq |\mathbb{N}|$  and  $|\mathbb{N}| \leq |A|$  then  $|A| = |\mathbb{N}|$

*Proof.*

(1) By Theorem 1.5

$$\begin{aligned} |A| < |\mathbb{N}| &\iff (|A| \leq |\mathbb{N}| \text{ and } |A| \neq |\mathbb{N}|) \\ &\iff (A \text{ is finite or countable and } A \text{ is not countable}) \\ &\iff A \text{ is finite} \end{aligned}$$

(2) By Theorem 1.7

$$\begin{aligned} |\mathbb{N}| < |A| &\iff (|\mathbb{N}| \leq |A| \text{ and } |\mathbb{N}| \neq |A|) \\ &\iff (A \text{ is not finite and } A \text{ is not countable}) \end{aligned}$$

(3) Suppose that  $|A| \leq |\mathbb{N}|$  and  $|\mathbb{N}| \leq |A|$ . Since  $|A| \leq |\mathbb{N}|$ , we know that  $A$  is finite or countable by Theorem 1.7. Since  $|\mathbb{N}| \leq |A|$ , we know that  $A$  is infinite by Theorem 1.5. Since  $A$  is finite or countable and  $A$  is not finite, it follows that  $A$  is countable. Thus  $|A| = |\mathbb{N}|$ . ■

**Definition 1.8 — at most countable, uncountable.**

Let  $A$  be a set. When  $A$  is countable we write  $|A| = \aleph_0$ . When  $A$  is finite we write  $|A| < \aleph_0$ . When  $A$  is infinite we write  $|A| \geq \aleph_0$ . When  $A$  is either finite or countable we write  $|A| \leq \aleph_0$  and we say that  $A$  is **at most countable**. When  $A$  is neither finite nor



countable we write  $|A| > \aleph_0$  and we say that  $A$  is **uncountable**.

### Theorem 1.9

- (1) If  $A$  and  $B$  are countable sets, then so is  $A \times B$
- (2) If  $A$  and  $B$  are countable sets, then so is  $A \cup B$
- (3) If  $A_0, A_1, A_2, \dots$  are countable sets, then so is  $\bigcap_{k=0}^{\infty} A_k$
- (4)  $\mathbb{Q}$  is countable

*Proof.*

- (1) Let  $A = \{a_0, a_1, a_2, \dots\}$  with the  $a_i$  distinct and let  $B = \{b_0, b_1, b_2, \dots\}$  with  $b_i$  distinct. Since every positive integer can be written uniquely in the form  $2^k(2l+1)$  with  $k, l \in \mathbb{N}$ , the map  $f : A \times B \rightarrow \mathbb{N}$  given by  $f(a_k, b_l) = 2^k(2l+1) - 1$  is bijective, and so  $|A \times B| = |\mathbb{N}|$
- (2) Similar to (1), since the map  $g : \mathbb{N} \rightarrow A \cup B$  given by  $g(k) = a_k$  is injective, we have  $|\mathbb{N}| \leq |A \cup B|$ . Since the map  $h : \mathbb{N} \rightarrow A \cup B$  given by  $h(2k) = a_k$  and  $h(2k+1) = b_k$  is surjective, we have  $|A \cup B| \leq |\mathbb{N}|$ . Since  $|\mathbb{N}| \leq |A \cup B|$  and  $|A \cup B| \leq |\mathbb{N}|$ , we have  $|A \cup B| = |\mathbb{N}|$  by Theorem 1.8
- (3) For each  $k \in \mathbb{N}$ , let  $A_k = \{a_{k0}, a_{k1}, a_{k2}, \dots\}$  with the  $a_{ki}$  distinct. Since the map  $f : \mathbb{N} \rightarrow \bigcap_{k=0}^{\infty} A_k$  given by  $f(k) = a_{0,k}$  is injective,  $|\mathbb{N}| \leq \left| \bigcap_{k=0}^{\infty} A_k \right|$ . Since  $\mathbb{N} \times \mathbb{N}$  is countable by (1), and since the map  $g : \mathbb{N} \times \mathbb{N} \rightarrow \bigcap_{k=0}^{\infty} A_k$  given by  $g(k, l) = a_{k,l}$  is surjective, we have  $\left| \bigcap_{k=0}^{\infty} A_k \right| \leq |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$ . By Theorem 1.8, we have  $\left| \bigcap_{k=0}^{\infty} A_k \right| = |\mathbb{N}|$ .
- (4) Since the map  $f : \mathbb{N} \rightarrow \mathbb{Q}$  given by  $f(k) = k$  is injective, we have  $|\mathbb{N}| \leq |\mathbb{Q}|$ . Since the map  $g : \mathbb{Q} \rightarrow \mathbb{Z} \times \mathbb{Z}$  given by  $g(\frac{a}{b}) = (a, b)$  for all  $a, b \in \mathbb{Z}$  with  $b > 0$  and  $\gcd(a, b) = 1$ , is injective, and since  $\mathbb{Z} \times \mathbb{Z}$  is countable, we have  $|\mathbb{Q}| \leq |\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N}|$ . Since  $|\mathbb{N}| \leq |\mathbb{Q}|$  and  $|\mathbb{Q}| \leq |\mathbb{N}|$ , we have  $|\mathbb{Q}| = |\mathbb{N}|$

■

**Exercise 1.1** Let  $A$  be a countable set. Show that the set of finite sequences with terms in  $A$  is countable. Show that the set of all finite subsets of  $A$  is countable.

### Definition 1.9 — power set.

For a set  $A$ , let  $\mathcal{P}(A)$  denote the **power set** of  $A$ , that is the set of all subsets of  $A$ , and let  $2^A$  denote the set of all functions from  $A$  to  $S_2 = \{0, 1\}$

### Theorem 1.10

- (1) For every set  $A$ ,  $\mathcal{P}(A) = 2^A$
- (2) For every set  $A$ ,  $|A| < \mathcal{P}(A)$
- (3)  $\mathbb{R}$  is uncountable

*Proof.*

- (1) Let  $A$  be any set. Define a map  $g : \mathcal{P}(A) \rightarrow 2^A$  as follows: given  $S \in \mathcal{P}(A)$ , that is given  $S \subseteq A$ , we define  $g(S) \in 2^A$  to be the map  $g(S) : A \rightarrow \{0, 1\}$  given by

$$g(S)(a) = \begin{cases} 1 & \text{if } a \in S, \\ 0 & \text{if } a \notin S. \end{cases}$$

Define map  $h : 2^A \rightarrow \mathcal{P}(A)$  as follows: given  $f \in 2^A$ , that is given a map:  $f : A \rightarrow \{0, 1\}$ , we define  $h(f) \in \mathcal{P}(A)$  to be the subset

$$h(f) = \{a \in A \mid f(a) = 1\} \subseteq A$$

This maps  $g$  and  $h$  are the inverses of each other because for every  $S \subseteq A$  and every  $f : A \rightarrow \{0, 1\}$  we have

$$\begin{aligned} f = g(S) &\iff \forall a \in A, f(a) = g(S)(a) \iff \forall a \in A, f(a) = \begin{cases} 1 & \text{if } a \in S, \\ 0 & \text{if } a \notin S. \end{cases} \\ &\iff \forall a \in A, (f(a) = 1 \iff a \in S) \iff \{a \in A \mid f(a) = 1\} = S \\ &\iff h(f) = S \end{aligned}$$

- (2) Let  $A$  be any set. Since the map  $f : A \rightarrow \mathcal{P}(A)$  given by  $f(a) = \{a\}$  is injective, we have  $|A| \leq |\mathcal{P}(A)|$ . We need to show that  $|A| \neq |\mathcal{P}(A)|$ . Let  $g : A \rightarrow \mathcal{P}(A)$  be any map. Let  $S = \{a \in A \mid a \notin g(a)\}$ . Note that  $S$  cannot be in the range of  $g$  because we could choose  $a \in A$  so that  $g(a) = S$  then, by the definition of  $S$ , we would have

$$a \in S \iff a \notin g(a) \iff a \notin S$$

which is impossible. Since  $S$  is not in the range of  $g$ , the map  $g$  is not surjective. Since  $g$  was an arbitrary map from  $A$  to  $\mathcal{P}(A)$ , it follows that there is no surjective map from  $A$  to  $\mathcal{P}(A)$ . Thus there is no bijective map from  $A$  to  $\mathcal{P}(A)$  and so we have  $|A| \neq |\mathcal{P}(A)|$ .

- (3) We prove  $\mathbb{R}$  is uncountable using the fact that every real number has a unique decimal expansion which does not end with an infinite string of 9's. Define a map  $g : 2^{\mathbb{N}} \rightarrow \mathbb{R}$  as follows: given  $f \in 2^{\mathbb{N}}$ , that is given a map  $f : \mathbb{N} \rightarrow \{0, 1\}$ , we define  $g(f)$  to be the real number of  $g(f) \in [0, 1)$  with the decimal expansion  $g(f) = 0.f(1)f(2)f(3)\dots$ , that is  $g(f) = \sum_{k=0}^{\infty} f(k)10^{-k-1}$ . By the uniqueness of decimal expansions, the map  $g$  is injective, so we have  $|2^{\mathbb{N}}| \leq |\mathbb{R}|$ . Thus  $|\mathbb{N}| < |\mathcal{P}(\mathbb{N})| = |2^{\mathbb{N}}| \leq |\mathbb{R}|$ , and so  $\mathbb{R}$  is uncountable by Theorem 1.8. ■

**Theorem 1.11 — Cantor-Schroeder-Bernstein.**

Let  $A$  and  $B$  be sets. Suppose that  $|A| \leq |B|$  and  $|B| \leq |A|$ . Then  $|A| = |B|$

*Proof.* We sketch a proof. Choose injective functions  $f : A \rightarrow B$  and  $g : B \rightarrow A$ . Since the functions  $f : A \rightarrow f(A)$ ,  $g : B \rightarrow g(B)$  and  $f : g(B) \rightarrow f(g(B))$  are bijective, we have  $|A| = |f(A)|$  and  $|B| = |g(B)| = |f(g(B))|$ . Also note that  $f(g(B)) \subseteq f(A) \subseteq B$ . Let  $X = f(g(B))$ ,  $Y = f(A)$  and  $Z = B$ . Then we have  $X \subseteq Y \subseteq Z$  and we have  $|x| = |z|$  and we need to show that  $|Y| = |Z|$ . The composite  $h = f \circ g : Z \rightarrow X$  is a bijective.

Define sets  $Z_n$  and  $Y_n$  for  $n \in \mathbb{N}$  recursively by

$$Z_0 = Z, Z_n = h(Z_{n-1}) \text{ and } Y_0 = Y, Y_n = h(Y_{n-1})$$

Since  $Y_0 = Y$ ,  $Z_0 = Z$ ,  $Z_1 = h(Z_0) = h(Z) = X$  and  $X \subseteq Y \subseteq Z$ , we have

$$Z_1 \subseteq Y \subseteq Z_0$$

Also note that for  $1 \leq n \in \mathbb{N}$ ,

$$Z_n \subseteq Y_{n-1} \subseteq Z_{n-1} \implies h(Z_n) \subseteq h(Y_{n-1}) \subseteq h(Z_{n-1}) \implies Z_{n+1} \subseteq Y_n \subseteq Z_n$$

By the Induction Principle, it follows that  $Z_n \subseteq Y_{n-1} \subseteq Z_{n-1}$  for all  $n \geq 1$ , so we have

$$Z_0 \supseteq Y_0 \supseteq Z_1 \supseteq Y_1 \supseteq Z_2 \supseteq Y_2 \supseteq \cdots$$

Let  $U_n = \frac{Z_n}{Y_n}$ ,  $U = \bigcup_{n=0}^{\infty} U_n$  and  $V = \frac{Z}{U}$ . Define  $H : Z \rightarrow Y$  by

$$H(x) = \begin{cases} h(x) & \text{if } x \in U \\ x & \text{if } x \in V \end{cases}$$

Verify that  $H$  is bijective. ■

**Exercise 1.2** Show that  $|\mathbb{R}| = |2^{\mathbb{N}}|$

*Solution.*  $g : 2^{\mathbb{N}} \rightarrow \mathbb{R}$  as follows: for  $f \in 2^{\mathbb{N}}$  we let  $g(f)$  be the real number  $g(f) \in [0, 1]$  with decimal expansion  $g(f) = 0.f(1)f(2)\cdots$ . Then  $g$  is injective so  $|2^{\mathbb{N}}| \leq |\mathbb{R}|$ . Define  $h : 2^{\mathbb{N}} \rightarrow [0, 1]$  as follows: for  $f \in 2^{\mathbb{N}}$  let  $h(f)$  be the real number  $h(f) \in [0, 1]$  with binary expansion  $h(f) = 0.f(0)f(1)f(2)\cdots$ . Then  $h$  is surjective so we have  $|[0, 1]| \leq |2^{\mathbb{N}}|$ . The map  $k : \mathbb{R} \rightarrow [0, 1]$  given by  $k(x) = \frac{1}{2} + \frac{1}{\pi} \tan^{-1}(x)$  is injective, so we have  $|\mathbb{R}| \leq |[0, 1]|$ . Since  $|\mathbb{R}| \leq |[0, 1]| \leq |2^{\mathbb{N}}|$  and  $|2^{\mathbb{N}}| \leq |\mathbb{R}|$ , we have  $|\mathbb{R}| = |2^{\mathbb{N}}|$  by the Cantor-Schroeder-Bernstein Theorem (1.11) ■

**Notation 1.1** For sets  $A$  and  $B$ , we write  $A^B$  to denote the set of functions  $f : B \rightarrow A$

**Theorem 1.12** Let  $A$  and  $B$  be finite sets and let  $\mathcal{P}(A)$  is the power set of  $A$  (that is the set of all subsets of  $A$ ). Then

- (1) if  $A$  and  $N$  are disjoint then  $|A \cup B| = |A| + |B|$
- (2)  $|A \times B| = |A| \cdot |B|$
- (3)  $|A^B| = |A|^{|B|}$
- (4)  $|\mathcal{P}| = 2^{|A|}$

*Proof.* The proof is left as an exercise ■

**Theorem 1.13** Let  $A, B, C$  and  $D$  be sets with  $|A| = |C|$  and  $|B| = |D|$ . Then

- (1) if  $A \cap B = \emptyset$  and  $C \cap D = \emptyset$  then  $|A \cup B| = |C \cup D|$
- (2)  $|A \times B| = |C \times D|$
- (3)  $|A^B| = |C^D|$

*Proof.* The proof is left as an exercise ■



It is possible to define certain specific sets called **cardinals** such that for every set  $A$  there exists a unique cardinal  $\kappa$  with  $|A| = |\kappa|$ . We can then define the **cardinality** of a set  $A$  to be equal to the unique cardinal  $\kappa$  such that  $|A| = |\kappa|$  and, in this case, we define the **cardinality** of the set  $A$  to be  $|A| = \kappa$ . In foundational set theory, the natural numbers are defined, formally, to be equal to the sets  $0 = \emptyset$ ,  $1 = \{0\} = \{\emptyset\}$ ,  $2 = \{0, 1\} = \{\emptyset, \{\emptyset\}\}$  and, in general,  $n + 1 = n \cup \{n\}$  so that the natural number  $n$  is equal to the set that we previously denoted by  $S_n$ , that is  $n = S_n = \{0, 1, \dots, n-1\}$ . The finite cardinals are equal to the natural numbers and the countable cardinal  $\aleph_0$  is equal to the set of natural numbers. The previous theorem allows us to define **arithmetic operations** on cardinals which extend the usual arithmetic operations on the natural numbers. Given cardinals  $\kappa$  and  $\lambda$  we define  $\kappa + \lambda$ ,  $\kappa \cdot \lambda$  and  $\kappa^\lambda$  to be the cardinals such that

$$\kappa + \lambda = |(\kappa \times \{0\}) \cup (\lambda \times \{1\})|$$

- $\kappa \cdot \lambda = |\kappa \times \lambda|$
- $\kappa^\lambda = |\kappa^\lambda|$

**Theorem 1.14** Let  $\kappa, \lambda$  and  $\mu$  be cardinals. Then

- (1)  $\kappa + \lambda = \lambda + \kappa$
- (2)  $(\kappa + \lambda) + \mu = \kappa + (\lambda + \mu)$
- (3)  $\kappa + 0 = \kappa$
- (4)  $\lambda \leq \mu \implies \kappa + \lambda \leq \kappa + \mu$
- (5)  $\kappa \cdot \lambda = \lambda \cdot \kappa$
- (6)  $(\kappa \cdot \lambda) \cdot \mu = \kappa \cdot (\lambda \cdot \mu)$
- (7)  $\kappa \cdot 1 = \kappa$
- (8)  $\kappa \cdot (\lambda + \mu) = (\kappa \cdot \lambda) + (\kappa \cdot \mu)$
- (9)  $\lambda \leq \mu \implies \kappa \cdot \lambda \leq \kappa \cdot \mu$
- (10)  $\kappa^{\lambda+\mu} = \kappa^\lambda \cdot \kappa^\mu$
- (11)  $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$
- (12)  $(\kappa \cdot \lambda)^\mu = \kappa^\mu \cdot \lambda^\mu$
- (13)  $\lambda \leq \mu \implies \kappa^\lambda \leq \kappa^\mu$
- (14)  $\kappa \leq \lambda \implies \kappa^\mu \leq \lambda^\mu$

*Proof.* We sketch a proof for (9) and (11) and leave the rest as an exercise.

- (9) Let  $A, B$  and  $C$  be sets with  $|A| = \kappa, |B| = \lambda$  and  $|C| = \mu$  and suppose that  $|B| \leq |C|$ .

We need to show that  $|A \times B| \leq |A \times C|$ . Let  $f : B \rightarrow C$  be an injective map. Define  $F : A \times B \rightarrow A \times C$  by  $F(a, b) = (a, f(b))$  then verify that  $F$  is injective.

- (11) Let  $A, B$  and  $C$  be sets with  $|A| = \kappa, |B| = \lambda$  and  $|C| = \mu$ . We need to show  $|(A^B)^C| = |A^{B \times C}|$ . Define  $F : (A^B)^C \rightarrow A^{B \times C}$  by  $F(f)(b, c) = f(c)(b)$ . Verify that  $F$  is bijective with inverse  $G : A^{B \times C} \rightarrow (A^B)^C$  given by  $G(g)(c)(b) = g(b, c)$

■

**Exercise 1.3** Show that  $\left| \bigcup_{n=0}^{\infty} \mathbb{R}^n \right| = 2^{\aleph_0}$

**Exercise 1.4** Find  $|\mathbb{R}^{[0,1]}|$

## 2. Metric Spaces

### Definition 2.1 — inner product, orthogonal, homomorphism, isomorphism.

Let  $F = \mathbb{R}$  or  $\mathbb{C}$ . Let  $U$  be a vector space over  $F$ . An **inner product** on  $U$  (over  $F$ ) is function  $\langle \cdot, \cdot \rangle : U \times U \rightarrow F$  (meaning that if  $u, v \in U$  then  $\langle u, v \rangle \in F$ ) such that for all  $u, v, w \in U$  and all  $t \in F$  we have

(1) (Sesquilinearity)

$$\begin{aligned}\langle u + v, w \rangle &= \langle u, w \rangle + \langle v, w \rangle, \langle tu, v \rangle = t \langle u, v \rangle \\ \langle u, v + w \rangle &= \langle u, v \rangle + \langle u, w \rangle, \langle u, tv \rangle = \bar{t} \langle u, v \rangle\end{aligned}$$

(2) (Conjugate Symmetry)

$$\langle u, v \rangle = \overline{\langle v, u \rangle}$$

(3) (Positiveness Definition)

$$\langle u, u \rangle \geq 0 \text{ with } \langle u, u \rangle = 0 \iff u = 0$$

For  $u, v \in U$ ,  $\langle u, v \rangle$  is called the **inner product** of  $u$  with  $v$ . We say that  $u$  and  $v$  are **orthogonal** when  $\langle u, v \rangle = 0$ . An **inner product space** (over  $F$ ) is a vector space over  $F$  equipped with an inner product. Given two inner product spaces  $U$  and  $V$  over  $F$ , a linear map  $L : U \rightarrow V$  is called a **homomorphism** of inner product spaces (or we say that  $L$  preserves inner product) when  $\langle L(x), L(y) \rangle = \langle x, y \rangle$  for all  $x, y \in U$ . A bijection homomorphism is called an **isomorphism**.

### Definition 2.2 — norm (length).

Let  $U$  be an inner product space over  $F = \mathbb{R}$  or  $\mathbb{C}$ . For  $u \in U$ , we define the **norm** (or **length**) of  $u$  to be

$$\|u\| = \sqrt{\langle u, u \rangle}$$

**Theorem 2.1** Let  $U$  be an inner product space over  $F = \mathbb{R}$  or  $\mathbb{C}$ . For  $u, v \in U$  and  $t \in F$  we have

(1) (Scaling)  $\|tu\| = |t|\|u\|$

(2) (Positive Definiteness)  $\|u\| \geq 0$  with  $\|u\| = 0 \iff u = 0$

(3)  $\|u + v\|^2 = \|u\|^2 + 2\operatorname{Re} \langle u, v \rangle + \|v\|^2$

- (4) (Polarization Identity) if  $F = \mathbb{R}$  then  $\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 - \|u - v\|^2)$  and if  $F = \mathbb{C}$  then  $\langle u, v \rangle = \frac{1}{4}(\|u + v\|^2 + i\|u + iv\|^2 - \|u - v\|^2 - i\|u - iv\|^2)$
- (5) (The Cauchy-Schwarz Inequality)  $|\langle u, v \rangle| \leq \|u\|\|v\|$  with  $|\langle u, v \rangle| = \|u\|\|v\|$  if and only if  $\{u, v\}$  is linearly dependent
- (6) (The Triangle Inequality)  $|\|u\| - \|v\|| \leq \|u\| + \|v\|$

*Proof.* The first 4 parts are easy to prove.

- (5) Suppose that  $\{u, v\}$  is linearly dependent. Then one of  $u$  and  $v$  is a multiple of the other, say  $v = tu$  with  $t \in F$ . Then we have  $|\langle u, v \rangle| = |\langle u, tu \rangle| = |\bar{t}\langle u, u \rangle| = |t|\|u\|^2 = \|u\|\|tu\| = \|u\|\|v\|$ . Next suppose that  $\{u, v\}$  is linearly independent. Then  $1 \cdot v + t \cdot u \neq 0$  for all  $t \in F$ , so in particular  $v - \frac{\langle v, u \rangle}{\|u\|^2}u \neq 0$ . Thus we have

$$\begin{aligned} 0 &< \|v - \frac{\langle v, u \rangle}{\|u\|^2}u\|^2 = \left\langle v - \frac{\langle v, u \rangle}{\|u\|^2}u, v - \frac{\langle v, u \rangle}{\|u\|^2}u \right\rangle \\ &= \langle v, v \rangle - \frac{\overline{\langle v, u \rangle}}{\|u\|^2} \langle v, u \rangle - \frac{\langle v, u \rangle}{\|u\|^2} \langle u, v \rangle + \frac{\langle v, u \rangle}{\|u\|^2} \frac{\overline{\langle v, u \rangle}}{\|u\|^2} \langle u, u \rangle \\ &= \|v\|^2 - \frac{|\langle v, u \rangle|^2}{\|u\|^2} \end{aligned}$$

So that  $\frac{|\langle v, u \rangle|^2}{\|u\|^2} < \|v\|^2$  and hence  $|\langle u, v \rangle| < \|u\|\|v\|$

- (6) Using (3) and (5), and the inequality  $|\operatorname{Re}(z)| \leq |z|$  for  $z \in \mathbb{C}$  (which follows Pythagoras' Theorem in  $\mathbb{R}^2$ ), we have

$$\begin{aligned} \|u + v\|^2 &= \|u\|^2 + \operatorname{Re} \langle u, v \rangle + \|v\|^2 \leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2 \end{aligned}$$

Taking the square root on both sides gives  $\|u + v\| \leq \|u\| + \|v\|$ . Finally note that  $\|u\| = \|(u + v) - v\| \leq \|u + v\| + \|-v\| = \|u + v\| + \|v\|$  so that we have  $\|u\| - \|v\| \leq \|u + v\|$ , and similarly  $\|v\| - \|u\| \leq \|u + v\|$ , hence  $|\|u\| - \|v\|| \leq \|u + v\|$  ■

### Definition 2.3 — norm, unit vector, normed linear space, homomorphism, isomorphism.

Let  $F = \mathbb{R}$  or  $\mathbb{C}$ . Let  $U$  be a vector space over  $F$ . A **norm** on  $U$  is a function  $\| \cdot \| : U \rightarrow \mathbb{R}$  (meaning that if  $u \in U$  then  $\|u\| \in \mathbb{R}$ ) such that for all  $u, v \in U$  and all  $t \in F$  we have

- (1) (Scaling)  $\|tu\| = |t|\|u\|$
- (2) (Positive Definiteness)  $\|u\| \geq 0$  with  $\|u\| = 0 \iff u = 0$
- (3) (Triangle Inequality)  $\|u + v\| \leq \|u\| + \|v\|$

For  $u \in U$  the real number  $\|u\|$  is called the **norm** (or **length**) of  $u$ , and we say that  $u$  is a **unit vector** when  $\|u\| = 1$ . A **normed linear space** (over  $F$ ) is a vector space equipped with a norm. Given two normed linear spaces  $U$  and  $V$  over  $F$ , a linear map  $L : U \rightarrow V$  is called a homomorphism of normed linear spaces (or we say that  $L$  preserves norm) when  $\|L(x)\| = \|x\|$  for all  $x \in U$ . A bijection homomorphism is called an **isomorphism**.

### Definition 2.4 — distance.

Let  $F = \mathbb{R}$  or  $\mathbb{C}$  and let  $U$  be a normed linear space over  $F$ . For  $u, v \in U$ , we define the **distance** between  $u$  and  $v$  to be

$$d(u, v) = \|v - u\|$$

**Theorem 2.2** Let  $U$  be a normed linear space over  $F = \mathbb{R}$  or  $\mathbb{C}$ . For all  $u, v, w \in U$

- (1) (Symmetry)  $d(u, v) = d(v, u)$
- (2) (Positive Definiteness)  $d(u, v) \geq 0$  with  $d(u, v) = 0 \iff u = v$
- (3) (Triangle Inequality)  $d(u, w) \leq d(u, v) + d(v, w)$

*Proof.* The proof is left as exercise ■

**Definition 2.5 — metric, distance, metric space, homomorphism, isomorphism.**

Let  $X$  be a non-empty set. A **metric** on  $X$  is a map  $d : X \times X \rightarrow \mathbb{R}$  such that for all  $a, b, c \in X$  we have

- (1) (Symmetry)  $d(a, b) = d(b, a)$
- (2) (Positive Definiteness)  $d(a, b) \geq 0$  with  $d(a, b) = 0 \iff a = b$
- (3) (Triangle Inequality)  $d(a, c) \leq d(a, b) + d(b, c)$

For  $a, b \in X$ ,  $d(a, b)$  is called the **distance** between  $a$  and  $b$ . A **metric space** is a set  $X$  which is equipped with a metric  $d$ , and we sometimes denote the metric space by  $X$  and sometimes by the pair  $(X, d)$ . Given two metric spaces  $(X, d_X)$  and  $(Y, d_Y)$ , a map  $f : X \rightarrow Y$  is called a **homomorphism of metric spaces** (or we say that  $f$  is **distance preserving**) when  $d_Y(f(a), f(b)) = d_X(a, b)$  for all  $a, b \in X$ . A bijective homomorphism is called an **isomorphism** or an **isometry**.

**Note 2.1** Every inner product space is also a normed linear space, using the induced norm given by  $\|u\| = \sqrt{\langle u, u \rangle}$ . Every normed linear space is also a metric space, using the induced metric given by  $d(u, v) = \|v - u\|$ . If  $U$  is an inner product space over  $F = \mathbb{R}$  or  $\mathbb{C}$  then every subspace of  $U$  is also an inner product space using (the restriction of) the same inner product used in  $U$ . If  $U$  is a normed linear space over  $F = \mathbb{R}$  or  $\mathbb{C}$  then every subspace of  $U$  is also a normed linear space using the same norm. If  $X$  is a metric space then so is every subset of  $X$  using the same metric.

■ **Example 2.1** Let  $F = \mathbb{R}$  or  $\mathbb{C}$ . The **standard inner product** on  $F^n$  is given by

$$\langle u, v \rangle = v * u = \sum_{i=1}^n u_i \overline{v_i}$$

The standard inner product induces the **standard norm** on  $F^n$ , which is also called the **2-norm** on  $F^n$ , given by

$$\|u\|_2 = \|u\| = \sqrt{\langle u, u \rangle} = \left( \sum_{i=1}^n |u_i|^2 \right)^{\frac{1}{2}}$$

The standard norm on  $F^n$  induces the **standard metric** on  $F^n$ , given by

$$d_2(u, v) = d(u, v) = \|v - u\| = \left( \sum_{i=1}^n |v_i - u_i|^2 \right)^{\frac{1}{2}}$$

The **1-norm** on  $F^n$  is given by

$$\|u\|_1 = \sum_{i=1}^n |u_i|$$

and it induces the **1-metric** on  $F^n$  given by  $d_1(u, v) = \|v - u\|_1$ . The **supremum norm** also called **infinity norm**, on  $F^n$  is given by

$$\|u\|_\infty = \max\{|u_1|, |u_2|, \dots, |u_n|\}$$

and it induces the **supremum metric** on  $\mathbf{F}^n$  given by  $d_\infty(u, v) = \|v - u\|_\infty$

■ **Example 2.2** For  $\mathbf{F} = \mathbb{R}$  or  $\mathbb{C}$ . We write

$$\begin{aligned}\mathbf{F}^\omega &= \{u = (u_1, u_2, u_3, \dots) \mid \text{each } u_i \in \mathbf{F}\} \\ \mathbf{F}^\infty &= \{u \in \mathbf{F}^\omega \mid \text{there exists } n \in \mathbb{Z}^+ \text{ such that } u_k = 0 \text{ for all } k \geq n\}\end{aligned}$$

Recall that  $\mathbf{F}^\infty$  is a countable-dimensional vector space with standard basis  $\{e_1, e_2, e_3, \dots\}$  where  $e_1 = (1, 0, 0, \dots)$ ,  $e_2 = (0, 1, 0, \dots)$  and so on. The **standard inner product** on  $\mathbf{F}^\infty$  is given by

$$\langle u, v \rangle = \sum_{i=1}^{\infty} u_i \overline{v_i}$$

and it induces the **standard norm**, also called the **2-norm**, on  $\mathbf{F}^\infty$  given by

$$\|u\|_2 = \sqrt{\langle u, u \rangle} = \left( \sum_{i=1}^{\infty} |u_i|^2 \right)^{\frac{1}{2}}$$

The **1-norm** on  $\mathbf{F}^\infty$  is given by

$$\|u\|_1 = \sum_{i=1}^{\infty} |u_i|$$

and it induces the **1-metric** on  $\mathbf{F}^\infty$  given by  $d_1(u, v) = \|v - u\|_1$ . The **supremum norm** also called the **infinity norm**, on  $\mathbf{F}^n$  is given by

$$\|u\|_\infty = \max\{|u_1|, |u_2|, \dots, |u_n|\}$$

and it induces the **supremum metric** on  $\mathbf{F}^n$  given by  $d_\infty(u, v) = \|v - u\|_\infty$

■ **Example 2.3** For  $\mathbf{F} = \mathbb{R}$  or  $\mathbb{C}$ , the standard inner product, the 1-norm, the 2-norm and the  $\infty$ -norm, which are well defined on the vector space  $\mathbf{F}^\infty$ , do not extend naturally to give a well defined inner product or well-defined norms on the vector space  $\mathbf{F}^\omega$  (because the relevant sums do not necessarily converge). But we can, and do, extend there definitions to various subspaces of  $\mathbf{F}^\omega$ . We define

$$\begin{aligned}\ell_1(\mathbf{F}) &= \{u \in \mathbf{F}^\omega \mid \sum_{i=1}^{\infty} |u_i| < \infty\} \\ \ell_2(\mathbf{F}) &= \{u \in \mathbf{F}^\omega \mid \sum_{i=1}^{\infty} |u_i|^2 < \infty\} \\ \ell_\infty(\mathbf{F}) &= \{u \in \mathbf{F}^\omega \mid \sup\{|u_1|, |u_2|, \dots\} < \infty\}\end{aligned}$$

Verify that  $\ell_1(\mathbf{F})$  is a normed linear space using **1-norm** given by  $\|u\|_1 = \sum_{i=1}^{\infty} |u_i|$ , hence  $\ell_1(\mathbf{F})$  is also a metric space using the **1-metric**  $d_1(u, v) = \|v - u\|_1$ . Verify that  $\ell_\infty(\mathbf{F})$  is a normed linear space using the **supremum norm**, also called the **infinity norm**, given by  $\|u\|_\infty = \sup\{|u_1|, |u_2|, \dots\}$ , hence  $\ell_\infty(\mathbf{F})$  is also a metric space using the **supremum metric**  $d_\infty = \|v - u\|_\infty$ . Verify that  $\ell_2(\mathbf{F})$  is an inner product space using the **standard inner product** given by  $\langle u, v \rangle = \sum_{i=1}^{\infty} u_i \overline{v_i}$ . The standard inner product on  $\ell_2(\mathbf{F})$  induces

the **standard norm**, also called the **2-norm**, on  $\ell_2(\mathbf{F})$  given by  $\|u\|_2 = \left( \sum_{i=1}^{\infty} |u_i|^2 \right)^{\frac{1}{2}}$  and

the **standard metric**, or the **2-metric**,  $d_2(u, v) = \|v - u\|_2$ .

Since we shall usually work with the field  $\mathbf{F} = \mathbb{R}$ , for  $p = 1, 2$  or  $\infty$  we shall write

$$\ell_p = \ell_p(\mathbb{R})$$



■ **Example 2.4** For  $\mathbf{F} = \mathbb{R}$  or  $\mathbb{C}$  and for  $a, b \in \mathbb{R}$  with  $a \leq b$ , we write

$$\begin{aligned}\mathcal{F}([a, b], \mathbf{F}) &= \mathbf{F}^{[a, b]} = \{f : [a, b] \rightarrow \mathbf{F}\} \\ \mathcal{B}([a, b], \mathbf{F}) &= \{f : [a, b] \rightarrow \mathbf{F} \mid f \text{ is bounded}\} \\ \mathcal{C}([a, b], \mathbf{F}) &= \{f : [a, b] \rightarrow \mathbf{F} \mid f \text{ is continuous}\}\end{aligned}$$

Recall that for  $f : [a, b] \rightarrow \mathbb{C}$  given by  $f = u + iv$  where  $u, v : [a, b] \rightarrow \mathbb{R}$ , the function  $f$  is continuous if and only if both  $u$  and  $v$  are continuous and, in this case,  $\int_a^b f = \int_a^b u + i \int_a^b v$ . In the space  $\mathcal{C}([a, b], \mathbf{F})$  we have the **1-norm**, the **2-norm**, and the **supremum norm**

$$\begin{aligned}\|f\|_1 &= \int_a^b |f| \\ \|f\|_2 &= \left( \int_a^b |f|^2 \right)^{\frac{1}{2}} \\ \|f\|_\infty &= \sup_{a \leq x \leq b} |f(x)|\end{aligned}$$

The supremum norm also gives a well-defined norm on the space  $\mathcal{B}([a, b], \mathbf{F})$ . The 2-norm on  $\mathcal{C}([a, b], \mathbf{F})$  is induced by the inner product  $\mathcal{C}([a, b], \mathbf{F})$  given by

$$\langle f, g \rangle = \int_a^b f \bar{g}$$

Since we shall usually work with the field  $\mathbf{F} = \mathbb{R}$  we shall write

$$\mathcal{F}[a, b] = \mathcal{F}([a, b], \mathbb{R}), \quad \mathcal{B}[a, b] = \mathcal{B}([a, b], \mathbb{R}) \text{ and } \mathcal{C} = \mathcal{C}([a, b], \mathbb{R})$$

Ⓡ For  $\mathbf{F} = \mathbb{R}$  or  $\mathbb{C}$  and for  $1 \leq p < \infty$ , one can show that we can define a norm on  $\mathbf{F}^n$  by

$$\|u\|_p = \left( \sum_{i=1}^n |u_i|^p \right)^{\frac{1}{p}}$$

and we can define a norm on  $\mathbf{F}^\omega$  or on the space  $\ell_\infty(\mathbf{F}) = \{u \in \mathbf{F}^\omega \mid \sum_{i=1}^\infty |u_i|^p < \infty\}$  by

$$\|u\|_p = \left( \sum_{i=1}^\infty |u_i|^p \right)^{\frac{1}{p}}$$

Also, we can define a norm on the space  $\mathcal{C}([a, b], \mathbf{F})$  by

$$\|f\|_p = \left( \int_{i=a}^b |f|^p \right)^{\frac{1}{p}}$$

■ **Example 2.5** For any set  $X \neq \emptyset$ , the **discrete metric** on  $X$  is given by  $d(x, y) = 1$  for all  $x, y \in X$  with  $x \neq y$  and  $d(x, x) = 0$  for all  $x \in X$ .

**Definition 2.6 — open ball, closed ball, punctured ball, bounded.**

Let  $X$  be a metric space. For  $a \in X$  and  $0 < r \in \mathbb{R}$ , the **open ball**, the **closed ball** and the (open) **punctured ball** in  $X$  centered at  $a$  of radius  $r$  are defined to be the sets

$$\begin{aligned}B(a, r) &= B_X(a, r) = \{x \in X \mid d(x, a) < r\} \\ \bar{B}(a, r) &= \bar{B}_X(a, r) = \{x \in X \mid d(x, a) \leq r\} \\ B^*(a, r) &= B_X^*(a, r) = \{x \in X \mid 0 < d(x, a) < r\}\end{aligned}$$

When the metric on  $X$  denoted by  $d_p$  with  $1 \leq p \leq \infty$ , we often write  $B(a, r)$ ,  $\overline{B}(a, r)$  and  $B^*(a, r)$  as  $B_p(a, r)$ ,  $\overline{B}_p(a, r)$  and  $B_p^*(a, r)$ . For  $A \subseteq X$ , we say that  $A$  is **bounded** when  $A \subseteq B(a, r)$  for some  $a \in X$  and some  $0 < r \in \mathbb{R}$ .

**Exercise 2.1** Draw a picture of the open balls  $B_1(0, 1)$ ,  $B_2(0, 1)$  and  $B_\infty(0, 1)$  in  $\mathbb{R}^2$  (using the metrics  $d_1$ ,  $d_2$  and  $d_\infty$ ).