PMATH 351: Real Analysis

Instructor Stephen New
LaTeXer Iris Jiang

Spring 2020

Contents

1. Cardinality

Definition 1.1 — domain, range, image, inverse image.
Let X and Y be sets and let $f : X \to Y$. 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 $f : X \to Y$ and let $g : Y \to Z$. 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 $f^{-1} : Y \to X$ 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 $f : X \to Y$ and let $g : Y \to Z$. 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 $g\left(f(x_1)\right) = g\left(f(x_2)\right)$ 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 injective. 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 $g\left(f(x)\right) = g(y) = z$. 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 $I_X : X \to X$ given by $I_X(x) = x$ for all $x \in X$. Note that for $f : X \to Y$ 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 $f : X \to Y$. A **left inverse** of f is a function $g : Y \to X$ given by $g \circ f = I_X$. Equivalently, a function $g : Y \to X$ is a left inverse of f when $g\left(f(x)\right) = x$ for all $x \in X$.
A **right inverse** of f is a function $h : Y \to X$ such that $f \circ h = I_Y$. Equivalently, a function $h : Y \to X$ is a right inverse of f when $f\left(h(y)\right) = y$ for all $y \in Y$.

Theorem 1.2 Let X and Y be nonempty sets and let $f : X \to Y$. 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 $X \neq \varnothing$ we can choose $a \in X$ and then define $g : Y \to X$ 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 $g\left(f(x)\right) = x$. 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 $x = g\left(f(x_1)\right) = g\left(f(x_2)\right) = x_2$.
(2) Suppose first that f is onto. For each $y \in Y$, choose $x_y \in X$ with $f(x_y) = y$, then define $g : X \to Y$ 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 $f^{-1} : Y \to X$ 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 $f : X \to Y$ if and only if there exists a surjective map $g : Y \to X$.

Proof. Suppose $f : X \to Y$ 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 $g : Y \to X$. Similarly, if $g : Y \to X$ is surjective, then it has a right inverse $f : X \to Y$ 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: $f : A \to B$ (or equivalently when there exists a bijective map $g : B \to A$).
We say that the cardinality of A is **less than or equal to** the cardinality of B, and write $|A| \le |B|$, when there exists an injective map $f : A \to B$ (or equivalently a surjective map $g : B \to A$).
We say that the cardinality of A is **less than** the cardinality of B, and write $|A| < |B|$, when $|A| \le |B|$ and $|A| \ne |B|$, (that is when there exists an injective map $f : A \to B$ but there does not exist a bijective map $g : A \to B$).
We also write $|A| \ge |B|$ when $|B| \le |A|$; and $|A| > |B|$ when $|B| < |A|$.

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

(1) The map $f : \mathbb{N} \to 2\mathbb{N}$ given by $f(k) = 2k$ is bijective, so $|2\mathbb{N}| = |\mathbb{N}|$.
(2) The map $g : \mathbb{N} \to \mathbb{Z}$ 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 $h : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ 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) $|A| \le |B| \iff (|A| = |B|$ or $|A| < |B|)$
(5) If $|A| \le |B|$ and $|B| \le |C|$, then $|A| \le |C|$

Proof.

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

■

Definition 1.7 — finite, infinite, countable.
Let A be a set. For each $n \in \mathbb{N}$, let $S_n = \{0, 1, 2, \cdots, n-1\}$. 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 $f : A \to S_n$ is a bijection, if we let $a_k = f^{-1}(k)$ for each $k \in S_n$ then we have $A = \{a_0, a_1, \cdots, a_{k-1}\}$ with the elements a_k distinct. Conversely, if $A = \{a_0, a_1, \cdots, a_{k-1}\}$ with the elements a_k all distinct, then we define a bijection $f : A \to S_n$ by $f(a_k) = k$. Thus we see that A is finite with $|A| = n$ if and only if A is of the form $A = \{a_0, a_1, \cdots, a_{n-1}\}$ with the elements a_k all distinct. Similarly, a set A is countable if and only if A is of the form $A = \{a_0, a_1, a_2, \cdots\}$ 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 $a \in |A \backslash \{a\}| = n$. Indeed, if $A = \{a_0, a_1, \cdots, a_n\}$ with the elements a_i distinct, and if $a = a_k$ so that we have $A \backslash \{a\} = \{a_0, a_1, \cdots, a_{k-1}, a_{k+1}, \cdots, a_n\}$, then we can define a bijection $f : S_n \to A \backslash \{a\}$ 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 $f : A \to A$ which is injective but not surjective

Proof.
(1) $\Longrightarrow$ (2) Suppose A is infinite. Since $A \neq \varnothing$ we can choose an element $a_0 \in A$. Since $A \neq \{a_0\}$ we can choose an element $a_1 \in A \backslash \{a_0\}$. Since $A \neq \{a_0, a_1\}$ we can choose $a_3 \in A \backslash \{a_0, a_1\}$. Continue this procedure: having chosen distinct elements $a_0, a_1, \cdots, a_{n-1} \in A$, since $A \neq \{a_0, a_1, \cdots, a_{n-1}\}$ we can choose $a_n \in A \backslash \{a_0, a_1, \cdots, a_{n-1}\}$. In this way we obtain $\{a_0, a_1, a_2, \cdots\} \subseteq A$.
(2) $\Longleftrightarrow$ (3) Suppose that A contains a countable subset, say $\{a_0, a_1, a_2, \cdots\} \subseteq A$ with the element a_i distinct. Since a_i are distinct, the map $f : \mathbb{N} \to A$ given by $f(k) = a_k$ is injective, and so we have $|\mathbb{N}| \leq |A|$. Conversely as a map from $\mathbb{N} \to f(\mathbb{N})$ where f is bijective, so we have $|\mathbb{N}| = |f(\mathbb{N})|$ hence $f(\mathbb{N})$ is a countable subset of A.
(2) $\Longrightarrow$ (4) Suppose that A has a countable subset, say $\{a_0, a_1, a_2, \cdots\} \subseteq A$ with the element a_i distinct. Define $f : A \to A$ by $f(a_k) = a_{k+1}$ for all $k \in \mathbb{N}$ and by $f(b) = b$ for all $b \in A \backslash \{a_0, a_1, a_2, \cdots\}$. Then f is injective but not surjective (the element a_0 is not in the range of f).
(4) $\Longrightarrow$ (1) To prove this we shall prove that if A is finite then every injective map $f : A \to A$ is surjective. We prove this by induction on the cardinality of A.
The only set A with $|A| = 0$ is the set $A \neq \varnothing$, and then the only function $f : A \to A$ 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 $f : A \to A$ 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 $f : A \to A$ is surjective. Let B be a set with $|B| = n+1$ and let $g : B \to B$ 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 $g : B \to B \backslash \{b\}$. Let $A = B\backslash\{b\}$ and let $f : A \to A$ be given by $f(x) = g(x)$ for all $x \in A$. Since $g : B \to A$ 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 $x \in B\backslash\{b\}$ 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 $f : A \to A$ 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 $I : S_n \to S_m$ 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 $f : S_n \to S_m$. Since $n > m$ we have $S_m \subsetneq S_n$ so we can consider f as a map $f : S_n \to S_m$, 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 $A \neq \varnothing$, we can set $a_0 = \min\{A\}$ (using the Well-Ordering Property of $\mathbb{N}$). Note that

$\{0, 10, \cdots, a_0\} \cap A = \{a_0\}$.
Since $A \neq \{a_0\}$ (so the set $A \backslash \{a_0\}$ is nonempty), we can set $a_1 = \min\{A \backslash \{a_0\}\}$. Then we have $a_0 < a_1$ and $\{0, 1, \cdots, a_1\} \cap A = \{a_0, a_1\}$.
Since $A \neq \{a_0, a_1\}$ we can set $a_2 = \min\{A \backslash \{a_0, a_1\}\}$. Then we have $a_0 < a_1 < a_2$ and $\{0, 1, 2, \cdots, a_3\} \cap A = \{a_0, a_1, a_2\}$
We continue the procedure: having chosen $a_0, a_1, \cdots, a_{n-1} \in A$ with $a_0 < a_1 < \cdots < a_{n-1}$ such that $\{0, 1, \cdots, a_{n-1}\} \cap A = \{a_0, a_1, \cdots, a_{n-1}\}$. Since $A \neq \{a_0, a_1, \cdots, a_{n-1}\}$, we can set $a_n = \min\{A \backslash \{a_0, a_1, \cdots, a_{n-1}\}\}$ and then we have $a_0 < a_1 < \cdots < a_{n-1} < a_n$ and $\{0, 1, \cdots, a_n\} \cap A = \{a_0, a_1, \cdots, a_n\}$.
In this way, we obtain a countable set $\{a_0, a_1, a_2, \cdots\} \subseteq A$ with $a_0 < a_1 < a_2 < \cdots$ with the property that for all $m \in \mathbb{N}$, $\{0, 1, 2, \cdots, a_m\} \cap A = \{a_0, a_1, \cdots, a_m\}$.
Since $0 \leq a_0 < a_1 < a_2 < \cdots$, it follows (by induction) that $a_k \geq k$ for all $k \in \mathbb{N}$. It follows in turn that $A \subseteq \{a_0, a_1, a_2, \cdots\}$ because given $m \in A$, since $m \leq a_m$ we have

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

Thus $A = \{a_0, a_1, a - 2, \cdots\}$ 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 $f : A \to \mathbb{N}$. Since f is injective, when we consider it as a map $f : A \to f(A)$, 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

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

(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 $|N| \leq |A|$, 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 $A_0, A_1, A_2, \cdots$ are countable sets, then so is $\bigcap\limits_{k=0}^{\infty} A_k$
(4) $\mathbb{Q}$ is countable

Proof.
(1) Let $A = \{a_0, a_1, a_2, \cdots\}$ with the a_i distinct and let $B = \{b_0, b_1, b_2, \cdots\}$ 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 $f : A \times B \to \mathbb{N}$ 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 $g : \mathbb{N} \to A \cup B$ given by $g(k) = a_k$ is injective, we have $|\mathbb{N}| \le |A \cup B|$. Since the map $h : \mathbb{N} \to A \cup B$ given by $h(2k) = a_k$ and $h(2k+1) = b_k$ is surjective, we have $|A \cup B| \le |\mathbb{N}|$. Since $|\mathbb{N}| \le |A \cup B|$ and $|A \cup B| \le |\mathbb{N}|$, we have $|A \cup B| = |\mathbb{N}|$ by Theorem 1.8
(3) For each $k \in \mathbb{N}$, let $A_k = \{a_{k0}, a_{k1}, a_{k2}, \cdots\}$ with the a_{ki} distinct. Since the map $f : \mathbb{N} \to \bigcap\limits_{k=0}^{\infty} A_k$ given by $f(k) = a_{0,k}$ is injective, $|\mathbb{N}| \le \left|\bigcap\limits_{k=0}^{\infty} A_k\right|$. Since $\mathbb{N} \times \mathbb{N}$ is countable by (1), and since the map $g : \mathbb{N} \times \mathbb{N} \to \bigcap\limits_{k=0}^{\infty} A_k$ given by $g(k, l) = a_{k,l}$ is surjective, we have $\left|\bigcap\limits_{k=0}^{\infty} A_k\right| \le |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$. By Theorem 1.8, we have $\left|\bigcap\limits_{k=0}^{\infty} A_k\right| = |\mathbb{N}|$.
(4) Since the map $f : \mathbb{N} \to \mathbb{Q}$ given by $f(k) = k$ is injective, we have $|\mathbb{N}| \le |\mathbb{Q}|$. Since the map $g : \mathbb{Q} \to \mathbb{Z} \times \mathbb{Z}$ 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 $|\mathbb{Q}| \le |\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N}|$. Since $|\mathbb{N}| \le |\mathbb{Q}|$ and $|\mathbb{Q}| \le |\mathbb{N}|$, 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 $g : \mathcal{P}(A) \to 2^A$ 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 $g(S) : A \to \{0, 1\}$ given by

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

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

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

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

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

(2) Let A be any set. Since the map $f : A \to \mathcal{P}(A)$ 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 $g : A \to \mathcal{P}(A)$ be any map. Let $S = \{a \in A | a \notin g(a)\}$. 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 $g : 2^{\mathbb{N}} \to \mathbb{R}$ as follows: given $f \in 2^{\mathbb{N}}$, that is given a map $f : \mathbb{N} \to \{0,1\}$, we define $g(f)$ to be the real number of $g(f) \in [0,1)$ with the decimal expansion $g(f) = 0.f(1)f(2)f(3)\cdots$, 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 $f : A \to B$ and $g : B \to A$. Since the functions $f : A \to f(A)$, $g : B \to g(B)$ and $f : g(B) \to f(g(B))$ 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 $h = f \circ g : Z \to X$ 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 $H : Z \to Y$ 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. $g : 2^{\mathbb{N}} \to \mathbb{R}$ 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 $h : 2^{\mathbb{N}} \to [0,1)$ 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 $k : \mathbb{R} \to [0,1]$ 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 $f : B \to A$

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 $A \cap B = \varnothing$ and $C \cap D = \varnothing$ 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 ∎

R It is possible to define certain specific sets called **cardinals** such that for every set A there exists a unique cardinal κ with $|A| = |\kappa|$. We can then define the **cardinality** of a set A to be equal to the unique cardinal κ such that $|A| = ||$ 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 $0 = \varnothing$, $1 = \{0\} = \{\varnothing\}$, $2 = \{0,1\} = \{\varnothing, \{\varnothing\}\}$ 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 $n = S_n = \{0, 1, \cdots, n-1\}$. 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 κ and λ we define $\kappa + \lambda$, $\kappa \cdot \lambda$ and κ^λ 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 κ, λ and μ 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) $\kappa \cdot (\lambda + \mu) = (\kappa \cdot \lambda) + (\kappa \cdot \lambda)$
(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 $f : B \to C$ be an injective map. Define $F : A \times B \to A \times C$ 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 $F : (A^B)^C \to A^{B \times C}$ by $F(f)(b, c) = f(c)(b)$. Verify that F is bijective with inverse $G : A^{B \times C} \to (A^B)^C$ 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 $\left|\mathbb{R}^{[0,1]}\right|$

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 $\langle\ ,\ \rangle U \times U \to F$ (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)
$$\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle, \langle tu, v\rangle = t\langle u, v\rangle$$
$$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle, \langle u, tv\rangle = \bar{t}\langle u, v\rangle$$

(2) (Conjugate Symmetry)
$$\langle u, v\rangle = \overline{\langle v, u\rangle}$$

(3) (Positie 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 $L : U \to V$ 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) $\|u+v\|^2 = \|u\|^2 + 2\mathrm{Re}\langle u, v\rangle + \|v\|^2$

(4) (Polarization Identity) if $F = \mathbb{R}$ then $\langle u, v\rangle = \frac{1}{4}(\|u+v\| - \|u-v\|)$ and if $F = \mathbb{C}$ then $\langle u, v\rangle = \frac{1}{4}(\|u+v\|^2 + i\|u+iv\|^2 - \|u-v\| - i\|u-iv\|^2)$
(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 x and y 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

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

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

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

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

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 $\| \ \| : U \to \mathbb{R}$ (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 Definitenes) $\|u\| \geq 0$ 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 $L : U \to V$ 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 as 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 $d : X \times X \to \mathbb{R}$ 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 $f : X \to Y$ 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 $d(u, u) = \|v - u\|$. If U is an inner product space over $\mathbf{F} = \mathbb{R}$ 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 $\mathbf{F} = \mathbb{R}$ 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 $\mathbf{F} = \mathbb{R}$ or $\mathbb{C}$. The **standard inner product** on $\mathbf{F}^n$ 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 $\mathbf{F}^n$, which is also called the **2-norm** on $\mathbf{F}^n$, 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 $\mathbf{F}^n$ induces the **standard metric** on $\mathbf{F}^n$, 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 $\mathbf{F}^n$ is given by

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

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

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

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

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

Recall that $\mathbf{F}^\infty$ is a countable-dimensional vector space with standard basis $\{e_1, e_2, e_3, \cdots\}$ where $e_1 = (1,0,0,\cdots), e_2 = (0,1,0,\cdots)$ 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

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

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

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

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 ∞-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

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

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 $\|u\|_\infty = \sup\{|u_1|, |u_2|, \cdots\}$, 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 ∞ 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

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

Recall that for $f : [a,b] \to \mathbb{C}$ given by $f = u + iv$ where $u, v : [a,b] \to \mathbb{R}$, the function f is continuous if and only if both u and v are continuous and, in this case, $\int_a^b f = \int_a^b + i \int_a^b v$. 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

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

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

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

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 $\mathbf{F}^\infty$ or on the space $\ell_\infty(\mathbf{F}) = \{u \in \mathbf{F}^\omega | \sum_{i=1}^{\infty} |u_i|^p < \infty\}$ 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 $X \neq \varnothing$, 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

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

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_∞).

Definition 2.7 — open, closed.
Let X be a metric space. For $A \subseteq X$, we say that A is **open** (in X) when for every $a \in A$ there exists $r > 0$ such that $B(a,r) \subseteq A$, and we say that A is closed (in X) when its complement $A^c = \mathbb{R}^n \backslash A$ is open in $\mathbb{R}^n$.

■ **Example 2.6** Let X be a metric space. Show that for $a \in X$ and $0 < r \in \mathbb{R}$, the set $B(a,r)$ is open and the set $\overline{B}(a,r)$ is closed.

Proof. Let $a \in X$ and Let $r > 0$. We claim that $B(a,r)$ is open. We need to show that for all $b \in B(a,r)$ there exists $s > 0$ such that $B(b,s) \subseteq B(a,r)$. Let $b \in B(a,r)$ and note that $d(a,b) < r$. Let $s = r - d(a,b)$ and note that $s > 0$. Let $x \in B(b,s)$, so we have $d(x,b) < s$. Then, by the Triangle Inequality, we have

$$d(x,a) \leq d(x,b) + d(b,a) < s + d(a,b) = r$$

and so $x \in B(a,r)$. This shows that $B(b,s) \subseteq B(a,r)$ and hence $B(a,r)$ is open.
Next we claim that $\overline{B}(a,r)$ is closed, that is $\overline{B}(a,r)^c$ is open. Let $b \in \overline{B}(a,r)^c$, that is let $b \in X$ with $b \notin \overline{B}(a,r)$. Since $b \notin \overline{B}(a,r)$ we have $d(a,b) > r$. Let $s = d(a,b) - r > 0$. Let $x \in B(b,s)$ and not that $d(x,b) < s$. Then, by the Triangle Inequality, we have

$$d(a,b) \leq d(a,x) + d(x,b) < d(x,a) + s$$

and so $d(x,a) > d(a,b) - s = r$. Since $d(x,a) > r$ we have $b \notin \overline{B}(a,r)$ and so $x \in \overline{B}(a,r)^c$. This shows that $B(b,s) \subseteq \overline{B}(a,r)^c$ and it follows that $\overline{B}(a,r)^c$ is open and hence that $\overline{B}(a,r)$ is closed. ■

Theorem 2.3 — Basic Properties of Open Sets.
Let X be a metric space
(1) The sets $\varnothing$ and X are open in X
(2) If S is a set of open sets in X then the union $\bigcup S = \bigcup_{U \in S} U$ is open in X
(3) If S is a finite set of open sets in X then the intersection $\bigcap S = \bigcap_{U \in S} U$ is open in X

Proof. (1) The empty set is open because any statement of the form "for all $x \in \varnothing, F$" (where F is any statement) is considered to be true (by convention). The set X is open because given $a \in X$ we can choose any value of $r > 0$ and then $B(a,r) \subseteq X$ by the definition of $B(a,r)$.
(2) Let S be any set of open sets in X. Let $a \in \bigcup S = \bigcup_{U \in S} U$. Choose an open set $U \in S$ such that $a \in U$. Since U is open we can choose $r > 0$ such that $B(a,r) \subseteq U$. Since $U \in S$ we have $U \subseteq \bigcup S$. Since $B(a,r) \subseteq U$ and $U \subseteq \bigcup S$ we have $B(a,r) \subseteq \bigcup S$. Thus $\bigcup S$ is open.
(3) Let S be a finite set of open sets in X. If $S \neq \varnothing$ then we use the convention that $\bigcap S = X$, which is open. Suppose that $S \neq \varnothing$. say $S = \{U_1, U_2, \cdots, U_m\}$ where

each U_k is an open set. Let $a \in \bigcap S = \bigcap_{k=1}^{m} U_k$. For each index k, since $a \in U_k$ we can choose $r_k > 0$ so that $B(a, r_k) \subseteq U_k$. Let $r = \min\{r_1, r_2, \cdots, r_m\}$. Then for each index k we have $B(a, r) \subseteq B(a, r_k) \subseteq U_k$. Since $B(a, r)U_k$ for every index k, it follows that $B(a, r) \subseteq \bigcap_{k=1}^{m} U_k = \bigcap S$

■

Theorem 2.4 — Basic Properties of Closed Sets.
Let X be a metric space

(1) The sets $\varnothing$ and X are closed in X
(2) If S is a set of closed sets in X then the union $\bigcap S = \bigcap\limits_{U \in S} U$ is open in X
(3) If S is a finite set of closed sets in X then the intersection $\bigcup S = \bigcup\limits_{U \in S} U$ is open in X

Proof. The proof is left as exercise ■

Definition 2.8 — topology, topology space, metric topology, finer, coarser.
A **topology** on a set X is a set T of subsets of X such that

(1) $\varnothing \in T$ and $X \in T$
(2) For every set $S \subseteq T$ we have $\bigcup S \in T$
(3) For every finite subset $S \subseteq T$ we have $\bigcap S \in T$

A **topology space** is a set X with a topology T. When X is a metric space, the set of all open sets in X is a topology on X. which we call the **metric topology** (or the topology **induced** by the metric). When X is any topological space, the sets in the topology T are called the **open sets** in X and their complements are called the **closed sets** in X. When S and T are both topologies on a set X with $S \subseteq T$, we say that the topology T is **finer** than the topology S, and the topology S is **coarser** than the topology T.

■ **Example 2.7** Show that in $\mathbb{R}^n$, the metrics d_1, d_2 and d_∞ all induce the same topology

Proof. For $a, x \in \mathbb{R}^n$ we have

$$\max_{a \le i \le n} |x_i - a_i| \le (\sum_{i=1}^{n} |x_i - a_i|^2)^{\frac{1}{2}} \le \sum_{i=1}^{n} |x_i - a_i| \le n \max_{a \le i \le n} |x_i - a_i|$$

and so

$$d_\infty(a, x) \le d_2(a, x) \le d_1(a, x) \le n d_\infty(a, x).$$

It follows that for all $a \in \mathbb{R}^n$ and $r > 0$ we have

$$B_\infty(a, x) \supseteq B_2(a, r) \supseteq B_1(a, r) \supseteq B_\infty(a, \frac{r}{n}).$$

Thus for $U \subseteq \mathbb{R}^n$, if U is open in $\mathbb{R}^n$ using d_∞ then it is open using d_2. and if U is open using d_2 then it is open using d_1, and if U is open using d_1 then it is open using d_∞. ■