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 $|A| \leq |B|$, 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 $|A| \leq |B|$ and $|A| \neq |B|$, (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 $|A| \geq |B|$ when $|B| \leq |A|$; and $|A| > |B|$ when $|B| < |A|$.

∎ **Example 1.1** Let $\mathbb{N} = \{n \in \mathbb{Z} | n \geq 0\} = \{0, 1, 2, \cdots\}$.
   (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) $|A| \leq |B| \iff (|A| = |B|$ or $|A| < |B|)$
>   (5) If $|A| \leq |B|$ and $|B| \leq |C|$, then $|A| \leq |C|$

Proof.
   (1) holds because the identity function $I_A : A \to A$ 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 $0 \le i < k$ and $f(i) = a_{i+1}$ for $k \le i < n$.

> **Theorem 1.5** Let A be a set. Then the following are equivalent:
> (1) A is infinite
> (2) A contains a countable subset
> (3) $|\mathbb{N}| \le |A|$
> (4) There exists a map $f : A \to A$ which is injective but not surjective

Proof.

$(1) \implies (2)$ Suppose A is infinite. Since $A \ne \varnothing$ we can choose an element $a_0 \in A$. Since $A \ne \{a_0\}$ we can choose an element $a_1 \in A \backslash \{a_0\}$. Since $A \ne \{a_0, a_1\}$ 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 $A \ne \{a_0, a_1, \cdots, a_{n-1}\}$ we can choose $a_n \in A \backslash \{a_0, a_1, \cdots, a_{n-1}\}$. In this way we obtain $\{a_0, a_1, a_2, \cdots\} \subseteq A$.

$(2) \iff (3)$ 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 $|\mathbb{N}| \le |A|$. 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.

$(2) \implies (4)$ 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).

$(4) \implies (1)$ 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 $A \ne \varnothing$, 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 $n \ge 1$ 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

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

(2) By Theorem 1.7

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

(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 $|\mathbb{N}| \leq |A \cup B|$. 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 $|A \cup B| \leq |\mathbb{N}|$. 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 $A_k = \{a_{k0}, a_{k1}, a_{k2}, \cdots\}$ 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, $|\mathbb{N}| \leq \left| \bigcap\limits_{k=0}^{\infty} A_k \right|$. 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 $\left| \bigcap\limits_{k=0}^{\infty} A_k \right| \leq |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$. 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 $|\mathbb{N}| \leq |\mathbb{Q}|$. 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 $|\mathbb{Q}| \leq |\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N}|$. 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 $g : \mathcal{P}(A) \to 2^A$ 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

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

$$\iff \forall a \in A, (f(a) = 1 \iff a \in S) \iff \{a \in A | f(a) = 1\} = S$$
$$\iff h(f) = S$$

(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 $\left| \bigcup\limits_{n=0}^{\infty} \mathbb{R}^n \right| = 2^{\aleph_0}$

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

2. Metric Spaces

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

    (Sesquilinearity)
        (1) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle, \langle tu, v \rangle = t \langle u, v \rangle$
           $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle, \langle u, tv \rangle = \bar{t} \langle u, v \rangle$
  (2) (Conjugate Symmetry)
        $\langle u, v \rangle = \overline{\langle v, u \rangle}$
  (3) (Positive Definition)
        $\langle u, u \rangle \geq 0$ 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

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

So that $\frac{|\langle u, v \rangle|^2}{|u|^2} < |v|^2$ 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 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.**
> distance 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
>   (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 $d(u, v) = \|u - v\|$. 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 $1 \le p \le \infty$, 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

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

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

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

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 $S \ne \varnothing$ then we use the convention that $\bigcap S = X$, which is open. Suppose that $S \ne \varnothing$. 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 $\bigcap S = \bigcap_{U \in S} U$ is open in X
> (3) If S is a finite set of closed sets in X then the intersection $\bigcup S = \bigcup_{U \in S} U$ 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

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

and so

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

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_∞.   ■

■ **Example 2.8** Show that on the space $\mathcal{C}[a, b]$, the topology induced by the metric d_∞ is strictly finer than the topology induced by the metric d_1

Proof. For $f, g \in \mathcal{C}[a, b]$ we have

$$d_1(f, g) = \int_a^b |f - g| \leq \int_a^b \max_{a \leq x \leq b} |f(x) - g(x)| = (b - a) d_\infty(f, g)$$

It follows that for $f \in \mathcal{C}[a,b]$ and $r > 0$ we have

$$B_\infty(f,r) \subseteq B_1(f, (b-a)r)$$

Thus for $U \subseteq \mathcal{C}[a,b]$, if U is open using d_1 then U is also open using d_∞, and so the topology induced by the metric d_∞ is finer (or equal to) the topology induced by d_1.

On the other hand, we claim that for $f \in \mathcal{C}[a,b]$ and $r > 0$, the set $B_\infty(f,r)$ is not open in the topology induced by d_1. Fix $g \in B_\infty(f,r)$ and let $s > 0$. Choose a bump function $h \in \mathcal{C}[a,b]$ with $h \geq 0$, $\int_a^b h < h$ and $\max\limits_{a \leq x \leq b} h(x) > 2r$. Then we have $g + h \in B_1(g,s)$ but $g + h \notin B_\infty(f,r)$. It follows that $B_\infty(f,r)$ is not open in the topology induced by d_1, as claimed. ∎

■ **Example 2.9** For any set X, the **trivial topology** on X is the topology in which the only open sets in X are the sets $\varnothing$ and X, and the **discrete topology** on X is the topology in which every subset of X is open. Note that the discrete metric on a nonempty set X induces the discrete topology on x.

> **Definition 2.9 — interior, closure, dense.**
> Let X be a metric space (or a topological space) and let $A \subseteq X$. The **interior** and the **closure** of A (in X) are the sets
>
> $$A^\circ = \bigcup \{U \subseteq X \mid U \text{ is open, and } U \subseteq A\}$$
>
> $$\overline{A} = \bigcap \{K \subseteq X \mid K \text{ is closed, and } A \subseteq K\}$$
>
> We say that A is **dense** in X when $\overline{A} = X$.

> **Theorem 2.5** Let X be a metric space (or a topological space) and let $A \subseteq X$
> (1) The interior of A is the largest open set which is contained in A. In other words, $A^\circ \subseteq A$ and A° is open, and for every open set U with $U \subseteq A$ we have $U \subseteq A^\circ$
> (2) The closure of A is the smallest closed set which contains A. In other words, $A \subseteq \overline{A}$ and $\overline{A}$ is closed, and for every closed set K with $A \subseteq K$ we have $\overline{A} \subseteq K$

Proof.
(1) Let $L = \{U \subseteq X \mid U \text{ is open, and } U \subseteq A\}$. Note that A° is open (by Part 2 of Theorem 2.3 or by Part 2 of Definition 2.8) because A° is equal to the union if S, which is a set of open sets. Also note that $A^\circ \subseteq A$ because A° is equal to the union of S, which is a set of subsets of A. Finally note that for any open set U with $U \subseteq A$ we have $U \in S$ so that $U \subseteq \bigcup S = A^\circ$.
(2) The proof is similar to (1) ∎

> **Corollary 2.6**
> Let X be a metric space (or a topological space) and let $A \subseteq X$
> (1) $(A^\circ)^\circ = A^\circ$ and $\overline{\overline{A}} = \overline{A}$
> (2) A is open $\iff A = A^\circ$
> (3) A is closed $\iff A = \overline{A}$

Proof. The proof is left as exercise ∎

Definition 2.10 — interior point, limit point, isolated point, boundary point, boundary.
isolated point Let X be a metric space and let $A \subseteq X$.
An **interior point** of A is a point $a \in A$ such that for some $r > 0$ we have $B(a, r) \subseteq A$.
A **limit point** of A is a point $a \in X$ such that for every $r > 0$ we have $B^*(a, r) \cap A \neq \varnothing$.
An **isolated point** of A is a point $a \in A$ which is not a limit point of A.
A **boundary point** of A is a point of A is a point $a \in X$ such that for every $r > 0$ we have $B(a, r) \cap A \neq \varnothing$ and $B(a, r) \cap A^c \neq \varnothing$.
The set of all limit points of A is denoted by A'. The **boundary** of A, is the set of all boundary points of A.

Theorem 2.7 — Properties of Interior, Limit and Boundary Points.
Let X be a metric space and let $A \subseteq X$
  (1) A° is equal to the set of all interior points of A
  (2) A is closed $\iff A' \subseteq A$
  (3) $\overline{A} = A \cup A'$
  (4) $\partial A = \overline{A} \setminus A^\circ$

Proof. We leave the proofs of (1) and (4) as exercise.
  (2) Note that when $a \notin A$ we have $B(a, r) \cap A = B^*(a, r) \cap A$ and so

$$
\begin{aligned}
A \text{ is closed} &\iff A^c \text{ is open} \\
&\iff \forall a \in A^c, \ \exists r > 0, \ B(a, r) \subseteq A^c \\
&\iff \forall a \in \mathbb{R}^n, \ (a \notin A \implies \exists r > 0, \ B(a, r) \subseteq A^c) \\
&\iff \forall a \in \mathbb{R}^n, \ (a \notin A \implies \exists r > 0, \ B(a, r) \cap A = \varnothing) \\
&\iff \forall a \in \mathbb{R}^n, \ (a \notin A \implies \exists r > 0, \ B^*(a, r) \cap A = \varnothing) \\
&\iff \forall a \in \mathbb{R}^n, \ (\forall r > 0, B^*(a, r) \cap A \neq \varnothing \implies a \in A) \\
&\iff \forall a \in \mathbb{R}^n, \ (a \in A' \implies a \in A) \\
&\iff A' \subseteq A.
\end{aligned}
$$

  (3) We shall prove that $A \cup A'$ is the smallest closed set which contains A. It is clear that $A \cup A'$ is closed, that is $(A \cup A')^c$ is open. Let $a \in (A \cup A')^c$ with, that is let $a \in X$ with $a \notin A$ and $a \notin A'$. Since $a \notin A'$ we can choose $r > 0$ so that $B(a, r) \cap A = \varnothing$. We claim that because $B(a, r) \cap A = \varnothing$ it follows that $B(a, r) \cap A'$. Since $b \in B(a, r)$ and $B(a, r)$ is open, we can choose $s > 0$ so that $B(b, s) \subseteq B(a, r)$. Since $b \in A'$ it follows that $B(b, s) \cap A \neq \varnothing$. Choose $x \in B(b, s) \cap A$. Then we have $x \in B(b, s) \subseteq B(a, r)$ and $x \in A$ and so $x \in B(a, r) \cap A$, which contradicts the fact that $B(a, r) \cap A = \varnothing$. Thus $B(a, r) \cap A' \neq \varnothing$ as claimed. Since $B(a, r) \cap A = \varnothing$ and $B(a, r) \cap A' = \varnothing$, it follows that $B(a, r) \cap (A \cup A') = \varnothing$, hence $B(a, r) \subseteq (A \cup A')^c$. Thus proves that $(A \cup A')^c$ is open, and hence $A \cup A'$ is closed.
  It remains to show that for every closed set K in X with $A \subseteq K$ we have $A \cup A' \subseteq K$. Let K be a closed set in X with $A \subseteq K$. Note that since $A \subseteq K$ it follows that $A' \subseteq K'$ because if $a \in A'$ then for all $r > 0$ we have $B(a, r) \cap A \neq \varnothing$ hence $B(a, r) \cap K \neq \varnothing$ and so $a \in K'$. Since K is closed we have $K' \subseteq K$ by (2). Since $A' \subseteq K'$ and $K' \subseteq K$, we have $A' \subseteq K$. Since AK and $A' \subseteq K$ we have $A \cup A' \subseteq K$ as required.

∎

Ⓡ Let X be a topological space and let $A \subseteq X$, An **interior point** of A is a point $a \in A^\circ$. A **limit point** of A is a point $a \in X$ such that for every open set U in X with $a \in U$ there

exists a point $b \in U \cap A$ with $b \neq a$. The **boundary** of A in X is the set $\partial A = \overline{A} \setminus A^\circ$, and a **boundary point** of A is a point $a \in \partial A$.

Note 2.2 Let X be a metric space and let $P \subseteq X$. Note that P is also a metric space using (the restriction of) the metric used in X. For $a \in P$ and $0 < r \in \mathbb{R}$, note that the open and closed balls in P, centered at a and of radius r, are related to the open and closed balls in X by

$$B_P(a, r) = \{x \in P \mid d(x, a) < r\} = B_X(a, r) \cap P$$

$$\overline{B}_P(a, r) = \{x \in P \mid d(x, a) < r\} = \overline{B}_X(a, r) \cap P$$

Theorem 2.8 Let X be a metric space and let $A \subseteq P \subseteq X$
  (1) A is open in $P \iff$ there exists an open set U in X such that $A = U \cap P$
  (2) A is closed in $P \iff$ there exists a closed set K in X such that $A = K \cap P$

Proof.
  (1) Suppose that A is open in P. For each $a \in A$, choose $r_a > 0$ so that $B_P(a, r_a) \subseteq A$, that is $B_X(a, r_a) \cap P \subseteq A$, and let $U = \bigcup_{a \in A} B_X(a, r_a)$. Since U is equal to the union of a set of open sets in X, it follows that U is open in X. Note that $A \subseteq U \cap P$, and, since $B_X(a, r_a) \cap P \subseteq A$ for every $a \in A$, we also have $U \cap P = (\bigcup_{a \in U} B_X(a, r_a)) \cap P = \bigcup_{a \in A}(B_X(a, r_a) \cap P) \subseteq A$. Thus $A = U \cap P$ as required.
  Conversely, suppose that $A = U \cap P$ with U open in X. Let $a \in A$. Since we have $a \in A = U \cap P$, we also have $a \in U$. Since $a \in U$ and U is open in X we can choose $r > 0$ so that $B_X(a, r) \subseteq U$. Since $B_X(a, r) \subseteq U$ and $U \cap P = A$ we have $B_P(a, r) = B_X(a, r) \cap P \subseteq U \cap P = A$. Thus A is open as required.
  (2) Suppose that A is closed in P. Let B be the complement of A in P, that is $B = P \setminus A$. Then B is open in P. Choose an open set U in X such that $B = U \cap P$. Let K be the complement of U in X, that is $K = X \setminus U$. Then $A = K \cap P$ since for $x \in X$ we have

$$x \in A \iff (x \in P \text{ and } x \notin U \cap P) \iff (x \in P \text{ and } x \notin U)$$

$$\iff (x \in P \text{ and } x \in K) \iff (x \in K \cap P)$$

Conversely, suppose that K is a closed set in P with $A = K \cap P$. Let B be the complement of A in P, that is $B = P \setminus A$, and let U be the complement of K in P, that is $U = P \setminus K$, and note that U is open in P. Then we have $B = U \cap P$ since for $x \in P$ we have

$$x \in B \iff (x \in P \text{ and } x \notin A) \iff (x \in P \text{ and } x \notin K \cap P)$$

$$\iff (x \in P \text{ and } x \notin K) \iff (x \in P \text{ and } x \in U) \iff (x \in U \cap P)$$

Since U us open in P and $B = U \cap P$ we know that B is open in P. Since B is open in P, its complement $A = P \setminus B$ is closed in P.

∎

Ⓡ Let X be a topological space and let $P \subseteq X$. Verify, as an exercise, that we can use the topology on X to define a topology on P as follows. Given a set $A \subseteq P$, we define A to be **open** in P when $A = U \cap P$ for some open set U in X. The resulting topology on P is called the **subspace topology**.

3. Limits and Continuity

Definition 3.1 — bounded, converge, limit, diverge, Cauchy.
Let $(x_n)_{n \geq p}$ be a sequence in a metric space X. We say that the sequence $(x_n)_{n \geq p}$ is **bounded** when the set $\{x_n\}_{n \geq p}$ is bounded, that is when there exists $a \in X$ and $r > 0$ such that $x_n \in B(a, r)$ for all indices $n \geq p$

For $a \in X$, we say that the sequence $(x_n)_{n \geq p}$ **converges** to a (or that the **limit** of x_n is equal to a) and we write $\lim_{n \to \infty} x_n = a$ (or write $x_n \to a$) when for every $\epsilon > 0$ there exists an index $m \geq p$ such that $d(x_n, a) < \epsilon$ for all indices $n \geq m$. We say that the sequence $(x_n)_{n \geq p}$ **converges** (in X) when it converges to some point $a \in X$, and otherwise we say that $(x_n)_{n \geq p}$ **diverges** (in X).

We say that the $(x_n)_{n \geq p}$ is **Cauchy** when for every $\epsilon > 0$ there exists an index $m \geq p$ such that $d(x_k, x_l) < \epsilon$ for all indices $k, l \geq m$.

(R) When $(x_n)_{n \geq p}$ is a sequence in a topological space X and $a \in X$, we say that $(x_n)_{n \geq p}$ **converges** to a (or we say the **limit** of $(x_n)_{n \geq p}$ is equal to a) and we write $\lim_{n \to \infty} x_n = a$ (or we write $x_n \to a$) when for every open set U in X with $a \in U$ there exists an index $m \geq p$ such that $x_n \in U$ for every index $n \geq m$.

Theorem 3.1 — Basic Properties of Limits of Sequences. Let $(x_n)_{n \geq p}$ be a sequence in a metric space X, and let $a \in X$

(1) If $(x_n)_{n \geq p}$ converges then its limit is unique
(2) If $q \geq p$ and $y_n = x_n$ for all $n \geq q$, then $(x_n)_{n \geq p}$ converges if and only if $(y_n)_{n \geq p}$ converges and, in this case, $\lim_{n \to \infty} y_n = \lim_{n \to \infty} x_n$
(3) If $(x_n)_{n \geq p}$ converges then it is bounded
(4) If $(x_n)_{n \geq p}$ converges then it is Cauchy
(5) We have $\lim_{n \to \infty} = a$ if and only if for every open set U in X with $a \in U$ there exists an index $m \geq p$ such that $x_n \in U$ for every index $n \geq m$

Proof. The proof is left as exercise. ∎

> **Note 3.1** Because of Part 2 of the above theorem, the initial index p of a sequence $(x_n)_{n \geq p}$ does not affect whether or not the sequence converges and it does not affect the limit. For this reason, we often omit the initial index p from our notation and write (x_n) for the sequence $(x_n)_{n \geq p}$. Also, we often choose a specific value of p, usually $p = 1$, in the statements or the proofs of various theorems with the understanding that any other initial value would work just as well.

> **Exercise 3.1** Foe each $n \in \mathbb{Z}^+$, let $x_n \in \mathbb{R}^\infty$ be the sequence given by $x_n = \frac{1}{n} \sum_{k=1}^{n} e_k$, that is by $(x_{n,k})_{k \geq 1} = (\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n}, 0, 0, 0, \cdots)$ with n non-zero terms. Show that (x_n) converges in $(\mathbb{R}^\infty, d_1)$.

> **Exercise 3.2** For each $n \in \mathbb{Z}^+$, let $f_n \in \mathcal{C}[0, 1]$ be given by $f_n(x) = \sqrt{n}x^n$. Show that $(f_n)_{n \geq 1}$ converges in $(\mathcal{C}[0, 1], d_1)$ but diverges in $(\mathcal{C}[0, 1], d_2)$.

> **Note 3.2** Recall that $\mathcal{B}[a, b]$ denotes the space of bounded functions $f : [a, b] \to \mathbb{R}$. Let (f_n) be a sequence of bounded functions in $\mathcal{B}[a, b]$ and let $g \in \mathcal{B}[a, b]$. Note that (f_n) converges in the metric space $(\mathcal{B}[a, b], d_\infty)$, if and only if (f_n) converges uniformly on $[a, b]$. Indeed for $\epsilon > 0$ we have $d_\infty(f_n, g) < \epsilon$ if and only if $\sup_{a \leq x \leq b} |f_n(x) - g(x)| < \epsilon$ if and only if $|f_n(x) - g(x)| < \epsilon$ for all $x \in [a, b]$. The same is true for a sequence (f_n) in $\mathcal{C}[a, b]$: (f_n) converges in the metric space $(\mathcal{C}[a, b], d_\infty)$ if and only if (f_n) converges uniformly on $[a, b]$.

> **Theorem 3.2 — The Sequential Characterization of Limit Points and Closed Sets.**
> Let X be a metric space, let $a \in X$, and let $A \subseteq X$.
>   (1) $a \in A'$ if and only if there exists a sequence (x_n) in $A \setminus \{a\}$ with $\lim_{n \to \infty} x_n = a$.
>   (2) $a \in \overline{A}$ if and only if there exists a sequence (x_n) in A with $\lim_{n \to \infty} x_n = a$.
>   (3) A is closed in X if and only if for every sequence (x_n) in A which converges in X, we have $\lim_{n \to \infty} x_n \in A$.

Proof.
  (1) Suppose that $a \in A'$ (which means that for every $r > 0$ we have $B^*(a, r) \cap A \neq \varnothing$). For each $n \in \mathbb{Z}^+$, choose $x_n \in B^*\left(a, \frac{1}{n}\right) \cap A$, that is choose $x_n \in A \setminus \{a\}$ with $d(x_n, a) < \frac{1}{n}$. Then $(x_n)_{n \geq 1}$ is a sequence in $A \setminus \{a\}$ with $\lim_{n \to \infty} x_n = a$.
  Suppose, conversely, that $(x_n)_{n \geq 1}$ is a sequence in $A \setminus \{a\}$ with $\lim_{n \to \infty} x_n = a$. Let $r > 0$. Choose $m \in \mathbb{Z}^+$ such that $d(x_n, a) < r$ for all $n \geq m$. Since $x_m \in A \setminus \{a\}$ with $d(x_m, a) < r$, we have $x_m \in B^*(a, r) \cap A$ and so $B^*(a, r) \cap A \neq \varnothing$.
  (2) Left as exercise
  (3) To prove Part 3, suppose that A is closed in X. Let $(x_n)_{n \geq 1}$ be a sequence in A which converges in X, and let $a = \lim_{n \to \infty} x_n \in X$. Suppose, for a contradiction, that $a \notin A$. Since $a \notin A$ we have $A = A \setminus \{a\}$ so in fact (x_n) is a sequence in $A \setminus \{a\}$. Since (x_n) is a sequence in $A \setminus \{a\}$ with $\lim_{n \to \infty} x_n = a$, it follows from Part 1 that $a \in A'$. Since A is closed we have $A' \subseteq A$ and so $a \in A$ giving the desired contradiction.
  Suppose, conversely, that for every sequence in A which converges in X, the limit

of the sequence lies in A. Let $a \in A'$. By Part 1, we can choose a sequence (x_n) in $A \setminus \{a\}$ with $\lim\limits_{n \to \infty} x_n = a$. Then (x_n) is a sequence in A which converges in X, so its limit lies in A, that is $a \in A$. Since $a \in A'$ was arbitrary, this shows that $A' \subseteq A$, and so A is closed.

$\blacksquare$

■ **Example 3.1** Note that $\mathcal{C}[a, b]$ is closed in the metric space $(\mathcal{B}[a, b], d_\infty)$. We can see this using Note 3.7 together with the above theorem. Indeed, given a sequence (f_n) with each $f_n \in \mathcal{C}[a, b]$, if the sequence (f_n) converges in $(\mathcal{B}[a, b], d_\infty)$ to the function $g \in \mathcal{B}[a, b]$, then (f_n) converges uniformly to g on $[a, b]$, and so (from MATH 148) we know that g must be continuous, hence $g \in \mathcal{C}[a, b]$.

> **Exercise 3.3** Let
>
> $$\mathcal{R}[a, b] = \big\{ f \in \mathcal{B}[a, b] \,\big|\, f \text{ is Riemann integrable} \big\},$$
> $$\mathcal{P}[a, b] = \big\{ f \in \mathcal{B}[a, b] \,\big|\, f \text{ is a polynomial} \big\},$$
> $$\mathcal{C}^1[a, b] = \big\{ f \in \mathcal{B}[a, b] \,\big|\, f \text{ is continuously differentiable} \big\}.$$
>
> Determine which of the above spaces are closed in the metric space $\mathcal{B}[a, b]$, using the supremum metric d_∞.

■ **Example 3.2** Recall that $\mathbb{R}^\infty$ denotes the set of sequences with only finitely many non-zero terms. Show that $\mathbb{R}^\infty$ is dense in the metric space (ℓ_1, d_1).

Proof. Since the closure of $\mathbb{R}^\infty$ in ℓ_1 is contained in ℓ_1 (by the definition of closure), it suffices to show that $\ell_1 \subseteq \overline{\mathbb{R}^\infty}$. Let $a = (a_n)_{n \geq 1} \in \ell_1$, so we have $\sum\limits_{n=1}^{\infty} |a_n| < \infty$. For each $n \in \mathbb{Z}^+$ let $x_n = (x_{n,k})_{k \geq 1}$ be the sequence given by $x_{n,k} = a_k$ for $1 \leq k \leq n$ and $x_{n,k} = 0$ for $k > n$, that is

$$(x_{n,k})_{k \geq 1} = (x_{n,1}, x_{n,2}, \cdots, x_{n,n}, x_{n,n+1}, \cdots) = (a_1, a_2, \cdots, a_n, 0, 0, 0, \cdots).$$

Then each $x_n \in \mathbb{R}^\infty$ and, in the metric space ℓ_1, we have $x_n \to a$ because given $\epsilon > 0$ we can choose an index m so that $\sum\limits_{k > m} |a_k| < \epsilon$ and then for all $n \geq m$ we have

$$\|x_n - a\|_1 = \sum_{k=1}^{\infty} |x_{n,k} - a_k| = \sum_{k > n} |a_k| \leq \sum_{k > m} |a_k| < \epsilon.$$

It follows, from Part 2 of Theorem 3.8, that $a \in \overline{\mathbb{R}^\infty}$ and so we have $\ell_1 \subseteq \overline{\mathbb{R}^\infty}$, as claimed. ■

> **Exercise 3.4** Find the closure of $\mathbb{R}^\infty$ in the metric space ℓ_2 using the metric d_2, and find the closure of $\mathbb{R}^\infty$ in the metric space ℓ_∞ using the metric d_∞.

Definition 3.2 — limit.
Let (X, d_X) and (Y, d_Y) be metric spaces. Let $A \subseteq X$, let $f : A \to Y$, let $a \in A'$, and let $b \in Y$. We say that the **limit** of $f(x)$ as x tends to a is equal to b, when for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in A$, if $0 < d_X(x, a) < \delta$ then $d_Y(f(x), b) < \epsilon$.

> **Theorem 3.3 — The Sequential Characterization of Limits.**
> Let X and Y be metric spaces, let $A \subseteq X$, let $f : A \to Y$, let $a \in A' \subseteq X$, and let $b \in Y$.
> Then $\lim\limits_{x \to a} f(x) = b$ if and only if for every sequence (x_n) in $A \setminus \{a\}$ with $x_n \to a$ we have
> $\lim\limits_{n \to \infty} f(x_n) = b$.

Proof. Suppose that $\lim\limits_{x \to A} f(x) = b$. Let (x_n) be a sequence in $A \setminus \{a\}$ with $x_n \to a$. Let $\epsilon > 0$.
Since $\lim\limits_{x \to a} f(x) = b$ we can choose $\delta > 0$ such that $0 < d(x, a) < \delta \implies d(f(x), b) < \epsilon$.
Since $x_n \to a$ we can choose $m \in \mathbb{Z}^+$ such that $n \geq m \implies d(x_n, a) < \delta$. For $n \geq m$
we have $d(x_n, a) < \delta$ and we have $x_n \neq a$ $(since\, (x_n)$ is a sequence in $A \setminus \{a\}, so that 0 <$
$d(x_n, a) < \delta$, and hence $d(f(x_n), b) < \epsilon$. Thus $\lim\limits_{n \to \infty} f(x_n) = b$, as required.

Suppose, conversely, that $\lim\limits_{x \to a} f(x) \neq b$. Choose $\epsilon > 0$ such that for every $\delta > 0$ there
exists $x \in A$ such that $0 < d(x, a) < \delta$ and $d(f(x), b) \geq \epsilon$. For each $n \in \mathbb{Z}^+$, choose $x_n \in A$
such that $0 < d(x_n, a) < \frac{1}{n}$ and $d(f(x_n), b) \geq \epsilon$. For each n, since $0 < d(x_n, a)$ we have
$x_n \neq a$ so the sequence (x_n) lies in $A \setminus \{a\}$. Since $d(x_n, a) < \frac{1}{n}$ for all $n \in \mathbb{Z}^+$, it follows
that $x_n \to a$. Since $d(f(x_n), b) \geq \epsilon$ for all $n \in \mathbb{Z}^+$, it follows that $\lim\limits_{x \to a} f(x) \neq b$. Thus we
have found a sequence (x_n) in $A \setminus \{a\}$ with $x_n \to a$ such that $\lim\limits_{x \to a} f(x_n) \neq b$.   ∎

> **Definition 3.3 — continuous, uniformly continuous, Lipschitz continuous, Lipschitz constant.**
> Let (X, d_X) and (Y, d_Y) be metric spaces, and let $f : X \to Y$. For $a \in X$, we say that
> f is **continuous** at a when for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in X$,
> if $d_X(x, a) < \delta$ then $d_Y(f(x), f(a)) < \epsilon$. We say that f is **continuous** (on X) when f
> is continuous at every point $a \in X$. We say that f is **uniformly continuous** (on X)
> when for every $\epsilon > 0$ there exists $\delta > 0$ such that for all $x, y \in X$, if $d_X(x, y) < \epsilon$ then
> $d_Y(f(x), f(y)) < \epsilon$. We say that f is **Lipschitz continuous** (on X) when there is a
> constant $\ell \geq 0$, called a **Lipschitz constant** for f, such that for all $x, y \in X$ we. have
> $d(f(x), f(y)) \leq \ell \cdot d(x, y)$. Note that if f is Lipschitz continuous then f is also uniformly
> continuous (indeed we can take $\delta = \frac{\epsilon}{\ell}$ in the definition of uniform continuity).

> **Note 3.3** Let X and Y be metric spaces and let $a \in X$. If a is a limit point of X then f
> is continuous at a if and only if $\lim\limits_{x \to a} f(x) = f(a)$. If a is an isolated point of X then f is
> necessarily continuous at a, vacuously.

> **Theorem 3.4 — The Sequential Characterization of Continuity.**
> Let X and Y be metric spaces using metrics d_X and d_Y, let $f : X \to Y$, and let $a \in X$.
> Then f is continuous at a if and only if for every sequence (x_n) in X with $x_n \to a$ we
> have $\lim\limits_{n \to \infty} f(x_n) = f(a)$.

Proof. The proof is left as an exercise.   ∎

> **Theorem 3.5 — Composition of Continuous Functions.**
> Let X, Y and Z be metric spaces, let $f : X \to Y$, let $g : Y \to Z$. If f is continuous at
> the point $a \in X$ and g is continuous at the point $f(a) \in Y$ then the composite function
> $g \circ f$ is continuous at a.

Proof. The proof is left as an exercise.   ∎

> **Theorem 3.6 — The Topological Characterization of Continuity.**
> Let X and Y be metric spaces and let $f : X \to Y$. Then f is continuous (on X) if an only if $f^{-1}(V)$ is open in X for every open set V in Y.

Proof. Suppose that f is continuous in X. Let V be open in Y. Let $a \in f^{-1}(V)$ and let $f(a) \in V$. Since V is open, we can choose $\epsilon > 0$ such that $B(f(a), \epsilon) \subseteq V$. Since f is continuous at a we can choose $\delta > 0$ such that for all $x \in X$ with $d(x, a) < \delta$ we have $d(f(x), f(a)) < \epsilon$. Then we have $f(B(a, \delta)) \subseteq B(f(a), \epsilon) \subseteq V$ and so $B(a, \delta) \subseteq f^{-1}(V)$. Thus $f^{-1}(V)$ is open in X, as required.

Suppose, conversely, that $f^{-1}(V)$ is open in X for every open set V in Y. Let $a \in X$ and let $\epsilon > 0$. Taking $V = B(f(a), \epsilon)$, which is open in Y, we see that $f^{-1}(B(f(a), \epsilon))$ is open in X. Since $a \in f^{-1}(B(f(a), \epsilon)$ and $f^{-1}(B(f(a), \epsilon))$ is open in X, we can choose $\delta > 0$ such that $B(a, \delta) \subseteq f^{-1}(B(f(a), \epsilon))$. Then we have $f(B(a, \delta)) \subseteq B(f(a), \epsilon)$ or, in other words, for all $x \in X$, if $d(x, a) < \delta$ then $d(f(x), f(a)) < \epsilon$. Thus f is continuous at a hence, since a was arbitrary, f is continuous on X.   ∎

> **Definition 3.4 — continuous.**
> Let X and Y be topological spaces and let $f : X \to Y$. We say that f is **continuous** (on X) when $f^{-1}(V)$ is open in X for every open set V in Y. A bijective map $f : X \to Y$ such that both f and f^{-1} are continuous is called a **homomorphism**.

> **Note 3.4** If U and V are inner product spaces and $L : U \to V$ is an inner product space isomorphism, then L and its inverse preserve distance so they are both continuous (we can take $\delta = \epsilon$ in the definition of continuity), hence L is a homomorphism.
>
> If U and V are finite-dimensional inner product spaces with say $\dim U = n$ and $\dim V = m$, and if $\phi : U \to \mathbb{R}^n$ and $\psi : V \to \mathbb{R}^m$ are inner product space isomorphisms (obtained by choosing orthonormal bases for U and V) then a map $F : U \to V$ is continuous if and only if the composite map $\psi F \phi^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ is continuous. In particular, if F is linear then F is continuous (since $\psi F \phi^{-1} : \mathbb{R}^n \to \mathbb{R}^m$ is linear, hence continuous).
>
> We shall see below that the same is true for finite dimensional normed linear spaces: every linear map between finite dimensional normed linear spaces is continuous. But this is not always true for infinite dimensional spaces.

■ **Example 3.3** Recall from Example 2.24 that every set $U \subseteq \mathcal{C}[a, b]$ which is open using the metric d_1 is also open using the metric d_∞, but not vice versa. It follows that the identity map $I : \mathcal{C} \to \mathcal{C}[a, b]$ given by $I(f) = f$ is continuous as a map from the metric space $(\mathcal{C}[a, b], d_\infty)$ to the metric space $(\mathcal{C}[a, b], d_1)$, but not vice versa.

> **Theorem 3.7** Let U and V be normed linear spaces and let $F : U \to V$ be a linear map. Then the following are equivalent:
>   (1) F is Lipschitz continuous on U,
>   (2) F is continuous at some point $a \in U$,
>   (3) F is continuous at 0, and
>   (4) $F(\overline{B}(0, 1))$ is bounded.
> In this case, if $m \geq 0$ with $F(\overline{B}(0, 1)) \subseteq B(0, m)$ then m is a Lipschitz constant for F.

Proof. It is clear that if F is Lipschitz continuous on U then F is continuous at some point $a \in U$ (indeed F is continuous at every point $a \in U$). Let us show that if F is continuous at some point $a \in U$ then F is continuous at 0. Suppose that F is continuous at $a \in U$.

Let $\epsilon > 0$. Since F is continuous at $a \in U$, we can choose $\delta_1 > 0$ such that for all $u \in U$ we have $\|u - a\| \le \delta_1 \implies \|F(u) - F(a)\| \le 1$. Choose $\delta = \delta_1\epsilon$. Let $x \in U$ with $\|x - 0\| < \delta$. If $x = 0$ then $\|F(x) - F(0)\| = \|0\| = 0$. Suppose that $x \ne 0$. Then for $u = a + \frac{\delta_1 x}{\|x\|}$ we have $\|u - a\| = \left\| \frac{\delta_1 x}{\|x\|} \right\| = \delta_1$ and so $\|F(u - a)\| = \|F(u) - F(a)\| \le 1$, that is $\left\| F\left(\frac{\delta_1 x}{\|x\|} \right) \right\| \le 1$ hence, by the linearity of F and the scaling property of the norm, we have

$$\|F(x) - F(0)\| = \|F(x)\| = \frac{\|x\|}{\delta_1} \left\| F\left(\frac{\delta_1 x}{\|x\|} \right) \right\| \le \frac{\|x\|}{\delta_1} < \frac{\delta_1 \epsilon}{\delta_1} = \epsilon.$$

Thus F is continuous at 0, as required

Next we show that if F is continuous at 0 then $F\left(\overline{B}(0,1) \right)$ is bounded. Suppose that F is continuous at 0. Choose $\delta > 0$ so that for all $u \in U$ we have $\|u\| \le \delta \implies \|F(u)\| \le 1$. Let $m = \frac{1}{\delta}$. For $x \in U$, when $x = 0$ we have $\|F(x)\| = 0 \le m$ and when $0 < \|x\| \le 1$ we have

$$\|F(x)\| = \left\| \frac{\|x\|}{\delta} F\left(\frac{\delta x}{\|x\|} \right) \right\| = \frac{\|x\|}{\delta} \left\| F\left(\frac{\delta x}{\|x\|} \right) \right\| \le \frac{\|x\|}{\delta} = m\|x\| \le m.$$

Thus $F\left(\overline{B}(0,1) \right)$ is bounded, as required.

Finally we show that if $F\left(\overline{B}(0,1) \right)$ is bounded then F is Lipschitz continuous. Suppose that $F\left(\overline{B}(0,1) \right)$ is bounded. Choose $m > 0$ so that $\|F(u)\| \le m$ for all $u \in U$ with $\|u\| \le 1$. Let $x, y \in U$. If $x = y$ then $\|F(x) - F(y)\| = 0$. Suppose that $x \ne y$. Then we have $\left\| \frac{x-y}{\|x-y\|} \right\| = 1$ so that $\left\| F\left(\frac{x-y}{\|x-y\|} \right) \right\| \le m$ and so

$$\|F(x) - F(y)\| = \|F(x - y)\| = \|x - y\| \left\| F\left(\frac{x - y}{\|x - y\|} \right) \right\| \le m\|x - y\|.$$

Thus F is Lipschitz continuous with Lipschitz constant m, as required. ∎

■ **Example 3.4** Define $L : (\mathcal{C}[a,b], d_\infty) \to (\mathcal{C}[a,b], d_\infty)$ by $L(f) = \int_a^x f(t)dt$. Show that L is Lipschitz continuous.

Proof. Let $f \in \mathcal{C}[a,b]$ with $\|f\|_\infty \le 1$, that is with $\max_{a \le x \le b} |f(x)| \le 1$. Then

$$\|F(f)\|_\infty = \max_{a \le x \le b} \left| \int_a^x f(t)\, dt \right| \le \max_{a \le x \le b} \int_a^x 1\, dt = \max_{a \le x \le b} |x - a| = |b - a|.$$

Thus $F\left(\overline{B}(0,1) \right)$ is bounded and so F is uniformly continuous. ∎

■ **Example 3.5** Define $D : \left(\mathcal{C}^1[0,1], d_\infty \right) \to (\mathcal{C}[0,1], d_\infty)$ by $D(f) = f'$. Show that D is not continuous.

Proof. For $n \in \mathbb{Z}^+$, define $f_n : [0,1] \to \mathbb{R}$ by $f_n(x) = x^n$. Then $f_n \in \mathcal{C}^1[a,b]$, and $\|f_n\|_\infty = \max_{0 \le x \le 1} |x^n| = 1$ so that $f_n \in B(0,1)$, and $\|D(f_n)\|_\infty = \max_{0 \le x \le 1} |n\, x^{n-1}| = n$. Thus $D\left(\overline{B}(0,1) \right)$ is not bounded, so D is not continuous (at any point $g \in \mathcal{C}[0,1]$). ∎

■ **Example 3.6** Let X be a metric space and let $\varnothing \ne A \subseteq X$. Define $F : X \to \mathbb{R}$ by

$$F(x) = \text{dist}(x, A) = \inf \left\{ d(x,a) \middle| a \in A \right\}.$$

Show that F is uniformly continuous.

Proof. Given $\epsilon > 0$, chose $\delta = \frac{\epsilon}{2}$. Let $x, y \in X$ with $d(x,y) < \delta = \frac{\epsilon}{2}$. Since $\text{dist}(y,A) = \inf \left\{ d(y,a) \middle| a \in A \right\}$ we can choose $a \in A$ such that $d(y,a) < \text{dist}(y,A) + \frac{\epsilon}{2}$. Then we have

$$\text{dist}(x,A) \le d(x,a) \le d(x,y) + d(y,a) < \frac{\epsilon}{2} + \text{dist}(y,A) + \frac{\epsilon}{2}$$

so that $\mathrm{dist}(x, A) - \mathrm{dist}(y, A) < \epsilon$. Similarly, we have $\mathrm{dist}(y, A) - \mathrm{dist}(x, A) < \epsilon$ and so

$$\left| F(y) - F(x) \right| = \left| \mathrm{dist}(x, A) - \mathrm{dist}(y, A) \right| < \epsilon.$$

∎

Theorem 3.8 Let U be an n-dimensional normed linear space over $\mathbb{R}$. Let $\{u_1, \cdots, u_n\}$ be any basis for U and let $F : \mathbb{R} \to U$ be the associated vector space isomorphism given by $F(t) = \sum_{k=1}^n t_k u_k$. Then both F and F^{-1} are Lipschitz continuous.

Proof. Let $M = \left(\sum_{k=1}^n \|u_k\|^2 \right)^{1/2}$. For $t \in \mathbb{R}^n$ we have

$$\|F(t)\| = \left\| \sum_{k=1}^n t_k u_k \right\| \leq \sum_{k=1}^n |t_k| \|u_k\| \text{ , by the Triangle Inequality,}$$

$$\leq \left(\sum_{k=1}^n t_k^2 \right)^{1/2} \left(\sum_{k=1}^n \|u_k\|^2 \right)^{1/2} \text{ , by the Cauchy-Schwarz Inequality,}$$

$$= M\|t\|.$$

For all $s, t \in \mathbb{R}^n$, $\|F(s) - F(t)\| = \|F(s-t)\| \leq M\|s-t\|$, so F is Lipschitz continuous.

Note that the map $N : U \to \mathbb{R}$ given by $N(x) = \|x\|$ is (uniformly) continuous, indeed we can take $\delta = \epsilon$ in the definition of continuity. Since F and N are both continuous, so is the composite $G = N \circ F : \mathbb{R}^n \to \mathbb{R}$, which given by $G(t) = \|F(t)\|$. By the Extreme Value Theorem, the map G attains its minimum value on the unit sphere $\{t \in \mathbb{R}^n | \|t\| = 1\}$, which is compact. Let $m = \min_{\|t\|=1} G(t) = \min_{\|t\|=1} \|F(t)\|$. Note that $m > 0$ because when $t \neq 0$ we have $F(t) \neq 0$ (since F is a bijective linear map) and hence $\|F(t)\| \neq 0$. For $t \in \mathbb{R}^n$, if $\|t\| > 1$ then we have $\left\| \frac{t}{\|t\|} \right\| = 1$ so, by the choice of m,

$$\|F(t)\| = \|t\| \left\| F\left(\frac{t}{\|t\|} \right) \right\| \geq \|t\| \cdot m > m.$$

It follows that for all $t \in \mathbb{R}^n$, if $\|F(t)\| \leq m$ then $\|t\| \leq 1$. Since F is bijective, it follows that for $x \in U$, if $\|x\| \leq m$ then $\|F^{-1}(x)\| \leq 1$. Thus for all $x \in U$, if $x = 0$ then $\|F^{-1}(x)\| = 0 = \frac{\|x\|}{m}$ and if $x \neq 0$ then since $\left\| \frac{mx}{\|x\|} \right\| = m$ we have

$$\left\| F^{-1}(x) \right\| = \frac{\|x\|}{m} \left\| F^{-1}\left(\frac{mx}{\|x\|} \right) \right\| \leq \frac{\|x\|}{m}.$$

For all $x, y \in U$, we have $\left\| F^{-1}(x) - F^{-1}(y) \right\| = \left\| F^{-1}(x-y) \right\| \leq \frac{1}{m} \|x-y\|$, so F^{-1} is Lipschitz continuous. ∎

Corollary 3.9
When U and V are finite-dimensional normed linear spaces, every linear map $F : U \to V$ is Lipschitz continuous.

Corollary 3.10
Any two norms on a finite-dimensional vector space U induce the same topology on U.

Index