Lecture Notes for PMATH 351, Real Analysis

by Stephen New

Chapter 1. Cardinality

1.1 Definition: Let X and Y be sets and let $f : X \to Y$. Recall that the **domain** of f and the **range** of f are the sets

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

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

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

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

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

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

1.3 Definition: We say that f is **injective** (or **one-to-one**, written as $1\!:\!1$) 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 bijective, we define the **inverse** of f to be the function $f^{-1} : Y \to X$ 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 $(f^{-1})^{-1} = f$.

1.4 Theorem: *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 $g \circ f$, and
(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: To prove Part 1, suppose that f and g are both injective. Let $x_1, x_2 \in X$. If $g(f(x_1)) = g(f(x_2))$ then since g is injective we have $f(x_1) = f(x_2)$, and then since f is injective we have $x_1 = x_2$. Thus $g \circ f$ is injective.

To prove Part 2, suppose that f and g are surjective. Given $z \in Z$, since g is surjective we can choose $y \in Y$ so that $g(y) = z$, then since f is surjective we can choose $x \in X$ so that $f(x) = y$, and then we have $g(f(x)) = g(y) = z$. Thus $g \circ f$ is surjective.

Finally, note that Part 3 follows from Parts 1 and 2.

1.5 Definition: 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 $I_Y \circ f = f$.

1.6 Definition: Let X and Y be sets and let $f : X \to Y$. A **left inverse** of f is a function $g : Y \to X$ such that $g \circ f = I_X$. Equivalently, a function $g : Y \to X$ is a left inverse of f when $g\big(f(x)\big) = x$ 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 $f\big(h(y)\big) = y$ for all $y \in Y$.

1.7 Theorem: *Let X and Y be nonempty sets and let $f : X \to Y$. Then*
(1) f is injective if and only if f has a left inverse,
(2) f is surjective if and only if f has a right inverse, and
(3) f is bijective if and only if f has a left inverse g and a right inverse h, and in this case we have $g = h = f^{-1}$.

Proof: To prove Part 1, suppose first that f is injective. Since $X \neq \emptyset$ 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 that f is 1:1) 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 $g\big(f(x)\big) = x$. Conversely, if f has a left inverse, say g, then f is 1:1 since for all $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$ then $x_1 = g\big(f(x_1)\big) = g\big(f(x_2)\big) = x_2$.

To prove Part 2, suppose first that f is onto. For each $y \in Y$, choose $x_y \in X$ with $f(x_y) = y$, then define $g : Y \to X$ 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 $f\big(g(y)\big) = f(x_y) = y$. 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 $f(x) = f\big(g(y)\big) = y$.

To prove Part 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 $f^{-1}\big(f(x)\big) = f^{-1}(y) = x$. 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 $f\big(f^{-1}(y)\big) = f(x) = y$. 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 Part 1. Since f has a right inverse, it is surjective by Part 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

$$g(y) = g\Big(f\big(f^{-1}(y)\big)\Big) = f^{-1}(y) = f^{-1}\Big(f\big(h(y)\big)\Big) = h(y)\,.$$

1.8 Corollary: *Let X and Y be nonempty 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.

1.9 Definition: Let A and B be sets. We say that A and B have the **same cardinality**, and we 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 we write $|A| \leq |B|$, when there exists an injective map $f : A \to B$ (or equivalently when there exists a surjective map $g : B \to A$). We say that the cardinality of A is **less than** the cardinality of B, and we 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 $|B| < |A|$.

1.10 Example: Let $\mathbf{N} = \big\{n \in \mathbf{Z} \,\big|\, n \geq 0\big\} = \{0, 1, 2, \cdots\}$. The map $f : \mathbf{N} \to 2\mathbf{N}$ given by $f(k) = 2k$ is bijective, so $|2\mathbf{N}| = |N|$. The map $g : \mathbf{N} \to \mathbf{Z}$ given by $g(2k) = k$ and $g(2k+1) = -k-1$ for $k \in \mathbf{N}$ is bijective, so we have $|\mathbf{Z}| = |\mathbf{N}|$. The map $h : \mathbf{N} \times \mathbf{N} \to \mathbf{N}$ given by $h(k, l) = 2^k(2l+1) - 1$ is bijective, so we have $|\mathbf{N} \times \mathbf{N}| = |\mathbf{N}|$.

1.11 Theorem: *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) $|A| \leq |B|$ if and only if ($|A| = |B|$ or $|A| < |B|$), and
(5) if $|A| \leq |B|$ and $|B| \leq |C|$ then $|A| \leq |C|$.

Proof: Part 1 holds because the identity function $I_A : A \to A$ is bijective. Part 2 holds because if $f : A \to B$ is bijective then so is $f^{-1} : B \to A$. Part 3 holds because if $f : A \to B$ and $g : B \to C$ are bijective then so is the composite $g \circ f : A \to C$. The rest of the proof is left as an exercise.

1.12 Definition: Let A be a set. For each $n \in \mathbf{N}$, let $S_n = \{0, 1, 2, \cdots, n-1\}$. For $n \in \mathbf{N}$, 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 $n \in \mathbf{N}$. We say that A is **infinite** when A is not finite. We say that A is **countable** when $|A| = |\mathbf{N}|$.

1.13 Note: 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.

1.14 Note: For $n \in \mathbf{N}$, if A is a finite set with $|A| = n+1$ and $a \in A$ then $|A \setminus \{a\}| = n$. 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 $A \setminus \{a\} = \{a_0, a_1, \cdots, a_{k-1}, a_{k+1}, \cdots, a_n\}$, then we can define a bijection $f : S_n \to A \setminus \{a\}$ by $f(i) = a_i$ for $0 \leq i < k$ and $f(i) = a_{i+1}$ for $k \leq i < n$.

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

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

Proof: To prove that (1) implies (2), suppose that A is infinite. Since $A \neq \emptyset$ we can choose an element $a_0 \in A$. Since $A \neq \{a_0\}$ we can choose an element $a_1 \in A \setminus \{a_0\}$. Since $A \neq \{a_0, a_1\}$ we can choose $a_3 \in A \setminus \{a_0, a_1\}$. 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 $a_n \in A \setminus \{a_0, a_1, \cdots, a_{n-1}\}$. In this way, we obtain a countable set $\{a_0, a_1, a_2, \cdots\} \subseteq A$.

Next we show that (2) is equivalent to (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 the a_i are distinct, the map $f : \mathbf{N} \to A$ given by $f(k) = a_k$ is injective, and so we have $|\mathbf{N}| \leq |A|$. Conversely, suppose that $|\mathbf{N}| \leq |A|$, and chose an injective map $f : \mathbf{N} \to A$. Considered as a map from $\mathbf{N}$ to $f(\mathbf{N})$, f is bijective, so we have $|\mathbf{N}| = |f(\mathbf{N})|$ hence $f(\mathbf{N})$ is a countable subset of A.

Next, let us show that (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 $k \in \mathbf{N}$ and by $f(b) = b$ for all $b \in A \setminus \{a_0, a_1, a_2, \cdots\}$. Then f is injective but not surjective (the element a_0 is not in the range of f).

Finally, to prove that (4) implies (1) 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 $A = \emptyset$, 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 $g : B \to B \setminus \{b\}$. Let $A = B \setminus \{b\}$ 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 $x \in B \setminus \{b\}$ with $g(x) = g(b)$, so there is no element $x \in A$ with $f(x) = g(b)$, and so f is not surjective. Since $|A| = n$ (by the above note), this contradicts the induction hypothesis. Thus g must be surjective. By the Principle of Induction, for every $n \in \mathbf{N}$ and for every set A with $|A| = n$, every injective function $f : A \to A$ is surjective.

1.16 Corollary: *Let A and B be sets.*

(1) If A is countable then A is infinite.
(2) When $|A| \leq |B|$, if B is finite then 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: Part 1 is immediate: if A is countable then A contains a countable subset (itself), so A is infinite, by Theorem 1.15.

To prove Part 2, suppose that $|A| \leq |B$ and that $|A|$ is infinite. Since A is infinite, we have $|\mathbf{N}| \leq |A|$ (by Theorem 1.15). Since $|\mathbf{N}| \leq |A|$ and $|A| \leq |B|$ we have $|\mathbf{N}| \leq |B|$ (by Theorem 1.11). Since $|\mathbf{N}| \leq |B|$, B is infinite (by Theorem 1.15 again).

To Prove Part 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 $S_m \subsetneqq S_n$. Since $|A| = |B|$ we have $|S_n| = |A| = |B| = |S_m|$ so we can choose a bijection $f : S_n \to S_m$. Since $S_m \subsetneqq S_n$, we can consider f as a function $f : S_n \to S_n$ which is injective but not surjective. This contradicts Theorem 1.16, and so we must have $n = m$. This proves Part 3.

To prove Part 4, we again suppose that $|A| = n$ 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 $S_m \subsetneqq S_n$ so we can consider f as a map $f : S_n \to S_n$, and this map is injective but not surjective. This contradicts Theorem 1.15, and so $n \leq m$.

Finally, to prove Part 5 we 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 $|B| \leq |A|$, Part 2 implies that B is finite. If B is finite then, since $|A| \leq |B|$, Part 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 Parts 3 and 4.

1.17 Theorem: *Let A be a set. Then $|A| \leq |\mathbf{N}|$ if and only if A is finite or countable.*

Proof: First we claim that every subset of $\mathbf{N}$ is either finite or countable. Let $A \subseteq \mathbf{N}$ and suppose that A is not finite. Since $A \neq \emptyset$, we can set $a_0 = \min A$ (using the Well-Ordering Property of $\mathbf{N}$). Note that $\{0, 1, \cdots, a_0\} \cap A = \{a_0\}$. Since $A \neq \{a_0\}$ (so the set $A \setminus \{a_0\}$ is nonempty) we can set $a_1 = \min A \setminus \{a_0\}$. Then we have $a_0 < a_1$ and $\{0, 1, 2, \cdots, a_1\} \cap A = \{a_0, a_1\}$. Since $A \neq \{a_0, a_1\}$ we can set $a_2 = \min A \setminus \{a_0, a_1\}$. Then we have $a_0 < a_1 < a_2$ and $\{0, 1, 2, \cdots, a_3\} \cap A = \{a_0, a_1, a_2\}$. 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 $A \cap \{0, 1, \cdots, a_{n-1}\} = \{a_0, a_1, \cdots, a_{n-1}\}$, since $A \neq \{a_0, a_1, \cdots, a_{n-1}\}$ we can set $a_n = \min A \setminus \{a_0, a_1, \cdots, a_{n-1}\}$, and then we have $a_0 < a_1 < \cdots < a_{n-1} < a_n$ and $A\{0, 1, 2, \cdots, a_n\} \cap A = \{a_0, a_1, \cdots, a_n\}$. 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 $m \in \mathbf{N}$, $\{0, 1, 2, \cdots, a_m\} \cap A = \{a_0, a_1, \cdots, a_m\}$. Since $0 \leq a_0 < a_1 < a_2 < \cdots$, it follows (by induction) that $a_k \geq k$ for all $k \in \mathbf{N}$. 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 $\mathbf{N}$ is either finite or countable.

Now suppose that $|A| \leq |N|$ and choose an injective map $f : A \to \mathbf{N}$. 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 $f(A) \subseteq \mathbf{N}$, 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 $|A| = |f(A)| = |\mathbf{N}|$. Thus A is finite or countable.

1.18 Theorem: *Let A be a set. Then*

(1) $|A| < |\mathbf{N}|$ if and only if A is finite,
(2) $|\mathbf{N}| < |A|$ if and only if A is neither finite nor countable, and
(3) if $|A| \leq |\mathbf{N}|$ and $|\mathbf{N}| \leq |A|$ then $|A| = |\mathbf{N}|$.

Proof: Part 1 follows from Theorem 1.15 because

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

and Part 2 follows from Theorem 1.17 because

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

To prove Part 3, suppose that $|A| \leq |\mathbf{N}|$ and $|\mathbf{N}| \leq |A|$. Since $|A| \leq |\mathbf{N}|$, we know that A is finite or countable by Theorem 1.17. Since $|\mathbf{N}| \leq |A|$, we know that that A is infinite by Theorem 1.15. Since A is finite or countable and A is not finite, it follows that A is countable. Thus $|A| = |\mathbf{N}|$.

1.19 Definition: 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**.

1.20 Theorem:

(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 $\bigcup_{k=0}^{\infty} A_k$.
(4) $\mathbf{Q}$ is countable.

Proof: To prove both Parts 1 and 2, 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 the b_i distinct. Since every positive integer can be written uniquely in the form $2^k(2l+1)$ with $k, l \in \mathbf{N}$, the map $f : A \times B \to \mathbf{N}$ given by $f(a_k, b_l) = 2^k(2l+1)-1$ is bijective, and so $|A \times B| = |\mathbf{N}|$. This proves Part 1. Since the map $g : \mathbf{N} \to A \cup B$ given by $g(k) = a_k$ is injective, we have $|\mathbf{N}| \leq |A \cup B|$. Since the map $h : \mathbf{N} \to A \cup B$ given by $h(2k) = a_k$ and $h(2k+1) = b_k$ is surjective, we have $|A \cup B| \leq |N|$. Since $|\mathbf{N}| \leq |A \cup B|$ and $|A \cup B| \leq |\mathbf{N}|$, we have $|A \cup B| = |\mathbf{N}|$ by Part 3 of Theorem 1.18. This proves 2.

To prove Part 3, for each $k \in \mathbf{N}$, let $A_k = \{a_{k0}, a_{k1}, a_{k2}, \cdots\}$ with the a_{ki} distinct. Since the map $f : \mathbf{N} \to \bigcup_{k=0}^{\infty} A_k$ given by $f(k) = a_{0,k}$ is injective, $|\mathbf{N}| \leq \left|\bigcup_{k=0}^{\infty} A_k\right|$. Since $\mathbf{N} \times \mathbf{N}$ is countable by Part (1), and since the map $g : \mathbf{N} \times \mathbf{N} \to \bigcup_{k=0}^{\infty} A_k$ given by $g(k,l) = a_{k,l}$ is surjective, we have $\left|\bigcup_{k=0}^{\infty} A_k\right| \leq |\mathbf{N} \times \mathbf{N}| = |\mathbf{N}|$. By Part 3 of Theorem 1.18, we have $\left|\bigcup_{k=0}^{\infty} A_k\right| = |\mathbf{N}|$, as required.

Finally, we prove Part 4. Since the map $f : \mathbf{N} \to \mathbf{Q}$ given by $f(k) = k$ is injective, we have $|\mathbf{N}| \leq |\mathbf{Q}|$. Since the map $g : \mathbf{Q} \to \mathbf{Z} \times \mathbf{Z}$, given by $g\left(\frac{a}{b}\right) = (a,b)$ for all $a, b \in \mathbf{Z}$ with $b > 0$ and $\gcd(a,b) = 1$, is injective, and since $\mathbf{Z} \times \mathbf{Z}$ is countable, we have $|\mathbf{Q}| \leq |\mathbf{Z} \times \mathbf{Z}| = |\mathbf{N}|$. Since $|\mathbf{N}| \leq |\mathbf{Q}|$ and $|\mathbf{Q}| \leq |\mathbf{N}|$, we have $|\mathbf{Q}| = |\mathbf{N}|$, as required.

1.21 Exercise: 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.

1.22 Definition: 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 $S_2 = \{0,1\}$.

1.23 Theorem:

(1) For every set A, $\left|\mathcal{P}(A)\right| = \left|2^A\right|$.
(2) For every set A, $|A| < \left|\mathcal{P}(A)\right|$.
(3) $\mathbf{R}$ is uncountable.

Proof: 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 a 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 $h(f) \in \mathcal{P}(A)$ to be the subset

$$h(f) = \left\{a \in A \middle| f(a) = 1\right\} \subseteq A.$$

The 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

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

This completes the proof of Part 1.

Let us prove Part 2. Again we let A be any set. Since the the map $f : A \to \mathcal{P}(A)$ given by $f(a) = \{a\}$ is injective, we have $|A| \leq \left|\mathcal{P}(A)\right|$. We need to show that $|A| \neq \left|\mathcal{P}(A)\right|$. Let $g : A \to \mathcal{P}(A)$ be any map. Let $S = \left\{a \in A \middle| a \notin g(a)\right\}$. Note that S cannot be in the range of g because if 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 not possible. 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 $|A| \neq \left|\mathcal{P}(A)\right|$, as desired.

Finally, we shall prove that $\mathbf{R}$ 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 $g : 2^{\mathbf{N}} \to \mathbf{R}$ as follows. Given $f \in 2^{\mathbf{N}}$, that is given a map $f : \mathbf{N} \to \{0, 1\}$, we define $g(f)$ to be the real number $g(f) \in [0, 1)$ with the decimal expansion $g(f) = 0.f(0)f(1)f(2)f(3)\cdots$, 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 $\left|2^{\mathbf{N}}\right| \leq |\mathbf{R}|$. Thus $|\mathbf{N}| < \left|\mathcal{P}(\mathbf{N})\right| = \left|2^{\mathbf{N}}\right| \leq |\mathbf{R}|$, and so $\mathbf{R}$ is uncountable, by Part 2 of Theorem 1.18.

1.24 Theorem: *(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 $f : g(B) \to f\big(g(B)\big)$ are bijective we have $|A| = |f(A)|$ and $|B| = |g(B)| = \big|f(g(B))\big|$. Also note that $f\big(g(B)\big) \subseteq f(A) \subseteq B$. Let $X = f\big(g(B)\big)$, $Y = f(A)$ 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 bijection. Define sets Z_n and Y_n for $n \in \mathbf{N}$ 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

$$Z_1 \subseteq Y_0 \subseteq Z_0.$$

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

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

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 $U_n = Z_n \setminus Y_n$, $U = \bigcup_{n=0}^{\infty} U_n$ and $V = Z \setminus U$. 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.

1.25 Example: Show that $|\mathbf{R}| = \left|2^{\mathbf{N}}\right|$.

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

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

1.27 Theorem: *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 B are disjoint then $|A \cup B| = |A| + |B|$,
(2) $|A \times B| = |A| \cdot |B|$,
(3) $\left|A^B\right| = |A|^{|B|}$, and
(3) $|\mathcal{P}(A)| = 2^{|A|}$.

Proof: The proof is left as an exercise.

1.28 Theorem: *Let A, B, C and D be sets with $|A| = |C|$ and $|B| = |D|$. Then*
(1) if $A \cap B = \emptyset$ and $C \cap D = \emptyset$ then $\big|A \cup B\big| = \big|C \cup D\big|$,
(2) $\big|A \times B\big| = \big|C \times D\big|$, and
(3) $\left|A^B\right| = \left|C^D\right|$.

Proof: The proof is left as an exercise.

1.29 Remark: 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 $|A| = |\kappa|$ and, in this case, we define the **cardinality** of the set A to be $|A| = \kappa$. In foundational set theory, the natural numbers are defined, formally, to be equal to the sets $0 = \emptyset$, $1 = \{0\} = \{\emptyset\}$, $2 = \{0,1\} = \big\{\emptyset, \{\emptyset\}\big\}$ 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 $n = S_n = \big\{0, 1, \cdots, n-1\big\}$. 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

$$\begin{aligned} \kappa + \lambda &= \big|(\kappa \times \{0\}) \cup (\lambda \times \{1\})\big|, \\ \kappa \cdot \lambda &= \big|\kappa \times \lambda\big|, \\ \kappa^\lambda &= \big|\kappa^\lambda\big|. \end{aligned}$$

1.30 Theorem: *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) $\lambda\leq\mu\Longrightarrow\kappa+\lambda\leq\kappa+\mu$,
(5) $\kappa\cdot\lambda=\lambda\cdot\kappa$,
(6) $(\kappa\cdot\lambda)\cdot\mu=\kappa\cdot(\lambda\cdot\mu)$,
(7) $\kappa\cdot 1=\kappa$,
(8) $\kappa\cdot(\lambda+\mu)=(\kappa\cdot\lambda)+(\kappa\cdot\mu)$,
(9) $\lambda\leq\mu\Longrightarrow\kappa\cdot\lambda\leq\kappa\cdot\mu$,
(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) $\lambda\leq\mu\Longrightarrow\kappa^\lambda\leq\kappa^\mu$, *and*
(14) $\kappa\leq\lambda\Longrightarrow\kappa^\mu\leq\lambda^\mu$.

Proof: We sketch a proof for Parts 9 and 11 and leave the rest as an exercise. To prove Part 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 $F(a,b)=\big(a,f(b)\big)$ then verify that F is injective.

To prove Part 11, let A, B and C be sets with $|A|=\kappa$, $|B|=\lambda$ and $|C|=\mu$. We need to show that $\big|(A^B)^C\big|=\big|A^{B\times C}\big|$. 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 $G(g)(c)(b)=g(b,c)$.

1.31 Exercise: Show that $\left|\bigcup_{n=0}^{\infty}\mathbf{R}^n\right|=2^{\aleph_0}$.

1.32 Exercise: Find $\big|\mathbf{R}^{[0,1]}\big|$.

Chapter 2. Metric Spaces

2.1 Definition: Let $F = \mathbf{R}$ or $\mathbf{C}$. Let U be a vector space over F. An **inner product** on U (over F) is a 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) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle\ ,\ \langle tu, v\rangle = t\,\langle u, v\rangle$,
$\langle u, v+w\rangle = \langle u, v\rangle + \langle u, w\rangle\ ,\ \langle u, tv\rangle = \overline{t}\,\langle u, v\rangle$,
(2) (Conjugate Symmetry) $\langle u, v\rangle = \overline{\langle v, u\rangle}$, and
(3) (Positive Definiteness) $\langle u, u\rangle \geq 0$ 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 $\big\langle L(x), L(y)\big\rangle = \langle x, y\rangle$ for all $x, y \in U$. A bijective homomorphism is called an **isomorphism**.

2.2 Theorem: *Let U be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$ and let $u, v \in U$. Then if $\langle x, u\rangle = \langle x, v\rangle$ for all $x \in U$, or if $\langle u, x\rangle = \langle v, x\rangle$ for all $x \in U$, then $u = v$.*

Proof: Suppose that $\langle x, u\rangle = \langle x, v\rangle$ for all $x \in U$. Then $\langle x, u-v\rangle = \langle x, u\rangle - \langle x, v\rangle = 0$ for all $x \in U$. In particular, taking $x = u - v$ we have $\langle u-v, u-v\rangle = 0$ so that $u = v$, by positive definiteness. Similarly, if $\langle u, x\rangle = \langle v, x\rangle$ for all $x \in U$ then $u = v$.

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

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

2.4 Theorem: *Let U be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. 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 $F = \mathbf{R}$ then $\langle u, v\rangle = \frac{1}{4}\big(||u+v|| - ||u-v||\big)$ and
if $F = \mathbf{C}$ then $\langle u, v\rangle = \frac{1}{4}\big(||u+v||^2 + i||u+iv||^2 - ||u-v||^2 - i||u-iv||^2\big)$,
(5) (The Cauchy-Schwarz Inequality) $\big|\langle u, v\rangle\big| \leq ||u||\,||v||$ with $|\langle u, v\rangle| = ||u||\,||v||$ if and only if $\{u, v\}$ is linearly dependent, and
(6) (The Triangle Inequality) $\big|\|u\| - \|v\|\big| \leq \|u+v\| \leq \|u\| + \|v\|$.

Proof: The first 4 parts are all easy to prove. To prove Part 5, suppose first that $\{u, v\}$ is linearly dependent. Then one of u and v is a multiple of the other, say $v = tu$ with $t \in F$. Then we have $|\langle u, v\rangle| = |\langle u, tu\rangle| = \big|\overline{t}\,\langle u, u\rangle\big| = |t|\,||u||^2 = ||u||\,||tu|| = ||u||\,||v||$. 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

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

so that $\frac{|\langle u,v\rangle|^2}{|u|^2} < |v|^2$ and hence $|\langle u,v\rangle| \leq |u|\,|v|$. This proves Part 5.

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

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

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 $\big|||u|| - ||v||\big| \leq ||u+v||$. This proves Part 6.

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

(1) (Scaling) $||tu|| = |t|\,||u||$,
(2) (Positive Definiteness) $||u|| \geq 0$ with $||u|| = 0 \iff u = 0$, and
(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 $\big||L(x)|\big| = ||x||$ for all $x \in U$. A bijective homomorphism is called an **isomorphism**.

2.6 Definition: Let $F = \mathbf{R}$ or $\mathbf{C}$ 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

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

2.7 Theorem: *Let U be as normed linear space over $F = \mathbf{R}$ or $\mathbf{C}$. For all $u, v, w \in U$,*

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

Proof: The proof is left as an easy exercise.

2.8 Definition: Let X be a non-empty set. A **metric** on X is a map $d : X \times X \to \mathbf{R}$ 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 $d(a,b) = 0 \iff a = b$, and
(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 $d_Y\big(f(a), f(b)\big) = d_X(a,b)$ for all $a, b \in X$. A bijective homomorphism is called an **isomorphism** or an **isometry**.

2.9 Note: Every inner product space is also a normed linear space, using the induced norm given by $\|u\| = \sqrt{\langle u, u\rangle}$. Every normed linear space is also a metric space, using the induced metric given by $d(u,v) = \|v-u\|$. If U is an inner product space over $\mathbf{F}=\mathbf{R}$ or $\mathbf{C}$ 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 $\mathbf{F}=\mathbf{R}$ or $\mathbf{C}$ 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.

2.10 Example: Let $\mathbf{F}=\mathbf{R}$ or $\mathbf{C}$. 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

$$\|u\|_2 = \|u\| = \sqrt{\langle u, u\rangle} = \Big(\sum_{i=1}^n |u_i|^2\Big)^{1/2}.$$

The standard norm on $\mathbf{F}^n$ induces the **standard metric** on $\mathbf{F}^n$, given by

$$d_2(u,v) = d(u,v) = \|v-u\| = \Big(\sum_{i=1}^n |v_i - u_i|^2\Big)^{1/2}.$$

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 the **infinity norm**, on $\mathbf{F}^n$ is given by

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

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

2.11 Example: Let $\mathbf{F}=\mathbf{R}$ or $\mathbf{C}$. We write

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

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

$$\|u\|_2 = \sqrt{\langle u, u\rangle} = \Big(\sum_{i=1}^\infty |u_i|^2\Big)^{1/2}.$$

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 $\mathbf{F}^\infty$ is given by

$$\|u\|_\infty = \max\big\{|u_1|, |u_2|, |u_3|\cdots\big\}$$

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

2.12 Example: For $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$, 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 these definitions to various subspaces of $\mathbf{F}^\omega$. We define

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

Verify that $\ell_1(\mathbf{F})$ is a normed linear space using the **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** $d_\infty(u,v) = \|v-u\|_\infty$. 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 $\|u\|_2 = \left(\sum_{i=1}^{\infty} |u_i|^2 \right)^{1/2}$ and the **standard metric**, or the **2-metric**, $d_2(u,v) = \|v-u\|_2$.

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

$$\ell_p = \ell_p(\mathbf{R}) \,.$$

2.13 Example: For $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$ and for $a, b \in \mathbf{R}$ with $a \le b$, we write

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

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

$$\begin{aligned}
\|f\|_1 &= \int_a^b |f| \,, \\
\|f\|_2 &= \left(\int_a^b |f|^2 \right)^{1/2} , \\
\|f\|_\infty &= \sup_{a \le x \le b} \big| f(x) \big| \,.
\end{aligned}$$

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

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

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

$$\mathcal{F}[a,b] = \mathcal{F}\big([a,b], \mathbf{R}\big) \;,\; \mathcal{B}[a,b] = \mathcal{B}\big([a,b], \mathbf{R}\big) \;\text{ and }\; \mathcal{C}[a,b] = \mathcal{C}\big([a,b], \mathbf{R}\big).$$

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

$$\|u\|_p=\Big(\sum_{i=1}^n |u_i|^p\Big)^{1/p},$$

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

$$\|u\|_p=\Big(\sum_{i=1}^\infty |u_i|^p\Big)^{1/p}.$$

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

$$\|f\|_p=\left(\int_a^b |f|^p\right)^{1/p}.$$

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

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

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

When the metric on X is 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 $0<r\in\mathbf{R}$.

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

2.18 Definition: 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 $A^c=X\setminus A$ is open in X.

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

Solution: 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 note 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 $x\notin\overline{B}(a,r)$ 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.

2.20 Theorem: *(Basic Properties of Open Sets) Let X be a metric Space.*

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

Proof: The empty set is open because any statement of the form "for all $x\in\emptyset$ 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 we have $B(a,r) \subseteq X$ by the definition of $B(a,r)$. This proves Part 1.

To prove Part 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, as required.

To prove Part 3, let S be a finite set of open sets in X. If $S = \emptyset$ then we use the convention that $\bigcap S = X$, which is open. Suppose that $S \neq \emptyset$, 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 $B(a,r) \subseteq U_k$ for every index k, it follows that $B(a,r) \subseteq \bigcap_{k=1}^m U_k = \bigcap S$. Thus $\bigcap S$ is open, as required.

2.21 Theorem: *(Basic Properties of Closed Sets) Let X be a metric space.*

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

Proof: The proof is left as an exercise

2.22 Definition: A **topology** on a set X is a set T of subsets of X such that

(1) $\emptyset \in T$ and $X \in T$,
(2) for every set $S \subseteq T$ we have $\bigcup S \in T$, and
(3) for every finite subset $S \subseteq T$ we have $\bigcap S \in T$.

A **topological 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 that the topology S is **coarser** than the topology T.

2.23 Example: Show that in $\mathbf{R}^n$, the metrics d_1, d_2 and d_∞ all induce the same topology.

Solution: For $a, x \in \mathbf{R}^n$ we have

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

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 $a \in \mathbf{R}^n$ and $r > 0$ we have

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

Thus for $U \subseteq \mathbf{R}^n$, if U is open in $\mathbf{R}^n$ 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_∞.

2.24 Example: 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.

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

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

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

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

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 $h\geq 0$, $\int_a^b h<s$ and $\max_{a\leq x\leq b}h(x)>2r$. 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.

2.25 Example: For any set X, the **trivial topology** on X is the the topology in which the only open sets in X are the sets $\emptyset$ 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.

2.26 Definition: 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

$$A^o=\bigcup\big\{U\subseteq X\,\big|\,U\text{ is open, and }U\subseteq A\big\},$$
$$\overline{A}=\bigcap\big\{K\subseteq X\,\big|\,K\text{ is closed and }A\subseteq K\big\}.$$

We say that A is **dense** in X when $\overline{A}=X$.

2.27 Theorem: *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, $A^o\subseteq A$ and A^o is open, and for every open set U with $U\subseteq A$ we have $U\subseteq A^o$.

(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: Let $S=\big\{U\subseteq X\,\big|\,U\text{ is open, and }U\subseteq A\big\}$. Note that A^o is open (by Part 2 of Theorem 2.20 or by Part 2 of Definition 2.22) because A^o is equal to the union of S, which is a set of open sets. Also note that $A^o\subseteq A$ because A^o 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 $U\subseteq\bigcup S=A^o$. This completes the proof of Part 1, and the proof of Part 2 is similar.

2.28 Corollary: *Let X be a metric space (or a topological space) and let $A\subseteq X$.*
(1) $(A^o)^o=A^o$ and $\overline{\overline{A}}=\overline{A}$.
(2) A is open if and only if $A=A^o$
(3) A is closed if and only if $A=\overline{A}$.

Proof: The proof is left as an exercise.

2.29 Definition: 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 $B^*(a,r) \cap A \neq \emptyset$. 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 $a \in X$ such that for every $r > 0$ we have $B(a,r) \cap A \neq \emptyset$ and $B(a,r) \cap A^c \neq \emptyset$. 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.

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

Proof: We leave the proofs of Parts 1 and 4 as exercises. To prove Part 2, note that when $a \notin A$ we have $B(a,r) \cap A = B^*(a,r) \cap A$ and so

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

To prove Part 3 we shall prove that $A \cup A'$ is the smallest closed set which contains A. It is clear that $A \cup A'$ contains A. We claim that $A \cup A'$ is closed, that is $(A \cup A')^c$ is open. Let $a \in (A \cup A')^c$, 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 $B(a,r) \cap A = \emptyset$. We claim that because $B(a,r) \cap A = \emptyset$ it follows that $B(a,r) \cap A' = \emptyset$. Suppose, for a contradiction, that $B(a,r) \cap A' \neq \emptyset$. Choose $b \in B(a,r) \cap A'$. 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 $B(b,s) \cap A \neq \emptyset$. 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 $B(a,r) \cap A = \emptyset$. Thus $B(a,r) \cap A' = \emptyset$, as claimed. Since $B(a,r) \cap A = \emptyset$ and $B(a,r) \cap A' = \emptyset$ it follows that $B(a,r) \cap (A \cup A') = \emptyset$ 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 $B(a,r) \cap A \neq \emptyset$ hence $B(a,r) \cap K \neq \emptyset$ and so $a \in K'$. Since K is closed we have $K' \subseteq K$ by Part 2. Since $A' \subseteq K'$ and $K' \subseteq K$ we have $A' \subseteq K$. Since $A \subseteq K$ and $A' \subseteq K$ we have $A \cup A' \subseteq K$, as required. This completes the proof of Part 3.

2.31 Remark: Let X be a topological space and let $A \subseteq X$. An **interior point** of A is a point $a \in A^o$. 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 $\partial A = \overline{A} \setminus A^o$, and a **boundary point** of A is a point $a \in \partial A$.

2.32 Note: 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 $0 < r \in \mathbf{R}$, note that the open and closed balls in P, centred at a and of radius r, are related to the open and closed balls in X by

$$B_P(a,r) = \big\{x \in P \,\big|\, d(x,a) < r\big\} = B_X(a,r) \cap P,$$
$$\overline{B}_P(a,r) = \big\{x \in P \,\big|\, d(x,a) \leq r\big\} = \overline{B}_X(a,r) \cap P.$$

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

Proof: To prove Part 1, suppose first 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 $U \cap P = \Big(\bigcup_{a \in U} B_X(a,r_a)\Big) \cap P = \bigcup_{a \in A}\big(B_X(a,r_a) \cap P\big) \subseteq A$. Thus $A = U \cap P$, as required.

Suppose, conversely, 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.

To prove Part 2, suppose first 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 $x \in A \iff \big(x \in P \text{ and } x \notin B\big) \iff \big(x \in P \text{ and } x \notin U \cap P\big) \iff \big(x \in P \text{ and } x \notin U\big) \iff \big(x \in P \text{ and } x \in K\big) \iff x \in K \cap P$.

Suppose, conversely, 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 $x \in B \iff \big(x \in P \text{ and } x \notin A\big) \iff \big(x \in P \text{ and } x \notin K \cap P\big) \iff \big(x \in P \text{ and } x \notin K\big) \iff \big(x \in P \text{ and } x \in U\big) \iff x \in U \cap P$. Since U is 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.

2.34 Remark: 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**.

Chapter 3. Limits and Continuity

3.1 Definition: 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 $\lim\limits_{n\to\infty} x_n=a$ (or we 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 sequence $(x_n)_{n\geq p}$ is **Cauchy** when for every $\epsilon>0$ there exists an index $m\geq p$ such that $d(x_k,x_\ell)<\epsilon$ for all indices $k,\ell\geq m$.

3.2 Remark: 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 that the **limit** of $(x_n)_{n\geq p}$ is equal to a) and we write $\lim\limits_{n\to\infty} x_n=a$ (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 $n\geq m$.

3.3 Theorem: *(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 $(y_n)_{n\geq q}$ converges and, in this case, $\lim\limits_{n\to\infty} y_n=\lim\limits_{n\to\infty} x_n$.
(3) If $(x_{n_k})_{k\geq q}$ is a subsequence of $(x_n)_{n\geq p}$, and $\lim\limits_{n\to\infty} x_n=a$, then $\lim\limits_{k\to\infty} x_{n_k}=a$.
(4) If $(x_n)_{n\geq p}$ converges then it is bounded.
(5) If $(x_n)_{n\geq p}$ converges then it is Cauchy.
(6) We have $\lim\limits_{n\to\infty} x_n=a$ in X if and only if $\lim\limits_{n\to\infty} d(x_n,a)=0$ in $\mathbf{R}$.
(7) We have $\lim\limits_{n\to\infty} x_n=a$ 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 an exercise.

3.4 Note: Because of Part 2 of the above theorem, the initial index p of a sequence $(x_n)_{n\geq p}$ does not effect whether or not the sequence converges and it does not effect 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.

3.5 Exercise: For each $n\in\mathbf{Z}^+$, let $x_n\in\mathbf{R}^\infty$ be the sequence given by $x_n=\frac{1}{n}\sum\limits_{k=1}^{n} e_k$, that is by $(x_{n,k})_{k\geq 1}=\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n},0,0,0,\cdots\right)$ with n non-zero terms. Show that (x_n) converges in $(\mathbf{R}^\infty,d_2)$ but diverges in $(\mathbf{R}^\infty,d_1)$.

3.6 Exercise: For each $n\in\mathbf{Z}^+$, 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 $\big(\mathcal{C}[0,1],d_1\big)$ but diverges in $\big(\mathcal{C}[0,1],d_2\big)$.

3.7 Note: Recall that $\mathcal{B}[a,b]$ denotes the space of bounded functions $f:[a,b]\to\mathbf{R}$. 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 $\big(\mathcal{B}[a,b],d_\infty\big)$, 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 $\sup\limits_{a\le x\le b}\big|f_n(x)-g(x)\big|<\epsilon$ if and only if $\big|f_n(x)-g(x)\big|<\epsilon$ 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 $\big(\mathcal{C}[a,b],d_\infty\big)$ if and only if (f_n) converges uniformly on $[a,b]$.

3.8 Theorem: *(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 $\lim\limits_{n\to\infty}x_n=a$.

(2) $a\in\overline{A}$ if and only if there exists a sequence (x_n) in A with $\lim\limits_{n\to\infty}x_n=a$.

(3) A is closed in X if and only if for every sequence (x_n) in A which converges in X, we have $\lim\limits_{n\to\infty}x_n\in A$.

Proof: We prove Parts 1 and 3 and leave the proof of Part 2 as an exercise. Suppose that $a\in A'$ (which means that for every $r>0$ we have $B^*(a,r)\cap A\neq\emptyset$). For each $n\in\mathbf{Z}^+$, choose $x_n\in B^*\big(a,\frac{1}{n}\big)\cap A$, that is choose $x_n\in A\setminus\{a\}$ with $d(x_n,a)<\frac{1}{n}$. Then $(x_n)_{n\ge1}$ is a sequence in $A\setminus\{a\}$ with $\lim\limits_{n\to\infty}x_n=a$.

Suppose, conversely, that $(x_n)_{n\ge1}$ is a sequence in $A\setminus\{a\}$ with $\lim\limits_{n\to\infty}x_n=a$. Let $r>0$. Choose $m\in\mathbf{Z}^+$ such that $d(x_n,a)<r$ for all $n\ge m$. 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 $B^*(a,r)\cap A\neq\emptyset$. This proves Part 1.

To prove Part 3, suppose that A is closed in X. Let $(x_n)_{n\ge1}$ be a sequence in A which converges in X, and let $a=\lim\limits_{n\to\infty}x_n\in X$. 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 $\lim\limits_{n\to\infty}x_n=a$, 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. This proves Part 3.

3.9 Example: Note that $\mathcal{C}[a,b]$ is closed in the metric space $\big(\mathcal{B}[a,b],d_\infty\big)$. 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 $\big(\mathcal{B}[a,b],d_\infty\big)$ 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 $g\in\mathcal{C}[a,b]$.

3.10 Exercise: Let

$$\begin{aligned}\mathcal{R}[a,b]&=\big\{f\in\mathcal{B}[a,b]\,\big|\,f\text{ is Riemann integrable}\big\},\\ \mathcal{P}[a,b]&=\big\{f\in\mathcal{B}[a,b]\,\big|\,f\text{ is a polynomial}\big\},\\ \mathcal{C}^1[a,b]&=\big\{f\in\mathcal{B}[a,b]\,\big|\,f\text{ is continuously differentiable}\big\}.\end{aligned}$$

Determine which of the above spaces are closed in the metric space $\mathcal{B}[a,b]$, using the supremum metric d_∞.

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

Solution: Since the closure of $\mathbf{R}^\infty$ in ℓ_1 is contained in ℓ_1 (by the definition of closure), it suffices to show that $\ell_1 \subseteq \overline{\mathbf{R}^\infty}$. Let $a = (a_n)_{n\geq 1} \in \ell_1$, so we have $\sum\limits_{n=1}^{\infty} |a_n| < \infty$. For each $n \in \mathbf{Z}^+$ 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 $x_n \in \mathbf{R}^\infty$ 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 $a \in \overline{\mathbf{R}^\infty}$ and so we have $\ell_1 \subseteq \overline{\mathbf{R}^\infty}$, as claimed.

3.12 Exercise: Find the closure of $\mathbf{R}^\infty$ in the metric space ℓ_2 using the metric d_2, and find the closure of $\mathbf{R}^\infty$ in the metric space ℓ_∞ using the metric d_∞.

3.13 Definition: 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 $d_Y(f(x), b) < \epsilon$.

3.14 Theorem: *(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 $0 < d(x, a) < \delta \Longrightarrow d\big(f(x), b\big) < \epsilon$. Since $x_n \to a$ we can choose $m \in \mathbf{Z}^+$ such that $n \geq m \Longrightarrow d(x_n, a) < \delta$. For $n \geq m$ we have $d(x_n, a) < \delta$ and we have $x_n \neq a$ $\big($since (x_n) is a sequence in $A \setminus \{a\}\big)$, so that $0 < d(x_n, a) < \delta$, and hence $d\big(f(x_n), b\big) < \epsilon$. 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 $d\big(f(x), b\big) \geq \epsilon$. For each $n \in \mathbf{Z}^+$, choose $x_n \in A$ such that $0 < d(x_n, a) < \frac{1}{n}$ and $d\big(f(x_n), b\big) \geq \epsilon$. 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 $n \in \mathbf{Z}^+$, it follows that $x_n \to a$. Since $d\big(f(x_n), b\big) \geq \epsilon$ for all $n \in \mathbf{Z}^+$, it follows that $\lim\limits_{n\to\infty} f(x) \neq b$. Thus we have found a sequence (x_n) in $A \setminus \{a\}$ with $x_n \to a$ such that $\lim\limits_{n\to\infty} f(x_n) \neq b$.

3.15 Definition: 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 $d_Y\big(f(x), f(a)\big) < \epsilon$. 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 $d_X(x, y) < \delta$ then $d_Y\big(f(x), f(y)\big) < \epsilon$. 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

$d\big(f(x), f(y)\big) \leq \ell \cdot d(x, y)$. 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).

3.16 Note: 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 $\lim_{x \to a} f(x) = f(a)$. If a is an isolated point of X then f is necessarily continuous at a, vacuously.

3.17 Theorem: *(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 $\lim_{n \to \infty} f(x_n) = f(a)$.*

Proof: The proof is left as an exercise.

3.18 Theorem: *(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.

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

3.20 Definition: 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 **homeomorphism**.

3.21 Note: 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 homeomorphism.

If U and V are finite-dimensional inner product spaces with say $\dim U = n$ and $\dim V = m$, and if $\phi : U \to \mathbf{R}^n$ and $\psi : V \to \mathbf{R}^m$ 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 $\psi F \phi^{-1} : \mathbf{R}^n \to \mathbf{R}^m$ is continuous. In particular, if F is linear then F is continuous (since $\psi F \phi^{-1} : \mathbf{R}^n \to \mathbf{R}^m$ 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.

3.22 Example: 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 $\big(\mathcal{C}[a, b], d_\infty\big)$ to the metric space $\big(\mathcal{C}[a, b], d_1\big)$, but not vice versa.

3.23 Theorem: *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) $F\big(\overline{B}(0,1)\big)$ is bounded.

In this case, if $m\geq 0$ with $F\big(\overline{B}(0,1)\big)\subseteq B(0,m)$ 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 $\|u-a\|\leq\delta_1\implies\big\|F(u)-F(a)\big\|\leq 1$. Choose $\delta=\delta_1\epsilon$. Let $x\in U$ with $\|x-0\|<\delta$. If $x=0$ then $\big\|F(x)-F(0)\big\|=\|0\|=0$. Suppose that $x\neq 0$. Then for $u=a+\frac{\delta_1 x}{\|x\|}$ we have $\|u-a\|=\big\|\frac{\delta_1 x}{\|x\|}\big\|=\delta_1$ and so $\big\|F(u-a)\big\|=\big\|F(u)-F(a)\big\|\leq 1$, that is $\big\|F\big(\frac{\delta_1 x}{\|x\|}\big)\big\|\leq 1$ hence, by the linearity of F and the scaling property of the norm, we have

$$\big\|F(x)-F(0)\big\|=\big\|F(x)\big\|=\tfrac{\|x\|}{\delta_1}\big\|F\big(\tfrac{\delta_1 x}{\|x\|}\big)\big\|\leq\tfrac{\|x\|}{\delta_1}<\tfrac{\delta_1\epsilon}{\delta_1}=\epsilon.$$

Thus F is continuous at 0, as required

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

$$\|F(x)\|=\Big\|\tfrac{\|x\|}{\delta}\,F\big(\tfrac{\delta x}{\|x\|}\big)\Big\|=\tfrac{\|x\|}{\delta}\Big\|F\big(\tfrac{\delta x}{\|x\|}\big)\Big\|\leq\tfrac{\|x\|}{\delta}=m\|x\|\leq m.$$

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

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

$$\big\|F(x)-F(y)\big\|=\big\|F(x-y)\big\|=\|x-y\|\big\|F\big(\tfrac{x-y}{\|x-y\|}\big)\big\|\leq m\|x-y\|.$$

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

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

Solution: Let $f\in\mathcal{C}[a,b]$ with $\|f\|_\infty\leq 1$, that is with $\max\limits_{a\leq x\leq b}|f(x)|\leq 1$. Then

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

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

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

Solution: For $n\in\mathbf{Z}^+$, define $f_n:[0,1]\to\mathbf{R}$ by $f_n(x)=x^n$. Then $f_n\in\mathcal{C}^1[a,b]$, and $\|f_n\|_\infty=\max\limits_{0\leq x\leq 1}|x^n|=1$ so that $f_n\in B(0,1)$, and $\big\|D(f_n)\big\|_\infty=\max\limits_{0\leq x\leq 1}\big|n\,x^{n-1}\big|=n$. Thus $D\big(\overline{B}(0,1)\big)$ is not bounded, so D is not continuous $\big($at any point $g\in\mathcal{C}[0,1]\big)$.

3.26 Example: Let X be a metric space and let $\emptyset \neq A \subseteq X$. Define $F : X \to \mathbf{R}$ by

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

Show that F is uniformly continuous.

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

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

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

$$\big| F(y) - F(x) \big| = \big| \text{dist}(x,A) - \text{dist}(y,A) \big| < \epsilon.$$

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

Proof: Let $M = \Big(\sum\limits_{k=1}^{n} \|u_k\|^2 \Big)^{1/2}$. For $t \in \mathbf{R}^n$ we have

$$\begin{aligned} \big\| F(t) \big\| &= \Big\| \sum_{k=1}^{n} t_k u_k \Big\| \leq \sum_{k=1}^{n} |t_k| \, \|u_k\| \quad \text{, by the Triangle Inequality,} \\ &\leq \Big(\sum_{k=1}^{n} {t_k}^2 \Big)^{1/2} \Big(\sum_{k=1}^{n} \|u_k\|^2 \Big)^{1/2} \quad \text{, by the Cauchy-Schwarz Inequality,} \\ &= M \|t\|. \end{aligned}$$

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

Note that the map $N : U \to \mathbf{R}$ 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 $G = N \circ F : \mathbf{R}^n \to \mathbf{R}$, which given by $G(t) = \big\| F(t) \big\|$. By the Extreme Value Theorem, the map G attains its minimum value on the unit sphere $\big\{ t \in \mathbf{R}^n \big| \|t\| = 1 \big\}$, which is compact. Let $m = \min\limits_{\|t\|=1} G(t) = \min\limits_{\|t\|=1} \big\| F(t) \big\|$. 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 $t \in \mathbf{R}^n$, if $\|t\| > 1$ then we have $\big\| \frac{t}{\|t\|} \big\| = 1$ so, by the choice of m,

$$\big\| F(t) \big\| = \|t\| \, \Big\| F\big(\tfrac{t}{\|t\|} \big) \Big\| \geq \|t\| \cdot m > m.$$

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

$$\big\| F^{-1}(x) \big\| = \tfrac{\|x\|}{m} \, \big\| F^{-1}\big(\tfrac{mx}{\|x\|} \big) \big\| \leq \tfrac{\|x\|}{m}.$$

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

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

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

Chapter 4. Separability and Completeness

4.1 Note: Let X be a metric space. Recall that for $A \subseteq X$ we say that A is **dense** in X when $\overline{A} = X$. Also recall that by Part 3 of Theorem 2.30 we have $\overline{A} = A \cup A'$ where A' is the set of limit points of A and so, by the definition of limit points, it follows that A is dense in X if and only if every open ball in X contains a point in A. By the sequential characterization of the closure (Part 2 of Therem 3.8) we can also say that A is dense in X if and only if for every $a \in X$ there exists a sequence (x_n) in A with $x_n \to a$ in X.

4.2 Definition: Let X be a metric space (or a topological space). We say that X is **separable** when it has a finite or countable dense subset.

4.3 Definition: Let X be a topological space. A **basis** (or a **base**) for the topology on X is a set $\mathcal{B}$ of open sets in X with the property that for every subset $A \subseteq X$, A is open if and only if for every point $a \in A$ there exists a basic set $U \in \mathcal{B}$ with $a \in U \subseteq A$.

4.4 Example: In a metric space X, the set of open balls $\mathcal{B} = \big\{B(a,r)\big|a \in X, 0 < r \in \mathbf{R}\big\}$ is a basis for the metric topology on X.

4.5 Theorem: *Let X be a metric space.*

(1) If X is separable then there is a finite or countable basis for the metric topology on X.
(2) If every infinite subset of X has a limit point then X is separable.
(3) If X is separable then every subspace of X is separable.

Proof: The proof is left as an exercise.

4.6 Example: Euclidean space $\big(\mathbf{R}^n, d_2\big)$ is separable with $\mathbf{Q}^n$ as a countable dense subset. Every subspace of Euclidean space is also separable.

4.7 Example: As an exercise, show that $\big(\ell_\infty, d_\infty\big)$ is not separable (consider characteristic functions χ_A for subsets $A \subseteq \mathbf{N}$.

4.8 Example: As an exercise, show that the set $\big(c, d_\infty\big)$ of convergent sequences of real (or complex) numbers is separable. Every subspace of c is also separable, for example the space c_0 of sequences which converge to 0.

4.9 Example: As an exercise, show that the space $\big(\mathcal{B}[a,b], d_\infty\big)$ of bounded functions on the interval $[a,b]$ is not separable $\big($consider characteristic functions χ_A for appropriate sets $A \subseteq [a,b]\big)$.

4.10 Example: Later (see Corollary 6.21 after the Weierstrass Approximation Theorem) we will show that the space $\big(\mathcal{C}[a,b], d_\infty\big)$ of continuous real valued functions on the interval $[a,b]$ is separable. Once we have proven this, it will follow that every subspace of $\mathcal{C}[a,b]$ is separable.

4.11 Definition: Recall that a sequence $(x_n)_{n\geq 1}$ in a metric space X is called a **Cauchy sequence** when it has the property that for all $\epsilon > 0$ there exists an index $m \in \mathbf{Z}^+$ such that for all indices $k, \ell \geq m$ we have $d(x_k, x_\ell) < \epsilon$.

4.12 Theorem: *Let X be a metric space.*

(1) Every Cauchy sequence in X is bounded.
(2) Every convergent sequence in X is Cauchy.
(3) If some subsequence of a Cauchy sequence (x_n) converges, then (x_n) converges.

Proof: To prove Part 1, let $(x_n)_{n\geq 1}$ be a Cauchy sequence in X. Choose $m \in \mathbf{Z}^+$ such that $k, \ell \geq m \Longrightarrow d(x_k, x_\ell) \leq 1$ and note that, in particular, we have $d(x_k, x_m) \leq 1$ for all $k \geq m$. Let $a = x_m$ and choose $r > \max\big\{d(x_1, a), d(x_2, a), \cdots, d(x_{m-1}, a), 1\big\}$. Then for all $n \in \mathbf{Z}^+$ we have $d(x_n, a) < r$ so the sequence (x_n) is bounded, as required.

We remark that Part 2 of this theorem was stated earlier, without proof, as Part 5 of Theorem 3.2. We give the proof here. Let $(x_n)_{n\geq 1}$ be a convergent sequence in X and let $a = \lim\limits_{n\to\infty} x_n$. Let $\epsilon > 0$. Choose $m \in \mathbf{Z}^+$ such that $n \geq m \Longrightarrow d(x_n, a) < \frac{\epsilon}{2}$. Then for all $k, \ell \geq m$ we have

$$d(x_k, x_\ell) \leq d(x_k, a) + d(a, x_\ell) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,$$

so the sequence (x_n) is Cauchy, as required.

To prove Part 3, let $(x_n)_{n\geq 1}$ be a Cauchy sequence in X, let $(x_{n_k})_{k\geq 1}$ be a subsequence of $(x_n)_{n\geq 1}$, suppose tha $(x_{n_k})_{k\geq 1}$ converges, and let $a = \lim\limits_{k\to\infty} x_{n_k}$. Let $\epsilon > 0$. Since (x_n) is Cauchy we can choose $m \in \mathbf{Z}^+$ so that $k, \ell \geq m \Longrightarrow d(x_k, x_\ell) < \frac{\epsilon}{2}$. Since $\lim\limits_{k\to\infty} n_k = \infty$ and $\lim\limits_{k\to\infty} x_{n_k} = a$, we can choose an index ℓ such that $n_\ell \geq m$ and $d(x_{n_\ell}, a) < \frac{\epsilon}{2}$. Then for all $k \geq m$ we have

$$d(x_k, a) \leq d(x_k, x_{n_\ell}) + d(x_{n_\ell}, a) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon.$$

4.13 Definition: A metric space X is called **complete** when every Cauchy sequence in X converges in X. A complete inner product space is called a **Hilbert space**, and a complete normed linear space is called a **Banach space**.

4.14 Theorem: *Let X be a complete metric space and let $A \subseteq X$. Then A is complete if and only if A is closed in X*

Proof: Suppose that A is closed in X. Let (x_n) be a Cauchy sequence in A. Since X is complete, (x_n) converges in X. Since A is closed in X and (x_n) is a sequence in A which converges in X, we have $\lim\limits_{n\to\infty} x_n \in A$ by Theorem 3.5 (The Sequential Characterization of Closed Sets). Thus every Cauchy sequence in A converges in A, so A is complete.

Suppose, conversely, that A is complete. Let $a \in A'$, that is let $a \in X$ be a limit point of A. Since $a \in A'$, by Theorem 3.5 (The Sequential Characterization of Limit Points) we can choose a sequence (x_n) in A (indeed in $A \setminus \{a\}$) with $\lim n \to \infty x_n = a$. Since (x_n) converges in X, it is Cauchy. Since (x_n) is Cauchy and A is complete, (x_n) converges in A, that is $a = \lim\limits_{n\to\infty} x_n \in A$.

4.15 Example: Recall, from MATH 247 or PMATH 333, that $\big(\mathbf{R}^n, d_2\big)$ is complete. It follows that every closed subset $A \subseteq \mathbf{R}^n$ is complete (using the standard metric d_2).

4.16 Example: Note that completeness is not invariant under homeomorphism. For example, $\mathbf{R}$ is homeomorphic to $(0,1) \subseteq \mathbf{R}$, but $\mathbf{R}$ is complete while $(0,1)$ is not.

4.17 Theorem: *Every finite-dimensional normed linear space is complete.*

Proof: Let U be an n-dimensional normed linear space. Let $\{u_1, \cdots, u_n\}$ be a basis for the vector space U and let $F : \mathbf{R}^n \to U$ be the associated vector space isomorphism given by $F(t) = \sum_{k=1}^{n} t_k u_k$. Recall, from Theorem 3.25, that both F and F^{-1} are Lipschitz continuous. Let L be a Lipshitz constant for F and let M be a Lipschitz constant for F^{-1}. Let $(x_n)_{n\geq 1}$ be a Cauchy sequence in U. For each $n \in \mathbf{Z}^+$, let $t_n = F^{-1}(x_n) \in \mathbf{R}^n$. Note that (t_n) is a Cauchy sequence in $\mathbf{R}^n$ because

$$\|t_k - t_\ell\| = \left\|F^{-1}(x_k) - F^{-1}(x_\ell)\right\| \leq M\|x_k - x_\ell\|.$$

Since (t_n) is a Cauchy sequence in $\mathbf{R}^n$ and $\mathbf{R}^n$ is complete, (t_n) converges in $\mathbf{R}^n$. Let $s = \lim_{n\to\infty} t_n \in \mathbf{R}^n$ and let $a = F(s) \in U$. Then we have $\lim_{n\to\infty} x_n = a$ because

$$\|x_n - a\| = \left\|F(t_n) - F(s)\right\| \leq \ L\|t_n - s\|.$$

4.18 Corollary: *The metric spaces* $(\mathbf{R}^n, d_1)$, $(\mathbf{R}^n, d_2)$ *and* $(\mathbf{R}^n, d_\infty)$ *are all complete.*

4.19 Theorem: *The metric spaces* (ℓ_1, d_1), (ℓ_2, d_2) *and* (ℓ_∞, d_∞) *are all complete.*

Proof: We prove that (ℓ_1, d_1) is complete and we leave the proof that (ℓ_2, d_2) and (ℓ_∞, d_∞) are complete as an exercise. Let $(a_n)_{n\geq 1}$ be a Cauchy sequence in ℓ_1. For each $n \in \mathbf{Z}^+$, write $a_n = (a_{n,k})_{k\geq 1} = (a_{n,1}, a_{n,2}, a_{n,3}, \cdots)$. Since $a_n \in \ell_1$ we have $\sum_{k=1}^{\infty} |a_{n,k}| < \infty$. Since $(a_n)_{n\geq 1}$ is Cauchy, for every $\epsilon > 0$ we can choose $N \in \mathbf{Z}^+$ such that for all $n, m \geq N$ we have $\|a_n - a_m\|_1 < \epsilon$, that is $\sum_{k=1}^{\infty} |a_{n,k} - a_{m,k}| < \epsilon$. For each fixed $k \in \mathbf{Z}^+$, note that for $n, m \geq N$ we have $|a_{n,k}, -a_{m,k}| \leq \sum_{j=1}^{\infty} |a_{n,j} - a_{m,j}| < \epsilon$, and so the sequence $(a_{n,k})_{n\geq 1}$ is Cauchy in $\mathbf{R}$, so it converges. For each $k \in \mathbf{Z}^+$, let $b_k = \lim_{n\to\infty} a_{n,k} \in \mathbf{R}$ and let $b = (b_k)_{k\geq 1}$.

We claim that $b \in \ell_1$. Since $(a_n)_{n\geq 1}$ is Cauchy, for every $\epsilon > 0$ we can choose $N \in \mathbf{Z}^+$ such that for all $n, m \geq N$ we have $\|a_n - a_m\|_1 < \epsilon$, that is $\sum_{k=1}^{\infty} |a_{n,k} - a_{m,k}| < \epsilon$. By the Triangle Inequality, for $n, m \geq N$ we have $\left|\|a_n\|_1 - \|a_m\|_1\right| \leq \|a_n - a_m\|_1 < \epsilon$ It follows that the sequence $\left(\|a_n\|\right)_{n\geq 1}$ is a Cauchy sequence in $\mathbf{R}$, so it converges. Let $M = \lim_{n\to\infty} \|a_n\|_1 \in \mathbf{R}$. For each fixed $K \in \mathbf{Z}^+$ we have

$$\sum_{k=1}^{K} |b_k| = \sum_{k=1}^{K} \left|\lim_{n\to\infty} a_{n,k}\right| = \lim_{n\to\infty} \sum_{k=1}^{K} |a_{n,k}| \leq \lim_{n\to\infty} \sum_{k=1}^{\infty} |a_{n,k}| = \lim_{n\to\infty} \|a_n\|_1 = M.$$

Since $\sum_{k=1}^{K} |b_k| \leq M$ for all $K \in \mathbf{Z}^+$ it follows that $\sum_{k=1}^{\infty} |b_k| \leq M$, so $b \in \ell_1$, as claimed.

Finally, we claim that $\lim_{n\to\infty} a_n = b$ in ℓ_1. Let $\epsilon > 0$. Choose $N \in \mathbf{Z}^+$ such that for all $n, m \geq N$ we have $\|a_n - a_m\|_1 < \epsilon$. Then for each $K \in \mathbf{Z}^+$ we have

$$\begin{aligned}\sum_{k=1}^{K} |a_{n,k} - b_k| &= \sum_{k=1}^{K} \left|a_{n,k} - \lim_{m\to\infty} a_{m,k}\right| = \lim_{m\to\infty} \sum_{k=1}^{K} |a_{n,k} - a_{m,k}| \\ &\leq \lim_{m\to\infty} \sum_{k=1}^{\infty} |a_{n,k} - a_{m,k}| = \lim_{m\to\infty} \|a_n - a_m\|_1 \leq \epsilon\end{aligned}$$

Since $\sum_{k=1}^{K} |a_{n,k} - b_k| \leq \epsilon$ for all $K \in \mathbf{Z}^+$ it follows that $\|a_n - b\|_1 = \sum_{k=1}^{\infty} |a_{n,k} - b_k| \leq \epsilon$.

4.20 Exercise: Show that (ℓ_1, d_∞) and (ℓ_2, d_∞) are not closed in (ℓ_∞, d_∞) and so they are not complete.

4.21 Exercise: Show that the metric spaces $\big(\mathcal{C}[a,b], d_1\big)$ and $\big(\mathcal{C}[a,b], d_2\big)$ are not complete. Hint: in the case $[a,b] = [-1,1]$, consider $f_n : [-1,1] \to \mathbf{R}$ given by $f_n(x) = x^{1/2n-1}$ for $n \in \mathbf{Z}^+$. Show that if (f_n) did converge, either in $\big(\mathcal{C}[-1,1], d_1\big)$ or in $\big(\mathcal{C}[-1,1], d_2\big)$, then it would necessarily converge to a function g with $g(x) = 1$ when $x > 0$ and $g(x) = -1$ when $x < 0$, but such a function g cannot be continuous.

4.22 Definition: Let $\mathbf{F} = \mathbf{R}$ or $\mathbf{C}$. For a metric space X, we define

$$\begin{aligned} \mathcal{F}(X,\mathbf{F}) &= \mathbf{F}^X = \big\{ f : X \to \mathbf{F} \big\} \\ \mathcal{B}(X,\mathbf{F}) &= \big\{ f : X \to \mathbf{F} \big| \ f \text{ is bounded} \big\} \\ \mathcal{C}(X,\mathbf{F}) &= \big\{ f : X \to \mathbf{F} \big| \ f \text{ is continuous} \big\}, \\ \mathcal{C}_b(X,\mathbf{F}) &= \big\{ f : X \to \mathbf{F} \big| \ f \text{ is bounded and continuous} \big\}. \end{aligned}$$

Since we usually take $\mathbf{F} = \mathbf{R}$ we write

$$\mathcal{F}(X) = \mathcal{F}(X,\mathbf{R}) \ , \ \mathcal{B}(X) = \mathcal{B}(X,\mathbf{R}) \ , \ \mathcal{C}(X) = \mathcal{C}(X,\mathbf{R}) \ \text{ and } \ \mathcal{C}_b(X) = \mathcal{C}_b(X,\mathbf{R}).$$

Note that $\mathcal{B}(X,\mathbf{F})$ is a normed linear space using the **supremum norm** given by

$$\|f\|_\infty = \sup_{x\in X} \big|f(x)\big|$$

and a metric space using the **supremum metric** given by $d_\infty(f,g) = \sup\limits_{x\in X} \big|f(x) - g(x)\big|$. These do not determine a well-defined norm and metric on $\mathcal{C}(X,\mathbf{F})$ since $\|f\|_\infty = \sup\limits_{x\in X} |f(x)|$ might not be finite, but they do determine a well-defined norm and metric on $\mathcal{C}_b(X,\mathbf{F})$.

4.23 Definition: For a sequence (f_n) in $\mathcal{F}(X)$ and for $g \in \mathcal{F}(X)$, we say that (f_n) **converges uniformly** to g on X, and write $f_n \to g$ uniformly on X, when for every $\epsilon > 0$ there exists $m \in \mathbf{Z}^+$ such that $|f_n(x) - g(x)| < \epsilon$ for every $n \geq m$ and every $x \in X$.

4.24 Note: For a sequence $(f_n) \in \mathcal{B}(X)$ and for $g \in \mathcal{B}(X)$, note that $|f_n(x) - g| < \epsilon$ for every $x \in X$ if and only if $\|f_n - g\|_\infty < \epsilon$. It follows that $f_n \to g$ uniformly on X if and only if $f_n \to g$ in the metric space $\big(\mathcal{B}(X), d_\infty\big)$.

4.25 Theorem: *Let X be a metric space. Then the metric spaces $\big(\mathcal{B}(X), d_\infty\big)$ and $\big(\mathcal{C}_b(X), d_\infty\big)$ are complete.*

Proof: Let $(f_n)_{n\geq 1}$ be a Cauchy sequence in $\big(\mathcal{B}(X), d_\infty\big)$. Note that for each $x \in X$, we have $\big|f_n(x) - f_m(x)\big| \leq \sup_{y\in X} \big|f_n(y) - f_m(y)\big| = \|f_n - f_m\|_\infty$, and so the sequence $\big(f_n(x)\big)_{n\geq 1}$ is a Cauchy sequence in $\mathbf{R}$, so it converges. Thus we can define a function $g : X \to \mathbf{R}$ by $g(x) = \lim\limits_{n\to\infty} f_n(x)$ and then we have $f_n \to g$ pointwise in X.

We claim that $g \in \mathcal{B}(X)$, that is we claim that g is bounded. Since (f_n) is a Cauchy sequence in $\mathcal{B}(X)$, it is bounded (by Part 1 of Theorem 4.12) so we can choose $M \geq 0$ such that $\|f_n\|_\infty \leq M$ for all indices n. Then for all $x \in X$ we have $|f_n(x)| \leq \|f_n\|_\infty \leq M$ and hence $|g(x)| = \lim\limits_{n\to\infty} |f_n(x)| \leq M$. Thus g is a bounded function, that is $g \in \mathcal{B}(X)$.

We know that $f_n \to g$ pointwise on X. We must show that $f_n \to g$ uniformly on X. Let $\epsilon > 0$. Since (f_n) is Cauchy we can choose $m \in \mathbf{Z}^+$ such that $\|f_k - f_\ell\|_\infty < \epsilon$ for all $k, \ell \geq m$. Then for all $k \geq m$ and for all $x \in X$ we have

$$\big|f_k(x) - g(x)\big| = \lim_{\ell\to\infty} \big|f_k(x) - f_\ell(x)\big| \leq \epsilon.$$

It follows that $f_n \to g$ uniformly on X, that is $f_n \to g$ in the metric space $\big(\mathcal{B}(X), d_\infty\big)$. Thus $\big(\mathcal{B}(X), d_\infty\big)$ is complete.

To show that $\big(\mathcal{C}_b(X), d_\infty\big)$ is complete, it suffices (by Theorem 4.14) to show that $\mathcal{C}_b(X)$ is closed in $\mathcal{B}(X)$. Let (f_n) be a sequence in $\mathcal{C}_b(X)$ which converges in $\big(\mathcal{B}(X), d_\infty\big)$. Let $g = \lim\limits_{n\to\infty} f_n \in \mathcal{B}(X)$. We need to show that g is continuous. Let $\epsilon > 0$ and let $a \in X$. Since $f_n \to g$ in $\big(\mathcal{B}(X), d_\infty\big)$ we know that $f_n \to g$ uniformly on X, so we can choose $m \in \mathbf{Z}^+$ such that $\big|f_m(x) - g(x)\big| < \frac{\epsilon}{3}$ for all $n \geq m$ and all $x \in X$. Since f_m is continuous at a we can choose $\delta > 0$ such that for all $x \in X$ with $d(x,a) < \delta$ we have $\big|f_m(x) - f_m(a)\big| < \frac{\epsilon}{3}$. Then for all $x \in X$ with $d(x,a) < \delta$ we have

$$\big|g(x) - g(a)\big| \leq \big|g(x) - f_m(x)\big| + \big|f_m(x) - f_m(a)\big| + \big|f_m(a) - g(a)\big| < \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3} = \epsilon.$$

Thus g is continuous at a. Since a was arbitrary, g is continuous on X, hence $g \in \mathcal{C}_b(X)$. By the Sequential Characterization of Closed Sets (Part 3 of Theorem 3.8) it follows that $\mathcal{C}_b(X)$ is closed in $\mathcal{B}(X)$, as required.

4.26 Corollary: *The metric space $\big(\mathcal{C}[a,b], d_\infty\big)$ is complete.*

Proof: Since every continuous function $f : [a,b] \to \mathbf{R}$ is bounded, we have $\mathcal{C}[a,b] = \mathcal{C}_b[a,b]$.

4.27 Example: In the metric space $\big(\mathcal{C}[a,b], d_\infty\big)$, the space $\mathcal{R}[a,b]$ of Riemann integrable functions is closed, hence complete, and the spaces $\mathcal{P}[a,b]$ of polynomial functions, and $\mathcal{C}^1[a,b]$ of continuously differentiable functions, are not closed, and hence not complete.

4.28 Theorem: *(Metric Completion) Every metric space X is isometric to a dense subspace of a complete metric space.*

Proof: Let X be a metric space. Fix $a \in X$. For each $x \in X$, define $f_x : X \to \mathbf{R}$ by $f_x(t) = d(t,x) - d(t,a)$. Note that f_x is bounded since, by the Triangle Inequality, $\left|f_x(t)\right| = \left|d(x,t) - d(a,t)\right| \leq d(a,x)$. Note that f_x is continuous (indeed f_x Lipschitz continuous) because for $s,t \in X$ we have

$$\begin{aligned}\left|f_x(s) - f_x(t)\right| &= \left|d(s,x) - d(s,a) - d(t,x) + d(t,a)\right| \\ &\leq \left|d(s,x) - d(t,x)\right| + \left|d(s,a) - d(t,a)\right| \\ &\leq d(s,t) + d(s,t) = 2\,d(s,t).\end{aligned}$$

Define $F : X \to \mathcal{C}_b(X)$ by $F(x) = f_x$. We claim that F preserves distance, using the d_∞ metric on $\mathcal{C}_b(X)$. For all $x,y,t \in X$ we have

$$\left|f_x(t) - f_y(t)\right| = \left|d(x,t) - d(a,t) - d(y,t) + d(a,t)\right| = \left|d(x,t) - d(y,t)\right| \leq d(x,y)$$

hence for all $x,y \in X$ we have

$$\|f_x - f_y\|_\infty = \sup_{t\in X}\left|f_x(t) - f_y(y)\right| \leq d(x,y).$$

On the other hand, for all $x,y \in X$ we also have

$$\|f_x - f_y\|_\infty = \sup_{t\in X}\left|f_x(t) - f_y(t)\right| \geq \left|f_x(y) - f_y(y)\right| = \left|d(x,y) - d(y,y)\right| = d(x,y),$$

and so F preserves distance, as claimed. Thus X is isometric to the image $F(X) \subseteq \mathcal{C}_b(X)$, which is dense in its closure $\overline{F(X)}$, which is complete because it is a closed subspace of the complete metric space $\mathcal{C}_b(X)$.

4.29 Remark: When X is a metric space and $F : X \to \mathcal{C}_b(X)$ is the distance preserving map in the proof of the above theorem, we often identify X with its isometric image $F(X)$ and think of X as a dense subspace of the complete metric space $Y = \overline{F(X)}$. Alternatively we can do some cutting and pasting operations on sets to obtain a complete metric space Y which actually contains X as a dense subspace. Here is an outline of one possible way of constructing such a set Y. Choose a set Z which is disjoint from X and has the same cardinality as $\mathcal{C}_b(X)$ (a bit of set theory is required to prove that such a set Z exists). Choose a bijection $G : \mathcal{C}_b(X) \to Z$ and give Z the metric which makes G an isometry. Then Z is complete and the composite $H = G \circ F : X \to Z$ is distance preserving so that X is isometric to the image $H(X)$, and $H(X)$ is dense in the complete space $\overline{H(X)}$, and $\overline{H(X)}$ is disjoint from X. Then let $Y = \left(\overline{H(X)} \setminus H(X)\right) \cup X$ so that we have $X \subseteq Y$. Let $K : Y \to \overline{H(X)}$ be the bijection given by $K(x) = h(x)$ if $x \in X$ and $K(y) = y$ if $h \notin X$, and give Y the metric for which K is an isometry. Then Y is complete and X is dense in Y.

4.30 Definition: When X and Y are metric spaces with $X \subseteq Y$ such that X is dense in Y and Y is complete, we say that Y is the **metric completion** of X. The metric completion of X is unique in the sense of the following theorem.

4.31 Theorem: *(Uniqueness of the Metric Completion) Let X, Y and Z be metric spaces with Y and Z complete such that $X \subseteq Y$ with $\overline{X} = Y$ and $X \subseteq Z$ with $\overline{X} = Z$. Then there is a (unique) isometry $F : Y \to Z$ with $F(x) = x$ for all $x \in X$.*

Proof: Let $a \in Y$. Since $\overline{X} = Y$ we can choose a sequence (x_n) in X with $x_n \to a$ in Y. Then (x_n) is Cauchy in Y, hence also in X, hence also in Z. Since (x_n) is Cauchy in Z, it converges in Z, say $x_n \to b$ in Z. In order for a map $F : Y \to Z$ to be continuous with $F(x) = x$ for every $x \in X$, we must have

$$F(a) = F\big(\lim_{n\to\infty} x_n\big) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} x_n = b.$$

This shows that if such a map F exists, it is unique, and it must be given by the following procedure: given $a \in Y$ we choose a sequence (x_n) in X with $x_n \to a$ and then we define $F(a) = \lim\limits_{n\to\infty} x_n \in Z$.

We claim that the above procedure does determine a well-defined map whose value $F(a)$ does not depend on the choice of the sequence (x_n). Let $a \in Y$ and let (x_n) and (y_n) be two sequences in X with $x_n \to a$ and $y_n \to a$ in Y. Let $b = \lim\limits_{n\to\infty} x_n$ in Z and let $c = \lim\limits_{n\to\infty}$ in Z. We need to show that $b = c$. Let $\epsilon > 0$. Choose $m \in \mathbf{Z}^+$ such that for all indices $n \geq m$ we have $d_Y(x_n, a) < \frac{\epsilon}{4}$, $d_Y(y_n, a) < \frac{\epsilon}{4}$, $d_Z(x_n, b) < \frac{\epsilon}{4}$. and $d_Z(y_n, c) < \frac{\epsilon}{4}$. Then since $d_Z(x_n, y_n) = d_X(x_n, y_n) = d_Y(x_n, y_n)$ we have

$$\begin{aligned} d_Z(b,c) &\leq d_Z(b,x_n) + d_Z(x_n,y_n) + d_Z(y_n,c) \\ &= d_Z(b,x_n) + d_Y(x_n,y_n) + d_Z(y_n,c) \\ &\leq d_Z(b,x_n) + d_Y(x_n,a) + d_Y(a,y_n) + d_Z(y_n,c) \\ &< \tfrac{\epsilon}{4} + \tfrac{\epsilon}{4} + \tfrac{\epsilon}{4} + \tfrac{\epsilon}{4} = \epsilon. \end{aligned}$$

Since $d_Z(b,c) < \epsilon$ for every $\epsilon > 0$ we must have $d_Z(b, c_{=}0$ hence $b = c$, as required.

Note that F is bijective with its inverse G given by the same construction: given $c \in Z$ we choose a sequence (x_n) in X with $x_n \to b$ in Z and define $G(c) = b = \lim\limits_{n\to\infty} x_n$ in Y.

It remains to prove that F preserves distance. Let $a, b \in Y$. Chooose sequences (x_n) and (y_n) in X with $x_n \to a$ and $y_n \to b$ in Y. Let $c, d \in Z$ with $x_n \to c$ and $y_n \to d$ in Z. We need to show that $d_Y(a,b) = d_Z(c,d)$. Since

$$\begin{aligned} d_Y(a,b) &\leq d_Y(a,x_n) + d_Y(x_n,y_n) + d_Y(y_n,b) \text{ , and} \\ d_Y(x_n,y_n) &\leq d_Y(x_n,a) + d_Y(a,b) + d_Y(b,y_n) \end{aligned}$$

it follows that

$$\big|d_Y(a,b) - d_Y(x_n,y_n)\big| \leq d_Y(a,x_n) + d_Y(y_n,b).$$

Taking the limit as $n \to \infty$ gives $\big|d_Y(a,b) - \lim\limits_{n\to\infty} d_Y(x_n,y_n)\big| = 0$ so that

$$d_Y(a,b) = \lim_{n\to\infty} d_Y(x_n,y_n) = \lim_{n\to\infty} d_X(x_n,y_n).$$

Similarly, we have $d_Z(c,d) = \lim\limits_{n\to\infty} d_X(x_n,y_n)$ and hence $d_Y(a,b) = d_Z(c,d)$, as required.

Chapter 5. Compactness

5.1 Definition: Let X be a metric space (or a topological space) and let $A \subseteq X$. An **open cover** for A (in X) is a set S of open sets in X such that $A \subseteq \bigcup S = \bigcup\limits_{U \in S} U$. When S is an open cover for A in X, a **subcover** of S for A is a subset $T \subseteq S$ such that $A \subseteq \bigcup T = \bigcup_{U \in T} U$. We say that A is **compact** (in X) when every open cover for A has a finite subcover.

5.2 Example: Recall (from MATH 247 or PMATH 333) that for $A \subseteq \mathbf{R}^n$, the Heine-Borel Theorem states that A is compact if and only if A is closed and bounded.

5.3 Example: When X is a metric space and $A \subseteq X$ is closed and bounded, it is *not* always the case that A is compact. For example, if X is any infinite set and d is the discrete metric on X, then every infinite subset $A \subseteq X$ is closed and bounded but not compact. In particular, closed unit balls are not compact, indeed for all $a \in X$ we have $\overline{B}(a, 1) = X$.

5.4 Theorem: *Let $A \subseteq X \subseteq Y$ where Y is a metric space (or a topological space). Then A is compact in X if and only if A is compact in Y.*

Proof: Suppose that A is compact in X. Let T be an open cover for A in Y. For each $V \in T$, let $U_V = V \cap X$. Note that each set U_V is open in X by Theorem 2.33 (or by Remark 2.34). Since $A \subseteq X$ and $A \subseteq \bigcup_{V \in T} V$, we also have $A \subseteq \bigcup_{V \in T}(V \cap X) = \bigcup_{V \in T} U_V$. Thus the set $S = \left\{U_V \middle| V \in T\right\}$ is an open cover for A in X. Since A is compact in X we can choose a finite subcover, say $\left\{U_{V_1}, \cdots U_{V_n}\right\}$ of S, where each $V_i \in T$. Since $A \subseteq \bigcup_{i=1}^n U_{V_i} = \bigcup_{i=1}^n (V_i \cap X)$, we also have $A \subseteq \bigcup_{i=1}^n V_i$ and so $\{V_1, \cdots, V_n\}$ is a finite subcover of T.

Suppose, conversely, that A is compact in Y. Let S be an open cover for A in X. For each $U \in S$, by Theorem 2.33 (or by Remark 2.34) we can choose an open set V_U in Y such that $U = V_U \cap X$. Then $T = \left\{V_U \middle| U \in S\right\}$ is an open cover of A in Y. Since A is compact in Y we con choose a finite subcover, say $\left\{V_{U_1}, \cdots, V_{U_n}\right\}$ of T, where each $U_i \in S$. Then we have $A \subseteq \bigcup_{i=1}^n (V_{U_i} \cap X) = \bigcup_{i=1}^n U_i$ and so $\{U_1, \cdots, U_n\}$ is a finite subcover of S.

5.5 Remark: Let $A \subseteq X$ where X is a metric space (or a topological space). By the above theorem, note that A is compact in X if and only if A is compact in itself. For this reason, we do not usually say that A is compact in X, we simply say that A is compact.

5.6 Theorem: *Let X be a metric space and let $A \subseteq X$. If A is compact then A is closed and bounded.*

Proof: Suppose that A is compact. We claim that A is closed. Let $a \in A^c$. For each $x \in A$, let $r_x = d(a, x) > 0$, let $U_x = B\left(a, \frac{r_x}{2}\right)$, and let $V_x = B\left(x, \frac{r_x}{2}\right)$ so that U_x and V_x are disjoint. Note that the set $S = \left\{V_x \middle| x \in A\right\}$ is an open cover for A. Since A is compact we can choose a finite subcover, say $\{V_{x_1}, \cdots, V_{x_n}\}$ where each $x_i \in A$. Let $r = \min\{r_{x_1}, \cdots, r_{x_n}\}$ so that $B\left(a, \frac{r}{2}\right) \subseteq U_{x_i}$ for all i, and hence $B\left(a, \frac{r}{2}\right)$ is disjoint from each set V_{x_i}. Since $B\left(a, \frac{r}{2}\right)$ is disjoint from each set V_{x_i} and the sets V_{x_i} cover A, it follows that $B\left(a, \frac{r}{2}\right)$ is disjoint from A, hence $B\left(a, \frac{r}{2}\right) \subseteq A^c$. Thus A^c is open, hence A is closed.

We claim that A is bounded. Let $a \in A$. For each $n \in \mathbf{Z}^+$, let $U_n = B(a, n)$. Then the set $S = \{U_1, U_2, U_3, \cdots\}$ is an open cover for A. Since A is compact, we can choose a finite subcover, say $\{U_{n_1}, U_{n_2}, \cdots, U_{n,\ell}\} \subseteq S$, with each $n_i \in \mathbf{Z}^+$. Let $m = \max\{n_1, n_2, \cdots, n_\ell\}$ so that $U_{n_i} \subseteq U_m$ for all indices i. Then we have $A \subseteq \bigcup_{i=1}^{\ell} U_{n_i} = U_m = B(a, m)$ and so A is bounded.

5.7 Theorem: *Let X be a metric space (or a topological space) and let $A \subseteq X$. If X is compact and A is closed in X, then A is compact.*

Proof: Suppose that X is compact and A is closed in X. Let S be an open cover for A. Then $S \cup \{A^c\}$ is an open cover for X. Since X is compact, we can choose a finite subcover T of $S \cup \{A^c\}$. Note that T may or may not contain the set A^c but, in either case, $T \setminus \{A^c\}$ is an open cover for A with $T \setminus \{A^c\} \subseteq S$, so that $T \setminus \{A^c\}$ is a finite subcover of S.

5.8 Corollary: *Let X be a metric space (or a topological space), let $A \subseteq X$ be closed, and let $K \subseteq X$ be compact. Then $A \cap K$ is compact.*

5.9 Theorem: *Let X and Y be metric spaces (or topological spaces) and let $f : X \to Y$. If X is compact and f is continuous then $f(X)$ is compact.*

Proof: Suppose that X is compact and f is continuous. Let T be an open cover for $f(X)$ in Y. Since f is continuous, so that $f^{-1}(V)$ is open in X for each $V \in T$, the set $S = \left\{f^{-1}(V)\middle|V \in T\right\}$ is an open cover for X. Since X is compact, we can choose a finite subcover, say $\left\{f^{-1}(V_1), f^{-1}(V_2), \cdots, f^{-1}(V_n)\right\}$ of S, with each $V_i \in T$. Then the set $\{V_1, V_2, \cdots, V_n\}$ is a finite subcover of T for $f(X)$.

5.10 Example: Note that continuous maps do not necessarily send closed sets to closed sets. For example, the map $f : \mathbf{R} \to \mathbf{R}$ given by $f(x) = \frac{2}{\pi} \tan^{-1}(x)$ sends the closed set $\mathbf{R}$ homeomorphically to the open interval $(-1, 1)$.

5.11 Theorem: *(The Extreme Value Theorem) Let X be a compact metric space (or topological space) and let $f : X \to \mathbf{R}$ be continuous. Then there exist $a, b \in X$ such that $f(a) \leq f(x) \leq f(b)$ for all $x \in X$.*

Proof: Since X is compact and f is continuous, it follows that $f(X)$ is compact in $\mathbf{R}$. Since $f(X)$ is compact, it is closed and bounded in $\mathbf{R}$. Since $f(X)$ is bounded in $\mathbf{R}$, it follows that $m = \inf f(X)$ and $M = \sup f(X)$ are both finite real numbers, and since $f(X)$ is closed in $\mathbf{R}$ it follows that $m \in f(X)$ and $M \in f(X)$ so that we can choose $a, b \in X$ such that $f(a) = m = \inf f(X)$ and $f(b) = M = \sup f(X)$.

5.12 Theorem: *Let X and Y be metric spaces with X compact. Let $f : X \to Y$ be continuous and bijective. Then f is a homeomorphism.*

Proof: Let $g = f^{-1} : Y \to X$. We need to prove that g is continuous. Let $A \subseteq X$ be closed in X. Since X is compact and $A \subseteq X$ is closed, it follows (from Theorem 5.7) that A is compact. Since the map $f : A \to Y$ is continuous and A is compact, it follows (from Theorem 5.9) that $f(A)$ is compact. Since $f(A)$ is compact it follows (from Theorem 5.6) that $f(A)$ is closed. Since $g = f^{-1}$ we have $g^{-1}(A) = f(A)$, which is closed. Since $g^{-1}(A)$ is closed in Y for every closed set A in X, it follows (by taking complements) that $g^{-1}(U)$ is open in Y for every open set U in X. Thus g is continuous, by the Topological Characterization of Continuity (Theorem 3.19).

5.13 Example: In the above theorem, the requirement that X is compact is necessary. For example, if X is the interval $X = [0, 2\pi)$ and Y is the unit circle $Y = \left\{z \in \mathbf{C}\middle|\|z\| = 1\right\}$, then the map $f : X \to Y$ given by $f(t) = e^{i\,t}$ is continuous and bijective, but the inverse map is not continuous at 1.

5.14 Theorem: *(The Lebesgue Number) Let X be a compact metric space and let S be an open cover for X. Then there exists a number $\lambda > 0$, which is called a* **Lebesgue number** *for the cover S, such that for all $a \in X$ there exists $U \in S$ such that $B(a, \lambda) \subseteq U$.*

Proof: For each $x \in X$, since S is an open cover for X we can choose $U_x \in S$ with $x \in U_x$ and then, since U_x is open we can choose $r_x > 0$ so that $B(a, 2r_x) \subseteq U_x$. Note that the set $T = \{B(x, r_x) \big| x \in X\}$ is an open cover for X. Since X is compact, we can choose a finite subcover, say $\{B(x_1, r_{x_1}), \cdots, B(x_n, r_{x_n})\}$ of T for X, with each $x_i \in X$. Let $\lambda = \min\{r_{x_1}, \cdots, r_{x_n}\}$. We claim that λ is a Lebesgue number for S. Let $a \in X$. Choose an index i such that $a \in B(x_i, r_{x_i})$, and let $U = U_{x_i} \in S$. For all $y \in B(a, \lambda)$ we have $d(y, x_i) \leq d(y, a) + d(a, x_i) \leq \lambda + r_{x_i} \leq 2r_{x_i}$ and hence $y \in B(x_i, 2r_{x_i}) \subseteq U_{x_i} = U$. This shows that $B(a, \lambda) \subseteq U$, as required.

5.15 Theorem: *Let X and Y be metric spaces with X compact and let $f : X \to Y$ be continuous. Then f is uniformly continuous.*

Proof: We leave the proof as an exercise.

5.16 Definition: Let X be a metric space. We say that X is **totally bounded** when for every $\epsilon > 0$ there exists a finite subset $\{a_1, a_2, \cdots, a_n\} \subseteq X$ such that $X = \bigcup_{i=1}^{n} B(a_i, \epsilon)$. We say that X has the **finite intersection property on closed sets** when for every set T of closed sets in X, if every finite subset of T has non-empty intersection, then T has non-empty intersection.

5.17 Theorem: *Let X be a metric space. Then the following are equivalent.*

(1) X is compact.
(2) X has the finite intersection property on closed sets.
(3) Every sequence (x_n) in X has a convergent subsequence.
(4) Every infinite subset $A \subseteq X$ has a limit point.
(5) X is complete and totally bounded.

Proof: First we prove that (1) implies (2). Suppose that X is compact. Let T be a set of closed sets in X. Suppose that T has empty intersection, that is suppose $\bigcap_{A \in T} A = \emptyset$. Then $\bigcup_{A \in T} A^c = X$ so the set $S = \{A^c \big| A \in T\}$ is an open cover for X. Since X is compact, we can choose a finite subcover, say $\{{A_1}^c, \cdots, {A_n}^c\}$ of S for X. Then we have $A_1 \cap A_2 \cap \cdots \cap A_n = \emptyset$, showing that some finite subset of T has empty intersection.

Next we prove that (2) implies (3). Suppose X has the finite intersection property on closed sets. Let $(x_n)_{n \geq 1}$ be a sequence in X. For each $m \in \mathbf{Z}^+$, let $A_m = \overline{\{x_n | n > m\}}$ and note that each A_m is closed with $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$. Let $T = \{A_m | m \in \mathbf{Z}^+\}$. Note that every finite subset of T has non-empty intersection because given $A_{m_1}, \cdots, A_{m_\ell} \in T$ we can let $m = \max\{m_1, \cdots, m_\ell\}$ and then we have $\bigcap_{i=1}^{\ell} A_{m_i} = A_m$ and we have $x_n \in A_m$. Since X has the finite intersection property on closed sets, it follows that T has non-empty intersection. Choose a point $a \in \bigcap_{m=1}^{\infty} A_m$. We construct a subsequence $(x_{n_k})_{k \geq 1}$ of $(x_n)_{n \geq 1}$ with $\lim_{k \to \infty} x_{n_k} = a$ as follows. Since $a \in A_1 = \overline{\{x_n | n > 1\}}$ we can choose $n_1 > 1$ such that $d(x_{n_1}, a) < 1$. Since $a \in A_{n_1} = \overline{\{x_n | n > n_1\}}$ we can choose $n_2 > n_1$ such that $d(x_{n_2}, a) < \frac{1}{2}$. Since $a \in A_{n_2} = \overline{\{x_n | n > n_2\}}$ we can choose $n_3 > n_2$ such that $d(x_{n_3}, a) < \frac{1}{3}$. Repeating this procedure, we can choose $1 < n_1 < n_2 < n_3 < \cdots$ such that $d(x_{n_k}, a) < \frac{1}{k}$ for all indices k, and then we have constructed a subsequence (x_{n_k}) such that $\lim_{k \to \infty} x_{n_k} = a$.

Next we prove that (3) implies (4). Suppose that every sequence (x_n) in X has a convergent subsequence. Let $A \subseteq X$ be an infinite subset. Choose a sequence (x_n) in A with the terms x_n all distinct. Choose a convergent subsequence (x_{n_k}) of (x_n) and let $a = \lim_{k\to\infty} x_{n_k}$. Then a is a limit point of the set A.

Now let us prove that (4) implies (5). Suppose that every infinite subset $A \subseteq X$ has a limit point. We claim that X is complete. Let (x_n) be a Cauchy sequence in X. We claim that (x_n) has a convergent subsequence. If the set $\{x_n | n \in \mathbf{Z}^+\}$ is finite, then some term in the sequence occurs infinitely often, so we can choose indices $n_1 < n_2 < n_3 < \cdots$ such that $x_1 = x_2 = x_3 = \cdots$, and so in this case (x_n) has a constant subsequence. Suppose the set $\{x_n | n \in \mathbf{Z}^+\}$ is infinite. Let a be a limit point of the infinite set $A = \{x_n | n \in \mathbf{Z}^+\}$. Since a is a limit point of the set $\{x_n\}$ we can choose indices n_k with $n_1 < n_2 < n_3 < \cdots$ such that $0 < d(x_{n_k}, a) < \frac{1}{k}$ for each index k. Then (x_{n_k}) is a subsequence of (x_n) with $\lim_{k\to\infty} x_{n_k} = a$. Since the sequence (x_n) Cauchy and has a convergent subsequence, it follows, from Part 3 of Theorem 5.11, that the sequence (x_n) converges. Thus X is complete, as claimed.

Continuing our proof that (4) implies (5), suppose that X is not totally bounded. Choose $\epsilon > 0$ such that there do not exist finitely many points $a_1, \cdots, a_n \in X$ for which $X = \bigcup_{i=1}^n B(a_i, \epsilon)$. Let $a_1 \in X$. Since $X \neq B(a_1, \epsilon)$ we can choose $a_2 \in X$ with $a_1 \notin B(a_1, \epsilon)$. Since $X \neq B(a_1, \epsilon) \cup B(a_2, \epsilon)$ we can choose $a_3 \in X$ with $a_3 \notin B(a_1, \epsilon) \cup B(a_2, \epsilon)$. Repeat this procedure to choose points $a_1, a_2, a_3, \cdots$ with $a_{n+1} \notin \bigcup_{k=1}^n B(a_k, \epsilon)$. Then the set $A = \{a_n | n \in \mathbf{Z}^+$ is an infinite subset of X which has no limit point.

Finally we prove that prove that (5) implies (1). Suppose that X is complete and totally bounded. Suppose, for a contradiction, that X is not compact, and choose an open cover S for X which has no finite subcover for X. Since X is totally bounded, we can cover X by finitely many balls of radius 1. Choose one of the balls, say $U_1 = B(a_1, 1)$ such that there is no finite subcover of S for U_1 (if there was a finite subcover for each ball, then the union of all these subcovers would be a finite subcover for X). Since X is totally bounded, we can cover X (hence also U_1) by finitely many balls of radius $\frac{1}{2}$. Choose one of these balls, say $U_2 = B(a_2, \frac{1}{2})$ such that there is no finite subcover of S for $U_1 \cap U_2$. Repeat the procedure to obtain balls $U_n = B\big(a_n, \frac{1}{n}\big)$ such that, for each n, there is no finite subcover of S for $\bigcap_{k=1}^n U_k$. In particular, each intersection $\bigcap_{k=1}^n U_k$ is nonempty so we can choose an element $x_n \in \bigcap_{k=1}^n U_k$. Since for all $k, \ell \geq m$ we have $x_k, x_\ell \in U_m = B\big(a_m, \frac{1}{m}\big)$ it follows that (x_n) is Cauchy. Since X is complete, it follows that (x_n) converges in X. Let $a = \lim_{n\to\infty} x_n$. Since S covers X we can choose $U \in S$ with $a \in U$. Since U is open we can choose $r > 0$ such that $B(a, r) \subseteq U$. Since $x_n \to a$ we can choose $m > \frac{3}{r}$ such that $d(x_m, a) < \frac{r}{3}$. Then for all $x \in U_m = B\big(a_m, \frac{1}{m}\big)$ we have $d(x, a) \leq d(x, a_m) + d(a_m, x_m) + d(x_m, a) < \frac{1}{m} + \frac{1}{m} + \frac{r}{3} < r$, and so $U_m \subseteq B(a, r) \subseteq U$. But then S has a finite subcover for U_m, namely the singleton $\{U\}$, which contradicts the fact that S has no finite subcover for $\bigcap_{k=1}^m U_k$.

5.18 Example: Show that in the metric space $\big(\mathcal{C}[0,1], d_\infty\big)$, the closed unit ball $\overline{B}(0,1)$ is not compact.

Solution: Let $f_n(x) = x^n$ for $n \in \mathbf{Z}^+$. Note that $\|f_n\|_\infty = 1$ so that each $f_n \in \overline{B}(0,1)$. Note that the pointwise limit of the sequence (f_n) is the function $g : [0,1] \to \mathbf{R}$ given by $g(x) = 0$ when $x < 1$ and $g(1) = 1$, which is not continuous. If some subsequence (f_{n_k}) of (f_n) were to converge in $\big(\mathcal{C}[0,1], d_\infty\big)$ then it would need to converge uniformly on $[0,1]$ to the function g. But this is not possible since the uniform limit of a sequence of continuous functions is always continuous. Thus (f_n) has no convergent subsequence and so $\overline{B}(0,1)$ is not compact.

Chapter 6. Some Applications

Contraction Maps and Picard's Theorem

6.1 Definition: Let X be a metric space. A map $f : X \to X$ is called a **contraction map** on X when there exists a constant $c \in [0,1)$ such thatt for all $x, y \in X$ we have

$$d\big(f(x), f(y)\big) \leq c\, d(x,y).$$

Such a constant c is called a **contraction constant** for f. Note that every contraction map is uniformly continuous.

6.2 Definition: For a map $f : X \to X$ (where X is any set), a point $a \in X$ such that $f(a) = a$ is called a **fixed point** of f.

6.3 Theorem: *(The Banach Fixed-Point Theorem) Every contraction map on a complete metric space has a unique fixed point.*

Proof: Let X be a complete metric space and let $f : X \to X$ be a contraction map on X with contraction constant $c \in [0,1)$. Let $x_0 \in X$ be any point. Let $x_1 = f(x_0)$ and $x_2 = f(x_1) = f^2(x_0)$ and so on, so that for $n \geq 1$ we have $x_n = f(x_{n-1}) = f^n(x_0)$. Note that the sequence $(x_n)_{n\geq 0}$ is Cauchy because for $n < m$ we have

$$\begin{aligned} d(x_n, x_m) &= d\big(f^n(x_0), f^n(x_{m-n})\big) \leq c^n\, d\big(x_0, x_{m-n}\big) \\ &\leq c^n \big(d(x_0,x_1) + d(x_1,x_2) + \cdots + d(x_{m-n-1}, x_{m-n})\big) \\ &\leq c^n d(x_0,x_1)\big(1 + c + c^2 + \cdots + c^{m-n-1}\big) \\ &\leq c^n d(x_0,x_1)\, \tfrac{1}{1-c} \;\; \longrightarrow 0 \text{ as } c \to 0^+. \end{aligned}$$

Since X is complete, the sequence $(x_n)_{n\geq 0}$ converges, so we can let $a = \lim\limits_{n\to\infty} x_n$. Note that $f(a) = a$ because f is continuous at a so that

$$f(a) = f\big(\lim_{n\to\infty} x_n\big) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n-1} = a.$$

Finally note that for $a, b \in X$, if $f(a) = a$ and $f(b) = b$ then since

$$d(a,b) = d\big(f(a), f(b)\big) \leq c\, d(a,b)$$

with $0 \leq c < 1$, it follows that $d(a,b) = 0$ so that $a = b$.

6.4 Definition: Let $A \subseteq \mathbf{R}^2$ and let $f : A \to \mathbf{R}$. We say that f satisfies a **Lipschitz condition** on A when there exists a constant $\ell \geq 0$ such that for all $x, y_1, y_2 \in \mathbf{R}$ for which $(x, y_1) \in A$ and $(x, y_2) \in A$, we have

$$\big|f(x, y_2) - f(x, y_1)\big| \leq \ell\,|y_2 - y_1|.$$

Such a constant ℓ is called a **Lipschitz constant** for f.

6.5 Theorem: *(Picard) Let U be an open set in $\mathbf{R}^2$, let $(a, b) \in U$, and let $F : U \to \mathbf{R}$ satisfy a Lipschitz condition on U. Then there exists $\delta > 0$ such that the differential equation $\frac{dy}{dx} = F(x, y)$ has a unique solution $y = f(x)$ with $f(a) = b$, defined for all $x \in [a-\delta, a+\delta]$.*

Proof: First note that $y = f(x)$ is a solution to the differential equation $\frac{dy}{dx} = F(x, y)$ with $f(a) = b$ if and only if $f(x)$ satisfies the integral equation

$$f(x) = b + \int_a^x F\big(t, f(t)\big)\,dt$$

for all $x \in [a-\delta, a+\delta]$. Let ℓ be a Lipschitz constant for F. Choose $r > 0$ such that $\overline{B}\big((a, b), r\big) \subseteq U$ and let $k = \max\limits_{(x,y)\in\overline{B}((a,b),r)} \big|F(x, y)\big|$. Choose δ with $0 < \delta < \frac{1}{\ell}$ small enough such that the rectangle

$$R = [a-\delta, a+\delta] \times [b-k\delta, b+k\delta]$$

is contained in $B\big((a, b), r\big)$. Verify as an exercise (Using the Mean Value Theorem) that if $f(x)$ is any solution to the given differential equation with $f(a) = b$ then the graph of f must be contained in the rectangle R. Let

$$X = \big\{f \in \mathcal{C}[a-\delta, a+\delta] \,\big|\, \text{Graph}(f) \subseteq R\big\}.$$

Verify that X is a closed subspace of the metric space $\mathcal{C}[a-\delta, a+\delta]$ (using the supremum metric) and so X is complete. Define $G : X \to \mathcal{C}[a-\delta, a+\delta]$ by

$$G(f)(x) = b + \int_a^x F\big(t, f(t)\big)\,dt.$$

Note that $G(X) \subseteq X$ because for all $f \in X$ and $x \in [a-\delta, a+\delta]$ we have

$$\big|G(f)(x) - b\big| = \left|\int_a^x F(t, f(t))dt\right| \leq \left|\int_a^x k\,dt\right| = k|x-a| \leq k\delta.$$

Note that G is a contraction map on X, with contraction constant $c = \ell\delta < 1$ because, for all $f, g \in X$ and all $x \in [a-\delta, a+\delta]$, we have

$$\begin{aligned}
\big|G(f)(x) - G(g)(x)\big| &= \left|\int_a^x \big(F(t, f(t)) - F(t, g(t))\big)dt\right| \leq \left|\int_a^x \big|F(t, f(t)) - F(t, g(t))\big|dt\right| \\
&\leq \left|\int_a^x \ell\big|f(t) - g(t)\big|\,dt\right| \leq \left|\int_a^x \ell\,\|f - g\|_\infty\,dt\right| \\
&= \ell|x-a|\,\|f-g\|_\infty \leq \ell\delta\,\|f-g\|_\infty.
\end{aligned}$$

By the Banach Fixed-Point Theorem, the map G has a unique fixed point $f \in X$, and this function $f \in X$ is the unique solution to the above integral equation, which is equivalent to the given differential equation.

6.6 Definition: Let X be a set and let $S \subseteq \mathcal{F}(X) = \mathcal{F}(X, \mathbf{R})$. We say that S is **pointwise bounded** when for every $x \in X$ there exists $m = m(x) > 0$ such that $|f(x)| \leq m$ for every function $f \in S$. We say that S is **uniformly bounded** when there exists $m > 0$ such that $|f(x)| \leq m$ for every $x \in X$ and every $f \in S$.

Let X be a metric space and let $S \subseteq \mathcal{C}(X) = \mathcal{C}(X, \mathbf{R})$. We say that S is **equicontinuous** when for every $\epsilon > 0$ there exists $\delta > 0$ such that for every $f \in S$ and for all $x, y \in X$, if $d(x, y) < \delta$ then $d\big(f(x), f(y)\big) < \epsilon$

6.7 Note: When X is a compact metric space, by the Extreme Value Theorem, every continuous function $f : X \to \mathbf{R}$ is also bounded, so we have $\mathcal{C}(X) = \mathcal{C}_b(X)$, which is a complete metric space using the supremum norm. Unless otherwise stated, when we refer to the metric space $\mathcal{C}(X)$ it is understood that we are using the supremum metric.

6.8 Note: When X is a compact metric space and $S \subseteq \mathcal{C}(X)$, note that S is uniformly bounded if and only if S is bounded as a subspace of the metric space $\mathcal{C}(X)$.

6.9 Theorem: *Let X be a compact metric space and let (f_n) be a sequence in $\mathcal{C}(X)$. If the sequence (f_n) converges in the metric space $\mathcal{C}(X)$ (equivalently, if the sequence (f_n) converges uniformly on X) then the set $\{f_n\}$ is equicontinuous.*

Proof: Suppose (f_n) converges in $\mathcal{C}(X)$. Let $\epsilon > 0$. Since (f_n) converges in $\mathcal{C}(X)$ we can choose $\ell \in \mathbf{Z}^+$ such that for all $n, m \geq \ell$ we have $\|f_n - f_m\|_\infty < \frac{\epsilon}{3}$. Since X is compact, each of the functions f_n is uniformly continuous on X. Choose $\delta > 0$ such that for all $x, y \in X$ with $d(x, y) < \delta$ we have $\big|f_n(x) - f_n(y)\big| < \epsilon$ for each $n < \ell$ and we have $\big|f_\ell(x) - f_\ell(y)\big| < \frac{\epsilon}{3}$. Then for all $n \geq \ell$ and all $x, y \in X$ with $d(x, y) < \delta$ we have

$$\big|f_n(x) - f_n(y)\big| \leq \big|f_n(x) - f_\ell(x)\big| + \big|f_\ell(x) - f_\ell(y)\big| + \big|f_\ell(y) - f_n(y)\big| < \epsilon.$$

6.10 Corollary: *Let X be a compact metric space. Then every compact set $S \subseteq \mathcal{C}(X)$ is equicontinuous.*

Proof: Let $S \subseteq \mathcal{C}(X)$. Suppose that S is not equicontinuous. Choose $\epsilon > 0$ such that for all $\delta > 0$ there exists $f \in S$ and there exist $x, y \in X$ with $d(x, y) < \delta$ such that $\big|f(x) - f(y)\big| \geq \epsilon$. For each $n \in \mathbf{Z}^+$, choose $f_n \in S$ such that there exist $x, y \in X$ with $d(x, y) < \frac{1}{2^n}$ such that $\big|f_n(x) - f_n(y)\big| \geq \epsilon$. Then no subsequence of (f_n) can possibly converge in S (using the supremum metric) and so S cannot be compact.

6.11 Theorem: *Let X be a compact metric space and let (f_n) be a sequence in $\mathcal{C}(X)$. If the set $\{f_n\}$ is pointwise bounded and equicontinuous then the set $\{f_n\}$ is uniformly bounded and the sequence (f_n) has a convergent subsequence in $\mathcal{C}(X)$.*

Proof: Suppose that the set $\{f_n\}$ is pointwise bounded and equicontinuous. We claim that the set $\{f_n\}$ is uniformly bounded. Since $\{f_n\}$ is equicontinuous, we can choose $\delta > 0$ such that for all $n \in \mathbf{Z}^+$ and for all $x, y \in X$ with $d(x, y) < \delta$ we have $\big|f_n(x) - f_n(y)\big| < 1$. Since X is compact, we can choose $a_1, a_2, \cdots, a_\ell \in X$ such that $X = B(a_1, \delta) \cup \cdots \cup B(a_\ell, \delta)$. Since $\{f_n\}$ is pointwise bounded, we can choose $m > 0$ such that for each index k with $1 \leq k \leq \ell$ we have $\big|f_n(a_k)\big| \leq m$. Let $n \in \mathbf{Z}^+$ and $x \in X$. Choose an index k with $1 \leq k \leq \ell$ such that $x \in B(a_k, \delta)$. Since $d(x, a_k) < \delta$ we have $\big|f_n(x) - f_n(a_k)\big| < 1$ and so $\big|f_n(x)\big| \leq \big|f_n(x) - f_n(a_k)\big| + \big|f_n(a_k)\big| < 1 + m$. Thus the set $\{f_n\}$ is uniformly bounded, as claimed.

It remains to show that the sequence (f_n) has a convergent subsequence in $\mathcal{C}(X)$. Since X is compact, and hence separable, we can choose a countable dense subset $A \subseteq X$, say $A = \{a_1, a_2, a_3, \cdots\}$. We claim that the sequence $(f_n)_{n\geq 1}$ has a subsequence $(f_{n_k})_{k\geq 1}$ which converges pointwise on A. Since the real-valued sequence $\big(f_n(a_1)\big)_{n\geq 1}$ is bounded, we can choose a subsequence, which we shall write as $\big(f_{1,k}\big)_{k\geq 1} = \big(f_{1,1}, f_{1,2}, f_{1,3}, \cdots\big)$, of the sequence of functions $(f_n)_{n\geq 1}$ such that the real-valued sequence $\big(f_{1,k}(a_1)\big)_{k\geq 1}$ converges. Since the real-valued sequence $\big(f_{1,k}(a_2)\big)_{k\geq 1}$ is bounded, we can choose a subsequence $\big(f_{2,k}\big)$ of the sequence of functions $\big(f_{1,k}\big)$ such that the real-valued sequence $\big(f_{2,k}(a_2)\big)$ converges. Note that since $\big(f_{2,k}(a_1)\big)$ is a subsequence of the convergent sequence $\big(f_{1,k}(a_1)\big)$, it also converges. By recursively repeating this procedure, we construct sequences $(f_{n,k})_{k\geq 1}$ for each $n \geq 1$, such that $\big(f_{n+1,k}\big)_{k\geq 1}$ is a subsequence of $\big(f_{n,k}\big)_{k\geq 1}$ and the real-valued sequences $\big(f_{n,k}(a_j)\big)_{k\geq 1}$ converge for all j with $1 \leq j \leq n$. Let $\big(f_{n_k}\big)_{k\geq 1}$ denote the sequence $\big(f_{1,1}, f_{2,2,}, f_{3,3,}, \cdots\big)$, note that this is a subsequence of the original sequence (f_n), and the real-valued sequences $\big(f_{n_k}(a_j)\big)_{k\geq 1}$ converge for all indices $j \in \mathbf{Z}^+$, so the subsequence (f_{n_k}) converges pointwise on A, as required.

Finally, we claim that the above subsequence (f_{n_k}) converges in $\mathcal{C}(X)$. Let $\epsilon > 0$. Since the set $\{f_n\}$ is equicontinuous we can choose $\delta > 0$ such that for all $n \in \mathbf{Z}^+$ and all $x, y \in X$ with $d(x,y) < \delta$ we have $\big|f_n(x) - f_n(y)\big| < \frac{\epsilon}{3}$. Since A is dense in X, the set $\mathcal{U} = \big\{B(a_n, \delta)\big|n \in]Z^+\big\}$ is an open cover of X. Since X is compact, we can choose a finite subcover of $\mathcal{U}$, so we can choose $a_1, a_2, \cdots, a_p \in X$ such that $X = B(a_1, \delta) \cup \cdots \cup B(a_p, \delta)$. Since the sequences $\big(f_{n_k}(a_j)\big)_{k\geq 1}$ all converge, we can choose $m \in \mathbf{Z}^+$ such that for all $j \in \mathbf{Z}^+$ with $1 \leq j \leq p$ and all $k, \ell \in \mathbf{Z}^+$ with $k, \ell \geq m$ we have $\big|f_{n_k}(a_j) - f_{n_\ell}(a_j)\big| < \frac{\epsilon}{3}$. Let $x \in X$ and let $k, \ell \in \mathbf{Z}^+$ with $k, \ell \geq m$. Choose an index j with $1 \leq j \leq p$ such that $x \in B(a_j, \delta)$. Then we have

$$\big|f_{n_k}(x) - f_{n_\ell}(x)\big| \leq \big|f_{n_k}(x) - f_{n_k}(a_j)\big| + \big|f_{n_k}(a_j) - f_{n_\ell}(a_j)\big| + \big|f_{n_\ell}(a_j) - f_{n_\ell}(x)\big| < \epsilon.$$

6.12 Theorem: *(The Arzela-Ascoli Theorem) Let X be a compact metric space and let $S \subseteq \mathcal{C}(X)$. Then S is compact if and only if S is closed, pointwise bounded, and equicontinuous.*

Proof: Suppose that S is compact. Then we know that S is closed and bounded and we know (from Corollary 6.9) that S is equicontinuous. Since S is bounded, using the supremum metric, it follows that S is uniformly bounded, hence also pointwise bounded.

Suppose, conversely, that S is closed, pointwise bounded, and equicontinuous. Let (f_n) be a sequence in S. Since S is pointwise bounded and equicontinuous, the subset $\{f_n\}$ is also pointwise bounded and equicontinuous. By the above theorem, the sequence (f_n) has a convergent subsequence (f_{n_k}) in $\mathcal{C}(X)$. Since S is closed, the limit of this subsequence lies in S. This proves that every sequence in S has a subsequence which converges in S, and so S is compact.

6.13 Theorem: *(Peano) Let $U \in \mathbf{R}^2$ be open, let $(a,b) \in U$, and let $F : U \to \mathbf{R}$ be continuous. Then there exists $\delta > 0$ such that the differential equation $\frac{dy}{dx} = F(x,y)$ has a solution $y = f(x)$ which is defined for all $x \in [a-\delta, a+\delta]$.*

Proof: I may include a proof later.

6.14 Definition: A (commutative) **algebra** over a field F is a vector space U with a binary multiplication operation such that for all $u, v, w \in U$ and all $t \in F$ we have $uv = vu$, $u(v+w) = uv + uw$, and $(tu)v = t(uv)$. A subspace $A \subseteq U$ is a **subalgebra** of U when it is an algebra using (the restriction of) the same operations used in U. Verify that a subset $A \subseteq U$ is a subalgebra of U when $0 \in A$ and for all $u, v \in A$ and all $t \in F$ we have $tu \in A$, $u + v \in A$ and $uv \in A$.

6.15 Example: When X is a metric space, $\mathcal{F}(X)$ is an algebra over $\mathbf{R}$ and $\mathcal{B}(X)$, $\mathcal{C}(X)$, and $\mathcal{C}_b(X)$ are all subalgebras.

6.16 Example: When $a \leq b$, the space $\mathcal{P}[a,b]$ of polynomial maps $f : [a,b] \to \mathbf{R}$ and the space $\mathcal{C}^1[a,b]$ of continuously differentiable maps are subalgebras of the algebra $\mathcal{C}[a,b]$ of continuous maps $f : [a,b] \to \mathbf{R}$, and the space $\mathcal{R}[a,b]$ of Riemann integrable functions is a subalgebra of the algebra $\mathcal{B}[a,b]$ of bounded functions $f : [a,b] \to \mathbf{R}$.

6.17 Example: Show that $f(x) = |x|$ lies in the closure of $\mathcal{P}[-1,1]$ in $\mathcal{C}[-1,1]$ (using the supremum metric).

Solution: Let $a \in \mathbf{R}$ with $0 < a \leq 1$ and let $g(x) = \sqrt{x + a^2}$. Then $g'(x) = \frac{1}{2}(x + a^2)^{-1/2}$, $g''(x) = -\frac{1}{2^2}(x + a^2)^{-3/2}$, $g'''(x) = \frac{1\cdot 3}{2^3}(x + a^2)^{-5/2}$ and in general

$$g^{(n)}(x) = \tfrac{(-1)^{n+1}1\cdot 3\cdot 5\cdots(2n-3)}{2^n}(x + a^2)^{-(2n-1)/2}.$$

Let $p_n(x)$ be the n^{th} Taylor polynomial for $g(x)$ centred at 1 $\big($we centre at 1 so the Taylor series converges in an open interval containing $[0,1]\big)$. For all $x \in [0,1]$, since $\binom{2n}{n} \leq 2^{2n}$ we have

$$\begin{aligned} \big|g(x) - p_n(x)\big| &= \left|\tfrac{g^{(n+1)}(t)}{(n+1)!}(x-1)^{n+1}\right| \text{ for some } t \in [x,1] \\ &\leq \tfrac{1\cdot 3\cdots(2n-1)\cdot a^{2n-1}}{2^{n+1}(n+1)!} = \tfrac{a^{2n-1}}{2^{2n+1}(n+1)}\binom{2n}{n} \leq \tfrac{1}{2(n+1)}. \end{aligned}$$

For all $x \in [-1,1]$ we have $x^2 \in [0,1]$ so

$$\Big|\sqrt{x^2 + a^2} - p_n(x^2)\Big| = \big|g(x^2) - p_n(x^2)\big| \leq \tfrac{1}{2(n+1)}.$$

Also note that for all x we have

$$\Big||x| - \sqrt{x^2 + a^2}\Big| = \sqrt{x^2 + a^2} - \sqrt{x^2} = \frac{a^2}{\sqrt{x^2 + a^2} + \sqrt{x^2}} \leq a.$$

Given $\epsilon > 0$ we can choose $a > 0$ with $a < 1$ and $a < \frac{\epsilon}{2}$ and we can choose $n \in \mathbf{Z}^+$ so that $\frac{1}{2(n+1)} < \frac{\epsilon}{2}$ and then for all $x \in [-1,1]$ we have

$$\big||x| - p_n(x^2)\big| \leq \Big||x| - \sqrt{x^2 + a^2}\Big| + \Big|\sqrt{x^2 + a^2} - p_n(x^2)\Big| \leq a + \tfrac{1}{2(n+1)} < \epsilon.$$

Thus for $f_n : [-1,1] \to \mathbf{R}$ given by $f_n(x) = p_n(x^2)$, we have each $f_n \in \mathcal{P}[-1,1]$ and $f_n \to f$ uniformly on $[-1,1]$, that is $f_n \to f$ in the metric space $\big(\mathcal{C}[-1,1], d_\infty\big)$.

6.18 Definition: Let $A \subseteq \mathcal{C}(X)$. We say that A **separates points** when for all $x, y \in X$ with $x \neq y$ there exist $f \in A$ with $f(x) \neq f(y)$. We say that A **vanishes nowhere** when for all $a \in X$ there exists $f \in A$ such that $f(x) \neq 0$. Note that if $1 \in A$ (where 1 denotes the constant function) the A vanishes nowhere.

6.19 Theorem: *(The Stone-Weierstrass Theorem) Let X be a compact metric space and let $A \subseteq \mathcal{C}(X)$ be an algebra. If A separates points and vanishes nowhere then $\overline{A} = \mathcal{C}(X)$.*

Proof: Note first that $\overline{A}$ is also a subalgebra of $\mathcal{C}(X)$. Indeed given $f, g \in \overline{A}$ and $c \in \mathbf{R}$, we can choose sequences (f_n) and (g_n) in A such that $f_n \to f$ and $g_n \to g$ in $\mathcal{C}(X)$ (that is $f_n \to f$ and $g_n \to g$ uniformly on X), and then we have $cf_n \to cf$, $f_n + g_n \to f + g$ and $f_n g_n \to fg$ uniformly on X, and hence $cf \in \overline{A}$, $f + g \in \overline{A}$ and $fg \in \overline{A}$. Also note that $\overline{A}$ separates points and vanishes nowhere, and so we may assume, without loss of generality, that A is closed.

Next we claim that if $f \in A$ then we also have $|f| \in A$. Let $f \in A \subseteq \mathcal{C}(X)$. Choose $m > 0$ with $m \geq \|f\|_\infty$. Let $g = \frac{1}{m} f$ and note that $g \in A$ with $\|g\|_\infty \leq 1$, that is $g(x) \in [-1, 1]$ for all $x \in X$. Let $\epsilon > 0$. By Example 6.17, we can choose a polynomial $p_0(x) = a_0 + a_1 x + \cdots + a_n x^n$ such that $\big|p_0(u) - |u|\big| \leq \frac{\epsilon}{2}$ for all $u \in [-1, 1]$. Let $p(x) = p_0(x) - a_0$ and note that $\big|p(u) - |u|\big| \leq \epsilon$ for all $u \in [-1, 1]$. For all $x \in X$, we have $g(x) \in [-1, 1]$ and so $\big|p\big(g(x)\big) - |g(x)|\big| < \epsilon$. Note that the function $h(x) = p\big(g(x)\big) = a_1 g(x) + a_2 g(x)^2 + \cdots a_n g(x)^n$ lies in A (because $g \in A$ and A is an algebra). This shows that for every $\epsilon > 0$ we can find $h \in A$ with $\big|h - |g|\big| < \epsilon$, and (since A is closed) it follows that $|g| \in A$ and hence $|f| = m|g| \in A$.

Next we note that if $f, g \in A$ then we also have $\max\{f, g\} \in A$ and $\min\{f, g\} \in A$ because

$$\max\{f, g\} = \frac{f + g}{2} + \frac{|f + g|}{2} \quad \text{and} \quad \min\{f, g\} = \frac{f + g}{2} - \frac{|f + g|}{2}$$

and it follows, inductively, that if $f_1, f_2, \cdots, f_n \in A$ then we have $\max\{f_1, \cdots, f_n\} \in A$ and $\min\{f_1, \cdots, f_n\} \in A$.

We claim that for all $r, s \in \mathbf{R}$ and for all $a, b \in X$ with $a \neq b$, there is a function $g \in A$ with $g(a) = r$ and $g(b) = s$. Let $r, s \in \mathbf{R}$ and let $a, b \in X$ with $a \neq b$. Since A separates points, we can choose $h \in A$ with $h(a) \neq h(b)$. Since A vanishes nowhere, we can choose $k, \ell \in A$ with $k(a) \neq 0$ and $\ell(b) \neq 0$. Define $u, v \in A$ by

$$u(x) = \big(h(x) - h(b)\big)k(x) \quad \text{and} \quad v(x) = \big(h(a) - h(x)\big)\ell(x)$$

and note that $u(a) \neq 0$ and $u(b) = 0$ while $v(a) = 0$ and $v(b) \neq 0$. Then define $g \in A$ by

$$g(x) = r\,\frac{u(x)}{u(a)} + s\,\frac{v(x)}{v(b)}$$

to obtain $g(a) = r$ and $g(b) = s$, as required.

We claim that for every $f \in \mathcal{C}(X)$, for every $a \in X$ and for every $\epsilon > 0$, there is a function $h \in A$ such that $h(a) = f(a)$ and $h(x) < f(x) + \epsilon$ for all $x \in X$. Let $f \in \mathcal{C}(X)$, let $a \in X$ and let $\epsilon > 0$. For each $b \in X$, by the previous claim we can choose $g_b \in A$ such that $g_b(a) = f(a)$ and $g_b(b) = f(b)$. For each $b \in X$, since f and g_b are continuous at b, we can choose $r_b > 0$ such that for all $x \in B(b, r_b)$ we have

$$\big|f(x) - f(b)\big| < \tfrac{\epsilon}{2} \ \text{ and } \ \big|g_b(x) - g_b(b)\big| < \tfrac{\epsilon}{2}\ \text{, hence } \big|g_b(x) - f(x)\big| < \epsilon.$$

Since X is compact and the set $\big\{B(b, r_b)\,\big|\,b \in X\big\}$ covers X, we can choose $b_1, b_2, \cdots, b_n \in X$ such that $X = \bigcup_{k=1}^{n} B(b_k, r_{b_k})$, and then we let

$$h = \min\big\{g_{b_1}, g_{b_2}, \cdots, g_{b_n}\big\} \in A.$$

For all $x \in X$ we can choose an index k such that $x \in B(b_k, r_{a_k})$ and then we have $h(x) \leq g_{b_k}(x) < f(x) + \epsilon$, as required.

Finally, we complete the proof by showing that for every $f \in \mathcal{C}[0,1]$ and every $\epsilon > 0$ there exists $g \in A$ such that $|g(x) - f(x)| < \epsilon$ for all $x \in X$. Let $f \in \mathcal{C}(X)$ and let $\epsilon > 0$. For each $a \in X$, by the previous claim we can choose $h_a \in A$ such that $h_a(a) = f(a)$ and $h_a(x) < f(x) + \epsilon$ for all $x \in X$. For each $a \in X$, since f and h_a iare continuous at a, we can choose $s_a > 0$ such that for all $x \in B(a, s_a)$ we have

$$\left|f(x) - f(a)\right| < \tfrac{\epsilon}{2} \;\text{ and }\; \left|h_a(x) - h_a(a)\right| < \tfrac{\epsilon}{2} \;\text{ hence }\; \left|h_a(x) - f(x)\right| < \epsilon.$$

Since X is compact and $\left\{B(a_k, s_k) \,\middle|\, a \in X\right\}$ covers X, we can choose $a_1, a_2, \cdots, a_m \in X$ such that $X = \bigcup\limits_{k=1}^{m} B(a_k, s_{a_k})$, and then we chose

$$g = \max\left\{h_{a_1}, h_{a_2}, \cdots, h_{a_m}\right\} \in A.$$

For all $x \in X$ we can choose an index k such that $x \in B(a_k, s_{a_k})$ and we can choose an index ℓ such that $g(x) = h_{a_\ell}(x)$ and then we have

$$g(x) \geq h_{a_k}(x) > f(x) - \epsilon \;\text{ and }\; g(x) = h_{a_\ell}(x) < f(x) + \epsilon.$$

6.20 Corollary: *(The Weierstrass Approximation Theorem) Let $X \subseteq \mathbf{R}^n$ be compact and let $f \in \mathcal{C}(X)$. Then for all $\epsilon > 0$ there exists a polynomial p in n variables such that $\left|p(x) - f(x)\right| < \epsilon$ for all $x \in X$.*

Proof: Each polynomial p in n-variables determines a continuous function $p : X \to \mathbf{R}$. The set $\mathcal{P}(X)$ of such polynomial functions is a subalgebra of $\mathcal{C}(X)$ which separates points (for $a, b \in X$, if $a \neq b$ then $a_k \neq b_k$ for some index k, and then the polynomial $p(x) = x_k$ separates a and b) and vanishes nowhere $\big($because $1 \in \mathcal{P}(X)\big)$, so $\mathcal{P}(X)$ is dense in $\mathcal{C}(X)$.

6.21 Corollary: *The space $\big(\mathcal{C}[a,b], d_\infty\big)$ is separable, where $a, b \in \mathbf{R}$ with $a < b$.*

Proof: Let P be the set of polynomials with coefficients in $\mathbf{Q}$. Note that P is countable by Theorem 1.20 $\big($indeed, $\mathbf{Q}$ is countable by Part 4 of Theorem 1.20, hence $\mathbf{Q}^2, \mathbf{Q}^3, \cdots, \mathbf{Q}^n$ are all countable by Part 1 of Theorem 1.20 and by induction, hence the space P_n of polynomials over $\mathbf{Q}$ of degree at most n is countable since the map $F : \mathbf{Q}^{n+1} \to P_n$ given by $F(a_0, a_1, \cdots, a_{n+1}) = \sum\limits_{k=0}^{n} a_k x^k$ is bijective, and hence $P = \bigcup\limits_{n=0}^{\infty} P_n$ is countable by Part 3 of Theorem 1.20$\big)$. We claim that P is dense in $\mathcal{C}[a,b]$. Let $f \in \mathcal{C}[a,b]$ and let $\epsilon > 0$. By the Weierstrass Approximation Theorem we can choose a polynomial p with coefficients in $\mathbf{R}$ such that $\|p - f\|_\infty < \frac{\epsilon}{2}$, say $p(x) = \sum\limits_{k=0}^{n} c_k x^k$ with each $c_k \in \mathbf{R}$. Let $m = \max\{|a|, |b|, 1\}$, for each index k, choose $a_k \in \mathbf{Q}$ with $|a_k - c_k| < \frac{\epsilon}{2(n+1)m^n}$ and let $g(x) = \sum\limits_{k=0}^{n} a_k x^k$. Then for all $x \in [a,b]$ we have $|x| \leq m$ $\big($since $m \geq \max\{|a|, |b|\}\big)$ and hence for all $0 \leq k \leq n$ we have $|x|^k \leq m^k \leq m^n$ (since $m \geq 1$). Thus for all $x \in [a,b]$ we have

$$\left|g(x) - p(x)\right| = \left|\sum_{k=0}^{n} (a_k - c_k) x^k\right| \leq \sum_{k=0}^{n} |a_k - c_k|\,|x|^k \leq \sum_{k=0}^{n} \tfrac{\epsilon}{2(n+1)m^n}\, m^n = \tfrac{\epsilon}{2}.$$

Thus $\|g - p\|_\infty \leq \frac{\epsilon}{2}$ and hence $\|g - f\|_\infty \leq \|g - p\|_\infty + \|p - f\|_\infty < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$.

6.22 Exercise: Let $A = \left\{b_0 + \sum\limits_{k=1}^{n} \big(a_k \sin(kx) + b_k \cos(kx)\big) \,\middle|\, n \in \mathbf{Z}^+, a_k, b_k \in \mathbf{R}\right\}$. Show that A is dense in $\mathcal{C}[0,\pi]$ but A is not dense in $\mathcal{C}[0, 2\pi]$.

7.1 Definition: When I is the bounded open interval $I=(a,b)$, where $a,b\in\mathbf{R}$ with $a\le b$, the diameter of I is $d(I)=b-a$. For a subset $A\subseteq\mathbf{R}$, we define the **Lebesgue outer measure** of A to be

$$\lambda(A)=\inf\left\{\sum_{k=1}^{\infty}d(I_k)\,\middle|\,\text{each } I_k \text{ is a bounded open interval in } \mathbf{R} \text{ and } A\subseteq\bigcup_{k=1}^{\infty}I_k\right\}$$

with $0\le\lambda(A)\le\infty$. We say that A has (Lebesgue) **measure zero** when $\lambda(A)=0$.

7.2 Note: Every finite or countable set $A\subseteq\mathbf{R}$ has measure zero. Indeed, if A is finite, say $A=\{a_1,a_2,\cdots,a_n\}$, then given $\epsilon>0$ then we can take $I_k=\left(a_k-\frac{\epsilon}{2n},a_k+\frac{\epsilon}{2n}\right)$ for $k\le n$, and we can take $I_k=\emptyset$ for $k>n$, to get $A\subseteq\bigcup\limits_{k=1}^{\infty}I_k$ and $\sum\limits_{k=1}^{\infty}d(I_k)=\sum\limits_{k=1}^{n}\frac{\epsilon}{n}=\epsilon$. And if A is infinite, say $A=\{a_1,a_2,a_3,\cdots\}$, then we can take $I_k=\left(a-\frac{\epsilon}{2^{k+1}},a_k+\frac{\epsilon}{2^{k+1}}\right)$ for all $k\ge1$ to get $A\subseteq\bigcup\limits_{k=1}^{\infty}I_k$ and $\sum\limits_{k=1}^{\infty}d(I_k)=\sum\limits_{k=1}^{\infty}\frac{\epsilon}{2^k}=\epsilon$. Perhaps surprisingly, it is not the case that every set of measure is at most countable.

7.3 Example: The (standard) **Cantor set** is the set $C\subseteq[0,1]$ constructed as follows. Let $C_0=[0,1]$. Let I_1 be the open middle third of C_0, that is let $I_1=\left(\frac{1}{3},\frac{2}{3}\right)$, and let $C_1=A_0\backslash U_1=\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]$. Let I_2 and I_3 be the open middle thirds of the two component intervals of C_1, that is let $I_2=\left(\frac{1}{9},\frac{2}{9}\right)$ and $I_3=\left(\frac{7}{9},\frac{8}{9}\right)$, and let $C_2=C_1\setminus(I_2\cup I_3)$. Having constructed the set C_n, which is the disjoint union of 2^n closed intervals each of length $\frac{1}{3^n}$, let $I_{2^n},I_{2^n+1},\cdots,I_{2^{n+1}-1}$ be the open middle thirds of these 2^n component intervals and let $C_{n+1}=C_n\setminus(I_{2^n},I_{2^n+1},\cdots,I_{2^{n+1}-1})$. Note that C_n is the set of all numbers $x\in[0,1]$ which can be written in base 3 such that the the first n digits of x are not equal to 1.

The Cantor set is the set

$$C=\bigcap_{n=0}^{\infty}C_n$$

or equivalently, C is the set of all numbers $x\in[0,1]$ which can be written in base 3 with none of the digits of x equal to 1.

Since $C=\bigcap\limits_{n=0}^{\infty}C_n$ with $C_0\supseteq C_1\supseteq C_2\supseteq\cdots$, it follows that $C\subseteq C_n$ for all $n\in\mathbf{N}$. Since C_n is the (disjoint) union of 2^n closed intervals each of size $\frac{1}{3^n}$, it follows that we can cover C_n (hence also C) by a union of 2^n open intervals each of size $\frac{2}{3^n}$, and so we have $\lambda(C)\le2^n\cdot\frac{2}{3^n}=\frac{2^{n+1}}{3^n}$. Since $\lambda(C)\le\frac{2^{n+1}}{3^n}$ for all $n\in\mathbf{N}$ and $\frac{2^{n+1}}{3^n}\to0$ as $n\to\infty$, it follows that $\lambda(C)=0$.

On the other hand, since C is the set of all real numbers $x\in[0,1]$ which can be written in base 3 using only the digits 0 and 2, it follows that $|C|=2^{\aleph_0}$.

7.4 Remark: Note that the set C of numbers $x\in[0,1]$ which *can* be written in base 3 without using the digit 1, is not equal to the complement of the set B of numbers $x\in[0,1]$ which *can* be written in base 3 using the digit 1 (at least once). For example, the number $x=\frac{1}{3}$ can be written in base 3 as $x=0.1$ so we have $x\in B$, but it can also be written in base 3 as $x=0.0222\cdots$, so we also have $x\in C$.

7.5 Exercise: Show that the set of all real numbers $x\in[0,1]$, which can be written in base 5 without using the digit 2, has measure zero.

7.6 Definition: Let X be a metric space and let $A \subseteq X$. Recall that A is **dense** (in X) when for every nonempty open ball $B \subseteq X$ we have $B \cap A \neq \emptyset$, equivalently when $\overline{A} = X$. We say A is **nowhere dense** (in X) when for every nonempty open ball $B \subseteq \mathbf{R}$ there exists a nonempty open ball $C \subseteq B$ with $C \cap A = \emptyset$, or equivalently when $\overline{A}^{\,0} = \emptyset$.

7.7 Exercise: Show that the Cantor set is nowhere dense in $[0,1]$ (or in $\mathbf{R}$).

7.8 Note: When $A \subseteq B \subseteq X$, note that if A is dense in X then so is B and, on the other hand, if B is nowhere dense in X then so is A.

7.9 Note: When $A, B \subseteq X$ with $B = A^c = X \setminus A$, note that A is nowhere dense $\iff \overline{A}^{\,0} = \emptyset \iff \overline{B^0} = X \iff$ the interior of B is dense.

7.10 Definition: Let $A \subseteq X$. We say that A is **first category** (or that A is **meagre**) when A is equal to a countable union of nowhere dense sets. We say that A is **second category** when it is not first category. We say that A **residual** when A^c is first category.

7.11 Note: Every countable set in $\mathbf{R}$ is first category since if $A = \{a_1, a_2, a_3, \cdots\}$ then we have $A = \bigcup_{k=1}^{\infty} \{a_k\}$. In particular $\mathbf{Q}$ is first category and $\mathbf{Q}^c = \mathbf{R} \setminus \mathbf{Q}$ is residual.

7.12 Note: If $A \subseteq X$ is first category then so is every subset of A.

7.13 Note: If $A_1, A_2, A_3, \cdots \subseteq X$ are are all first category then so is $\bigcup_{k=1}^{\infty} A_k$.

7.14 Theorem: *(The Baire Category Theorem) Let X be a complete metric space.*
(1) Every first category set in X has an empty interior.
(2) Every residual set in X is dense.
(3) Every countable union of closed sets with empty interiors in X has an empty interior.
(4) Every countable intersection of dense open sets in X is dense.

Proof: Parts (1) and (2) are equivalent by taking complements, and Parts (3) and (4) are special cases of Parts (1) and (2), so it suffices to prove Part (1). We sketch a proof.

Let $A \subseteq X$ be first category, say $A = \bigcup_{n=1}^{\infty} C_n$ where each C_n is nowhere dense. Suppose, for a contradiction, that A has nonempty interior, and choose an open ball $B_0 = B(a_0, r_0)$ with $0 < r_0 < 1$ such that $\overline{B}_0 \subseteq A$. Since each C_n is nowhere dense, we can chose a nested sequence of open balls $B_n = B(a_n, r_n)$ with $0 < r_n < \frac{1}{2^n}$ such that $\overline{B}_n \subseteq B_{n-1}$ and $\overline{B}_n \cap C_n = \emptyset$. Because $r_n \to 0$, it folows that the sequence $\{a_n\}$ is Cauchy. Because X is complete, it follows that $\{a_n\}$ converges in X, say $a = \lim_{n\to\infty} a_n$. Note that $a \in \overline{B}_n$ for all n since $a_k \in \overline{B}_n$ for all $k \geq n$. Since $a \in \overline{B}_0$ and $\overline{B}_0 \subseteq A$ we have $a \in A$. But since $a \in \overline{B}_n$for al $n \geq 1$, and $\overline{B}_n \cap C_n = \emptyset$, we have $a \notin C_n$ for all $n \geq 1$ hence $a \notin \bigcup_{n=1}^{\infty} C_n$, that is $a \notin A$.

7.15 Example: Recall that $\mathbf{Q}$ is first category and $\mathbf{Q}^c$ is residual. The Baire Category Theorem shows that $\mathbf{Q}^c$ cannot be first category because if $\mathbf{Q}$ and $\mathbf{Q}^c$ were both first category then $\mathbf{R} = \mathbf{Q} \cup \mathbf{Q}^c$ would also be first category, but this is not possible since $\mathbf{R}$ does not have empty interior.

7.16 Example: Let $f \in \mathcal{C}^\infty(\mathbf{R})$ and suppose that for all $x \in \mathbf{R}$ there exists $n_x \in \mathbf{Z}^+$ such that $f^{(n_x)}(x) = 0$. Show that there exists a nonempty open interval $(a,b) \subseteq \mathbf{R}$ such that the restriction of f to (a,b) is a polynomial.

Solution: For each $n \in \mathbf{Z}^+$, let $A_n = \left\{x \in \mathbf{R} \middle| f^{(n)}(x) = 0\right\}$. Since we are assuming that for all $x \in \mathbf{R}$ there exists n_x such that $f^{(n_x)}(x) = 0$, it follows that $\bigcup_{n=1}^{\infty} A_n = \mathbf{R}$. Note that each set A_n is closed because $f \in \mathcal{C}^\infty(\mathbf{R})$ so that $f^{(n)}$ is continuous, and A_n is the inverse image under $f^{(n)}$ of the closed set $\{0\}$. Since $\mathbf{R} = \bigcup_{n=1}^{\infty} A_n$ and each A_n is closed, it follows from the Baire Category Theorem that at least one of the sets A_n must have a nonempty interior. Choose n such that A_n has a nonempty interior, and choose a nonempty open interval $(a,b) \subseteq A_n$. Then we have $f^{(n)}(x) = 0$ for all $x \in (a,b)$, and so the restriction of f to (a,b) is a polynomial of degree at most n.

7.17 Exercise: For each $n \in \mathbf{Z}^+$, let $f_n : \mathbf{R} \to \mathbf{R}$ be continuous. Suppose that for all $x \in \mathbf{R}$ there exists $n \in \mathbf{Z}^+$ such that $f_n(x) \in \mathbf{Q}$. Prove that there exists $n \in \mathbf{Z}^+$ such that f_n is constant in some nondegenerate interval.

7.18 Remark: Let $\mathcal{C}_1 = \left\{A \subseteq \mathbf{R} \middle| A \text{ is finite or countable}\right\}$, $\mathcal{C}_2\left\{A \subseteq \mathbf{R} \middle| \lambda(A) = 0\right\}$ and $\mathcal{C}_3 = \left\{A \subseteq \mathbf{R} \middle| A \text{ is first category}\right\}$. Note that if $\mathcal{C} = \mathcal{C}_k$ for some $k \in \{1,2,3\}$, then $\mathcal{C}$ has the following properties:

(1) if $A \subseteq B$ and $B \in \mathcal{C}$ then $A \in \mathcal{C}$,

(2) if $A_1, A_2, A_3, \cdots \in \mathcal{C}$ then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{C}$, and

(3) if $A \in \mathcal{C}$ then $A^0 = \emptyset$.

Because of this, it seems reasonable to consider each set $\mathcal{C}_k$ to be, in some sense, "small". Perhaps surprisingly, the following theorem states that every set in $\mathbf{R}$ is the union of two such small sets.

7.19 Theorem: *Every subset of* $\mathbf{R}$ *is equal to the disjoint union of a set of measure zero and a set of first category.*

Proof: Let $\mathbf{Q} = \{a_1, a_2, a_3, \cdots\}$. For $k, \ell \in \mathbf{Z}^+$, let $I_{k,\ell} = \left(a_\ell - \frac{1}{2^{k+\ell}}\,,\, a_\ell + \frac{1}{2^{k+\ell}}\right)$ and for $k \in \mathbf{Z}^+$, let $U_k = \bigcup_{\ell=1}^{\infty} I_{k,\ell}$. Note that each U_k is open with $\mathbf{Q} \subseteq U_k$, so each U_k is a dense open set. Also note that for each $k \in \mathbf{Z}^+$ we have $\lambda(U_k) \leq \sum_{\ell=1}^{\infty} d(I_{k,\ell}) = \frac{1}{2^{k-1}}$. Let $B = \bigcap_{k=1}^{\infty} U_k$ and note that B is residual, since it is a countable intersection of dense open sets. Since $B = \bigcap_{k=1}^{\infty} U_k$ and $U_1 \supseteq U_2 \supseteq U_3 \supseteq \cdots$, we have $B \subseteq U_k$ for all k, hence $\lambda(B) \leq \lambda(U_k) \leq \frac{1}{2^{k-1}}$ for all $k \in \mathbf{Z}^+$, and it follows that $\lambda(B) = 0$. Thus $\mathbf{R}$ is the disjoint union of the set B, which has measure zero, and its complement B^c which is first category (since B is residual). Finally note that any set $A \subseteq \mathbf{R}$ is equal to the disjoint union $A = (A \cap B) \cup (A \cap B^c)$, and we have $\lambda(A \cap B) = 0$ and the set $A \cap B^c$ is first category.

7.20 Remark: At first glance, it might appear that the set B constructed in the above proof might simply be equal to $\mathbf{Q}$. But in fact, B must be uncountable, because if B was countable then B would be first category, but then B and B^c would both be first category, and hence $\mathbf{R} = B \cup B^c$ would also be first category. But $\mathbf{R}$ is not first category by the Baire Category Theorem.

7.21 Example: Most students will have seen that it is possible to construct a continuous function $f:[0,1]\to\mathbf{R}$ such that f is nowhere differentiable. Show that the set of nowhere differentiable functions is residual (hence dense) in $\mathcal{C}[0,1]$.

Solution: Let A be the complement of the set of nowhere differentiable functions in $\mathcal{C}[0,1]$, that is

$$A=\Big\{f\in\mathcal{C}[0,1]\,\Big|\,f\text{ is differentiable at some point }a\in[0,1]\Big\}.$$

For each $k,\ell\in\mathbf{Z}^+$, let

$$A_{k,\ell}=\Big\{f\in\mathcal{C}[0,1]\,\Big|\,\exists a\in[0,1]\ \forall x\in[0,1]\ \ 0<|x-a|<\tfrac{1}{k}\implies\left|\tfrac{f(x)-f(a)}{x-a}\right|\le\ell\Big\}.$$

We shall show that $A=\bigcup\limits_{k,\ell}A_{k,\ell}$, and that each $A_{k,\ell}$ is closed in $\mathcal{C}[0,1]$ with an empty interior and so A is first category. Thus the set of nowhere differentiable functions is residual, and hence dense by the Baire Category Theorem.

We claim that $A=\bigcup\limits_{k,\ell}A_{k,\ell}$. Let $f\in A$. Choose $a\in[0,1]$ such that f is differentiable at a. Choose $\ell\in\mathbf{Z}^+$ such that $|f'(a)|\le\ell$. Choose $\delta>0$ such that for all $x\in[0,1]$ we have $0<|x-a|<\delta\implies\left|\frac{f(x)-f(a)}{x-a}-f'(a)\right|<\ell-|f'(a)|$. Choose $k\in\mathbf{Z}^+$ with $\frac{1}{k}\le\delta$. Then for all $x\in[0,1]$, if $0<|x-a|<\frac{1}{k}$ then we have $\left|\frac{f(x)-f(a)}{x-a}-f'(a)\right|<\ell-|f'(a)|$ and hence

$$\left|\tfrac{f(x)-f(a)}{x-a}\right|\le\left|\tfrac{f(x)-f(a)}{x-a}-f'(a)\right|+|f'(a)|\le\big(\ell-|f'(a)|\big)+|f'(a)|=\ell$$

so that $f\in A_{k,\ell}$. Thus $A=\bigcup\limits_{k,\ell}A_{k,\ell}$, as claimed.

We claim that each set $A_{k,\ell}$ is closed in $\mathcal{C}[0,1]$. Let $(f_n)_{n\ge1}$ be a sequence in $A_{k,\ell}$ which converges in $\mathcal{C}[0,1]$, and let $g=\lim\limits_{n\to\infty}f_n$ in $\mathcal{C}[0,1]$. Then $f_n\to g$ uniformly in $[0,1]$, and we need to show that $g\in A_{k,\ell}$. For each $n\in\mathbf{Z}^+$, since $f_n\in A_{k,\ell}$ we can choose $a_n\in[0,1]$ such that for all $x\in[0,1]$ we have $0<|x-a_n|<\frac{1}{k}\implies\left|\frac{f_n(x)-f(a_n)}{x-a_n}\right|\le\ell$. Since $[0,1]$ is compact, we can choose a convergent subsequence $(a_{n_k})_{k\ge1}$ of the sequence $(a_n)_{n\ge1}$ and let $a=\lim\limits_{k\to\infty}a_{n_k}\in[0,1]$. Note that the corresponding subsequence $(f_{n_k})_{k\ge1}$ of $(f_n)_{n\ge1}$ converges in $\mathcal{C}[0,1]$ with the same limit $g=\lim\limits_{k\to\infty}f_{n_k}$ in $\mathcal{C}[0,1]$. Note that when $0<|x-a|<\frac{1}{k}$, since $a_{n_k}\to a$ it follows that we also have $0<|x-a_{n_k}|<\frac{1}{k}$ for sufficiently large $k\in\mathbf{Z}^+$. Since $f_{n_k}\to g$ uniformly on $[0,1]$ and $a_{n_k}\to a$ in $[0,1]$, recall (or verify) that $\lim\limits_{k\to\infty}f_{n_k}(a_{n_k})=g(a)$ and so, for all $x\in[0,1]$ with $0<|x-a|<\frac{1}{k}$

$$\left|\frac{g(x)-g(a)}{x-a}\right|=\lim_{k\to\infty}\left|\frac{f_{n_k}(x)-f_{n_k}(a_{n_k})}{x-a_{n_k}}\right|\le\ell.$$

This proves that $g\in A_{k,\ell}$ and so $A_{k,\ell}$ is closed in $\mathcal{C}[0,1]$, as claimed.

We claim that each set $A_{k,\ell}$ has empty interior in $\mathcal{C}[0,1]$. Let $f\in A_{k,\ell}$. We need to show that for all $r>0$ there there is a function $g\in B(f,r)$ with $g\notin A_{k,\ell}$. Our strategy is to first find a piecewise linear function p with $\|p-f\|_\infty<\frac{r}{2}$ and then to add a rapidly oscillating sine function to obtain a function $g=p+\frac{r}{2}\sin(wx)$ with $g\notin A_{k,\ell}$ and with $\|g-f\|_\infty<r$. Let $r>0$. Since f is uniformly continuous on $[0,1]$ we can choose $\delta>0$ such that $|x-y|<\delta\implies|f(x)-f(y)|<\frac{r}{4}$. we can choose $n\in\mathbf{Z}^+$ such that $\frac{1}{n}<\delta$. Let $x_i=\frac{i}{n}$ for $0\le i\le n$ and let $p\in\mathcal{C}[0,1]$ be the piecewise linear function whose graph has vertices at $\big(x_i, f(x_i)\big)$ for $0 \le i \le n$. Then for all i and for all $x \in [x_{i-1}, x_i]$, we have

$$\begin{aligned}\big|f(x)-p(x)\big| &\le \big|f(x)-f(x_i)\big| + \big|f(x_i)-p(x)\big| = \big|f(x)-f(x_i)\big| + \big|p(x_i)-p(x)\big| \\ &\le \big|f(x)-f(x_i)\big| + \big|p(x_i)-p(x_{i-1})\big| < \tfrac{r}{4}+\tfrac{r}{4} = \tfrac{r}{2}\end{aligned}$$

and hence $\|f-p\|_\infty < \frac{r}{2}$. Let $m = \max\limits_{t \ne x_i} \big|p'(t)\big| = \max\limits_{1\le i\le n} n\big|f(x_i)-f(x_{i-1})\big|$. Choose $\omega \in \mathbf{R}$ such that $\frac{2\pi}{\omega} < \frac{1}{k}$ and $\frac{2\pi}{w} < \frac{r}{2(\ell+m)}$, and consider the function $g = p + \frac{r}{2}\sin(\omega x)$. Note that $\|g-f\|_\infty \le \|g-p\|_\infty + \|p-f\|_\infty < \frac{r}{2}+\frac{r}{2} = r$, so it remains only to show that $g \notin A_{k,\ell}$. Let $a \in [0,1]$. By our choice of ω we can choose $x \in [0,1]$ with $0 < |x-a| < \frac{1}{k}$ such that $|x-a| < \frac{r}{2(\ell+m)}$ and such that $\sin(\omega x) = \pm 1$ with $\sin(\omega x) = 1 \iff \sin(\omega a) \le 0$ so that $\big|\sin(\omega x) - \sin(\omega a)\big| \ge 1$. Then we have

$$\begin{gathered}\tfrac{r}{2}\,\big|\sin(\omega x)-\sin(\omega a)\big| = \big|\big(g(x)-g(a)\big) - \big(p(x)-p(a)\big)\big| \le \big|g(x)-g(a)\big| + \big|p(x)-p(a)\big| \\ \big|g(x)-g(a)\big| \ge \tfrac{r}{2}\big|\sin(\omega x)-\sin(\omega a)\big| - \big|p(x)-p(a)\big| \ge \tfrac{r}{2} - \big|p(x)-p(a)\big| \\ \left|\tfrac{g(x)-g(a)}{x-a}\right| \ge \frac{r}{2|x-a|} - \left|\frac{p(x)-p(a)}{x-a}\right| \ge \frac{r}{2\cdot\frac{2(\ell+m)}{r}} - m = \ell\end{gathered}$$

so that $g \notin A_{k,\ell}$, as required.

7.22 Notation: Let X be a set. For any set $\mathcal{C}$ of subsets of X we write

$$\mathcal{C}_\sigma = \left\{ \bigcup_{k=1}^{\infty} A_k \,\middle|\, \text{each } A_k \in \mathcal{C} \right\} \text{ and } \mathcal{C}_\delta = \left\{ \bigcap_{k=1}^{\infty} A_k \,\middle|\, \text{each } A_k \in \mathcal{C} \right\}.$$

Note that $\mathcal{C}_{\sigma\sigma} = \mathcal{C}_\sigma$ and $\mathcal{C}_{\delta\delta} = \mathcal{C}_\delta$.

7.23 Definition: Let X be a set. A **σ-algebra** in X is a set $\mathcal{C}$ of subsets of X such that
(1) $\emptyset \in \mathcal{C}$,
(2) if $A \in \mathcal{C}$ then $A^c = X \setminus A \in \mathcal{C}$, and
(3) if $A_1, A_2, A_3, \dots \in \mathcal{C}$ then $\bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{C}$.

Note that when $\mathcal{C}$ is a σ-algebra in X we have $\mathcal{C}_\sigma = \mathcal{C}$ and $\mathcal{C}_\delta = \mathcal{C}$.

7.24 Notation: In a metric space (or topological space) X, we let $\mathcal{G}$ denote the set of all open sets in X and we let $\mathcal{F}$ denote the set of all closed subsets of X. Note that $\mathcal{G}_\sigma = \mathcal{G}$ and $\mathcal{F}_\delta = \mathcal{F}$.

7.25 Example: For any set X, the set $\{\emptyset, X\}$ and the set $\mathcal{P}(X)$ of all subsets of X are σ-algebras in X,

7.26 Note: Note that given any set $\mathcal{C}$ of subsets of a set X there exists a unique smallest σ-algebra in X which contains $\mathcal{C}$, namely the intersection of all σ-algebras in X which contain $\mathcal{C}$.

7.27 Definition: In a metric space (or topological space) X, the **Borel** σ-algebra $\mathcal{B}$ is the smallest σ-algebra in X which contains $\mathcal{G}$ (hence also $\mathcal{F}$). The elements of $\mathcal{B}$ are called **Borel** sets. Note that $\mathcal{B}$ contains all of the sets $\mathcal{G}, \mathcal{G}_\delta, \mathcal{G}_{\delta\sigma}, \mathcal{G}_{\sigma\delta\sigma}, \cdots$ and all of the sets $\mathcal{F}, \mathcal{F}_\sigma, \mathcal{F}_{\sigma\delta}, \mathcal{F}_{\sigma\delta\sigma}, \cdots$.

7.28 Exercise: Using the Baire Category Theorem, show that in $\mathbf{R}$ we have $\mathcal{F} \subseteq \mathcal{G}_\delta$ (equivalently $\mathcal{G} \subseteq \mathcal{F}_\sigma$), $\mathcal{F}_\sigma \ne \mathcal{G}_\delta$, and $\mathcal{G}_\delta \cup \mathcal{F}_\sigma \subsetneqq \mathcal{G}_{\delta\sigma} \cap \mathcal{F}_{\sigma\delta}$.