Contents

1 Fun Facts from Set Theory

Definition 1.1. A **Cartesian product** of two sets X and Y is the set of ordered pairs

$$\{(x, y)|x \in X, y \in Y\}.$$

Definition 1.2. A **function** f between two sets X and Y is a subset of $X \times Y$ such that for every $x \in X$, there exists exactly one $y \in Y$ such that $(x, y) \in f$. In this case, we often denote this unique y as $f(x)$.
The set of functions from X to Y is denoted X^Y.

Definition 1.3. Two sets have the same **cardinality** iff there exists a bijection between the them.
We say a set E is **countable** if there exists an injective function from E into the set of natural numbers $\mathbb{N} = \{1, 2, ...\}$.

Propostion 1.1. *To verify that a set E is countable, it suffices to establish an injection from E to any other countable set.*

Propostion 1.2. *The following sets are countable:*

- $\mathbb{N} \cup F$ *where F is any finite set.*
- $A \cup B$ *where A and B are countable.*
- $\mathbb{N} \times \mathbb{N}$
- $A \times B$ *where A and B are countable.*
- $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup \{-n|n \in \mathbb{N}\}$
- $\mathbb{Q} = \{\frac{p}{q}|p, q \in \mathbb{Z}$ *are relatively prime* $, q \neq 0\}$

We'll prove it for $\mathbb{N} \times \mathbb{N}$:

Proof. By the Fundamental Theorem of Arithmetic, every natural number has a prime decomposition. Hence, the function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by $f(n, m) = 2^n 3^m$ is well defined and injective. □

Exercise 1.1. Show that $\mathbb{Q}$ is countable.

Propostion 1.3. $\mathbb{R}$ *is uncountable. So is the* **Cantor set** $\{0, 1\}^{\mathbb{N}}$.

The proof is Cantor's diagonalization argument, which I suggest you look up if you have not seen it before.

1.1 Orders and AC

Definition 1.4. A **partial ordering** on a set X is a binary relation $\geq$ (i.e. a subset of $X \times X$ where we denote (a, b) by $a \leq b$) if for every $a, b \in X$

1. $a \leq a$
2. $a \leq b$ and $b \leq a \Rightarrow a = b$
3. $a \leq b$ and $b \leq c \Rightarrow a \leq c$

A set X has a **total ordering** if it has a partial ordering and additionally, for any two pair of elements $a, b \in X$ either $a \leq b$ or $b \leq a$.
A set X is **well-ordered** if it has a total ordering under which every non-empty subset has a least element.

Example 1.1. An example of a partial order would be inclusion on the power set of some set X. The real numbers have a total ordering. The natural numbers are well-ordered.

Exercise 1.2. Any countable set is well-ordered. (Note: this means it has an order with respect to which it is well-ordered, but the order need not be canonical.)

Definition 1.5. A **sequence** is a function $f : I \to X$ to a countable well-ordered set I into a set X. It is usually identified with an ordered subset of X, which is denoted by $(x_n)_{n \in I}$. I is usually $\mathbb{N}$, which is often omitted from subscripts.
A **net** is a function $f : A \to X$ from a **directed set** A into a set X. It is usually identified with a partially ordered subset of X, which is denoted by $(x_\alpha)_{\alpha \in A}$.

Definition 1.6. A **directed set** D is a partially ordered set such that for any two elements $\alpha, \beta \in D$, there exists $\gamma \in D$ such that $\gamma \geq \alpha$ and $\gamma \geq \beta$.

Definition 1.7. A function between two partially ordered sets X and Y is **nondecreasing** (**nonincreasing**) if $x \leq y \Rightarrow f(x) \leq f(y)$ $(x \leq y \Rightarrow f(x) \geq f(y))$ for all $x, y \in X$. It is strictly monotone if the non-strict inequality can be replaced with a strict inequality.

Example 1.2. A sequence $(x_n)_{n \in \mathbb{N}}$ is monotone if $x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$.

Theorem 1.1. *The following are equivalent:*

1. *(Choice Axiom) For any nonempty collection X of sets, there exists a function $f : X \to \bigcup X$ such that $f(Y) \in Y$ for each $Y \in X$.*
2. *(Zorn's Lemma) If X is a partially ordered set such that every chain in X has an upper bound in X, then X has a maximal element. i.e. Suppose that X is a partially ordered set such that for every totally ordered subset Y of X there exists an $x \in X$ such that $x \geq y$ for all $y \in X$. Then there exists an $x \in X$ such that if $z \geq x$ then $z = x$ for all $z \in X$.*

3. (Well-Ordering Principle) Every set X can be well-ordered.

We can work through (3) $\Rightarrow$ (1) pretty easily. Indeed, suppose $X = \{X_\alpha\}_{\alpha \in A}$ is a collection of nonempty sets. Then there exists a well-ordering on $\bigcup_\alpha X_\alpha$. Define $f : X \to \bigcup_\alpha X_\alpha$ by $f(X_\alpha) = \min\{x \in X_\alpha \subseteq \bigcup_\alpha X_\alpha\}$.

Invoking the choice axiom is very subtle. Note that this argument makes only one arbitrary choice: the ordering relation. Since the well-ordering theorem only says that there exists a well-ordering for each X_α, assuming that we could choose a well-ordering for each X_α would amount to assuming the axiom of choice.

Nonetheless, the finite version of the choice axiom is a theorem in ZF set theory:

Theorem 1.2. *A finite collection of nonempty sets has a choice function.*

We implicitly used this when proving that $\mathbb{N} \times \mathbb{N}$ is countable.

Let's have one more formulation of the choice axiom in terms of a Cartesian product.

Definition 1.8. Let $\mathcal{A}$ be a partially ordered set, and suppose $\{A_\alpha\}_{\alpha \in \mathcal{A}}$ is a family of sets. The Cartesian product of this family is the set

$$\prod_{\alpha \in \mathcal{A}} A_\alpha = \{\phi : \mathcal{A} \to \bigcup_\alpha A_\alpha | \phi(\alpha) \in A_\alpha \ \forall \ \alpha \in \mathcal{A}\}.$$

The Axiom of Choice states that if each A_α and $\mathcal{A}$ are non-empty, then

$$\prod_{\alpha \in \mathcal{A}} A_\alpha \neq \varnothing.$$

(So innocuous.) This fact does not require the choice axiom if $|\mathcal{A}| < \infty$.

Propostion 1.4. *Let X be a collection of countable sets. If X is finite, then $\bigcup X$ is countable. If X is countable, then $\bigcup X$ is countable, if one assumes the choice axiom (a countable version will suffice).*

We'll just prove the claim for a countable union of countable sets.

Proof. Suppose $(X_n)_{n \in \mathbb{N}}$ is a countable collection of countable sets, and let $X = \bigcup_n X_n$. Since each X_n is countable, there exists at least one injection $f_n : X_n \to \mathbb{N}$ for each n. That is, the sets $F_n = \{f_n : X_n \to \mathbb{N}$ an injection$\}$ are nonempty. By the axiom of choice, $\prod_n F_n$ is nonempty. Let $(f_n) \in \prod_n F_n$. Now, define $f : X \to \mathbb{N} \times \mathbb{N}$ by

$$f(x) = (\min\{n \in \mathbb{N} | x \in X_n\}, f_n(x)).$$

Since $\mathbb{N}$ is well-ordered, this is well-defined and moreover injective. Since $\mathbb{N} \times \mathbb{N}$ is countable, we are done. $\square$

Many fundamental results in functional analysis, such as the Hahn-Banach Theorem, Baire Category Theorem (and its consequences), Banach-Alaoglu's Theorem, the existence of an orthonormal basis for any Hilbert space, the existence of non-trivial ultrafilters, etc. rely on the choice axiom, which makes it seem like something a mathematician would obviously want to incorporate into their axiomatic framework. However, it also leads to paradoxes, such as the Banach-Tarski paradoxical decompositions, which (assuming the choice axiom) exist for any nonamenable group and for measurable subsets of Euclidean space. Colloquially speaking, the paradox says we can take finitely many (5 actually) subsets of the unit sphere in $\mathbb{R}^2$ and rearrange them, without altering their size (measure), into two copies of the unit sphere. (This would be a good presentation topic. If you are algebraically inclined, you could talk about a paradoxical decomposition of a nonabelian free group!)

2 Properties of $\mathbb{R}$ and $\mathbb{C}$

2.1 Properties of $\mathbb{R}$

Algebraically speaking, $\mathbb{R}$ is a field, meaning it has commutative multiplication and addition, and every element is invertible except the additive identity, 0. $\mathbb{R}$ is not algebraically closed, i.e. there exist polynomails with real coefficients that are never zero on $\mathbb{R}$, e.g. $x^2 + 1$. $\mathbb{R}$ contains the rational numbers

$$\mathbb{Q} = \{p/q | p, q \in \mathbb{Z}, q \neq 0\},$$

but it also includes many more. Here is a classic proof that $\sqrt{2} \notin \mathbb{Q}$:

Propostion 2.1. $\sqrt{2} \notin \mathbb{Q}$

Proof. Suppose $p, q \in \mathbb{Z}$ are relatively prime, and $(p/q)^2 = 2$. Then $p^2 = 2q^2$ is even and so p must be as well, i.e. $p = 2k$ for some $k \in \mathbb{Z}$. Then $4k^2 = 2q^2$, and so $2k^2 = q^2$, and q is likewise even. This contradicts the assumption that p and q were relatively prime. $\square$

Definition 2.1. The **absolute value** is a function $|\cdot| : \mathbb{R} \to [0, \infty)$ defined by $|a| = \sqrt{a^2}$.

It is sometimes defined to be $|a| = a$ if $a \geq 0$ and $|a| = -a$ if $a < 0$. This is an example of a **norm**, which we will define and discuss later. Suffice to say that it is symmetric, non-negative definite, and subadditive. Moreover, it is **multiplicative**, i.e. $|ab| = |a||b|$.

Definition 2.2. Let $S \subseteq \mathbb{R}$.

- $x \in \mathbb{R}$ is a **supremum** of S if x is an upper bound for S, i.e. $x \geq s\ \forall s \in S$, and if it is the least such number, i.e. if $y \in \mathbb{R}$ is any number such that $y \geq x\ \forall\ s \in S$, then $x \leq y$.

- $x \in \mathbb{R}$ is a **infimum** of S if x is an lower bound for S, i.e. $x \leq s \ \forall s \in S$, and if it is the greatest such number, i.e. if $y \in \mathbb{R}$ is any number such that $y \leq x \ \forall \ s \in S$, then $x \geq y$.

The following are axiomatic properties of $\mathbb{R}$, which you likely often take for granted.

- **Algebraic Properties:** $\mathbb{R}$ is a field.
- **Order Properties:** There is a non-empty subset P of $\mathbb{R}$ called the set of strictly positive real numbers such that for all $a, b \in \mathbb{R}$,
  1. $a, b \in P \Rightarrow a + b \in P$
  2. $a, b \in P \Rightarrow ab \in P$
  3. Exactly one of the following relations hold: $a \in P$, $a = 0$, or $-a \in P$.
- **Supremum Property** (or Completeness): Every non-empty set of real numbers which has an upper bound has a supremum.
- **Infimum Property**: Every non-empty set of real numbers which has a lower bound has an infimum.

Definition 2.3. We define order relations $<$ and $\leq$ on $\mathbb{R}$ as follows: Let $a, b \in \mathbb{R}$

1. $a > b$ if $a - b \in P$
2. $a \geq b$ if $a - b \in P \cup \{0\}$
3. $a < b$ if $-(a - b) \in P$
4. $a \leq b$ if $-(a - b) \in P \cup \{0\}$

Exercise 2.1. Using the order properties of $\mathbb{R}$, prove the following for any $a, b, c \in \mathbb{R}$:

1. If $a > b$ and $b > c$, then $a > c$.
2. Exactly one of the following holds: $a > b$, $a = b$, or $a < b$.
3. If $a \geq b$ and $b \geq a$, then $a = b$.

Theorem 2.1. *1. For any $a \in \mathbb{R} \backslash \{0\}$, $a^2 > 0$.*

2. $1 > 0$

3. $n > 0$ *for all* $n \in \mathbb{N}$

Theorem 2.2. *Let $a, b, c, d \in \mathbb{R}$.*

1. If $a > b$, then $a + c > b + c$

2. If $a > b$ and $c > d$ then $a + c > b + d$

3. *If* $a > b$ *and* $c > 0$*, then* $ac > bc$

4. *If* $a > b$ *and* $c < 0$*, then* $ac < bc$

5. *If* $a > 0$*, then* $1/a > 0$

6. *If* $a < 0$*, then* $1/a < 0$

Exercise 2.2. Show that if $a > b$, then $a > \frac{1}{2}(a+b) > b$. Show also that there is no smallest strictly positive number.

Exercise 2.3. Show that if $ab > 0$, then either $a > 0$ or $b > 0$ or we have $a < 0$ and $b < 0$. Show that if $ab < 0$, then we either have $a > 0$ and $b < 0$ or we have $a < 0$ and $b > 0$.

Theorem 2.3. *For any* $a, b \in \mathbb{R}$

1. $|a| = 0$ *iff* $a = 0$

2. $|-a| = |a|$

3. $|ab| = |a||b|$

4. *If* $b \geq 0$*, then* $|a| \leq b$ *iff* $-b \leq a \leq b$

5. $-|a| \leq a \leq |a|$

Exercise 2.4. Prove the "reverse triangle inequality", i.e. for any $a, b \in \mathbb{R}$,

$$||a| - |b|| \leq |a - b|$$

Exercise 2.5. Show that for $a, b \in \mathbb{R}$, $|a + b| = |a| + |b|$ iff $ab \geq 0$.

Exercise 2.6. Show that a number $x \in \mathbb{R}$ is the supremum of a nonempty set $S \subseteq \mathbb{R}$ iff it satisfies the following two properties:

1. There are no elements $s \in S$ with $x < s$.

2. If $y < x$, then there exists an $s \in S$ with $y < s$.

Exercise 2.7. Let $S_0 \subseteq S \subseteq \mathbb{R}$ where S is bounded (i.e. it has upper and lower bounds) and S_0 is nonempty. Show that

$$\inf S \leq \inf S_0 \leq \sup S_0 \leq \sup S.$$

Exercise 2.8. Let $f : X \times Y \to S$ where $S \subseteq \mathbb{R}$ is bounded. Define $f_1 : X \to \mathbb{R}$ and $f_2 : Y \to \mathbb{R}$ by

$$f_1(x) = \sup_{y \in Y} f(x, y) \text{ and } f_2(y) = \sup_{x \in X} f(x, y)$$

Show that

$$\sup_{x \in X, y \in Y} f(x, y) = \sup_{x \in X} f_1(x) = \sup_{y \in Y} f_2(y)$$

Exercise 2.9. Prove the Infimum Property (see above) using the Supremum Property.

Exercise 2.10. Suppose $S, T \subseteq \mathbb{R}$ are bounded nonempty subsets. Determine whether the following is necessary, sufficient, or both for showing that $\sup_{s\in S} s = \sup_{t\in T} t$:

$$\forall\ s \in S\ \exists\ t \in T : t \geq s \quad \text{and} \quad \forall\ t \in T \exists\ s \in S : s \geq t$$

Theorem 2.4 (Archimedean Property)**.** *If $x \in \mathbb{R}$, then there is a natural number $n_x \in \mathbb{N}$ such that $x < n_x$.*

Proof. Suppose $x \geq n$ for all $n \in \mathbb{N}$. Then $\mathbb{N} \subseteq \mathbb{R}$ has an upper bound and hence a supremum, call it $s \in \mathbb{R}$. Then by exercise 2.6 there exists $n \in \mathbb{N}$ such that $s - 1 < n \leq s$. Then $s \leq n + 1$, but $n + 1 \in \mathbb{N}$. □

As a corollary to this and Theorem 2.2, we have the following:

Corollary 2.1. *Let $a, b \in P$.*

1. $\exists\ n \in \mathbb{N}$ *such that* $na > b$.
2. $\exists\ n \in \mathbb{N}$ *such that* $0 < 1/n < a$.
3. $\exists\ n \in \mathbb{N}$ *such that* $n - 1 \leq b < n$.

Similar arguments also get us the following theorem.

Theorem 2.5. *Let $x, y \in \mathbb{R}$ with $x < y$.*

1. *There exists $r \in \mathbb{Q}$ such that $x < r < y$.*
2. *If $\xi \in \mathbb{R}\backslash\mathbb{Q}$, then there exists $s \in \mathbb{Q}$ such that $x < s\xi < y$.*

2.2 Properties of $\mathbb{C}$

$\mathbb{C}$ is also a field. However, unlike $\mathbb{R}$, $\mathbb{C}$ is algebraically closed:

Theorem 2.6. *Fundamental Theorem of Algebra $\mathbb{C}$ is algebraically closed, i.e. any nonconstant polynomial $p(z)$ with coeffients in $\mathbb{C}$ has at least one zero.*

In fact, $\mathbb{C}$ is the algebraic closure of $\mathbb{R}$, i.e. the zeros of all polynomials with coefficients in $\mathbb{R}$ lie in $\mathbb{C}$. This property makes calculus on $\mathbb{C}$ incredibly elegant. We will not discuss this too heavily in this course because it is the focus of your Complex Analysis course in the fall.
For $z \in \mathbb{C}$, there exist $x, y \in \mathbb{R}$ such that $z = x + iy$. The real part of z is $\mathrm{Re}(z) = x$ and the imaginary part is $\mathrm{Im}(z) = y$.

Definition 2.4. The **modulus** is a function $|\cdot| : \mathbb{C} \to [0, \infty)$, defined by $|z| = \sqrt{x^2 + y^2}$.

The modulus is also an example of a multiplicative norm. We define the **conjugate** of z to be

$$\bar{z} = x - iy.$$

The conjugate enjoys the following properties for all $z, w \in \mathbb{C}$:

- $\bar{\bar{z}} = z$
- $\overline{zw} = \bar{z}\bar{w}$
- $\overline{z+w} = \bar{z} + \bar{w}$
- $|z| = \sqrt{z\bar{z}}$
- $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$
- $\mathrm{Re}(z) = \frac{1}{2}(z + \bar{z})$ and $\mathrm{Im}(z) = \frac{1}{2i}(z - \bar{z})$.

2.2.1 Polar Representation of a Complex Number

A point $(x, y) \neq (0, 0)$ in the plane can be described by polar coordinates r and θ where $r = \sqrt{x^2 + y^2}$ and θ is the angle subtended by (x, y) and the x-axis, which is determined up to adding an integer multiple of 2π. The Cartesian coordinates can be recovered by the formulas

$$x = r\cos\theta \qquad\qquad y = r\sin\theta.$$

If we set $z = x + iy \in \mathbb{C}$, then we can also express z in polar coordinates as

$$z = x + iy = r(\cos\theta + i\sin\theta).$$

Here $r = |z|$ and θ is often called the **argument** of z, denoted $\arg z$. Note that $\arg z$ is a multi-valued function. The **principal value** of $\arg z$, denoted by $\mathrm{Arg}\, z$ is chosen to be the value of θ that lies in $(-\pi, \pi]$. So, we have for $z \neq 0$,

$$\arg z = \{\mathrm{Arg} z + 2k\pi |\ k \in \mathbb{Z}\}.$$

Example 2.1. $\mathrm{Arg}(i) = \frac{\pi}{2}$.

Recall that

$$e^{i\theta} = \cos\theta + i\sin\theta.$$

Definition 2.5. The **polar representation** of a complex number $z \neq 0$ is

$$z = re^{i\theta}$$

where $r = |z|$ and $\theta = \arg z$.

Propostion 2.2. *We have the following properties due to the periodicity of sine and cosine and to properties of the exponential function:*

1. $|e^{i\theta}| = 1$

2. $\overline{e^{i\theta}} = e^{-i\theta}$

3. $\frac{1}{e^{i\theta}} = e^{-i\theta}$

4. $e^{i(\theta+2\pi k)} = e^{i\theta}$

5. $e^{i(\theta+\phi)} = e^{i\theta}e^{i\phi}$

(1) follows from the identity $\cos^2\theta+\sin^2\theta = 1$. (2) follows from the fact that cosine is an even function and sine an odd function. (3) follows from "cleverly multiplying by 1".

Remark 2.1. *One can manipulate (5) to recover the sum formulae for sine and cosine.*

From the properties of $e^{i\theta}$, we can derive the following properties of arg.

Propostion 2.3. *We have the following properties for the argument function (modulo adding integer multiples of 2π):*

1. $arg\bar{z} = -argz$

2. $arg(1/z) = -argz$

3. $argzw = argz + argw$.

Theorem 2.7 (de Moivre's formulae)**.** *For $n \geq 1$,*

$$\cos(n\theta) + i\sin(n\theta) = e^{in\theta} = (e^i n\theta)^n.$$

Definition 2.6. A complex number z is an **nth root of** w if $z^n = w$. For $w = 1$, we call z an **nth root of unity**.

The roots of w are the solutions to the polynomial $z^n = w$. If $w \neq 0$, it has n distinct roots. To find the roots, first write

$$w = se^{i\phi}$$

Then, the equation becomes

$$r^n e^{in\theta} = se^{i\phi}.$$

So, $r = \sqrt[n]{s}$ and $n\theta = \phi + 2\pi k$ for some $k \in \mathbb{Z}$, i.e.

$$\theta = \frac{\phi + 2\pi k}{n}, k = 0, 1, ..., n-1.$$

Applying this to $w = 1$, we get a formula for the nth roots of unity:

$$\omega_k = e^{2\pi ik/n}, k = 0, ..., n-1.$$

Going back to $w \neq 0$, we can find the nth roots of $w = se^{i\phi}$ by first finding the root

$$z_0 = s^{1/n}e^{i\phi/n}$$

and then multiplying this by the n roots of unity:

$$z_k = z_0\omega_k = s^{1/n}e^{i\phi/n}e^{2\pi ik/n}, k = 0, 1, ..., n-1.$$

Exercise 2.11. Plot the 2nd, 3rd, 4th, 5th, and 6th roots of unity.

Exercise 2.12. Check that you believe the processes for finding the nth roots of unity.

3 Basic Pointset Topology

Definition 3.1. A **topological space** is a set X with a collection τ of subsets of X such that

1. $\varnothing, X \in \tau$
2. $U, V \in \tau \Rightarrow U \cap V \in \tau$
3. $W \subseteq \tau \Rightarrow \bigcup W \in \tau$

Definition 3.2. Members of τ are called **open** sets. $C \subseteq X$ is **closed** if $X \backslash C = C^c$ is open.
For $x \in X$, a **neighborhood** of x is a set $N \subseteq X$ such that there exists an open set $U \subseteq X$ such that $x \in U \subseteq N$.

Exercise 3.1. Show that a subset $U \subseteq X$ is open iff for all $x \in U$, U contains an open neighborhood of x.

Exercise 3.2. Suppose $U \subseteq X$ is open and $C \subseteq X$ is closed. Show that $U \backslash C = \{u \in U |\ u \notin C\}$ is open and $C \backslash U$ is closed.

Exercise 3.3. Suppose $\{\tau_\alpha\}_{\alpha \in \mathcal{A}}$ is a collection of topologies on X.

1. Is $\bigcup_\alpha \tau_\alpha$ a topology on X?
2. Is $\bigcap_\alpha \tau_\alpha$ a topology on X?

Propostion 3.1. *For a topological space X:*

1. *Arbitrary intersections of closed sets are closed.*
2. *Finite unions of closed sets are closed.*
3. *$\varnothing$ and X are closed.*

The proof utilizes De Morgan's Laws (from set theory– no topology assumed). In case you haven't seen them:

Theorem 3.1. *De Morgan's Laws For a collection $\{X_\alpha\}_{\alpha \in \mathcal{A}}$ of subsets of a set X,*

- $(\bigcup_\alpha X_\alpha)^c = \bigcap_\alpha X_\alpha^c$
- $(\bigcap_\alpha X_\alpha)^c = \bigcup_\alpha X_\alpha^c$

Proof. We will just prove (1) using DeMorgan's Law: Suppose $\{A_\alpha\}_{\alpha\in\mathcal{A}}$ are closed sets. Then

$$(\bigcap_\alpha A_\alpha)^c = \bigcup_\alpha A_\alpha^c$$

The set on the right is open. □

Definition 3.3. For a subset $Y \subseteq X$, the **interior** of Y, denoted $\mathrm{Int}(Y)$ is the union of all open sets contained in Y. It is the largest open set contained in Y. The closure of Y, denoted $\bar{Y}$ is the intersection of all closed sets in X that contain Y. It is the smallest closed set of X containing Y.

Definition 3.4. A point x is a **limit point** for a set $Y \subseteq X$ if every open neighborhood of x also contains at least one element of Y that is distinct from x.

Propostion 3.2. *Let $Y \subseteq X$. Then*

1. *$x \in \bar{Y}$ iff every open neighborhood of x in X intersects Y.*
2. *$\bar{Y} = Y \cup \{x \in X | \ x$ is a limit point of $Y\}$.*

Exercise 3.4. Prove Proposition 3.2.

Proof.
1. Suppose $x \notin \bar{Y}$. Then, $\bar{Y}^c$ is an open set containing x that does not intersect Y. Suppose there exists an open set U containing x such that $U \cap Y = \emptyset$. Then U^c is a closed set containing Y, which means it also contains $\bar{Y}$. Hence $x \notin \bar{Y}$.
2. That $Y \cup \{x \in X | \ x$ is a limit point of $Y\} \subseteq \bar{Y}$ follows from the definitions. Suppose $x \in \bar{Y}$, and assume $x \notin Y$. Since every neighborhood of x contains some element in Y, and since $x \notin Y$, x satisfies the definition of limit point.

□

Corollary 3.1. *A set Y in X is closed iff it contains all of its limit points.*

Exercise 3.5. Supppose $A, B \subseteq X$. Determine whether the following equalities hold.

1. $\overline{A \cup B} = \bar{A} \cup \bar{B}$
2. $\overline{A \cap B} = \bar{A} \cap \bar{B}$
3. $\overline{A \backslash B} = \bar{A} \backslash \bar{B}$

(Note, $A\backslash B = \{a \in A | \ a \notin B\}$.)

Example 3.1. Any set X admits the following two topologies:

- trivial topology: $\tau = \{\varnothing, X\}$
- discrete topology: $\tau = 2^X$ (power set of X)

Definition 3.5. Given two topologies τ_1 and τ_2 on a space X, if $\tau_1 \subseteq \tau_2$, we say that τ_1 is **weaker** and τ_2 is **stronger**.

Definition 3.6. A collection $\mathcal{B}$ of subsets of X is a **basis** for the topology τ if

- $\mathcal{B} \subseteq \tau$ and
- for every $U \in \tau$, there exists $W \subseteq \mathcal{B}$ such that $U = \bigcup W$.

Exercise 3.6. Let X be a set.

1. Given a topology τ on X, show that a collection $\mathcal{B} \subseteq \tau$ forms a basis for τ if for each nonempty $U \in \tau$ and any $x \in U$, there exists a $B \in \mathcal{B}$ such that $x \in B \subseteq U$.
2. Suppose τ_1 and τ_2 are topologies with bases $\mathcal{B}_1$ and $\mathcal{B}_2$ respectively. Show that τ_2 is stronger than τ_1 iff for each $x \in X$ and each $B_1 \in \mathcal{B}_1$ containing x, there exists a $B_2 \in \mathcal{B}_2$ such that $x \in B_2 \subseteq B_1$.

Definition 3.7. Without an ambient topology, we define a **basis** to be a collection $\mathcal{B}$ of subsets of X such that

1. For each $x \in X$ there exists a $B \in \mathcal{B}$ such that $x \in U$, and
2. if $B_1, B_2 \in \mathcal{B}$ and $x \in B_1 \cap B_2$, then there exists a $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

With this, we can define the **topology generated by** $\mathcal{B}$ by saying a subset $U \subseteq X$ is open if for each $x \in U$ there exists a $B \in \mathcal{B}$ with $x \in B \subseteq U$.

Example 3.2. Open intervals form a basis for the standard topology on $\mathbb{R}$. Would a proper subset of this collection suffice?

Exercise 3.7. Does the collection $\{(a, b]|\ a, b \in \mathbb{Q}, a < b\}$ form a basis for the topology generated by $\{(a, b]|\ a, b \in \mathbb{R}, a < b\}$?

Exercise 3.8. Is the topology generated by the basis $\mathcal{B} = \{(a, b]|\ a, b \in \mathbb{R}, a < b\}$ stronger or weaker than the standard topology on $\mathbb{R}$?

Example 3.3. Given a set X, the **finite complement topology** on X is defined as

$$\tau = \varnothing \cup \{U \subseteq X : U^c \text{ is a finite set}\}.$$

Definition 3.8. A G_δ set is a countable intersection of open sets. An F_σ set is the countable union of closed sets.

Exercise 3.9. Must a G_δ set be open? closed? Is the compliment of a G_δ set an F_σ set? Could a set be both a G_δ set and an F_σ set? Show that the rational numbers are an F_σ set. After we cover the Baire Category Theorem, we will be able to show that they do not form a G_δ set.

Definition 3.9. A topological space X is **Hausdorff** if for every pair of distinct points $x, y \in X$, there exist disjoint open subsets $U, V \subseteq X$ such that $x \in U$ and $y \in V$.

Exercise 3.10. Determine (with proof) whether or not the finite complement topology is Hausdorff.

Exercise 3.11. If X is Hausdorff, then singleton sets $\{x\}$ are closed.

We will, unless otherwise stated, assume our topological space is Hausdorff.

Definition 3.10. If X is a topological space and $Y \subseteq X$, then the **subspace topology** on Y is the collection

$$\tau_Y = \{U \cap Y : U \text{ is open in } X\}.$$

Exercise 3.12. Show that a subspace of a Hausdorff space is Hausdorff.

Definition 3.11. A subset $Y \subseteq X$ is **dense** in X if $\bar{Y} = X$. Y is **nowhere dense** in X if $\text{Int}(\bar{Y}) = \varnothing$.

Exercise 3.13. Show that a subspace $Y \subseteq X$, a subset $Z \subseteq Y$ is closed in Y iff $Z = Y \cap C$ for some closed subset $C \subseteq X$.

3.1 Continuous Functions on Topological Spaces

Definition 3.12. A function $f : X \to Y$ between topological spaces is **continuous** if $f^{-1}(U)$ is open in X for every open set $U \subseteq Y$.
A function $f : X \to Y$ is continuous at a point $x \in X$ if for every open neighborhood N of $f(x)$, there exists an open neighborhood U of x such that $f(U) \subseteq N$.
A function $f : X \to Y$ is called **open** if it maps open sets to open sets.
A function $f : X \to Y$ is called **closed** if it maps closed sets to closed sets.
A function $f : X \to Y$ is a **homeomorphism** if f is bijective and f is continuous and open.

Counterexamples in Analysis 12.13-18

Exercise 3.14. Are the following properties equivalent to continuity for $f : X \to Y$:

1. For every $U \subseteq f(X)$ that is open in Y, $f^{-1}(U)$ is open in X.
2. For every $U \subseteq f(X)$ that is open in the subspace topology of $f(X)$, $f^{-1}(U)$ is open in X.

Example 3.4. For any $a, b, c, d \in \mathbb{R} \cup \{\pm\infty\}$ with $a \neq b$ and $c \neq d$,

- (a, b) is homeomorphic to (c, d)
- $[a, b]$ is homeomorphic to $[c, d]$

Exercise 3.15. Show the following:

1. $f : X \to Y$ is continuous iff $f^{-1}(C)$ is closed for every closed set $C \subseteq Y$.

2. $f : X \to Y$ is continuous iff f is continuous at every point $x \in X$.

Exercise 3.16. Prove that to show that a function $f : X \to Y$ between two topological spaces is continuous if the inverse image of any basis element is open.

Exercise 3.17. Show that for any topological spaces X and Y and any function $f : X \to Y$:

1. If X is discrete (i.e. has the discrete topology), then f is continuous.

2. If Y has the trivial topology, then f is continuous.

Exercise 3.18.

1. Suppose $f : X \to Y$ is a continuous function and $A \subseteq X$. Is the restriction $f|_A : A \to Y$ (where $f|_A(a) = f(a)$ for all $a \in A$) also continuous?

2. Suppose $A, B \subseteq X$ are closed subspaces of X such that $X = A \cup B$. Show that $f : X \to Y$ is continuous if $f|_A$ and $f|_B$ are continuous.

Exercise 3.19. Give a homeomorphism mapping $(0, 1)$ to $\mathbb{R}$ (with the usual topologies). Show that this is a homeomorphism. (Let's call this a running exercise, i.e. as we cover more and more material, it may become easier to answer fully.)

Exercise 3.20. Suppose τ_1 and τ_2 are two topologies on X, and suppose τ_2 is stronger than τ_1. Show that the identity function $(X, \tau_1) \to (X, \tau_2)$ (mapping $x \mapsto x$) is continuous. Is it a homeomorphism? What is required for it to be a homeomorphism?

3.2 Sequences

Definition 3.13. Suppose X is a topological space, and let (x_n) be a sequence of points in X. A point $x_0 \in X$ is a **limit** of the sequence (x_n) any open neighborhood of x_0 contains x_n for all but finitely many n.
Let $(x_\alpha)_{\alpha \in A}$ be a net in X. A point $x_0 \in X$ is a **limit** of the net (x_α) if for any open neighborhood U of x_0 there exists an $\alpha \in A$ such that $x_\beta \in U$ for all $\beta \geq \alpha$.

We will likely not rely too much on nets, but it is a good term to see defined.

Remark 3.1. *Although some topologies may seem contrived, you would be surprised at how important pointset is to analysts (as opposed to topologists). If you are confronted with a rogue topology, one good way to develop an intuition for it is to determine what it takes for a sequence to converge. In particular, different types of convergence will lead to different topologies.*

Exercise 3.21. Suppose τ_1 and τ_2 are two topologies on X, and suppose τ_2 is stronger than τ_1. Let (x_n) be a sequence in X. If (x_n) converges with respect to τ_2, must it converge with respect to τ_1? If (x_n) converges with respect to τ_1, must it converge with respect to τ_2?

Definition 3.14. A **subsequence** (a_{k_n}) of (a_n) is a sequence of terms from (a_n) preserving their relative order.

Exercise 3.22. Show that any sequence in a Hausdorff space has at most one limit. If (a_n) is a sequence in a Hausdorff space X and the set $\{a_n\}_{n\in\mathbb{N}}$ has a limit point a, must a be a limit of the sequence (a_n)?

Propostion 3.3. *A sequence (a_n) converges to a limit a iff every subsequence also converges to a.*

3.3 Product Spaces

Recall that the Cartesian product for a family of sets $\{A_\alpha\}_{\alpha\in\mathcal{A}}$ is

$$\prod_{\alpha\in\mathcal{A}} A_\alpha = \{\phi : \mathcal{A} \to \bigcup_\alpha A_\alpha | \phi(\alpha) \in A_\alpha \ \forall \ \alpha \in \mathcal{A}\}.$$

In case $\mathcal{A}$ is not finite, we assume the axiom of choice to conclude that this product is nonempty.

Definition 3.15. The **product topology** on $\prod_{\alpha\in\mathcal{A}} A_\alpha$ has the basis

$$\mathcal{B} = \{\prod_{\alpha\in\mathcal{A}} U_\alpha | U_\alpha \subset A_\alpha \text{ open and } U_\alpha = A_\alpha \text{ for all but finitely many } \alpha\}.$$

Example 3.5.
- $\mathbb{R}^d$ and $\mathbb{C}^d$, where the product topology coincides with the standard topology.
- $\{0,1\}^{\mathbb{N}}$ where $\{0,1\}$ has the discrete topology.

Definition 3.16. For each $\beta \in \mathcal{A}$, we define the **projection map** $\pi_\beta : \prod_\alpha X_\alpha \to X_\beta$ by

$$\pi_\beta(\phi) = \phi(\beta).$$

Exercise 3.23. Show that these maps are continuous and moreover open.

Remark 3.2. *The product topology is the weakest topology for which these maps are all continuous.*

Exercise 3.24. Suppose $A \subseteq X$ and $B \subseteq Y$. Show the following

1. If A and B are closed in their respective spaces, then $A \times B$ is closed in $X \times Y$.
2. $\overline{A \times B} = \overline{A} \times \overline{B}$.
3. If X and Y are Hausdorff, then so it $X \times Y$.
4. X is Hausdorff iff the diagonal

$$\Delta = \{(x,x) | \ x \in X\} \subseteq X \times X$$

is closed.

3.4 Metric Spaces

When there is no significant difference between the two fields, we let $\mathbb{K}$ denote $\mathbb{R}$ or $\mathbb{C}$.

Definition 3.17. A **metric** on a set X is a function $\rho : X \times X \to [0, \infty)$ such that $\forall\ x, y, z \in X$

- $\rho(x, y) \geq 0$ and $\rho(x, y) = 0 \Leftrightarrow x = y$,
- $\rho(x, y) = \rho(y, x)$, and
- $\rho(x, y) + \rho(x, z) \geq \rho(y, z)$

Definition 3.18. For a set X with a metric ρ, the **ball of radius** r about a point $y \in X$ is the set $B_r(y) = \{x \in X | \rho(x, y) < r\}$

Definition 3.19. If X has a metric ρ, then the collection

$$\{B_r(x)|\ r \in (0, \infty), x \in X\}$$

forms a basis for a topology on X called the **metric topology**. With this topology X is called a **metric space**.
We say a topological space X is **metrizable** if there exists a metric that generates the topology of X.

Example 3.6. The metric given by $d(x, y) = 1$ if $x \neq y$ and $d(x, x) \neq 0$ yields the discrete topology.

Example 3.7. For $d < \infty$, the **Euclidean metric** on $\mathbb{K}^d$ is given by

$$\text{dist}(x, y) = \sqrt[2]{\sum_{k=1}^{n} |x_k - y_k|^2}$$

for $x = (x_1, ..., x_n), y = (y_1, ..., y_n) \in \mathbb{K}^d$. The topology induced by this metric agrees with the standard topology on $\mathbb{K}$.

Exercise 3.25. Suppose ρ is a metric on X. Show that the function

$$\hat{\rho}(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)}$$

for all $x, y \in X$ is a metric on X such that $\hat{\rho}(x, y) \leq 1$ for all $x, y \in X$.

Definition 3.20. Let (X, ρ) be a metric space and $A \subseteq X$. The **diameter** of A is

$$\text{diam}(A) = \sup_{x,y \in A} \rho(x, y).$$

We say A is **bounded** if $\text{diam}(A) < \infty$.

Exercise 3.26. Show that a sequence (x_n) in a metric space (X, ρ) converges to $x \in X$ iff for any $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $\rho(x, x_n) < \epsilon$ for all $n \geq N$.

Exercise 3.27. A function $f : (X, \rho_X) \to (Y, \rho_Y)$ between metric spaces is continuous at a point $x_0 \in X$ iff for every $\epsilon > 0$, there exists a $\delta > 0$ such that for all $x \in X$,

$$\rho_X(x, x_0) < \delta \Rightarrow \rho_Y(f(x), f(x_0)) < \epsilon.$$

Definition 3.21. A topological space is **first countable** if ever point has a countable neighborhood basis. (A **neighborhood basis** of $x \in X$ is a collection $\mathcal{B}_x$ of open neighborhoods of X so that any open neighborhood of x contains an element of $\mathcal{B}_x$.)
A topological space is **second countable** if it has a countable basis.
A topological space is **separable** if it has a countable, dense subset.

Exercise 3.28. If (X, ρ) is a metric space, then we have the following

- X is first countable.
- X is second countable iff it is separable. (Which direction requires that X is a metric space?)

Exercise 3.29. Give an example of a metric space that is not second countable.

Propostion 3.4. *Suppose (X, ρ) is a metric space, $A \subseteq X$, and $x \in X$. Then $x \in \bar{A}$ iff there exists a sequence of points (a_n) in A converging to x.*

Exercise 3.30. A function $f : (X, \rho_X) \to (Y, \rho_Y)$ between metric spaces is continuous iff for any convergent sequence (x_n) in X with limit x, the sequence $(f(x_n))$ converges in Y to $f(x)$.

Remark 3.3. *In fact, these hold in any first countable space. If a space is not first countable, then nets will serve as a replacement for sequences.*

Definition 3.22. A map $f : (X, \rho_X) \to (Y, \rho_Y)$ between metric spaces is called an **isometry** if $\rho_X(x, y) = \rho_Y(f(x), f(y))$ for all $x, y \in X$.

Exercise 3.31. Show that an isometry is injective.

3.4.1 Cauchy Sequences

Definition 3.23. A sequence (x_n) in a metric space (X, ρ) is *Cauchy* if for every $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \geq N$, $\rho(x_n, x_m) < \epsilon$.

Exercise 3.32. Every convergent sequence is Cauchy.
A Cauchy sequence is convergent iff it has a convergent subsequence.

Definition 3.24. A metric space is *complete* if every Cauchy sequence converges.

Exercise 3.33. Show that $\mathbb{R}$ is complete with respect to the Euclidean metric. Show that $(0, 1)$ is not complete with respect to the standard metric. Is there a metric on $\mathbb{R}$ for which $\mathbb{R}$ is not complete?

Exercise 3.34. Prove Cantor's Intersection Theorem for Complete Metric spaces, i.e.

Theorem 3.2. *Suppose (X, ρ) is a metric space. Then (X, ρ) is complete iff for every nested sequence of bounded nonempty closed subsets $C_1 \supseteq C_2 \supseteq \ldots$ with diam$(C_n) \to 0$, their intersection is a single point.*

3.5 Connected Spaces, Compact Spaces

3.5.1 Connectedness and the Intermediate Value Theorem

Definition 3.25. A topological space X is **connected** if there is no decomposition $X = Y \cup Z$ into two nonempty disjoint open sets.

Exercise 3.35. Prove the following:

1. A topological space X is connected iff the only sets that are both closed and open are X and $\varnothing$.

2. A topological space X is connected iff for any topological space Y with the discrete topology, every continuous function $f : X \to Y$ is constant.

3. Let $\{X_\alpha\}_{\alpha \in \mathcal{A}}$ be a family of connected subspaces of a topological space X such that $X_\alpha \cap X_\beta \neq \emptyset$ for all $\alpha, \beta \in \mathcal{A}$. Then

$$\bigcup_{\alpha \in \mathcal{A}} X_\alpha$$

   is connected.

Theorem 3.3. *A nonempty interval $I \subseteq \mathbb{R}$ is connected.*

Proof. Suppose A and B are nonempty disjoint open sets so that $I = A \cup B$. Suppose without loss of generality, there exits $a, b \in \text{Int}(I)$ with $a \in A$, $b \in B$ and $a < b$. Let $c = sup\{x \in A : x < b\}$. Then $c \in \text{Int}(I) \subseteq A \cup B$, but since A and B is open, neither can contain c. □

Exercise 3.36. If $Y \subseteq \mathbb{R}$ is connected, then Y is an interval. (This is an "if and only if" by the Theorem 3.3.)

Exercise 3.37. Prove the following:

1. The image of a connected set under a continuous function is connected. In fact, connectedness is a **topological invariant**.

2. Let $I \subseteq \mathbb{R}$ be an interval, and $f : I \to \mathbb{R}$ a continuous function. Then $f(I)$ is an interval.

3. Prove the Intermediate Value Theorem from calculus:

   Theorem 3.4 (Intermediate Value Theorem)**.** *Let $a, b \in \mathbb{R}$ with $a < b$, and assume $f : [a, b] \to \mathbb{R}$ is continuous. Then there exists a $d \in \mathbb{R}$ such that d lies between $f(a)$ and $f(b)$.*

4. Suppose $f : [a, b] \to [a, b]$ is continuous. Then f has a fixed point, i.e. a point $x \in [a, b]$ for which $f(x) = x$.

Exercise 3.38. A space X is **path connected** if for any points $x, y \in X$, there exists a continuous function $f : [0, 1] \to X$ such that $f(0) = x$ and $f(1) = y$. Is the topologist's sine curve

$$X = \{(x, \sin(1/x))| x \in (0, 1]\} \cup \{(0, y)|\ y \in [-1, 1]\}$$

is connected? path connected?

3.5.2 Compactness and the Extreme Value Theorem

Definition 3.26. A space X is **compact** if every open cover of X has a finite subcover.

Exercise 3.39. Prove the following:

1. A closed subset of a compact set is compact.
2. Suppose X is compact and Y is Hausdorff. Then a continuous bijection $f : X \to Y$ is a homeomorphism.

Propostion 3.5. *The interval* $[0, 1] \subseteq \mathbb{R}$ *is compact.*

Proof. Let $\mathcal{U}$ be an open cover of $[0, 1]$, and define a set

$$I = \{s \in [0, 1]|\ [0, s] \text{ is covered by finitely many members of } \mathcal{U}\}.$$

Then $I \neq \varnothing$ since $0 \in I$. Further, since I is bounded above by 1, I has a supremum, $b \leq 1$. Then, I is an interval of the form $[0, b)$ or $[0, b]$. We now need only to show that $b \in I$ and that $b = 1$. To see that $b \in I$, suppose not. Let $U \in \mathcal{U}$ such that $b \in U$. Then, there exists an $a \in I$ so that $[a, b] \subset U$. But $a \in I$ means $[0, a]$ is covered by finitely many members of $\mathcal{U}$, and so $[0, b] = [0, a] \cup [a, b]$ is covered by finitely many members of $\mathcal{U}$. Now, suppose $b \neq 1$, and let $V \in \mathcal{U}$ such that $b \in V$. Then, there exist a $c \in (b, 1)$ such that $[b, c] \subset V$. Then V along with the finite cover of $[0, b]$ covers $[0, c]$, contradicting b being the supremum of I. ☐

Exercise 3.40. Prove the following:

1. Suppose X is Hausdorff and $Y \subseteq X$ is compact. Then Y is closed.
2. Show that a compact subset of a metric space is bounded.

3. The continuous image of a compact set is compact. (Compactness is also a topological invariant.)

4. Prove the Extreme Value Theorem from calculus

   Theorem 3.5 (Extreme Value Theorem)**.** *Suppose $K \subseteq \mathbb{R}$ is a compact subset and $f : K \to \mathbb{R}$ is continuous. Then there exist $a, b \in K$ such that $f(a) = \max_{k \in K} f(k)$ and $f(b) = \min_{k \in K} f(k)$.*

Theorem 3.6 (Tychonoff)**.** *If each X_α is compact, then so is $\prod_{\alpha \in \mathcal{A}} X_\alpha$.*

Remark 3.4. *This is equivalent to the Choice Axiom.*

Definition 3.27. A map $f : X \to Y$ is **proper** if $f^{-1}(K)$ is compact for every compact set $K \subseteq Y$.

Propostion 3.6. *Compact Hausdorff spaces are* **normal***, i.e. for any two closed subsets $A, B \subseteq X$, there exist disjoint open sets $U, V \subseteq X$ such that $A \subseteq U$ and $B \subseteq V$.*

Exercise 3.41. That the intervals $(0, 1)$ and $[0, 1]$ (with the standard subspace topologies) are not homeomorphic follows quickly from a compactness argument. How would you prove this using a connectedness argument?

3.5.3 Heine-Borel

Propostion 3.7. *If X is compact, then the projection $\pi_Y : X \times Y \to Y$ is closed for any Y.*

Proof. Let $C \subseteq X \times Y$ and $y_0 \in Y \backslash \pi_Y(C)$. Then for any $x \in X$, $(x, y_0) \notin C$. Then there exists an open neighborhood of (x, y_0) missing C and hence open sets $U_x \subseteq X$ and $V_x \subseteq Y$ with $x \in U_x$ and $y_0 \in V_x$ such that $U_x \times V_x \cap C = \varnothing$. Then $\{U_x | x \in X\}$ forms an open cover of X. Let $\{U_{x_1}, ... U_{x_n}\}$ be a finite subcover. Then $\bigcap_{j=1}^n V_{x_j} = V$ is an open neighborhood of y_0. Moreover, $X \times V \cap C = \varnothing$. Then $y_0 \in V \subseteq Y \backslash \pi_Y(C)$. Hence $\pi_Y(C)$ is closed. □

Exercise 3.42. Let's prove the Heine-Borel Theorem using the theory of product spaces: i.e.

Theorem 3.7 (Heine-Borel)**.** *For $d < \infty$, a subset of $\mathbb{R}^d$ is compact iff it is closed and bounded.*

By Exercise 3.40, you may assume the sufficient implication is done. For the necessary direction, first show the following corollaries to Proposition 3.7 (assume all spaces are Hausdorff):

1. If X is compact, then $\pi_Y : X \times Y \to Y$ is proper.

2. If X and Y are compact, then $X \times Y$ is compact. (Use part 1, not Tychonoff.) Conclude that if $X_1, ..., X_d$ are compact, then $\prod_{i=1}^{d} X_i$ is compact.

3. For any $d < \infty$, the unit cube $[0,1]^d \subseteq \mathbb{R}^d$ is compact. Conclude that $[-n,n]^d$ is compact for any $n \in \mathbb{N}$.

Now, try to prove the rest.

Remark 3.5. *This theorem does not hold in general metric spaces. A metric space has the* **Heine-Borel Property** *if the Heine-Borel theorem holds on that space. It is worth noting that, a normed vector space satisfies the Heine-Borel property iff it is finite-dimensional. More on this later.*

3.5.4 Bolzano-Weierstrass

Definition 3.28. A space is *sequentially compact* if every sequence has a convergent subsequence.

Definition 3.29. A metric space is *totally bounded* if for any $\epsilon > 0$, there exists a finite cover of X by balls of radius ϵ.

Theorem 3.8. *Let (X, ρ) be a metric space. The following are equivalent:*

1. *X is compact.*

2. *X is sequentially compact.*

3. *X is complete and totally bounded.*

Remark 3.6. *This is an extension of the Heine-Borel Theorem to general metric spaces.*

Exercise 3.43. Prove the Bolzano-Weierstrass Theorem for $\mathbb{R}^d$ ($d < \infty$) along with Corollary 3.2 and Cauchy's Convergence Criterion.

Theorem 3.9 (Bolzano-Weierstrass)**.** *Every bounded infinite set in $\mathbb{R}^d$ has a limit point.*

Corollary 3.2. *Every bounded sequence (i.e. $\exists\ M > 0$ such that $|x_n| < M$ for all n) in $\mathbb{R}^n$ has a convergent subsequence.*

Corollary 3.3 (Cauchy's Convergence Criterion)**.** *A sequence in $\mathbb{R}^n$ is convergent iff it is a Cauchy sequence.*

Recall that A sequence (x_n) in $\mathbb{R}$ is **monotone** if $a_n \geq a_{n+1}$ for all n or $a_n \leq a_{n+1}$ for all n.

Theorem 3.10. *If (a_n) is a monotone sequence in $\mathbb{R}$, then (a_n) is convergent iff it is bounded.*

3.5.5 The Cantor Set

Cantor's Intersection theorem sometimes appears in the following form:

Theorem 3.11. *Let $K_1 \supset K_2 \supset K_3 \supset ...$ be a nested sequence of nonempty compact sets. Then*

$$\bigcap_n K_n \neq \varnothing.$$

Example 3.8 (The Cantor Set)**.** For each $n \in \mathbb{N}$, define $C_n \subseteq [0,1]$ as follows:

$$C_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}\right] \tag{1}$$

$$C_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right] \tag{2}$$

$$... \tag{3}$$

$$\tag{4}$$

In general, C_n consists of the disjoint union of 2^n closed intervals of the form $[\frac{k}{3^n}, \frac{k+1}{3^n}]$.
The cantor set $\mathcal{C}$ is the intersection

$$\bigcap_n C_n.$$

By Cantor's theorem, we know $\mathcal{C}$ is nonempty. This will prove to be an interesting example, especially in measure theory.

Exercise 3.44. Prove that the Cantor set has the following properties:

1. $\mathcal{C}$ is compact.
2. $\mathcal{C}$ contains no non-empty open interval.
3. $\mathcal{C}$ is nowhere dense.

Exercise 3.45. Prove that the Cantor set $\mathcal{C}$ is homeomorphic to the space $\{0,1\}^{\mathbb{N}}$ where each copy of $\{0,1\}$ carries the discrete topology. (Cantor's intersection theorem will be helpful.)

4 Real Analysis Toolkit

4.1 A few more continuity theorems from pointset

Theorem 4.1 (Brouwer Fixed Point Theorem)**.** *Let $r > 0$ and $B_r(0) \subseteq \mathbb{R}^d$. Then any continuous $f : \overline{B_r(0)} \to \overline{B_r(0)}$ has at least one fixed point.*

Remark 4.1. *There exists a version for convex compact sets.*

Theorem 4.2. *$\mathbb{K}^d$ is a complete metric space with respect to the Euclidean norm.*

We give now Uryshon's Lemma, which essentially says that if X is normal, then it has "plenty" of continuous functions into $\mathbb{R}$.

Theorem 4.3 (Uryshon's Lemma)**.** *Suppose X is a normal space and $A, B \subseteq X$ are closed subsets. Then there exists a function $f : X \to [0, 1]$ such that $f(A) \equiv 0$ and $f(B) \equiv 1$, i.e. A and B can be separated by a continuous function.*

This is equivalent (and usually used to prove) Tietze's Extension Theorem.

Theorem 4.4 (Tietze Extension Theorem)**.** *Suppose X is a normal space, $A \subseteq X$ closed, and $f : A \to \mathbb{R}$ continuous. Then there exists $g : X \to \mathbb{R}$ continuous such that $g|_A = f$ and*

$$\sup_{x \in X} |g(x)| = \sup_{a \in A} |f(a)|.$$

4.2 Miscellaneous Topics

4.2.1 limsups and liminfs

Unfortunately, not every bounded sequence will have a limit. For that reason, in analysis, we will often make use of limit supremum and limit infimum of sequences.

Definition 4.1. Let (x_n) be a (bounded) sequence of real numbers. The **limit superior** of the sequence is

$$\limsup_n x_n = \overline{\lim}_n x_n = \inf\{y \in \mathbb{R} | y < x_n \text{ for at most finitely many } n\},$$

and the **limit inferior** is

$$\liminf_n x_n = \underline{\lim}_n x_n = \sup\{y \in \mathbb{R} | y > x_n \text{ for at most finitely many } n\}.$$

Remark 4.2. *A bounded sequence need not have a limit, but it will always have a unique limsup (and a unique liminf). This follows from the fact that*

$$\sup\{x_n : n \in \mathbb{N}\} \in \{y \in \mathbb{R} | y < x_n \textit{ for at most finitely many } n\}$$

and the fact that

$$\inf\{x_n : n \in \mathbb{N}\} - 1$$

is a lower bound for

$$\{y \in \mathbb{R} | y < x_n \textit{ for at most finitely many } n\}$$

.

Exercise 4.1. Suppose (x_n) is a bounded sequence in $\mathbb{R}$. Show that the following are equivalent for $x \in \mathbb{R}$:

1. $x = \limsup x_n$
2. If $\epsilon > 0$, there are at most a finite number of $n \in \mathbb{N}$ such that $x + \epsilon < x_n$ but there are an infinite number of n such that $x - \epsilon < x_n$.
3. If $v_m = \sup\{x_m : n \geq m\}$, then $x = \inf\{v_m : m \in \mathbb{N}\}$.
4. If $v_m = \sup\{x_m : n \geq m\}$, then $x = \lim_m v_m$.
5. If L is the set of $v \in R$ such that there exists a subsequence of (x_n) which converges to v, then $x = \sup L$.

Proof. Suppose (x_n) is as above, and let $x \in \mathbb{R}$. (Note that L is nonempty by Bolzano-Weierstrass.)

(1) ⇒ (2) Let $\epsilon > 0$, and let $X = \inf\{y \in \mathbb{R} | y < x_n$ for at most finitely many $n\}$. Then, by definition, there exists $y \in X$ such that $y \in [x, x + \epsilon]$. Since $y \in X$, it follows that $x + \epsilon \in X$. On the other hand, $x - \epsilon \notin X$ since $x = \inf X$, but that meas that there are an infinite number of x_n such that $x - \epsilon < x_n$.

(2) ⇒ (3) Let $\epsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $x_n \leq x + \epsilon$ for all $n \geq N$. Hence, $v_N \leq x + \epsilon$, and so $\inf\{v_m : m \in \mathbb{N}\} \leq x^* + \epsilon$. On the other hand, since there is an infinite number of $n \in \mathbb{N}$ for which $x - \epsilon < x_n$, $x - \epsilon < v_m$ for all $m \in \mathbb{N}$, and so $\inf\{v_m : m \in \mathbb{N}\} \geq x - \epsilon$. Since $\epsilon > 0$ was arbitrary, we are done.

(3) ⇒ (4) Notice that the sequence (v_m) is monotone decreasing. Hence, its infimum is its limit.

(4) ⇒ (5) Assume $x = \lim v_m$, and suppose (x_{n_k}) is a subsequence of (x_n) with limit v. Note that $n_k \geq k$ for each $k \in \mathbb{N}$, and so $x_{n_k} \leq v_k$, which means $v \leq x$. So, x is an upper bound for L. To see that it is the least upper bound, we show that it is actually the limit of a subsequence of (x_n). Choose $n_1 \in \mathbb{N}$ such that $v_1 - 1 < x_{n_1} \leq v_1$. Choose $n_2 > n_1$ such that $v_2 - \frac{1}{2} < x_{n_2} \leq v_2$. Inductively, choose $n_k > n_{k-1}$ such that $v_k - \frac{1}{k} < x_{n_k} \leq v_k$. Then, for any $\epsilon > 0$, we can choose k sufficiently large such that $|x_{n_k} - v_k| \leq \epsilon/2$ and $|v_k - x| < \epsilon/2$. Using the triangle inequality, it follows that $x = \lim_k x_{n_k}$.

(5) ⇒ (1) Let $X = \inf\{y \in \mathbb{R} | y < x_n$ for at most finitely many $n\}$, and let $x = \sup L$. If for some $\epsilon > 0$ there exists infinitely many $n \in \mathbb{N}$ for which $x_n > x + \epsilon$, then we would have a bounded sequence (x_{n_k}) with each $x_{n_k} > \epsilon$. By Bolzano-Weierstrass, there would exist a convergent subsequence $(x_{n_{k_j}})$ whose limit is strictly greater than x, which contradicts $x = \sup L$. Hence, for any $\epsilon > 0$, there exists at most finitely many $n \in \mathbb{N}$ with $x_n > x + \epsilon$. Then, $x + \epsilon \in X$, and so $\limsup x_n \leq x$.

On the other hand, since there exists a subsequence of (x_n) converging to some $v > x - \epsilon$ for any $\epsilon > 0$, $x - \epsilon \notin X$ for any $\epsilon > 0$. Hence, $x - \epsilon \leq \limsup x_n$ for all $\epsilon > 0$. Since $\epsilon > 0$ was arbitrary, we are done.

□

Theorem 4.5. *Let* (x_n) *and* (y_n) *be bounded sequences of real numbers and* $c \in (0, \infty)$. *Then we have the following:*

1. $\liminf x_n \leq \limsup x_n$
2. $\liminf(cx_n) = c \liminf x_n$ *and* $\limsup(cx_n) = c \limsup x_n$
3. $\liminf(-cx_n) = -c \limsup x_n$ *and* $\limsup(-cx_n) = -c \liminf x_n$
4. $\liminf x_n + \liminf y_n \leq \liminf(x_n + y_n)$
5. $\limsup x_n + \limsup y_n \geq \limsup(x_n + y_n)$
6. *If* $x_n \leq y_n$ *for all* n, *then* $\liminf x_n \leq \liminf y_n$ *and* $\limsup x_n \leq \limsup y_n$.
7. (x_n) *is convergent iff* $\limsup x_n = \liminf x_n$, *in which case* $\lim x_n$ *is the common value.*

Definition 4.2. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a bounded function, and $a \in \mathbb{R}$.

$$\limsup_{x \to a} f(x) = \inf_{\epsilon > 0} \sup_{x \in (a-\epsilon, a+\epsilon) \setminus \{a\}} f(x)$$
$$\lim_{\epsilon \to 0} \sup_{x \in (a-\epsilon, a+\epsilon) \setminus \{a\}} f(x)$$
$$Limsup_{x \to a} f(x) = \inf_{\epsilon > 0} \sup_{x \in (a-\epsilon, a+\epsilon)} f(x)$$
$$\lim_{\epsilon \to 0} \sup_{x \in (a-\epsilon, a+\epsilon)} f(x)$$

$$\begin{aligned} \limsup_{x \to \infty} f(x) &= \inf_{y \in \mathbb{R}} \sup_{x \geq y} f(x) \\ &= \lim_{y \to \infty} \sup_{x \geq y} f(x) \end{aligned}$$

Remark 4.3. *These can also be defined for* $f : D \to \mathbb{R}$ *where* $D \subseteq \mathbb{R}$ *by instead taking* $\sup\{f(x) | \ x \in D \cap B_\epsilon(a) \setminus \{a\}\}$.

As before, we have the following proposition for a function $f : \mathbb{R} \to \mathbb{R}$.

Propostion 4.1. *If* $M > \limsup_{x \to a} f(x)$, *then there exists an* $\epsilon > 0$ *such that* $f(x) < M$ *for all* $x \in (a - \epsilon, a + \epsilon) \setminus a$.

If we define

$$L = \{\lim f(x_n) : x_n \to a, \lim f(x_n) \text{ exists}\}$$

, *then* $\limsup_{x \to a} f(x) = \sup L$.

We have analogous claims for $a = \infty$.

If $M > \limsup_{x \to \infty} f(x)$, then there exists $N \in \mathbb{R}$ such that $f(x) < M$ for all $x \geq N$.

4.2.2 Landau's Symbols

Suppose $D \subseteq \mathbb{R}$ and $f, g : D \to \mathbb{R}$.

Definition 4.3. We say $f(x) = O(g(x))$ ("Big Oh") as $x \to a$ for some $a \in \mathbb{R}$ if there exists an $M > 0$ and $\delta > 0$ such that for $x \in (a - \delta, a + \delta)$,

$$|f(x)| \leq M|g(x)|.$$

We say $f(x) = O(g(x))$ as $x \to \infty$ if there exists an $M > 0$ and $x_0 \in \mathbb{R}$ such that for all $x \geq x_0$,

$$|f(x)| \leq M|g(x)|.$$

Remark 4.4. *If unstated, one usually assumes $a = \infty$.*

Example 4.1.

1. Obviously $x^2 = O(2x^2)$.
2. For a degree n polynomial $p(x) = \sum_{k=0}^{n} a_k x^m$, we have that $p(x) = O(x^n)$ as $x \to \infty$.
3. To say $f(x) = O(1)$ as $x \to a$ is to say that f is bounded near a.
4. Possibly the most common use is to describe the growth rate of the tail of a series of functions, e.g.

$$e^x = 1 + x + \frac{1}{2}x^2 + O(x^3)$$

as $x \to 0$ or

$$\sin(x) = x - \frac{1}{6}x^3 + O(x^5)$$

as $x \to 0$.

Exercise 4.2. Suppose $g \neq 0$ for values of x sufficiently close to a (or sufficiently large if $a = \infty$. Show that $f(x) = O(g(x))$ as $x \to a$ iff

$$\limsup_{x \to a} \frac{|f(x)|}{|g(x)|} < \infty.$$

Remark 4.5. *We think of Landau notation as giving a rough description of the "order" of one function relative to another.*

$O(1)$	constant
$O(\log(x))$	logarithmic
$O(x)$	linear
$O(x^2)$	quadratic
$O(x^n)$	polynomial
$O(a^x)$	exponential

Definition 4.4. We say $f(x) = o(g(x))$ ("Little oh") as $x \to a$ for some $a \in \mathbb{R}$ if for any $\epsilon > 0$ there exists a $\delta > 0$ such that for $x \in (a - \delta, a + \delta)$,

$$|f(x)| \leq \epsilon |g(x)|.$$

We say $f(x) = o(g(x))$ as $x \to \infty$ if for any $\epsilon > 0$ there exists an $x_0 \in \mathbb{R}$ such that for $x \geq x_0$,

$$|f(x)| \leq \epsilon |g(x)|.$$

Remark 4.6. *Similarly, if $g(x) \neq 0$ for x close to a, then we say $f(x) = o(g(x))$ iff*

$$\lim_{x \to a} \frac{|f(x)|}{|g(x)|} = 0.$$

4.2.3 Oscillation

We will want this notion in the integration section, but we will define it here so you will know it when you see it later.

Definition 4.5. Let $I \subseteq \mathbb{R}$ be a bounded interval and $f : I \to \mathbb{R}$. The **oscillation of** f on I is

$$\omega_f(I) = \sup_{x,y \in I} |f(x) - f(y)|.$$

For $x \in I$, the **oscillation of** f **at** x is

$$\omega_f(x) = \inf_{B_\delta(x) \subset I} \omega_f(B_\delta(x)).$$

Exercise 4.3. For any $f : I \to \mathbb{R}$, check that the following hold:

1. $\omega_f(I) = \sup_{x \in I} f(x) - \inf_{y \in I} f(y)$.
2. $\omega_f(x) = \mathrm{Limsup}_{y \to x} f(y) - \mathrm{Liminf}_{y \to x} f(y)$.

4.2.4 Semi-Continuity

Suppose $f : \mathbb{R} \to \mathbb{R}$.

Definition 4.6. For $\alpha \in \mathbb{R}$, the α-sublevel set of f is $\{f^{-1}((-\infty, \alpha])\}$.

Definition 4.7. $f : \mathbb{R} \to \mathbb{R}$ is lower semi-continuous if for each $\alpha \in \mathbb{R}$, the α-sublevel set $\{f^{-1}((-\infty, \alpha])\}$ is closed. f is upper semi-continuous if $\{f^{-1}([\alpha, \infty))\}$ is closed for each $\alpha \in \mathbb{R}$.

Propostion 4.2. *A function $f : \mathbb{R} \to \mathbb{R}$ is lower semi-continuous at x_0 if*

$$\liminf_{x \to x_0} f(x) \geq f(x_0)$$

(equivalently if Liminf$_{x \to x_0} f(x) = f(x_0)$). f is upper semi-continuous at x_0 if

$$\limsup_{x \to x_0} f(x) \leq f(x_0)$$

(similarly for Liminf).

Exercise 4.4. Show that a continuous function is both upper and lower semi-continuous.

In essence, if f is continuous at x_0, then $f(x_0)$ is close to $f(x)$ for all x near x_0. If f is lower semi-continuous near x_0, then $f(x_0) \leq f(x)$ for all x near x_0, and if f is upper semi-continuous at x_0, then $f(x_0) \geq f(x)$ for all x near x_0.

Semi-continuity allows you to define local maximums/minimums without continuity.

Remark 4.7. *Lower semi-continuity is the requirement for rate functions for large deviations principles.*

Exercise 4.5. Suppose $I = [a, b] \subseteq \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$. Show that the oscillation function $\omega_f(x)$ is upper semi-continuous.

4.3 Continuity on $\mathbb{R}^d$

First, let's recall some characterizations for continuity of a function from $\mathbb{K}^d$ to $\mathbb{K}^m$ for $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$:

Theorem 4.6. *A function $f : D \subseteq \mathbb{K}^d \to \mathbb{K}^m$, is continuous iff*

1. *f is continuous at every point $x \in D$, i.e. for every open neighborhood U of $f(x)$, there exists an open neighborhood V of x such that $f(V) \subseteq U$.*

2. *f is continuous at every point $x \in D$, i.e. for every $\epsilon > 0$ there exists a $\delta > 0$ such that $f(B_\delta(x)) \subseteq B_\epsilon(f(x))$.*

3. *If (x_n) is any sequence in D that converges to x, then the sequence $(f(x_n))$ converges to $f(x)$.*

Exercise 4.6. Show that $f(x) = e^x$ is continuous.

Definition 4.8. Suppose $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^m$. For $a \in D$, we say $\lim_{x\to a} f(x) = b$ if for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\|x - a\| < \epsilon \Rightarrow \|f(x) - b\| < epsilon.$$

For $f : D \subseteq \mathbb{R} \to \mathbb{R}$, we say $\lim_{x\to a^+} f(x) = b$ if for any $\epsilon > 0$ there exists a $\delta > 0$ such that for any $x > a$, if $|x - a| < \delta$, then $|f(x) - b| < \epsilon$. We define $\lim_{x\to a^-} f(x) = b$ in a similar fashion.

Remark 4.8. *You may recognize an application of this third criteria from calculus, e.g. for any $a \in [-\infty, \infty]$,*

$$\lim_{x\to a} e^x = e^{\lim_{x\to a} x}.$$

Exercise 4.7. (You can give this to your calculus students to fight their "don't pick up your pencil" intuition for continuity.) Define a **Dirichlet function** $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & : x \in \mathbb{Q} \\ x & : x \notin \mathbb{Q} \end{cases}$$

Prove that f is continuous at exactly one point. (Which point?)

Theorem 4.7. *Suppose $f, g : D \subseteq \mathbb{R}^d \to \mathbb{R}$ are continuous at a point x, then so are $|f|, f+g, f-g, f \cdot g$, and αf for some $\alpha \in \mathbb{R}$. If $g(x) \neq 0$, then f/g is continuous at x.*
If h is continous at $f(x)$, then $h \circ f$ is continuous at x.

Corollary 4.1. *If $f, g : D \subseteq \mathbb{K}^d \to \mathbb{R}$ are continuous, then so are $f \vee g$ and $f \wedge g$ where*

$$(f \vee g)(x) = \max\{f(x), g(x)\} \qquad \textit{and} \qquad (f \wedge g)(x) = \min\{f(x), g(x)\}$$

for all $x \in D$.

Proof. We can write $(f \vee g)(x) = \frac{1}{2}(f(x)+g(x)+|f(x)-g(x)|)$ and $(f \wedge g)(x) = \frac{1}{2}(f(x)+g(x)-|f(x)-g(x)|)$. □

4.3.1 Types of Continuity

Definition 4.9. In general we say a function $f : (X, \rho_X) \to (Y, \rho_Y)$ between metric spaces is **uniformly continuous on** $X_0 \subseteq X$ if for any $\epsilon > 0$ there exists a $\delta > 0$ such that $\rho_Y(f(x), f(y))$ for all $x, y \in X_0$ with $\rho_X(x, y) < \delta$. We say f is **uniformly continuous** if it is uniformly continuous on X.

Exercise 4.8. Give an example of a subset $D \subseteq \mathbb{R}$ and a function $f : D \to \mathbb{R}$ that is continuous but not uniformly continuous. Give an example on a bounded domain and an example on an unbounded domain.

Exercise 4.9. Show that if $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^m$ is not uniformly continuous, then there exists an $\epsilon > 0$ and two sequences (x_n), $(y_n) \in D$ such that if $n \in \mathbb{N}$, then

$$\|x_n - y_n\| \leq \frac{1}{n} \qquad \text{and} \qquad \|f(x_n) - f(y_n)\| > \epsilon.$$

Theorem 4.8. *Let $D \subseteq \mathbb{R}^d$ be a compact subset and $f : D \to \mathbb{R}^m$ continuous. If $K \subseteq D$ is compact, then f is uniformly continuous on K.*

Proof. Suppose f is not uniformly continuous on K. Then there exists an $\epsilon > 0$ and two sequences (x_n), $(y_n) \in K$ such that if $n \in \mathbb{N}$, then

$$\|x_n - y_n\| \leq \frac{1}{n} \qquad \text{and} \qquad \|f(x_n) - f(y_n)\| > \epsilon. \tag{5}$$

Since K is compact, both sequences are bounded. By Bolzano-Weierstrass, there exists a subsequences (x_{n_k}) and (y_{n_k}) each with the same limit $z \in K$. Since f is continuous,

$$\lim_{k\to\infty} f(x_{n_k}) = f(z) = \lim_{k\to\infty} f(y_{n_k}).$$

However, this contradicts (5). □

Definition 4.10. Let $f : D \subseteq \mathbb{R}^d \to \mathbb{R}^m$. We say f is **Lipschitz continuous** if there exists a constant $C > 0$ such that

$$\|f(x) - f(y)\| \leq C\|x - y\|$$

for all $x, y \in D$. If C can be chosen to be no more than 1, then we call f a **contraction**.

Definition 4.11. A function $f : [a,b] \to \mathbb{R}$ is **absolutely continuous** if for every $\epsilon > 0$ there exists a $\delta > 0$ such that for any finite set of disjoint intervals $(a_1, b_1), (a_2, b_2), ..., (a_N, b_N) \subset [a,b]$,

$$\sum_{n=1}^{N}(b_n - a_n) < \delta \Longrightarrow \sum_{n=1}^{N} |f(b_n) - f(a_n)| < \epsilon.$$

Exercise 4.10. Rank uniform, Lipschitz, and absolute continuity from strongest to weakest for a function $f : [a,b] \to \mathbb{R}$. (Counterexamples are not necessary.)

The following is an example you will likely see in a measure theory course. Here it will provide an example of a function that is uniformly but not absolutely continuous.

Example 4.2 (The Cantor-Lebesgue Function)**.** For each $k \in \mathbb{N}$, let $\mathcal{O}_k$ be the union of the $2^k - 1$ intervals which have been removed during the first k stages of the Cantor deletion process, i.e. $C_k = [0,1] \backslash \mathcal{O}_k$. Define $\mathcal{O} = \bigcup_{k=1}^{\infty} \mathcal{O}_k$. Then $\mathcal{C} = [0,1] \backslash \mathcal{O}$. We will define ϕ on $\mathcal{O}$ and then on $\mathcal{C}$.

Fix $k \in \mathbb{N}$. Define ϕ on $\mathcal{O}_k$ to be the increasing piecewise linear function on $\mathcal{O}_k$ that is constant on each of its $2^k - 1$ open intervals and takes the $2^k - 1$ values $\{\frac{j}{2^k}\}_{j=1}^{2^k-1}$.

Extend ϕ to all of $[0,1]$ by defining it on $\mathcal{C}$ as $\phi(0) = 0$ and for $x \in \mathcal{C} \backslash \{0\}$,

$$\phi(x) = \sup_{t \in \mathcal{O} \cap [0,x)} \phi(t).$$

To see that ϕ is not absolutely continuous, pick $\epsilon < 1$. Then, for any $\delta > 0$, there exists a finite collection of open intervals $(a_1, b_n), ..., (a_N, b_N)$ that cover $\mathcal{C}$ with $\sum_1^N |b_k - a_k| < \delta$, but $\sum_1^N |f(b_k) - f(a_k)| = 1$.

4.3.2 Discontinuities and Monotone Functions

Theorem 4.9. *A monotone function $f : (a,b) \subseteq \mathbb{R} \to \mathbb{R}$ has at most countably many discontinuities.*

Proof. Assume f is nondecreasing. First, we show that $\lim_{t \to x^{\pm}} f(t)$ exists for every $x \in (a,b)$. In particular, for all $x \in (a,b)$

$$\sup_{a<t<x} f(t) = \lim_{t \to x^-} f(t) \le f(x) \le \lim_{t \to x^+} f(x) = \inf_{x<t<b} f(x).$$

Since f is nondecreasing, the set $\{f(t) | t \in (a,x)\}$ is bounded above by $f(x)$ and hence has a supremum, call it M, and the set $\{f(t) | t \in (x,b)\}$ is bounded below by $f(x)$ and hence has an infimum. It remains to show that

$$\sup_{a<t<x} f(t) = \lim_{t \to x^-} f(t)$$

(a similar argument will show the rest). Let $\epsilon > 0$ and $\delta > 0$ so that if $x - \delta \in (a, x)$, then $f(x - \delta) \in (M - \epsilon, M]$ (such a δ exists by the fact that M is a supremum). Since f is nondecreasing, $f(x-\delta) \leq f(t) \leq M$ for all $x \in (x-\delta, x)$, and so $|f(t) - M| < \epsilon$ for all $t \in (x - \delta, x)$, i.e. $M = \lim_{t \to x^-} f(t)$.

Moreover, if f is not continuous at $x \in (a, b)$, then $\lim_{t\to x^+} f(t) \neq \lim_{t \to x^-} f(t)$. (Note, since f is nondecreasing, it cannot be that $\lim_{t\to x^+} f(t) = \lim_{t\to x^-} f(t) \neq f(x)$.) Let D be the set of points in (a, b) at which f is not continuous. Then for each $x \in D$, assign a unique exists $r_x \in \mathbb{Q}$ such that

$$\lim_{t\to x^+} f(t) < r_x < \lim_{t\to x^-} f(t).$$

□

Similar results hold for more general classes of functions, such as those of bounded variation.

4.4 Sequences of Functions, Spaces of Functions

4.4.1 Types of Convergence

Definition 4.12. Suppose X and Y are metric spaces and $D \subseteq X$. A sequence of functions on D to Y is a collection of functions $f_n : D \to Y$.

For each $x \in D$, we obtain a sequence $(f_n(x))$ of points in Y by evaluating each f_n at x. These sequences may or may not converge. Note that, if $(f_n(x))$ does converge to some $y \in Y$, then y must be unique.

Definition 4.13. We say the sequence of functions (f_n) **converges pointwise** to some function $f : D \to Y$ if $f_n(x) \to f(x)$ for every point $x \in D$.

Exercise 4.11. For each $n \in \mathbb{N}$, let $f_n : [0, 1] \to \mathbb{R}$ be given by $f_n(x) = x^n$. Define $f(x)$ on $[0, 1]$ by

$$f(x) = \begin{cases} 0 & : x \in [0,1) \\ 1 & : x = 1 \end{cases}$$

Show that $f_n \to f$ pointwise.
Show also that f is not continuous.

Definition 4.14. Suppose X and Y are metric spaces and (f_n) is a sequence of functions on X. We say (f_n) **converges uniformly** to a function $f : D \to Y$ if for each $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and all $x \in X$, $\rho_Y(f_n(x), f(x)) < \epsilon$.
In case $Y = \mathbb{K}$, that means there exists $N \in \mathbb{N}$ such that for all $n \geq N$ and all $x \in D$, $|f_n(x) - f(x)| < \epsilon$.

Exercise 4.12. Show that the sequence in the previous example does not converge uniformly.
Give a (nontrivial) example of a sequence of functions on a subset of $\mathbb{R}$ that does converge uniformly.

Remark 4.9. *If a sequence* (f_n) *converges to a function* f *uniformly, then the must converge pointwise. Hence, whenever you want to find a uniform limit, it must agree with the pointwise limit at every point, i.e.* $f(x) = \lim_n f_n(x)$ *for all* $x \in X$*.*

Exercise 4.13. Suppose a sequence (f_n) of functions from a space X into a metric space Y converge uniformly to some function $f : X \to Y$. Show that f is continuous if each f_n is continuous.

4.4.2 Spaces of Bounded, Continuous Functions

Definition 4.15. Let X and Y be metric spaces. We denote the set of continuous functions from X to Y by $C(X, Y)$. We denote the set of bounded functions from X into $\mathbb{K}$ by $B(X, \mathbb{K})$. If $Y = \mathbb{K} = \mathbb{C}$, then we write $C(X)$ and $B(X)$ respectively.

Definition 4.16. If X is compact, then $C(X) \subseteq B(X)$. In general, we denote the set of bounded continuous functions by $BC(X)$ or $C_b(X)$. (Similarly for $\mathbb{R}$-valued functions.)

Definition 4.17. We define a the **uniform metric** on $B(X, \mathbb{K})$ by

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)|$$

for $f, g \in B(X, \mathbb{K})$. This metric will be suggestively denoted as $d(f, g) = \|f - g\|_X$, and it will be inherited by $C_b(X, \mathbb{K})$ and other subspaces.

Remark 4.10. *Notice that uniform convergence of a sequence of bounded functions is exactly convergence with respect to this metric.*

Exercise 4.14. Show that if (f_n) is a Cauchy sequence in $B(X, \mathbb{K})$, then (f_n) converges uniformly to some function $f \in B(X, \mathbb{K})$. How is this function defined? (Make sure you explain why this function is defined on all of X, i.e. why $f(x)$ exists for all x.)

This exercise essentially proves the Cauchy Criteria for Uniform Convergence:

Theorem 4.10 (Cauchy Criteria for Uniform Convergence)**.** *Let* (f_n) *be a sequence of (bounded) functions from a space* X *into* $\mathbb{K}$*. Then there is a (bounded) function* $f : X \to \mathbb{K}$ *to which* (f_n) *converges uniformly iff for every* $\epsilon > 0$ *there exists an* $N \in \mathbb{N}$ *such that* $|f_n(x) - f_m(x)| < \epsilon$ *for all* $x \in X$ *and all* $m, n \geq N$*.*

In other words, $B(X, \mathbb{K})$ is a complete metric space!

Remark 4.11. *Sometimes "bounded" is omitted, but it is implied by the fact that a Cauchy sequence is bounded and also by the fact that a uniformly convergent sequence must be a sequence of bounded functions.*

Corollary 4.2. $C_b(X, \mathbb{K})$ *is a complete metric space.*

Proof. Since $B(X,\mathbb{K})$ is complete, every Cauchy sequence in $C_b(X,\mathbb{K}) \subseteq B(X,\mathbb{K})$ has a limit in $B(X,\mathbb{K})$. This limit must also be in $C(X,\mathbb{K})$ since it is the uniform limit of continuous functions. $\square$

Remark 4.12. *If you are more familiar with the series version of the Cauchy Criteria, don't worry, it's coming.*

Definition 4.18. Let X a space. The space $C_b(X,\mathbb{K})$ has the following subspaces:

- $C_c(X,\mathbb{K}) = \{f \in C(X,\mathbb{K}) |\ f \text{ has compact support}\}$
  (The **support of** f is the closure of the set $\{x \in X |\ f(x) \neq 0\}$.)
- $C_0(X,\mathbb{K}) = \{f \in C(X,\mathbb{K}) |\ \{x : |f(x)| \geq \epsilon\} \text{ is compact } \forall\ \epsilon > 0\}$
  We say such a function **vanishes at infinity**.

Example 4.3. $f \in C_0(\mathbb{R},\mathbb{R})$ if $\lim_{x\to\pm\infty} f(x) = 0$.
More generally, if $X = (0,1]$ and $Y = \mathbb{R}$, $C_0((0,1],\mathbb{R})$ would be all continuous functions $f : (0,1] \to \mathbb{R}$ such that $\lim_{x\to 0} f(x) = 0$. "Infinity" here just means an endpoint not contained in your interval.

Exercise 4.15. Check that $C_c(X,\mathbb{K}) \subseteq C_b(X,\mathbb{K})$ and $C_0(X,\mathbb{K}) \subseteq C_b(X,\mathbb{K})$.

Exercise 4.16. Show that the closure of $C_c(\mathbb{R},\mathbb{R})$ in the uniform metric is $C_0(\mathbb{R},\mathbb{R})$.

Example 4.4. Consider the special case where $X = \mathbb{N}$ with the discrete topology. Then, for $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$,

- $C(\mathbb{N},\mathbb{K}) = \mathbb{K}^{\mathbb{N}} = \{(x_n)_{n\in\mathbb{N}} |\}\ x_n \in \mathbb{K}\ \forall\ n \in \mathbb{N}\}$
- $C_0(\mathbb{N},\mathbb{K}) = \{(x_n) \in \mathbb{K}^{\mathbb{N}} |\ \lim\limits_{n\to\infty} x_n = 0\}$
  This space is usually referred to as $c_0(\mathbb{K})$ or just c_0 when $\mathbb{K} = \mathbb{C}$.
- $C_b(\mathbb{N},\mathbb{K}) = \{(x_n) \in \mathbb{K}^{\mathbb{N}} |\ \exists\ M > 0 : |x_n| \leq M\ \forall\ n \in \mathbb{N}\}$
  This space is usually referred to as $\ell^\infty(\mathbb{K})$ or just ℓ^∞ when $\mathbb{K} = \mathbb{C}$.

Example 4.5 (Cantor-Lebesgue Function, again)**.** Define a sequence (f_n) of bounded continuous functions $f_n : [0,1] \to [0,1]$ as follows: for each n, define f_n on $\mathcal{O}_n$ as before, and on $[0,1]\backslash\mathcal{O}_n$ to be linear and such that f_n is continuous.

Note that $|f_{n+1}(x) - f_n(x)| < 2^{-n}$ for all $x \in [0,1]$ for each $n \in \mathbb{N}$, and hence (f_n) converge uniformly to a bounded, continuous function $\phi : [0,1] \to [0,1]$ (why?). Moreover, since $[0,1]$ is compact, ϕ is uniformly continuous.

4.4.3 Arzelà-Ascoli

Here we give a generalization of the Bolzano-Weierstass Theorem, which applies to sequences of functions in $C(X,\mathbb{K})$. However, we will need to require a little more of our sequences to get the result.

Definition 4.19. sequence (f_n) in $B(X, \mathbb{K})$ is **uniformly bounded** if there exists an $M \in \mathbb{N}$ such that $|f_n(x)| \leq M$ for all $x \in X$ and $n \in \mathbb{N}$.

In other words, (f_n) is bounded in the uniform metric.

Definition 4.20. For a metric space X, we say a collection $\mathcal{F} \subseteq C(X, \mathbb{K})$ is **equicontinuous at a point** $x \in X$ if for each $\epsilon > 0$ there exists a $\delta > 0$ such that for every $f \in \mathcal{F}$ and $y \in X$,

$$\rho(x, y) < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

$\mathcal{F}$ is equicontinuous if it is equicontinuous at every point $x \in X$.

In other words, every $f \in \mathcal{F}$ is continuous at $x \in X$, *and* given $\epsilon > 0$ and $x \in X$, the same δ works for every $f \in \mathcal{F}$.

Propostion 4.3. *If X is compact and $\mathcal{F} \subseteq C(X, \mathbb{K})$ is equicontinuous, then it is* **uniformly equicontinuous**, *i.e. for each $\epsilon > 0$, there exists a $\delta > 0$ such that for any $f \in \mathcal{F}$ and any $x, y \in X$,*

$$\rho(x, y) < \delta \Rightarrow |f(x) - f(y)| < \epsilon.$$

Theorem 4.11 (Arzelà-Ascoli)**.** *Let X be a compact metric space and (f_n) a uniformly bounded equicontinuous sequence in $C(X, \mathbb{K})$. Then (f_n) has a subsequence that converges uniformly on X to a continuous function f on X.*

The proof uses a technique often called a "Cantor diagonalization argument."

Proof. Since X is compact, $C(X, \mathbb{K})$ is complete in the uniform metric. Hence, it suffices to prove that (f_n) has a Cauchy subsequence.

Since X is a compact metric space, it is separable. Let $S = \{x_n\}_{n \in \mathbb{N}}$ be a countable dense subset. Since (f_n) are uniformly bounded, the sequence $(f_n(x_1))_{n \in \mathbb{N}}$ is bounded in $\mathbb{K}$ and hence, by Bolzano-Weierstrass, has a convergent subsequence, which we denote $(f_{1,m}(x_1))_{m \in \mathbb{N}}$. Similarly, the sequence $(f_{1,m}(x_2))_{m \in \mathbb{N}}$ is bounded and hence has a convergent subsequence, which we denote $(f_{2,m}(x_2))_{m \in \mathbb{N}}$. Continue this process to get a collection $\{f_{k,m}\} \subseteq f_n$.

Define $g_n = f_{n,n}$. It remains to show that g_n is Cauchy. Let $\epsilon > 0$ and $\delta > 0$ such that for every $x, y \in X$ with $\rho(x, y) < \delta$ and every $n \in \mathbb{N}$, $|f_n(x) - f_n(y)| < \epsilon$ (which is guaranteed by uniform equicontinuity). Since X is compact, there exist $y_1, ..., y_N \in S$ such that $\{B_\delta(y_k)\}$ cover X. For each y_k, the sequence $(g_n(y_k))_{n \in \mathbb{N}}$ is Cauchy, and so there exists an $M \in \mathbb{N}$ such that for all $n, m \geq M$,

$$|g_n(y_k) - g_m(y_k)| < \epsilon.$$

Now, let $x \in X$, and choose y_k such that $\rho(x, y_k) < \delta$. Then for any $n, m \geq M$,

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(y_k)| + |g_n(y_k) - g_m(y_k)| + |g_m(y_k) - g_m(x)| < 3\epsilon,$$

and we are done. □

****Presentation topic: The proof for Arzelà-Ascoli.****

The theorem has several applications such as Montel's Theorem in complex analysis and the Peano existence theorem (for local solutions to certain first order diff eq's).

4.4.4 Series of Functions

If you recall from series of real numbers, a series $\sum_n x_n$ is said to converge if its sequence of partial sums converges, i.e. $\lim_{N\to\infty}\sum_{n=1}^{N} x_n$ exists. In this case, the x_n's are replaced with functions.

Definition 4.21. Suppose (f_n) is a sequence of functions from a metric space X into $\mathbb{K}$.

1. The series $\sum f_n$ is said to **converge** if the sequence of partial sums

$$F_N(x) = \sum_{n=1}^{N} f_n(x)$$

   converges pointwise.

2. The series $\sum f_n$ is said to **converge uniformly** if the sequence of partial sums

$$F_N = \sum_{n=1}^{N} f_n$$

   converges uniformly.

3. The series $\sum f_n$ is said to **converge absolutely** if the sequence of partial sums

$$\sum_{n=1}^{N} |f_n|$$

   converges pointwise.

Perhaps you have seen the Cauchy Criteria for Uniform Convergence in the following form:

Theorem 4.12. *Suppose (f_n) is a sequence of functions from a space X into $\mathbb{K}$. The series $\sum f_n$ is uniformly convergent iff for every $\epsilon > 0$ there exists an $k \in \mathbb{N}$ such that for $M \geq N \geq K$,*

$$\|\sum_{n=1}^{M} f_n - \sum_{n=1}^{N} f_n\|_X = \sup_{x\in X} |f_{N+1}(x) + ... + f_M(x)| < \epsilon.$$

Note that this follows from the previous version of the theorem and that this insures that the sum $f = \sum f_n$ is continuous.
While we are here, we might as well prove Weierstrass's M-Test. We will have to save most of the section on series of functions for later.

Theorem 4.13 (Weierstrass's M-Test)**.** *Let* (M_n) *be a sequence of non-negative real numbers such that* $\sup_{x\in X} \|f_n(x)\| \leq M_n$ *for each* n*. If the series* $\sum M_n$ *is convergent, then the series* $\sum f_n$ *is uniformly convergent on* X*.*

The proof follows quickly from Cauchy's Criteria; however the language requires norms, which we have yet to cover.

Exercise 4.17. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \sum_{n=1}^{\infty} \frac{e^{inx} + e^{-inx}}{2^n}.$$

Is f continuous?

4.4.5 Power Series over $\mathbb{R}$

Definition 4.22. A **power series** around $x = c$ is a series of functions $f_n : \mathbb{R} \to \mathbb{R}$ of the form

$$f_n(x) = a_n(x - c)^n$$

where $a_n, c \in \mathbb{R}$, and $n \in \mathbb{N} \cup \{0\}$.

Definition 4.23. For a power series $\sum a_n(x - c)^n$, define

$$R = \lim_n \frac{|a_n|}{|a_{n+1}|}.$$

Provided the limit exists, we call R the **radius of convergence** of the power series.
The **interval of convergence** is the open interval $(c - R, c + R)$.

Theorem 4.14 (Cauchy-Hadamard)**.** *If* R *is the radius of convergence of the power series* $\sum a_n(x - c)^n$*, then the series is absolutely convergent for* $|x| < R$ *and divergent for* $|x| > R$*.*

Theorem 4.15. *The limit of a power series is continuous on the interval of convergence. Moreover, if* $\sum a_n(x-c)^n$ *and* $\sum b_n(x-c)^n$ *converge on* $(c-r, c+r)$ *to the same function, then* $a_n = b_n$ *for all* n*.*

Theorem 4.16 (Differentiation of Power Series)**.** *A power series may be differentiated term-by-term within the interval of convergence. Moreover, the derivative will have the same radius of convergence.*

Theorem 4.17 (Integration of Power Series)**.** *A power series can be integrated term-by-term on any compact interval contained in the interval of convergence.*

4.5 Differentiation on $\mathbb{R}$

Suppose $f : D \subseteq \mathbb{R} \to \mathbb{R}$.

Definition 4.24. If $c \in D$ is a limit point of D, we say $L \in \mathbb{R}$ is the **derivative of** f at c if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $x \in D$ and $0 < |x - c| < \delta$, then

$$\left| \frac{f(x) - f(c)}{x - c} - L \right| < \epsilon.$$

In this case, we write $f'(c) = L$.

In other words,

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = 0.$$

If f is differentiable at every point in a set D, we say f is differentiable on D.

Theorem 4.18 (Rolle's Theorem)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and $f(a) = f(b) = 0$. Then there exists a $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. Assume f is not constant. Then by the extreme value theorem, there exists (WLOG or use $-f$) a maximum at some point $c \in (a, b)$. If $f'(c) \neq 0$, then WLOG $f'(c) > 0$. Let $\epsilon > 0$ such that $f'(c) > \epsilon$. Let $\delta > 0$ such that for all $x \in D$ with $c < x < c + \delta$,

$$-\epsilon < \frac{f(x) - f(c)}{x - c} - f'(c) < \epsilon.$$

But then since $x - c > 0$,

$$0 < (f'(c) - \epsilon)(x - c) < f(x) - f(c),$$

i.e. $f(x) > f(c)$, contradicting f a maximum. □

Corollary 4.3 (Mean Value Theorem)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Then there exists a $c \in (a, b)$ such that*

$$f(b) - f(a) = f'(c)(b - a).$$

Proof. Consider $\phi : [a, b] \to \mathbb{R}$ defined by

$$\phi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Applying Rolle's Theorem to ϕ gives a $c \in (a, b)$ for which

$$0 = \phi'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

□

Propostion 4.4. *If $f : I \to \mathbb{R}$ is differentiable, then it is also continuous.*

Definition 4.25. We define the space $C^1(X, \mathbb{R})$ (and $C^1(X, \mathbb{C})$) to be the space of continuously differentiable (i.e. f is differentiable and f' is continuous) functions from X to $\mathbb{R}$ ($\mathbb{C}$).

Exercise 4.18. Suppose $g \in C^2(\mathbb{R}, \mathbb{R})$, $g(x) \geq 0$ for all $x \in \mathbb{R}$, and $g(0) = g(1) = 0$. Prove (pictures wont suffice) that there exists a point $x \in (0, 1)$ such that $g''(x) \leq 0$.

Exercise 4.19. Let $f(x)$ be continuous on $[-1, 1]$ and differentiable except possibly at $x = 0$. Prove that if

$$\lim_{x \to 0} f'(x) = L$$

exists, then $f'(0)$ exists and is equal to L.

4.5.1 Convexity

Definition 4.26. We say a subset $C \subseteq \mathbb{K}^d$ is **convex** if for any $x, y \in C$ and any $t \in [0, 1]$,

$$tx + (1 - t)y \in C.$$

Note that the set $\{tx + (1 - t)y | \ t \in [0, 1]\}$ is exactly the line segment connecting x and y in C. In particular, if $C \subseteq \mathbb{R}$, the C is convex iff C is path connected iff C is an interval.

Remark 4.13. *This definition also holds in any vector space over $\mathbb{K}$.*

Definition 4.27. We say that a function $f : K \to \mathbb{R}$ on a convex subset K of $\mathbb{K}^d$ is convex if for all $x, y \in K$,

$$f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

Example 4.6. $f(x) = e^x$ is convex, but this is not so easy to see algebraically. It is best justified graphically.

For a function $f : \mathbb{R} \to \mathbb{R}$, being convex is sitting below ones secant lines (i.e. above ones tangent lines).

Propostion 4.5. *If f is convex on (a, b), then f is Lipschitz on any closed subinterval of (a, b) at is differentiable at all but countably many points in (a, b).*

Propostion 4.6 (Jensen's Inequality)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is a convex function and $\sum_{n=1}^N \lambda_n x_n$ is a convex combination of $x_1, ..., x_n \in [a, b]$. Then*

$$f(\sum_1^N \lambda_n x_n) \leq \sum_1^N \lambda_n f(x_n).$$

Exercise 4.20. Prove this.

Remark 4.14. *This is only the discrete version. There is a much more useful version that arises in probability theory.*

Example 4.7. $f(x) = e^x$ is convex. Hence, for $\lambda_1, ..., \lambda_n \in [0,1]$ with $\sum_{k=1}^n \lambda_n = 1$ and any $x_1, ..., x_n \in \mathbb{R}$,

$$e^{\sum \lambda_k x_k} \leq \sum_k \lambda_k e^{x_k}.$$

4.5.2 Higher Derivatives and Taylor's Theorem

Definition 4.28. For $n \in \mathbb{N}$ and $D \subseteq \mathbb{R}$, we say $f \in C^n(D, \mathbb{R})$ if $f, f', ..., f^(n)$ are defined and continuous on D. We say $f \in C^\infty(D, \mathbb{R})$ if $f^(n)$ is defined and continuous for all $n \in \mathbb{N}$.

Theorem 4.19 (Taylor's Theorem)**.** *Suppose $n \in \mathbb{N}$, $I \subseteq \mathbb{R}$ is a closed interval, $f \in C^{n-1}(I, \mathbb{R})$, and that $f^(n)$ exists on Int(I). If $\varnothing \neq (a,b) \subseteq [a,b]$, then there exists a $c \in (a,b)$ such that*

$$f(b) = \left[\sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(b-a)^k\right] + \frac{f^{(n)}(c)}{n!}(b-a)^n.$$

The following is another version of Taylor's Theorem:

Theorem 4.20 (Taylor's Theorem)**.** *Suppose $f(x) = \sum a_n x^n$ with the series converging for all $x \in (-R, R)$.If $c \in (-R, R)$, then f can be expanded in a power series about the point $x = c$ which converges for all x such that $|x - c| < R - |c|$, and*

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(c)}{n!}(x-c)^n.$$

There do exist infinitely differentiable functions $f : \mathbb{R} \to \mathbb{R}$ whose Taylor series converges but not to f. In contrast, the Taylor series of a holomorphic (or analytic) function $f : \mathbb{C} \to \mathbb{C}$ always converges to f. We will leave this topic for complex analysis.

Example 4.8. Here are some Taylor series for some common functions along with their intervals of convergence:

1. A polynomial is its Taylor series. Its interval of convergence is $(-\infty, \infty)$.
2. $\dfrac{1}{1-x} = \sum_{n=0}^{\infty} x^n$ on $(-1, 1)$.
3. $e^x = \sum_{n=0}^{\infty} \dfrac{x^n}{n!}$ on $(-\infty, \infty)$.
4. $\ln(x) = \sum_{n=1}^{\infty} \dfrac{(-1)^{n+1}}{n}(x-1)^n$ on $(0, 1]$.

Exercise 4.21. Determine the Taylor Series and radius of convergence for the following functions:

1. $\frac{1}{x}$ centered at $c = 1$
2. xe^x centered at $c = 0$
3. $\frac{1}{(1-x)^2}$ centered at $c = 0$.

4.6 Approximations of Continuous Functions

Despite their nice behavior, continuous functions are often difficult to evaluate. However, they can often be approximated by much nicer functions, e.g. polynomials.

Theorem 4.21 (Weierstrass Approximation Theorem). *Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then for each $\epsilon > 0$ there is a polynomial p for which*

$$\|f - p\|_X < \epsilon.$$

In other words, the polynomials are dense in $C([a, b], \mathbb{R})$.[1] But, there is a more general theorem, which highlights the properties of the polynomials that underlies this. Notice that for any compact Hausdorff space X, $C(X, \mathbb{R})$ is an algebra, i.e. it is closed with respect to scalar multiplication, addition, and multiplication.

Definition 4.29. A collection functions $\mathcal{F}$ between metric spaces X and Y is said to separate points if for any distinct $x, y \in X$, there exists an $f \in \mathcal{F}$ for which $f(x) \neq f(y)$.

Recall that Uryshon's Lemma implies that for a compact Hausdorff space X, $C(X, \mathbb{R})$ separates the points of X. However, we can often find a subcollection that will do this, e.g. for $X = [a, b]$, the polynomials are such a subcollection.

Theorem 4.22 (Stone-Weierstrass Theorem). *Let X be a compact Hausdorff space. Suppose $\mathcal{F} \subseteq C(X, \mathbb{R})$ is a subalgebra that separates points in X and contains the constant functions on X. Then $\mathcal{F}$ is dense in $C(X, \mathbb{R})$.*

Remark 4.15. *Note that we need only require that the constant function $1 \in \mathcal{F}$ because $\mathcal{F}$ being an algebra will require all scalar multiples of 1 (i.e. all other constant functions)be contained in $\mathcal{F}$.*

Exercise 4.22. Suppose X and Y are compact Hausdorff spaces. Show that the algebra $\mathcal{F}$ generated by the functions of the form $f(x, y) = g(x)h(y)$ where $g \in C(X, \mathbb{R})$ and $h \in C(Y, \mathbb{R})$ is dense in $C(X \times Y, \mathbb{R})$.

Exercise 4.23. Let X be a compact Hausdorff space. Suppose $\mathcal{F} \subseteq C(X, \mathbb{R})$ is a subalgebra that separates points in X. Show that either $\overline{\mathcal{F}} = C(X, \mathbb{R})$ or $\overline{\mathcal{F}} = \{f \in C(X, \mathbb{R}) : f(x_0) = 0\}$ for some $x_0 \in X$.

[1]You need trigonometric polynomials to get this for $\mathbb{C}$.

Polynomial approximation is extremely useful for continuous functions. However, we may also approximate continuous functions by so-called step functions. These will be more useful for approximating a wider variety of functions that are not necessarily continuous but are still "nice enough" for some calculus.

Definition 4.30. Let $A \subseteq \mathbb{R}^d$. The **indicator (characteristic) function** for A is a function $\mathbb{1}_A : \mathbb{R}^d \to \mathbb{R}$ defined by

$$\mathbb{1}_A(x) = \begin{cases} 1 & : x \in A \\ 0 & : x \notin A \end{cases}$$

Example 4.9. The Dirichlet function from before can be expressed as $x\mathbb{1}_{\mathbb{Q}}$.

Definition 4.31. A **step function** $f : \mathbb{R} \to \mathbb{R}$ is a finite linear combination of characteristic functions for intervals, i.e. a function of the form

$$\sum_{k=1}^{n} \alpha_k \mathbb{1}_{I_k}$$

where $I_1, ..., I_n \subseteq \mathbb{R}$ are intervals and $\alpha_1, ..., \alpha_n \in \mathbb{R}$.

Exercise 4.24. Let $[a, b] \subseteq \mathbb{R}$ and $f \in C([a, b], \mathbb{R})$. Show that f can be uniformly approximated on $[a, b]$ by step functions.

Remark 4.16. *This should remind you of Riemann Rectangles.*

Exercise 4.25. To prepare us Lebesgue's criteria for Riemann Integrability, check that the following is true:

For a function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ and $x_0 \in D$, f is continuous at x_0 iff $\omega_f(x_0) = 0$.

In a sense, oscillation quantifies discontinuity at a point.

4.7 Integration over $\mathbb{R}$

Let $f \in B([a, b], \mathbb{R})$, and let $P = [x_0, ..., x_n]$ be a partition of $[a, b]$, i.e.

$$a = x_0 < x_1 < ... < x_{n-1} < x_n = b.$$

Then, we have disjoint subintervals (closed or open, it doesn't matter) $I_1 = [a, x_1]$, $I_i = (x_i, x_{i+1}]$ for $i > 1$ whose union is $[a, b]$. For each $1 \leq i \leq n$, let

$$m_i = \inf\{f(x) |\ x \in I_i\}$$

and

$$M_i = \sup\{f(x) |\ x \in I_i\}.$$

For each $1 \leq i \leq n$, we denote the length of I_i by

$$\lambda(I_i) = x_i - x_{i-1}.$$

Define the **lower and upper Darboux sums** for f with respect to P by

$$L(f, P) = \sum_{i=1}^{n} m_i \lambda(I_i)$$

and

$$U(f, P) = \sum_{i=1}^{n} M_i \lambda(I_i).$$

We define the **lower and upper Riemann integrals** of f over $[a, b]$ by

$$(R)\underline{\int_a^b} f = \sup\{L(f, P) |\ P \text{ a partition of } [a, b]\}$$

and

$$(R)\overline{\int_a^b} f = \inf\{U(f, P) |\ P \text{ a partition of } [a, b]\}.$$

Since f is bounded and $\lambda([a, b]) < \infty$, both of these values will be finite. Moreover,

$$(R)\underline{\int_a^b} \leq (R)\overline{\int_a^b}.$$

When the two integrals are equal, we say that f is **Riemann integrable** over $[a, b]$ and call the common value the Riemann integral of f over $[a, b]$, denoted by

$$(R)\int_a^b f.$$

Notice that for each $1 \leq i \leq n$,

$$M_i - m_i = \omega_f(I_i).$$

Hence, we can say that f is Riemann integrable if for any $\epsilon > 0$ there exists a partition of $[a, b]$ into intervals $I_1, ..., I_n$ such that

$$U(f, P) - L(f, P) = \sum_{i=1}^{n} \omega_f(I_i)\lambda(I_i) < \epsilon.$$

Since we know that a continuous function has zero oscillation at any point, that means for any $x \in [a, b]$,

$$\lim_{\delta \to 0} \omega_f(x - \delta, x + \delta) = 0.$$

Moreover, on a closed interval $[a, b]$, a continuous function is uniformly continuous, and so for any $\epsilon > 0$, we can find a partition P of $[a, b]$ so that for each $I_i \in P$,

$$\omega_f(I_i) < \frac{\epsilon}{b - a},$$

and so

$$\sum_{i=1}^{n} \omega_f(I_i)\lambda(I_i) < \epsilon.$$

Thus, we have proved the following:

Theorem 4.23. *If $f \in C([a,b],\mathbb{R})$, then f is integrable on $[a,b]$.*

However, there is another way to make the difference $U(f,P) - L(f,P)$ small. Note that at any point $x \in X$ of discontinuity, $\omega_f(x) < \infty$ since f is bounded. Hence, if we are able to cover the points of discontinuity of a function with arbitrarily small open intervals, then the function could still be integrable. Let's try to quantify this idea.

4.7.1 Lebesgue Measure Zero Sets

In real analysis, you will be introduced to the notion of measures, specifically Lebesgue measures, for a basis of calculus on a more general class of functions. We will not venture deep into this territory, but one notion we can and will borrow is the notion of Lebesgue measure zero sets.

In general, not every set is Lebesgue measurable, but any set we will encounter is (e.g. open, closed sets, basic combinations of these, G_δ and F_σ sets, etc.) We define the measure of an interval to be its length. For any other set $E \subseteq \mathbb{R}$, we define its measure to be

$$\inf\{\sum_n (b_n - a_n) |\ E \subseteq \bigcup_n (a_n, b_n)\}.$$

Now, we can define what it means for a set to have Lebesgue measure zero:

Definition 4.32. A set $E \subseteq \mathbb{R}$ has Lebesgue measure zero if for any $\epsilon > 0$ there exists a collection of open intervals $\{(a_n, b_n)\}_{n\in\mathbb{N}}$ such that $E \subseteq \bigcup_n (a_n, b_n)$ and $\sum_n (b_n - a_n) < \epsilon$.

Exercise 4.26. Show that any countable subset of $\mathbb{R}$ has Lebesgue measure zero.

Example 4.10. The Cantor set has Lebesgue measure zero.

Exercise 4.27. Must all measure zero sets have empty interior?

4.7.2 Lebesgue's Criteria for Riemann Integrability

So, it turns out the criteria that we want to handle the points of discontinuity is exactly that the collection of such points has Lebesgue measure zero.

Theorem 4.24. *A function $f : [a,b] \to \mathbb{R}$ is Riemann integrable iff it is bounded and its set of points of discontinuity for f has Lebesgue measure zero.*

Remark 4.17. *The proof I will leave open (with resources) for a possible presentation topic. However, I will remark that the proof of the necessary direction uses the fact that ω_f is upper semi-continuous to conclude that for any $\epsilon > 0$, the set $\omega_f^{-1}([\epsilon, \infty))$ containing some discontinuities of f is closed (+ bounded by the boundedness of f) and hence compact. Since it has measure zero (as a subset of a set with measure zero), it can be covered with finitely many intervals of arbitrarily small length.*

Remember how we showed that a monotone function $f : [a, b] \to \mathbb{R}$ has at most countably many discontinuities?

Corollary 4.4. *A monotone function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.*

Example 4.11. The Cantor-Lebesgue function is Riemann integrable as is the indicator function $\mathbb{1}_{\mathcal{O}}$.

4.7.3 Once more with step functions

Recall that to define a step function on $[a, b]$, we give a partition $P = \{I_1, ..., I_n\}$ of $[a, b]$ along with values $c_1, ..., c_n$ and let

$$\phi(x) = \sum_{i=1}^{n} c_i \mathbb{1}_{I_i}.$$

Then we have that

$$L(f, P) = \sum_{i=1}^{n} c_i \lambda(I_i) = U(f, P).$$

And so, we conclude the following:

Propostion 4.7. *Any step function $\phi = \sum_{k=1}^{n} \alpha_k \mathbb{1}_{I_k}$ on $[a, b]$ is Riemann integrable with Riemann integral*

$$(R) \int_a^b \phi = \sum_{k=1}^{n} \alpha_k \lambda(I_k).$$

Let's lose the (R).

Using step functions allows us to reformulate the definition of lower and upper Riemann integrals as follows:

$$\underline{\int_a^b} f = \sup \left\{ \int_a^b \phi |\ \phi \text{ a step function and } \phi \leq f \text{ on } [a, b] \right\}$$

and

$$\overline{\int_a^b} f = \inf \left\{ \int_a^b \psi |\ \psi \text{ a step function and } \psi \geq f \text{ on } [a, b] \right\}.$$

Example 4.12. Let $Q = \mathbb{Q} \cap [0,1]$. Then $\mathbb{1}_Q$ is not Riemann integrable. Indeed, for any partition P of Q, $L(\mathbb{1}_Q, P) = 0$ and $U(\mathbb{1}_Q, P) = 1$.

Moreover, notice that $\mathbb{1}_Q$ is the pointwise limit of an increasing sequence of bounded step functions. Let q_k be an enumeration of Q. For each $n \in \mathbb{N}$, define $f_n = \mathbb{1}_{\{q_1,\ldots,q_n\}}$.

Exercise 4.28. Suppose (f_n) is a sequence of bounded integrable functions on $[a,b]$ that converge uniformly of f on $[a,b]$. Show that if each f_n is Riemann integrable then f is also Riemann integrable. Is

$$\lim_n \int_a^b f_n = \int_a^b f?$$

4.7.4 Facts about the Riemann Integral

Propostion 4.8. *Suppose $c \in [a,b]$. f is integrable on $[a,b]$ iff it is integrable on both $[a,c]$ and $[c,b]$; in either case, the following holds:*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Theorem 4.25. *Suppose $f, g : [a,b] \to \mathbb{R}$ are Riemann integrable.*

1. *For any $\alpha, \beta \in \mathbb{R}$, $\alpha f + \beta g$ is integrable, and*

$$\int_a^b \alpha f + \beta g = \alpha \int_a^b f + \beta \int_a^b g.$$

2. *$|f|$ is Riemann integrable, and*

$$\left| \int_a^b f \right| \le \int_a^b |f|.$$

3. *f^2 is integrable on $[a,b]$.*

4. *fg is integrable on $[a,b]$.*

Remark 4.18. *The integrability of $|f|$ follows from the triangle inequality. To see that f^2 is integrable, note that if $\sup_{x\in[a,b]} |f(x)| \le M$, then for all $x, y \in [a,b]$, $|f(x)^2 - f(y)^2| \le 2M|f(x) - f(y)|$. Hence, we know f^2 is integrable for any integrable function. Then to see that fg is integrable, we note that $2fg = (f+g)^2 - f^2 - g^2$.*

Theorem 4.26 (Mean Value Theorem for Riemann Integrals)**.** *Suppose $f \in C([a,b], \mathbb{R})$. Then there exists a $c \in [a,b]$ such that*

$$\int_a^b f = f(c)(b-a).$$

Proof. By definition, we know that

$$(b-a)\inf_{x\in[a,b]} f(x) \le \int_a^b f \le (b-a)\sup_{x\in[a,b]} f(x).$$

Actually,

$$\min_{x\in[a,b]} f(x) \le \frac{1}{b-a}\int_a^b f \le \max_{x\in[a,b]} f(x).$$

By the continuity of f and the Intermediate Value Theorem, there exists a $c \in [a,b]$ such that

$$f(x) = \frac{1}{b-a}\int_a^b f.$$

□

Theorem 4.27 (Differentiation Theorem for Riemann Integrals)**.** *Suppose $f \in C([a,b],\mathbb{R})$, then for any $c \in [a,b]$, the function F defined for $x \in [a,b]$ by*

$$F(x) = \int_a^x f$$

has a derivative at c and $F'(c) = f(c)$.

Proof. We want to show that

$$\lim_{h\to 0} \frac{F(c+h)-F(c)}{h} = 0.$$

If $h > 0$ is such that $c + h \in [a,b]$, then by the preceding theorem, there exists a $c' \in [c, c+h]$ such that

$$\frac{1}{h}(F(c+h)-F(c)) = \frac{1}{h}\int_c^{c+h} f = \frac{1}{h} f(c')(c+h-c) = f(c').$$

There exists a similar c' for $h < 0$. Since f is continuous, $F'(c)$ exists and is equal to $f(c)$. □

Aside from checking that two antiderivatives of a continuous function differ only by a constant, this proves the FToC.

Theorem 4.28 (Fundamental Theorem of Calculus)**.** *Let $f \in C([a,b],\mathbb{R})$. A functions F on $[a,b]$ satisfies*

$$F(x) - F(a) = \int_a^x f$$

for $x \in [a,b]$ iff $F' = f$ on $[a,b]$.

There are two more techniques of integration from calculus that are used again and again. So, they deserve to be cited here:

Theorem 4.29 (Integration by Parts)**.** *If $f, g \in C^1([a,b],\mathbb{R})$, then*

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g.$$

Theorem 4.30 (Change of Variables)**.** *Suppose $g \in C^1([c,d],\mathbb{R})$, $g(c) = a$, $g(d) = b$, and $f \in C(g([c,d]),\mathbb{R})$. Then*

$$\int_a^b f(x)\ dx = \int_c^d f(g(t))g'(t)\ dt.$$

4.8 Series (i.e. Integration over $\mathbb{N}$)

Though these concepts generalize to $\mathbb{K}^{\mathbb{N}}$, we will stick to $\mathbb{R}$.

Definition 4.33. For a sequence $(a_n) \in \mathbb{R}^{\mathbb{N}}$, we say the **series** $\sum_{n=1}^{\infty} a_n$ is convergent if the sequence $S_N = \sum_{n=1}^{N} a_n$ of partial sums converges. If so, then we say the limit is the **sum** of the series.

We say the series **converges absolutely** if the corresponding series $\sum_{n=1}^{\infty} |a_n|$ converges. We say the series **converges conditionally** if $\sum_{n=1}^{\infty} a_n$ converges but does not converge absolutely.

Theorem 4.31 (Cauchy-Criterion for Series)**.** *The series (x_n) in $\mathbb{R}^d$ converges iff for each $\epsilon > 0$ there exists $M > 0$ such that for $M \leq n \leq m$,*

$$\|S_m - S_n\| = \|x_{n+1} + x_{n+2} + ... + x_m\| < \epsilon.$$

Theorem 4.32. *If $\sum x_n$ converges absolutely, then it converges.*

Theorem 4.33. *A series $\sum a_n$ in $\mathbb{R}^d$ converges absolutely iff it converges unconditionally, i.e. the series sums to the same value under any permutation.*

Remark 4.19. *Absolute convergence always implies unconditional convergence, but the converse can fail in infinite dimensions.*

Theorem 4.34 (Riemann's Conditional Convergence Theorem)**.** *If a series $\sum a_n$ in $\mathbb{R}^d$ is conditionally convergent, then for any $M \in [-\infty, \infty]$ there exists a permutation σ of $\mathbb{N}$ so that*

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = M.$$

Exercise 4.29. Give an argument for this. I don't necessarily want a proof, just thoughts on why it's true.

Exercise 4.30. Assume $a_n \geq 0$ for all $n \in \mathbb{N}$ and the series $\sum_n a_n$ converges. Then, if σ is any permutation of $\mathbb{N}$, then

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma(n)}.$$

Exercise 4.31. For $(x_n) \in \mathbb{K}^{\mathbb{N}}$,

$$\left|\sum_{n=1}^{\infty} x_n\right| \leq \sum_{n=1}^{\infty} |x_n|.$$

Theorem 4.35. *Suppose the series $\sum x_n$ and $\sum y_n$ converge and $c \in \mathbb{K}$. Then,*

- *$\sum(x_n + y_n)$ converges with sum $\sum x_n + \sum y_n$*
- *$\sum(cx_n)$ converges with sum $c\sum x_n$*

Exercise 4.32. Suppose $\sum x_n$ and $\sum y_n$ converge in $\mathbb{R}$, must $\sum x_n y_n$ converge? Proof or counterexample.

Propostion 4.9. *If $\sum x_n$ converges in $\mathbb{R}$, then $\lim x_n = 0$.*

Theorem 4.36. *Let (x_n) be a sequence of positive real numbers, and let $S_N = \sum_{n=1}^{N} x_n$ be the corresponding sequence of partial sums. Then $\sum x_n$ is convergent iff (S_N) is bounded. In which case the sum is $\sup S_n$.*

Example 4.13. An important example is the **geometric series**. Let $a \in \mathbb{C}$. Then the sequence (a^n) converges to 0 iff $|a| < 1$. Moreover, for $a \neq 1$ and $N \in \mathbb{N}$,

$$S_N = \sum_{n=0}^{N} a^n = \frac{1 - a^{N+1}}{1 - a}.$$

So, when $|a| < 1$,

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1 - r}.$$

A particularly useful instance is for $a = \frac{1}{2}$.

Remark 4.20. *There are a plethora of tests for convergence of series. This is more of a calculus tool, so I shall only list them in case you see one you have not seen before. These can be found in [1]:*

p-series test, comparison test, limit comparison test, root test, ratio test, Raabe's test, integral test, Abel's Lemma, Abel's Test, Dirichlet's Test, alternating series test, etc.

It will be more beneficial to see sequential versions of a few topics that will arise in real analysis next spring.

4.8.1 Convergence Theorems

Sequences will can give a preview of some material that you will see in the spring.

Definition 4.34. A **double sequence** is a function $(x_{m,n}) : \mathbb{N}\times\mathbb{N} \to X$ where X is some topological space. Suppose X is a metric space. Then we say $(x_{m,n})$ has the limit $x \in X$ if for each $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|x_{m,n} - x\| < \epsilon$ for all $m, n \geq N$.

Theorem 4.37. *If the double limit $x = \lim_{m,n} x_{m,n}$ exists and for each fixed $m \in \mathbb{N}$ the limit $y_m = \lim_n x_{m,n}$ exists, then the iterated limit $\lim_m \lim_n x_{m,n}$ exists and equals x.*

There are two ways to think of these double sequences. The first, is to think of them as a sequence of sequences:

Consider $f_m : \mathbb{N} \to \mathbb{R}$ defined by $f_m(n) = a_{m,n} \in \mathbb{R}$ for each $n \in \mathbb{N}$. Suppose $a_{m,n} \to a_n$ for all n, i.e. f_m converges pointwise. When can we say $\lim_m \sum_n a_{m,n} = \sum_n \lim_m a_{m,n}$?

Theorem 4.38 (Fatou's Lemma)**.** *Let (a_{m_n}) be a double sequence of non-negative terms with $\liminf m \to \infty a_{m,n} = a_n$ for each $n \in \mathbb{N}$. Then*

$$\sum_n a_n = \sum_n \liminf_m a_{m,n} \leq \liminf_m \sum_n a_{m,n}.$$

Theorem 4.39 (Monotone Convergence Theorem)**.** *Suppose $a_{m,n} \geq 0$ and $a_{m,n} \leq a_{m+1,n}$ for all m (i.e the sequences f_m are monotone increasing pointwise). Then, if $a_n = \lim_m a_{m,n}$ for each $n \in \mathbb{N}$, then*

$$\lim_m \sum_n a_{m,n} = \sum_n \lim_m a_{m,n} = \sum_n a_n.$$

Exercise 4.33. Using the MCT, prove the non-increasing version:

Theorem 4.40. *Suppose $a_{m,n} \geq 0$ and $a_{m,n} \geq a_{m+1,n}$ for all m. Assume furthermore that $\sum_{n=1}^{\infty} a_{1,n} < \infty$. Then if $a_n = \lim_m a_{m,n}$ for each $n \in \mathbb{N}$, then*

$$\lim_m \sum_n a_{m,n} = \sum_n \lim_m a_{m,n} = \sum_n a_n.$$

Hint: $(-a_{m,n})_{m\in\mathbb{N}}$ is increasing to $(-a_n)$, but these are negative. Instead, take the sequence $b_{m,n} = a_{1,n} - a_{m,n}$ for each $n \in \mathbb{N}$. Then you have an increasing double sequence of non-negative terms.

Exercise 4.34. Determine whether

$$f(x) = \sum_{n=1}^{\infty} e^{-n^2 x}$$

is continuous on $(0, \infty)$. *Hint: Recall the "calculus" version of continuity at a point x_0: $\lim_{x\to x_0^+} f(x) = f(x_0) = \lim_{x\to x_0^-} f(x)$.*

Theorem 4.41 (Dominated Convergence Theorem)**.** *Suppose* $\lim_m a_{m,n} = a_n$ *for each* $n \in \mathbb{N}$ *and that there exists a sequence* $\{b_n\}$ *such that* $b_n \geq |a_{m,n}|$ *for all* m, n*. Moreover, assume* $\sum_n b_n$ *converges. Then* $\sum a_n$ *converges absolutely, and*

$$\lim_m \sum_n a_{m,n} = \sum_n \lim_m a_{m,n} = \sum_n a_n.$$

Exercise 4.35. Assume $\sum_n |a_n|$ converges. Prove that

$$\lim_{\epsilon \searrow 0} \sum_n e^{-\epsilon n} a_n = \sum_n a_n.$$

Now, we switch gears and consider double sequences as functions $f : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ where $a_{m,n} = f(m,n)$, and we ask whether and how we can find the sum $\sum_{m,n} a_{m,n}$. Now instead of interchanging limits with sums, we will want to interchange sums with sums. To do this, we will first want to define what we mean by a double sum.

Definition 4.35. For a double sequence (x_{nm}) in $\mathbb{R}^d$, we say the double series $\displaystyle\sum_{n,m} x_{n,m}$ converges if the double sequence of (n,m)th partial sums converge.

Theorem 4.42 (Tonelli's Theorem for Series)**.** *If* $x_{m,n} \geq 0$ *for all* m, n*, then*

$$\sum_{m,n} x_{m,n} = \sum_m \sum_n x_{m,n} = \sum_n \sum_m x_{m,n},$$

even when these sums are infinite.

Exercise 4.36. Prove that if each $a_{m,n} \geq 0$, then

$$\sum_m \sum_n x_{m,n} = \sum_n \sum_m x_{m,n}.$$

This is an old prelim problem, so feel free to skip. If you want to go for it, try using MCT on partial sums

Theorem 4.43 (Fubini's Theorem for Series)**.** *Suppose* $\sum_{m,n} x_{m,n}$ *is absolutely convergent. Then,*

$$\sum_{m,n} x_{m,n} = \sum_m \sum_n x_{m,n} = \sum_n \sum_m x_{m,n}.$$

Propostion 4.10. *If the iterated series* $\displaystyle\sum_{m=1}^{\infty} \sum_{n=1}^{\infty} |x_{nm}|$ *converges, then the double series* $\sum x_{nm}$ *is absolutely convergent.*

4.8.2 ℓ^p spaces

For $1 \leq p < \infty$, we define the space $\ell^p(\mathbb{K})$ (or just ℓ^p when $\mathbb{K} = \mathbb{C}$) as the space of "p-summable sequences":

$$\ell^p(\mathbb{K}) = \{(a_n) \in \mathbb{K}^{\mathbb{N}} | \sum_{n=1}^{\infty} |a_n|^p < \infty\}.$$

We define ℓ^∞ to be the space of bounded sequences:

$$\ell^\infty(\mathbb{K}) = \{(a_n) \in \mathbb{K}^{\mathbb{N}} | \sup_n |a_n| < \infty\}.$$

5 Normed Vector Spaces

5.1 Vector Spaces

A **vector space** or **linear space** is a group V associated with a scalar field $\mathbb{K}$ (usually $\mathbb{R}$ or $\mathbb{C}$)[2] with two binary operations, vector addition and scalar multiplication such that the following hold for all $x, y, z \in V$ and $\alpha, \beta \in \mathbb{K}$

- $x + y = y + x$
- $(x + y) + z = x + (y + z)$
- There exists a $0 \in V$ such that $x + 0 = x$ for all $x \in V$.
- Given $x \in V$, there exists an element $-x \in V$ such that $x + (-x) = 0$.
- $1x = x$ for all $x \in V$.
- $\alpha(\beta x) = (\alpha\beta)x$.
- $\alpha(x + y) = \alpha x + \alpha y$.

Example 5.1. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^d$, $\mathbb{C}^d$, $M_{m,n}(\mathbb{R})$, $M_{m,n}(\mathbb{C})$, $\bigoplus_{n\in\mathbb{N}} M_n(\mathbb{C})$,...

Definition 5.1. An **algebra** over a field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ is a vector space that also has the multiplication binary operation $V \times V \to V$.

Note that $\mathbb{R}, \mathbb{C}, M_{n,n}(\mathbb{C}), M_{n,n}(\mathbb{R})$, and $\bigoplus_{n\in\mathbb{N}} M_n(\mathbb{C})$ are all algebras as well.

Definition 5.2. A basis of a vector space V is a collection of linearly independent vectors that span V. The **dimension** of V is the size of any basis of V.

Theorem 5.1. *Every vector space has a basis. (Requires choice axiom.)*

Example 5.2. The standard basis for $\mathbb{K}^d$ is $e_1, ..., e_d$ where e_k is the vector whose $k^t h$ component is 1 and all other components are zero.

[2]We will assume by default that the scalar field is $\mathbb{C}$. Algebraic closure is nice.

Definition 5.3. Let W, Y be subspaces of a vector space V. The **sum** of W and Y is

$$W + Y = \{w + y | w \in W, y \in Y\},$$

and the **direct sum** is

$$W \oplus Y = \{(w, y) | w \in W, y \in Y\}.$$

Note that $\dim(W + Y) \leq \dim(V)$, and $\dim(W \oplus Y) = \dim(W) + \dim(V)$.

Remark 5.1. *For finitely many spaces, direct sum $\oplus$ and direct product $\times$ are the same. For an infinite family of vector spaces $\{V_\alpha\}_{\alpha \in A}$,*

$$\prod_\alpha V_\alpha = \{(v_\alpha) | v_\alpha \in V_\alpha \text{ for each } \alpha \in A\},$$

and

$$\bigoplus_\alpha V_\alpha = \{(v_\alpha) \in \prod_\alpha V_\alpha | v_\alpha = 0 \text{ for all but finitely many } \alpha \in A\}.$$

Definition 5.4. Suppose W is a subspace of V. The quotient vector space V/W is the set of equivalence classes

$$[v] = \{y \in V : v - y \in W\}$$

for all $v \in V$. This has a natural vector space structure:

- $[v] + [y] = [v + y]$
- $[\alpha v] = \alpha[v]$.

Moreover, if $\dim(V) < \infty$, then $\dim(V/W) = \dim(V) - \dim(W)$.

5.1.1 Linear Transforms

Definition 5.5. A linear transform between vector spaces X and Y (over the same scalar field, which we assume to be $\mathbb{C}$) is a function $T : X \to Y$ that preserves linear structure, i.e. for any $x, y \in X$ and $\alpha, \beta \in \mathbb{K}$

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y).$$

Much like with matrices, we normally denote $T(x)$ by Tx.

In fact, let's see how we represent T using matrices. We'll stick to finite-dimensions for the moment as we will need to revise our definition of a basis for the infinite-dimensional topological case.

Let $\{x_1, ..., x_n\}$ and $\{y_1, ..., y_m\}$ be bases for vector spaces X and Y, respectively, and suppose $T : X \to Y$ is a linear transform. Then for each $1 \leq j \leq n$, there exist $a_{1,j}, ..., a_{m,j}$ such that

$$Tx_j = \sum_{i=1}^{m} a_{i,j} y_i.$$

By linearity, the values that T takes on $\{x_1, ..., x_n\}$ determine T and so, we may represent T as $T = \begin{pmatrix} a_{1,1} & a_{1,2} & ... & a_{1,n} \\ a_{2,1} & a_{2,2} & ... & a_{2,n} \\ ... & ... & ... & ... \\ a_{m,1} & a_{m_2} & ... & a_{m,n} \end{pmatrix}$

Notice that because the underlying field does have well-defined multiplication, the composition of two linear transforms $T : X \to Y$ and $S : Y \to Z$ is just the multiplication of the two linear transforms according to standard matrix multiplication.

Exercise 5.1. Show that the collection of linear transforms from a vector space X to itself forms an algebra.

Example 5.3. Again, $M_n(\mathbb{K})$ is a key example here.

Definition 5.6. A linear transform $T : X \to Y$ is **left invertible** if there exists a transform $S : Y \to X$ such that $ST = 1_X$; it is **right invertible** if there exists a transform $S : Y \to X$ such that $TS = 1_Y$. If T is left and right invertible, then its left and right inverses coincide, and we call denote the inverse transform by T^{-1}.

Exercise 5.2. Show that an injective function has a left inverse and a surjective function has a right inverse.

Propostion 5.1. *A linear transform $T : X \to X$ maps the basis of X to another basis iff T is invertible.*

We will actually be concerned with a stronger notion of change of basis, so no more will be said on this.

Theorem 5.2. *The* **kernel** *of a linear transform $T : X \to Y$ is the set*

$$\{x \in X |\ Tx = 0\}.$$

Propostion 5.2. *A linear transform $T : X \to Y$ is injective iff $ker(T) = \{0\}$.*

Theorem 5.3 (Rank Plus Nullity)**.** *Let $T : X \to Y$ be a linear transformation between vector spaces. The* **rank** *of T is the dimension of the range of T, and the* **nullity** *of T is the dimensionl of the kernel.*
If X is a finite-dimensional vector and $T : X \to Y$ any linear transformation, then ran(T)+ nul(T) =dim(X).

Exercise 5.3. Show that for X and Y finite-dimensional vector spaces, the following are equivalent:

1. T is injective.

2. T is surjective.

3. T is bijective.

Definition 5.7. The **spectrum** of a linear transformation from X to X is the set $\{\lambda \in \mathbb{K} | T - \lambda 1_X$ is not invertible$\}$.

Theorem 5.4 (First Isomorphism Theorem)**.** *For vector spaces X and Y, each linear transformation $T : X \to Y$ induces a linear isomorphism $\hat{T} : X/ker(T) \to Im(T)$.*

Definition 5.8. A **linear functional** on a vector space X is a linear transform from X to its underlying scalar field. The **algebraic dual** of a vector space X is the space of all linear functionals on X. Algebraists often denote this as X^*, but we reserve that notation for the continuous dual space (for which we will need a topology) and denote the algebraic dual space by X'.

Exercise 5.4. Check that the dual space X' of X is also a vector space.

5.2 Norms

Definition 5.9. A **norm** on a vector space V (over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$) is a function $\|\cdot\| : V \to \mathbb{R}$ satisfying the following properties:

1. $\|x\| \geq 0$ for all $x \in V$. (non-negative)
2. $\|x\| = 0$ iff $x = 0$. (definite)
3. $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{K}$ and $x \in V$. (absolute homogeneity)
4. $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$. (sub-additive)

A vector space with a norm is a **normed space**.

Remark 5.2. *(1) and (2) together yield "positive definiteness". If you only know that $x = 0 \Rightarrow \|x\| = 0$, then $\|\cdot\|$ is called a* **semi-norm***. (4) is sometimes called the triangle-inequality, but "sub-additivity" will help prime you for additivity/sub-additivity criteria that will come up in measure theory.*

A norm on a vector space V induces a metric:

$$\rho(x, y) = \|x - y\| \ \forall \ x, y \in V$$

Example 5.4. The natural norm on $\mathbb{R}$ is the absolute value function. Then norm on $\mathbb{C}$ is the modulus function. These are specific examples of the Euclidean norm on $\mathbb{K}^d$ where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$:

$$\|(x_1, ..., x_d)\|_2 = \sqrt[2]{\sum_{k=1}^{d} |x_k|^2}$$

for $(x_1, ..., x_d) \in \mathbb{K}^d$.

We will consider the Euclidean norm to be the standard norm and its induced topology the standard topology on $\mathbb{R}^d$ and $\mathbb{C}^d$ (i.e. if no topology is mentioned, this is the one assumed), and we will often drop the subscript. (If we get far enough, $\|\cdot\|_2$ will come to stand for a different norm entirely.)

Exercise 5.5. With respect to this metric space, $\|\cdot\|$ is a continuous function, i.e. if $x_n \to x$ then $\|x_n\| \to \|x\|$. Moreover, vector addition

$$+ : V \times V \to V$$

and scalar multiplication

$$\cdot : \mathbb{K} \times V \to V$$

are jointly continuous, i.e. if $x_n \to x$ and $y_n \to y$ in X, then $x_n + y_n \to xy$ and if $a_n \to a$ in $\mathbb{K}$, then $a_n x_n \to ax$. Moreover, if V is also an algebra, then $x_n y_n \to xy$.

Propostion 5.3. *Just as with $|\cdot|$ on $\mathbb{C}$ and $\mathbb{R}$, we have a for any normed vector space, i.e. for any normed vector space V and any $x, y \in V$,*

$$\|x - y\| \geq |\|x\| - \|y\||.$$

Propostion 5.4. *Prove that for normed vector spaces X and Y, a compact subset $K \subseteq X$, and a continuous function $f : K \to Y$, there exists a $k \in K$ such that $\|f(k)\| = \max_{x \in K} \|f(x)\|_Y$.*

Proof. Since the composition $\|\cdot\|_Y \circ f$ is continuous, the image of K under $\|\cdot\|_Y \circ f$ is a compact subset of $\mathbb{R}$ and hence has a maximum. □

Exercise 5.6. Verify that the following spaces are normed vector spaces:

1. $B(X, \mathbb{R})$ where X is some metric space with norm $\|f\|_X = \sup_{x \in X} |f(x)|$. This is often called the **uniform** or **sup** norm.

2. $C_b(X, \mathbb{R})$ with the same norm. (You don't have to verify that $\|\cdot\|_X$ is a norm twice.)

Definition 5.10. If endowed with a norm, an algebra may have an additional property of **sub-multiplicativity**: $\|xy\| \leq \|x\|\|y\|$ for all $x, y \in V$.

Example 5.5. The key example we have right now are $n \times n$ matrices over $\mathbb{K}$: $M_n(\mathbb{K})$. We shall see soon enough that these are actually better realized as a class of continuous functions on a finite-dimensional Hilbert space.
$\mathbb{C}$ and $\mathbb{R}$ are actually multiplicative.

Example 5.6. There exist other norms on $\mathbb{K}^{\mathbb{N}}$ where $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ given by

- $\|(x_n)\|_\infty = \sup\limits_{k \in \mathbb{N}} |x_k|$ for all $(x_1, ..., x_n) \in \mathbb{K}^{\mathbb{N}}$

- $\|(x_n)\|_p = \sqrt[p]{\sum\limits_{k=1}^{\infty} |x_k|^p}$ for $1 \leq p < \infty$ $(x_n) \in \mathbb{K}^{\mathbb{N}}$

The subspace of $\mathbb{K}^{\mathbb{N}}$ that is closed with respect to $\|\cdot\|_p$ is $\ell^p(\mathbb{K})$

Exercise 5.7. Show that $\|\cdot\|_\infty$ and $\|\cdot\|_1$ are both norms on $\mathbb{K}^{\mathbb{N}}$.

Exercise 5.8. Prove Hölder's inequality for sequences:
For $p \in (1, \infty)$, define $p' \in [1, \infty]$ to be the unique number so that $\frac{1}{p} + \frac{1}{p'} = 1$. For $p = 1$, let $p' = \infty$, and for $p = \infty$, let $p' = 1$. Now, show that for any $p \in [1, \infty]$ and any $x, y \in \mathbb{K}^{\mathbb{N}}$

$$\|xy\|_1 \leq \|x\|_p \|y\|_{p'}.$$

Hint: Easy if $p = 1$ or $p = \infty$. If $1 < p < \infty$, first consider the case where $\|x\|_p = \|y\|_{p'} = 1$ and use Young's Inequality[3]: for $a, b \in [0, \infty)$,

$$ab \leq \frac{1}{p} a^p + \frac{1}{p'} b^{p'}.$$

Argue that you may assume $\|x\|_p$ and $\|y\|_{p'} \neq 0$ and consider $\tilde{x} = \frac{x}{\|x\|_p}$ and $\tilde{y} = \frac{y}{\|y\|_{p'}}$.

Exercise 5.9. Prove that the map $\| \cdot \|_p$ on $\ell^p(\mathbb{K})$ is a norm for $1 < p < \infty$. *Hint: To prove subadditivity, Use Hölder's inequality and the inequality[4]*

$$|a + b|^p \leq |a||a + b|^{p-1} + |b||a + b|^{p-1}$$

for all $a, b \in \mathbb{K}$.

Definition 5.11. We say two norms $\| \cdot \|_\alpha$ and $\| \cdot \|_\beta$ are **equivalent** if they induce the same topology.

Propostion 5.5. *Two norms $\| \cdot \|_\alpha$ and $\| \cdot \|_\beta$ are equivalent iff there exist constants $a, b \in \mathbb{R}$ such that $\|x\|_\alpha \leq a\|x\|_\beta$ and $\|x\|_\beta \leq b\|x\|_\alpha$ for all $x \in X$.*

Theorem 5.5. *Any normed vector space of dimension $n < \infty$ is homeomorphic to $\mathbb{K}^n$ with the Euclidean topology.*

The proof relies on the Heine-Borel and fails in infinite dimensions.

Remark 5.3. *If $0 < p < 1$, then the above definition does not yield a norm. However, we can still define a metric on $\mathbb{R}^d$ by*

$$d_p(x, y) = \sum_n |x_n - y_n|^p$$

for all $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n) \in \mathbb{R}^d$.

Example 5.7. If we have only $\mathbb{K}^d$ for some $d < \infty$, then we define the $\| \cdot \|_p$ norm for $1 \leq p \leq \infty$ as before, only now these are all norms on the same space, $\mathbb{K}^d$.

When endowed with the $\| \cdot \|_p$ norm for $1 \leq p \leq \infty$, we denote $(\mathbb{K}^d, \| \cdot \|_p)$ as $\ell^p_d(\mathbb{K})$. Notice that $\ell^2_d(\mathbb{K})$ is $\mathbb{K}^d$ with the standard Euclidean norm.

[3] This estimate is a consequence of the convexity of the exponential function, and may be used without proof for this exercise.

[4] again without proof

Exercise 5.10. Try sketching the "ball" of radius 1 centered at the origin in $\ell^p_2(\mathbb{R})$ for $p = 1, 2$, and ∞, i.e.

1. $B_1^\infty = \{(x_1, x_2) \in \mathbb{R}^2 | \max_{k=1,2} |x_k| \leq 1\}$
2. $B_1^2 = \{(x_1, x_2) \in \mathbb{R}^2 | \sqrt{x^2 + y^2} \leq 1\}$
3. $B_1^1 = \{(x_1, x_2) \in \mathbb{R}^2 | |x_1| + |x_2| \leq 1\}$
4. What do you think the unit ball looks like for $p \in (0, 1)$?
5. Which of these spaces have a convex unit ball?

Now that we have a topology, many of our algebraic notions can be adapted with closures with respect to this topology. For instance, we will now want to take quotients of closed subspaces as opposed to just subspaces[5].

Definition 5.12. For a family $\{V_\alpha\}_{\alpha \in A}$ of normed vector spaces, the closure of the direct sum $\bigoplus_\alpha V_\alpha$ with respect to the norm is

$$\sum_\alpha^\oplus V_\alpha = \{(v_\alpha) \in \prod_\alpha V_\alpha : \|v_\alpha\| \to 0\}.$$

Definition 5.13. A normed vector space that is complete as a metric space is called a **Banach Space**.

Example 5.8. The following are some examples of complete vector spaces

- Any finite-dimensional normed vector space.
- $\ell^p(\mathbb{K})$ (This is a special case of a theorem due to Riesz.)
- $B(X, \mathbb{K})$
- $C_b(X, \mathbb{K})$

The standard proof that $\ell^p(\mathbb{K}))$ is complete uses the following characterization.

Theorem 5.6. *A normed vector space X is complete iff every absolutely convergent series in X converges in X.*

For a proof, see Folland Theorem 5.1

5.3 Linear Operators

Definition 5.14. Let X and Y be normed vector spaces (over the same field $\mathbb{K}$). We call a linear transform $T : X \to Y$ **bounded** if there exists a constant $c \in (0, \infty)$ such that for all $x \in X$,

$$\|Tx\|_y \leq c\|x\|_X.$$

[5]This guarantees that the quotient will be a normed vector space.

Theorem 5.7. *Let $T : X \to Y$ be a linear transform. Let $B_1 = \{x \in X | \; \|x\| \leq 1\}$ be the unit ball in X and $S_1 = \{x \in X | \; \|x\| = 1\}$ be the unit sphere in X. Then the following are equivalent:*

1. *T is bounded.*
2. *$T(S_1)$ is bounded.*
3. *$T(B_1)$ bounded.*
4. *T is continuous at 0.*
5. *T is Lipschitz continuous.*

We will prove (1) ⇔ (2) and leave some other pieces for homework.

Proof. Suppose $\|Tx\| \leq c\|x\|$ for all $x \in X$. Then $T(S_1) \subseteq B_c(0)$. On the other hand, suppose $T(S_1) \subseteq B_c(0)$ for some $c > 0$, and let $x \in X$. Clearly $T0 = 0$, so we assume $x \neq 0$. Then $\frac{x}{\|x\|} \in S_1$, and we have

$$\begin{aligned} \left\| T \frac{x}{\|x\|} \right\| &\leq c \\ \left\| \frac{1}{\|x\|} Tx \right\| &\leq c \\ \frac{1}{\|x\|} \|Tx\| &\leq c \\ \|Tx\| &\leq c\|x\|. \end{aligned}$$

Since this holds for any $x \in X$, we are done. □

Exercise 5.11. Prove (1) ⇔ (3) and (1) ⇔ (5)

Remark 5.4. *T being bounded does* **not** *mean Im(T) is a bounded subset. In fact, if T is nonzero, this will never hold. This means that T maps bounded sets to bounded sets.*

We denote the space of bounded linear transforms between X and Y by $B(X, Y)$. (Some texts use $L(X, Y)$.)

Definition 5.15. We define a norm on the space of of linear transforms between two normed vector spaces X and Y by

$$\|T\| = \inf\{c \in (0, \infty) | \; \|Tx\| \leq c \; \forall \; x \in X\}.$$

A bounded linear transform is usually called a **linear operator**.

Remark 5.5. *This is often expressed as*

$$\|Tx\| \leq \|T\|\|x\|$$

for all $x \in X$. We call T a **contraction** *if $\|T\| \leq 1$, in which case, $\|Tx\| \leq \|x\|$ for all $x \in X$.*

Propostion 5.6. *For any $T \in B(X,Y)$, we have the following*

$$\begin{aligned} \|T\| &= \sup_{x \in B_1^X(0)} \|Tx\| \\ &= \sup_{x \in S_1^X} \|Tx\| \end{aligned}$$

Theorem 5.8. *Given any normed linear space X and $d < \infty$, any linear operator $T : \mathbb{K}^d \to X$ is bounded (w.r.t. the Euclidean norm on $\mathbb{K}^d$).*

Proof. Let $\{e_1, ..., e_d\}$ be the standard basis of $\mathbb{K}^d$, and let

$$c = (\sum_{k=1}^{d} \|Te_k\|^2)^{1/2}.$$

Let $x \in \mathbb{K}^d$. Then $x = \sum_{k=1}^{d} a_k e_k$ for some $a_1, ..., a_d \in \mathbb{K}$,[6] and moreover $\|x\| = (\sum_{k=1}^{d} |a_k|^2)^{1/2}$. Then,

$$\begin{aligned} \|Tx\| &= \|\sum_{k=1}^{d} a_k T e_k\| \\ &\le \sum_{k=1}^{d} |a_k| \|Te_k\| \\ &\le (\sum_{k=1}^{d} |a_k|^2)^{1/2} (\sum_{k=1}^{d} \|Te_k\|^2)^{1/2} \\ &= \|x\| c \end{aligned}$$

The last line following from Hölder's Inequality for $p = p' = 2$. ☐

We will see examples of unbounded operators after we get to Hilbert spaces. In particular, note that

$$\mathcal{L}(\mathbb{K}^n, \mathbb{K}^m) \simeq M_{m,n}(\mathbb{K}).$$

Example 5.9. Find the norms of the following matrices in $M_n(\mathbb{K})$

1. $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M_2(\mathbb{C})$

2. $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in M_2(\mathbb{R})$

We don't yet have the tools to handle 2 in the case of $M_2(\mathbb{C})$, but the norm is still $\sqrt{2}$ for reference.

[6] This is just a linear expression of $x = (a_1, ..., a_d)$.

Exercise 5.12. Determine the norm of $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$ in $M_2(\mathbb{R})$.

Hint: Lagrange Multipliers

Exercise 5.13. Suppose $A \in M_n(\mathbb{C})$ is a diagonal matrix with diagonal entries $(\lambda_1, \lambda_2, ..., \lambda_n)$. What is $\|A\|$?

5.4 Dual Spaces

Definition 5.16. A **bounded linear functional** is a linear operator $\phi : X \to \mathbb{K}$ where $\mathbb{K}$ is the underlying scalar field. The **dual space** of X is the normed vector space consisting of all bounded linear functionals on X and endowed with the operator norm. We denote this by X^*.

Definition 5.17. Suppose $\{x_\alpha\}_{\alpha\in\mathcal{A}}$ is a basis for X. This basis corresponds with a linearly independent set $\{x^*_\alpha\}_{\alpha\in\mathcal{A}}$ in X^* where $x^*_\alpha(x_\beta) = \delta_{\alpha,\beta}$. Where $\delta_{\alpha,\beta}$ is the Kroneker delta symbol, i.e $\delta_{\alpha,\beta} = 1$ when $\alpha = \beta$ and 0 when $\alpha \neq \beta$. When this set is spanning, we call it a **dual basis** of X^*.

Note that this implies that the dimension of X^* is at least as large as the dimension of X.

In case $X = \mathbb{K}^d$ for some $d < \infty$ with standard basis $\{e_1, ..., e_d\}$, the corresponding $e^*_k \in (\mathbb{K}^d)^*$ are the dual basis of $(\mathbb{K}^d)^*$ is the family of k^{th}-coordinate projections, i.e. $e^*_k = \pi_k$.

Exercise 5.14. Show that $\mathbb{K}^d$ is isomorphic to its dual, i.e. show that any $\phi : \mathbb{K}^d \to \mathbb{K}$ can be written as

$$\sum_{k=1}^{d} a_k \pi_k$$

for some $a_1, ..., a_d \in \mathbb{K}$.

Because all finite-dimensional normed vector spaces are isomorphic, we can conclude the following theorem:

Theorem 5.9. *If X is a finite-dimensional normed vector space, then $X = X^*$.*

Theorem 5.10. *If X is a Banach space, then so is X^*.*

5.4.1 Hahn-Banach (Corollaries)

It is not so clear that normed vector space should have many nontrivial linear functionals. That they exist in abundance is due to the Hahn-Banach Theorem, one of the fundamental results in functional analysis. The theorem itself is used much less frequently than its many corollaries.

Theorem 5.11 (Real Hahn-Banach)**.** *Let X be a vector space over $\mathbb{R}$, $p : X \to \mathbb{R}$ be positively homogeneous and subadditive, $Y \subseteq X$ a linear subspace, and $f \in Y^*$ such that $f(x) \leq p(x)$ for all $x \in Y$. Then, there exists $F \in X^*$ such that $F(x) \leq p(x)$ for all $x \in X$ and $F|_Y = f$.*

It takes more shape in the complex version:

Theorem 5.12 (Complex Hahn-Banach)**.** *Let X be a vector space over $\mathbb{C}$, p a semi-norm on X, $Y \subseteq X$ a linear subspace, and $f \in Y^*$ such that $|f(x)| \leq p(x)$ for all $x \in Y$. Then there exists an $F \in X^*$ such that $|F(x)| \leq p(x)$ for all $x \in X$ and $F|_Y = f$.*

Corollary 5.1. *Let X be a normed vector space over $\mathbb{C}$ and $Y \subseteq X$ a linear subspace of X.*

1. *If $f \in Y^*$, then there exists a $g \in X^*$ such that $g|_Y = f$ and $\|g\|_{X^*} = \|f\|_{Y^*}$.*
2. *Suppose $x \in X \backslash \{0\}$. Then there exists a $g \in X^*$ with $\|g\|_{X^*} = 1$ and $g(x) = \|x\|$.*
3. *If $x \in X \backslash Y$ is such that $d = \inf_{y \in Y} \|x - y\| > 0$, then there exists a $g \in X^*$ such that $g(x) = d$ and $g|_Y = 0$.*
4. *The bounded linear functionals on X separate points.*
5. *If $x \in X$, define $\hat{x} : X^* \to \mathbb{C}$ by $\hat{x}(f) = f(x)$ for all $f \in X^*$. Then the map $x \mapsto \hat{x}$ is a linear isometry from X into X^{**}. If X is a Banach space, then $\hat{X}$ is complete in X^{**}.*

Exercise 5.15. Assuming 1-4, prove 5. (You wont need all of 1-4).

Definition 5.18. We say a Banach space is **reflexive** if the map $x \mapsto \hat{x}$ is surjective, i.e. if $X \simeq X^{**}$.

5.4.2 $(\ell^p)^*$

Theorem 5.13. *For $p \in [1, \infty)$, let p' denote the number in $(1, \infty)$ for which $\frac{1}{p} + \frac{1}{p'} = 1$ (with the convention that $\frac{1}{\infty} = 0$). Then,*

$$(\ell^p)^* \simeq \ell^{p'}.$$

We will not prove this, but we will indicate the map $T : \ell^{p'} \to (\ell^p)^*$. For each $y = (y_n) \in \ell^{p'}$ we define $T(y)$ for each $x = (x_n) \in \ell^p$ by

$$T(y)x = \sum_n x_n y_n.$$

It turns out T is a bijective bounded linear operator with $\|T(y)\| = \|y\|$ for each $y \in \ell^{p'}$..

Remark 5.6. *However, $(\ell^\infty)* \neq \ell^1$. In fact, $\ell^1 = c_0^*$.*

Remark 5.7. *Notice that ℓ^2 is the only ℓ^p space that is self-dual, which is due to the fact that $\frac{1}{2} + \frac{1}{2} = 1$.*

5.4.3 Weak*-topology

Definition 5.19. We can now define a new topology on X as a (subspace of a) dual space. The **weak topology** is cumbersome to define via basic open sets. Instead, we "define" it be giving convergence criteria: We say a sequence (x_n) **converges in the weak topology on** X (or "weakly") iff the sequence $(x_n(\phi))$ converges in $\mathbb{K}$ for any $\phi \in X^*$.

Exercise 5.16. Prove that if (x_n) converges in norm, then it converges weakly. Recalling some earlier exercises, remark on why the norm topology is sometimes called the "strong" topology on a Banach space.

Remark 5.8. *The "dual" to this topology defined on X^* is called the weak*-topology. We say (ϕ_n) converges weak* in X^* iff $(\phi_n(x))$ converges for any $x \in X$. Think of this as a topology of pointwise convergence.*

Why do we care so much about these weaker topologies? Well, sometimes the weaker topology is just nicer. For example, note that a strongly compact set is also weakly compact. However, we have the two following theorems.

Theorem 5.14 (Riesz's Lemma)**.** *Let X be a normed vector space and $Y \subseteq X$ a closed proper subspace and $\alpha \in (0,1)$. Then there exists an $x \in B_1^x$ such that $\|x - y\| \geq \alpha$ for all $y \in Y$.*

Exercise 5.17. Use Riesz's Lemma to prove that no infinite-dimensional vector space has a compact unit ball by inductively building a sequence (x_n) in B_1^X so that $\|x_m - x_n\| > \frac{1}{2}$ for every $m \neq n$. *Hint: Build the sequence inductively. Take x_1 to be any element in B_1. For $n \geq 1$, let Y_n be the linear span of $x_1, ..., x_n$. (Note that since dim$(Y) < \infty$, Y is a closed subspace of X – remember how all norms on finite-dimensional spaces are the same and finite dimensional spaces are complete?) Recall that on a metric space compactness is the same as sequential compactness. Reach your contradiction accordingly.*

Theorem 5.15 (Banach-Alaoglu)**.** *Let X be a normed vector space. Then the unit ball in X^* is weak* compact.*

Remark 5.9. *Recall that if X is a finite-dimensional normed vector space, then $X \simeq X^*$. In particular, the strong and weak topologies coincide.*

5.5 Inner Product Spaces

Definition 5.20. An **inner product** (**dot product**) on a vector space V over a field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$ is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ satisfying the properties:

1. $\langle x, x \rangle \geq 0$ for all $x \in V$
2. $\langle x, x \rangle = 0$ iff $x = 0$
3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$ for all $x, y \in V$ (conjugate symmetry)

4. $\langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle$ and $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$ for all $\alpha \in \mathbb{K}$ and $x, y \in V$ (linearity in first argument)

5. $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$ and $\langle x, \alpha y\rangle = \overline{\alpha}\langle x, y\rangle$ for all $\alpha \in \mathbb{K}$ and $x, y \in V$ (conjugate linearity in second argument)

A vector space with an inner product is an inner product space.

Remark 5.10. *The "conjugate" bits are redundant over* $\mathbb{R}$*.*

Example 5.10. $\mathbb{R}^d$ and $\mathbb{C}^d$ are inner product spaces with

$$\langle x, y\rangle = \sum_{i=1}^{d} x_i \bar{y_i}$$

for all $x = (x_1, ..., x_d), y = (y_1, ..., y_d) \in \mathbb{K}^d$.

Example 5.11. $\ell^2(\mathbb{K})$ is an inner product space with inner product

$$\langle x, y\rangle = \sum_n x_n \bar{y_n}.$$

An inner product induces a norm on V by

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

Example 5.12. The norm induced by the aforementioned inner product on $\mathbb{R}^d$ and $\mathbb{C}^d$ is the Euclidean norm:

$$\|x\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

for all $x = (x_1, ..., x_n) \in \mathbb{K}^d$.

The norm induced on $\ell^2(\mathbb{K})$ by the inner product is the $\|\cdot\|_2$ norm.

Theorem 5.16 (Cauchy-Schwarz Inequality)**.** *If* V *is an inner product space over* $\mathbb{K} = \mathbb{R}$ *or* $\mathbb{C}$*, then for any* $x, y \in V$*,*

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

Moreover, if $y \neq 0$*, then equality holds iff there exists* $c \in \mathbb{K}$ *such that* $x = cy$*.*

Exercise 5.18. Prove the Cauchy-Schwarz inequality for the inner product on ℓ^2.

Exercise 5.19. Use the Cauchy-Schwarz inequality to show that $\sqrt{\langle x, x\rangle}$ gives a norm on an inner product space V over $\mathbb{R}$. The same argument will work for V over $\mathbb{C}$, but one needs to first note that $Re(\alpha) \leq |\alpha|$ for any $\alpha \in \mathbb{C}$, where $Re(\alpha)$ is the real part of $\alpha = a + ib$.

Corollary 5.2. *The inner product is a jointly continuous function with respect to the norm topology on V and the standard topology on $\mathbb{R}$.*

Definition 5.21. A complete inner product space is called a **Hilbert space**.

Example 5.13. Euclidean spaces are Hilbert spaces. ℓ^2 is also a Hilbert space.

Theorem 5.17 (Riesz Representation Theorem)**.** *A Hilbert space is (conjugate-linearly) isomorphic to its own dual by identifying $x \in \mathcal{H}$ with $x^* \in \mathcal{H}^*$ defined by*

$$x^*(y) = \langle y, x \rangle$$

for all $y \in \mathcal{H}$.

From this and the fact that for $1 \leq p < \infty$, $(\ell^p)^* = \ell^{p'}$, we see that ℓ^p is a Hilbert space exactly when $p = 2$.

5.5.1 An orthonormal basis

Definition 5.22. Two vectors $x, y \in V$ are **orthogonal** if $\langle x, y \rangle = 0$.
If $W \subseteq V$ is a subspace, then we say that the **orthogonal complement of** W is

$$W^\perp = \{v \in V : \langle v, w \rangle = 0 \ \forall w \in W\}.$$

Two orthogonal vectors are **orthonormal** if $\|x\| = \|y\| = 1$.

Definition 5.23. An **orthonormal basis** for a Hilbert space is a subset $\{e_a\}_{a \in \mathcal{A}} \subseteq \mathcal{H}$ satisfying the following properties:

1. Orthogonality: If $a, b \in \mathcal{A}$ with $a \neq b$, then

$$\langle e_a, e_b \rangle = 0.$$

2. Normalization: $\|e_a\| = 1$ for each $a \in \mathcal{A}$, and
3. Completeness: $\overline{\text{Span}(\{e_a\}_{a \in \mathcal{A})}} = \mathcal{H}$.

Completeness can be rephrased to state that for any $x \in \mathcal{H}$.

Example 5.14. The standard basis vectors for $\mathbb{K}^d$ where $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ are $e_j = (\delta_{j1}, \delta_{j2}, ..., \delta_{jn})$ where $\delta_{jk} = 1$ when $j = k$ and 0 when $j \neq k$. We can similarly define basis vectors for $M_n(\mathbb{C})$ and $M_n(\mathbb{R})$, but these are just isomorphic to $\mathbb{C}^{n^2}$ and $\mathbb{R}^{n^2}$ anyway.

Example 5.15. The basis for ℓ^2 is the collection $\{e_j\}_{j \in \mathbb{N}}$ defined by $e_j = (\delta_{j1}, \delta_{j2}, ...)$. Since this basis is countable, we can still give a matrix representation for linear functionals mapping $\ell^2 \to \ell^2$.

With respect to this basis, we may represent an linear operator $T : \ell^2 \to \ell^2$ as an infinite matrix:

$$T = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots \\ a_{2,1} & a_{2,2} & \cdots \\ \vdots & \vdots & \cdots \end{pmatrix}$$

Exercise 5.20. Using a matrix representation, it's pretty easy to give an unbounded operator on ℓ^2. Give an example of one.

Exercise 5.21. A Hilbert space is separable iff it has a countable orthonormal basis.

Exercise 5.22. Assume $\|\cdot\|$ is the norm induced by the inner product on V.

1. Show that $x, y \in V$ are orthogonal iff $\|x+y\|^2 = \|x\|^2 + \|y\|^2$.
2. If $x_n \to x$ and $y_n \to y$ in norm (i.e. $\|x_n - x\| \to 0$), then the sequence $(\langle x_n, y_n\rangle)$ converges to 0 in $\mathbb{R}$. What about the double sequence $\langle x_n, y_m\rangle$?
3. Parallelogram Law: ("The sum of the squares of the diagonals of a parallelogram is the sum of the squares of the four sides.") For all $x, y \in V$:

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2).$$

   (Hint: Over $\mathbb{R}$ add the two formulas $\|x \pm y\|^2 = \|x\|^2 \pm 2\langle x, y\rangle + \|y\|^2$. Over $\mathbb{C}$ add the two formulas $\|x \pm y\|^2 = \|x\|^2 \pm 2\mathrm{Re}\langle x, y\rangle + \|y\|^2$.)

Theorem 5.18 (Gram-Schmidt Theorem)**.** *If $v_1, ..., v_k \in V$ are linearly independent, then there are k orthonormal vectors $v_1', ..., v_k' \in V$ such that span$\{v_1', ..., v_k'\} =$ span$\{v_1, ..., v_k\}$.*

Theorem 5.19. *Let $\{e_\alpha\}_{\alpha \in A}$ be an orthonormal basis of a inner product space V.*

1. *(Completeness) $\langle v, e_\alpha\rangle = 0$ for all $\alpha \in A$ iff $v = 0$.*
2. *(Parseval's Identity)*

$$\sum_{\alpha \in A} |\langle v, e_\alpha\rangle|^2 = \|v\|^2$$

   In particular, $\{\alpha : \langle v, e_\alpha\rangle \neq 0\}$ is countable.
3. *For each $v \in V$,*

$$v = \sum_{\alpha \in A} \langle v, e_\alpha\rangle e_\alpha,$$

   where the sum converges in the norm topology no matter how these terms are ordered.

Definition 5.24. The series in (3) is called the **Fourier series decomposition** of the vector v, and the coefficients $\langle v, e_\alpha\rangle$ are called the **Fourier coefficients** of v.

Remark 5.11. *We can also define a basis for an infinite-dimensional Banach space X as being a set of linearly independent elements whose span is dense in X. For instance*

- *The set $\{x^n\}_{n\geq 0}$ forms a basis for $C([a,b],\mathbb{R})$.*
- *The set $\{z^n\}_{n\in\mathbb{Z}}$ forms a basis for $C(\mathbb{T},\mathbb{C})$ where $\mathbb{T}$ is the unit circle in $\mathbb{C}$ (a.k.a. the 1-torus).*[7]

They form an orthonormal basis for the Hilbert space $L^2([0,1],\mathbb{R})$ of "square-integrable functions", i.e. $\{f : [a,b] \to \mathbb{R} | \int |f|^2 < \infty\}$,[8] *with inner product*

$$\langle f, g\rangle = \int_{[0,1]} fg.$$

Given any $f \in L^2([0,1],\mathbb{R}$, the Fourier coefficients are

$$\langle f, x^n\rangle = \int_{[0,1]} fx^n$$

for each $n \geq 0$, and the **Fourier series** *is*

$$\sum_{n=0}^{\infty} \langle f, x^n\rangle x^n.$$

A similar construction can be done for square integrable functions on $\mathbb{T}$, but we would first need to normalize to ensure orthogonality.

5.5.2 Bounded Operators on $\mathcal{H}$

Some of the following results do not hold over $\mathbb{R}$, so we assume henceforth that the underlying scalar field is $\mathbb{C}$.

These can be defined for linear operators between normed vector spaces X and Y, but they are easiest to describe and work with when $X = Y$ is a Hilbert space.

We denote the Banach space of bounded linear operators from $\mathcal{H}$ to itself by $B(\mathcal{H})$ or $L(\mathcal{H})$.

Definition 5.25. If $T \in B(\mathcal{H})$, then there exists a unique $T^* \in B(\mathcal{H})$ for which, for all $x, y \in \mathcal{H}$,

$$\langle Tx, y\rangle = \langle x, T^* y\rangle.$$

T^* is called the **adjoint** of T.

Example 5.16. For $T \in M_n(\mathbb{C})$, T^* is the conjugate transpose of T.

[7] If we switch to the polar form, i.e. $z = re^{i\theta} = r(\cos\theta + i\sin\theta)$ for $\theta \in [0, 2\pi)$, we see why linear combinations of these are called **trigonometric polynomials**.

[8] This is Lebesgue integrable, not Riemann integrable.

For operators on a Hilbert space, we have an involution operation $T \mapsto T^*$, which gives us the following properties of the adjoint:

Propostion 5.7. *Properties of the adjoint of $S, T \in B(\mathcal{H})$:*

1. $(T^*)^* = T$
2. $(T+S)^* = T^* + S^*$
3. $(TS)^* = S^*T^*$
4. $(\lambda T)^* = \bar{\lambda} T^*$ *for* $\lambda \in \mathbb{K}$.

Definition 5.26. $T \in B(\mathcal{H})$ is called an **isometry** if $T^*T = 1$; it is called a *co-isometry* if $TT^* = 1$.
$U \in B(\mathcal{H})$ is called a **unitary** if $U^*U = UU^* = 1$.

Example 5.17. The unilateral shift operator S on ℓ^2 is defined for all $(x_n) \in \ell^2$ by

$$S(x_1, x_2, ...) = (0, x_1, x_2, ...).$$

Its adjoint is defined by

$$S^*(x_1, x_2, ...) = (x_2, x_3, ...).$$

Then $S^*S = I$, but $SS^* \neq I$. In particular, S is left invertible but not right invertible.

Theorem 5.20. *For $U \in B(\mathcal{H})$, the following are equivalent:*

1. *U is a unitary.*
2. *U is an orthonormal change of basis operator.*
3. $\langle Ux, Uy \rangle = \langle x, y \rangle$ *for all* $x, y \in \mathcal{H}$.

Moreover, if dim$(\mathcal{H}) < \infty$, $U \in B(\mathcal{H})$ is a unitary iff it is an isometry.

In $M_n(\mathbb{C})$ a matrix U is unitary iff the columns of U form an orthonormal basis of $\mathbb{C}^n$.

Definition 5.27. An operator $T \in B(\mathcal{H})$ is **normal** if $T^*T = TT^*$. An operator $T \in B(\mathcal{H})$ is called **self-adjoint** (Hermitian) if $T = T^*$.

Note that T^*T is self-adjoint for any operator T.

Remark 5.12. $T \in M_n(\mathbb{R})$ *is Hermitian iff it is symmetric.*

Theorem 5.21. *For $T \in B(\mathcal{H})$ we have the following:*

1. $\|T\| = \|T^*\|$
2. $\|T^*T\| = \|T\|^2$

3. $\sup_{\lambda \in \sigma(T)} |\lambda| \leq \|T\|$

4. *If T is normal, then* $\sup_{\lambda \in \sigma(T)} |\lambda| = \|T\|$

Remark 5.13. *If $T \in M_n(\mathbb{C})$, then its spectrum is the set of its eigenvalues.*

Putting the pieces of the previous theorem together, we have for any $T \in M_n(\mathbb{C})$

$$\|T\| = \sup_{\lambda \in \sigma(T^*T)} \sqrt{\lambda}.$$

Exercise 5.23. Find the norms for the following matrices in $M_2(\mathbb{C})$.

1. $\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$

2. $\begin{pmatrix} 2 & -1 \\ 1 & 2 \end{pmatrix}$

Exercise 5.24. Could we define a norm $\| \cdot \|_*$ on $M_n(\mathbb{C})$ be defining

$$\|T\|_* = \sup_{\lambda \in \sigma(T)} \sqrt{\bar{\lambda}\lambda} = \sup_{\lambda \in \sigma(T)} |\lambda|?$$

Theorem 5.22. *If $T \in M_n(\mathbb{C})$, then T is normal iff it is unitarily diagonalizable, i.e. it is a diagonal matrix up to an orthonormal change of basis of $\mathbb{C}^n$. Any matrix is unitarily equivalent to an upper triangular matrix.*

Remark 5.14. *A matrix is only* **similar** *to its Jordan Canonical form. Conjugation by an invertible matrix does not necessarily preserve things like the adjoint or the norm of an operator. However, conjugation by a unitaries does.*

Theorem 5.23. *Recall that the spectrum of $T \in B(\mathcal{H})$ is*

$$\sigma(T) = \{\lambda \in \mathbb{C} | \ \lambda I - T \textit{ does not have a bounded inverse}\}.$$

Then for any $T \in B(\mathcal{H})$, $\sigma(T)$ is a compact subset of $\mathbb{C}$, and hence a compact Hausdorff space.

Theorem 5.24 (The Functional Calculus for a Normal Operator)**.** *Suppose $T \in B(\mathcal{H})$ is a normal operator with spectrum $\sigma(T)$, and let*

$$\mathfrak{A} := \overline{\{p(T, T^*) | \ p \in \mathbb{C}[x_1, x_2]\}}.$$

Then

$$\mathfrak{A} \simeq C(\sigma(T), \mathbb{C}),$$

where $\sigma(T)$ is a compact subset of $\mathbb{C}$.

In fact, for any $T \in B(\mathcal{H})$, and any polynomial p,

$$\sigma(p(T)) = p(\sigma(T)).$$

Since any matrix is unitarily equivalent to an upper triangular matrix and since unitarily equivalence (even regular similarity) preserves eigenvalues, it suffices to check this fact for upper triangular matrices.

Exercise 5.25. Suppose T is an upper triangular matrix and p is a polynomial. Show that $\sigma(p(T)) = p(\sigma(T))$. *Hint: An upper triangular matrix can be written as* $D+N$ *where* D *is a diagonal matrix and* N *is a strictly upper triangular (and hence nilpotent) matrix. Recall also that the eigenvalues for an upper triangular matrix lie on the diagonal.*

References

[1] Robert G. Bartle, The Elements of Real Analysis. Second Edition. John Wiley and Sons, New York, 1964.

[2] G. B. Folland, Real Analysis, 2nd ed., Wiley, New York, 1999.

[3] Theodore W. Gamelin, Complex Analysis. Springer, New York, 2001.

[4] J. Kirkwood, An Introduction to Analysis. PWS-KENT, Boston, 1989.

[5] J. R. Munkres, Topology. Prentice Hall, Inc., 2000.

[6] H. Royden and P. Fitzpatrick, Real Analysis. Prentice Hall, 2010.

[7] Walter Rudin, Principles of Mathematical Analysis. International Series in Pure and Applied Mathematics, McGraw-Hill, New York, 1964.