Lecture Notes in Analysis

Alex Rutar*
University of Waterloo

Summer 2020†

*public@rutar.org
†Last updated: January 12, 2021

Contents

Preface

This collection of notes is my collection of personal notes on various topics in analysis, or important to analysis. The notes are linearly ordered, in the sense that any fact used in the proof of a result is contained in the prior pages. However, as this is intended as a personal reference, the notes are often written in an order which, while logically sound, is not conducive to understanding the material presented within. Often the most general form or definition will be presented earlier within the book, and many special cases (which are most reasonably learned before the general case) are obtained as corollaries. Occasionally, proofs of these special cases are given in the notes, but only for interest in the technique used to prove the result. Naturally, there are no exercises, though I have attempted to organize collections of examples at the end of each section which may be of interest.

The content of these notes are based of a large number of references and sources. Most notably, substantial portions of these notes are based on courses at the University of Waterloo taught by (alphabetically): Kathryn Hare, Rahim Moosa, Blake Madill, Matthew Satriano, Yi Shen, and Nico Spronk.

These notes are currently in draft form and will likely remain so for a long time. If you find errors (typographical or logical), you can contact me at alex@rutar.org.

I. Set Theory and Topology

1 Naïve Set Theory

1.1 Sets and Functions

We take for granted the existence of "sets", which are collections of objects. If a set A contains some object a, we write $a \in A$; conversely, write $a \notin A$. If every object in some set A is an element of some set B, we say $A \subseteq B$ is a *subset*; if there is some element of B not in A, we say A is a *proper subset* of B. We denote by $\emptyset$ the *empty set*, or the set with no elements. Note that $\emptyset \subset A$ for any set A.

Let $\{A_\alpha\}_{\alpha \in A}$ be an indexed family of sets. We define the *union* as $\bigcup_{\alpha \in I} A_\alpha = \{x : x \in A_\alpha \text{ for some } \alpha\}$, the *intersection* $\bigcap_{\alpha \in I} A_\alpha = \{x : x \in A_\alpha \text{ for all } \alpha\}$. Given two sets A, B, we denote the *difference* $A \setminus B = \{x : x \in A, x \notin B\}$. If some set A is contained in some universe U, i.e. $A \subset U$, then we define $A^c = U \setminus A$ to denote the ($U-$dependent) *complement*. Often, the definition of U is clear from context.

Given a set A, we denote the *power set* $\mathcal{P}(A)$ as the set of all subsets of A. Given a finite list of sets $A_1, \ldots, A_n$, we define the *(cartesian) product* $A_1 \times \cdots \times A_n = \prod_{i=1}^n A_i = \{(a_1, \ldots, a_n) : a_i \in A_i \text{ for each } i\}$.

Fix some pair of sets A, B. A *binary relation* R on A and B is any subset of $A \times B$. Typically, we write $a\,R\,b$ to denote $(a, b) \in R$. A particularly useful type of relation is an *equivalence relation* on $A \times A$, typically denoted by $\sim$, which satisfies three axioms:

- *(reflexive)*: $x \sim x$.
- *(symmetric)*: if $x \sim y$, then $y \sim x$.
- *(transitive)*: if $x \sim y$ and $y \sim z$, then $x \sim z$.

Given $a \in A$, we say $[a] = \{x \in A : x \sim a\}$. By the equivalence relation axioms, equivalence classes are either identical or disjoint. We denote by $A/\sim$ the *quotient* of A by $\sim$, where $A/\sim = \{[a] : a \in A\}$.

A *function* $f : A \to B$ is a relation on $A \times B$ where for any $a \in A$, there is some unique $b \in B$ so that $(a, b) \in f$. For $a \in A$, we denote by $f(a)$ the unique element $(a, f(a)) \in f$ and write $a \mapsto f(a)$. We denote the *image* of f by $\operatorname{im} f = \{f(a) : a \in A\} \subset B$. We say that f is *injective* if $x \neq y$ implies $f(x) \neq f(y)$, and *surjective* if for any $b \in B$, there is some $a \in A$ so $f(a) = b$. Then f is *bijective* if it is injective and surjective. If $f : A \to B$ and $g : B \to C$, we say $g \circ f : A \to C$ is the *composition* of f with g, and assign the rule $g \circ f(a) = g(f(a))$. If $A' \subseteq A$, we denote the *restriction* $f|_{A'} : A' \to B$ by $f|_{A'}(a') = f(a')$ for $a' \in A' \subseteq A$.

A function $f : A \to B$ induces natural map $f : \mathcal{P}(A) \to \mathcal{P}(B)$ and $f^{-1} : \mathcal{P}(B) \to \mathcal{P}(A)$ by $f(S) = \{f(s) : s \in S\}$ and $f^{-1}(T) = \{s \in A : f(s) \in T\}$. We denote by f^{-1} the *preimage*. Importantly, the preimage commutes with the standard set operations:

$$f^{-1}\left(\bigcup_{\alpha \in I} A_\alpha\right) = \bigcup_{\alpha \in I} f(A_\alpha) \qquad\qquad f^{-1}\left(\bigcap_{\alpha \in I} A_\alpha\right) = \bigcap_{\alpha \in I} f(A_\alpha).$$

As a result, f^{-1} also commute with complementation and set difference. One may also verify that $(f \circ g)^{-1}(A) = g^{-1}(f^{-1}(A))$.

A *sequence* on a set X is any function $a : \mathbb{N} \to X$. We typically write $a(n) = a_n$.

1.2 CARDINALITY

Cardinality is a way of thinking about the size of a set.

Definition. Two sets A and B have the same *cardinality* if there is a bijection between the sets. If this is the case, we say that $|A| = |B|$. If there exists an injection, then we say $|A| \leq |B|$.

It is easy to see that cardinality is an equivalence relation. We say that A is *denumerable* if $|A| = |\mathbb{N}|$, and *countable* if it is denumerable or finite. Denumerable sets can be written as a sequence by taking some bijection $\phi : \mathbb{N} \to A$.

1.1 Proposition. *The following hold:*

(i) Every infinite subset of $\mathbb{N}$ *is countably infinite.*

(ii) If A *is infinite and* $|A| \leq |\mathbb{N}|$*, then* $|\mathbb{N}| = |A|$.

PROOF (i) We use the well-ordering property of $\mathbb{N}$: every non-empty subset of $\mathbb{N}$ has a least element. Let B be an infinite subset of $\mathbb{N}$, so it is non-empty. Thus B has some least element b_1. But then, $B \setminus \{b_1\}$ is also non-empty, so we can repeat this process to create an increasing sequence

$$b_1 < b_2 < b_3 < \cdots <$$

I claim that every element of B is in this set. Let $b \in B$ and consider $\{n \in B : n \leq b\}$. This set is finite with, say, k elements, so $b = b_k$. We then get our bijection by the standard map $b_i \mapsto i$.

(ii) Assume $j : A \to \mathbb{N}$ is an injection. Let $B = j(A) \subseteq \mathbb{N}$. Notice $j : A \to B$ is a bijection, so $|A| = |B|$ and B is infinite. By (1), B is countably infinite, so $|B| = |\mathbb{N}|$, and the result follows by transitivity. ■

We say that a set is *uncountable* if it is not countable.

Example. The set $\mathcal{B}$ of all binary sequences $f : \mathbb{N} \to \{0, 1\}$ is uncountable.

This argument is a classical argument known as *Cantor diagonalization*. Certainly $\mathcal{B}$ is not finite; suppose it is denumerable and write $\mathcal{B} = \{b^i : i \in \mathbb{N}\}$. Define $a : \mathbb{N} \to \{0, 1\}$ by

$$a_n = \begin{cases} 0 & : b_n^n = 1 \\ 1 & : b_n^n = 0 \end{cases}$$

so that for any i, $a \neq b^i$ since $a_i \neq b_i^i$, a contradiction.

1.2 Theorem. (Cantor) *For any set* A, $|A| < |\mathcal{P}(A)|$*, where* $|A| < |B|$ *if* $|A| \leq |B|$ *and* $|A| \neq |B|$.

PROOF The argument here is, in essence, identical to the previous example. We certainly have an injection given by the map $a \mapsto \{a\}$, so $|A| \leq |\mathcal{P}(A)|$. Suppose for contradiction we have some bijection $g : A \to \mathcal{P}(A)$ and define

$$B = \{a \in A : a \notin g(a)\} \subseteq A.$$

Since $B \subseteq A$, we have $B \in \mathcal{P}(A)$, so there exists some $x \in A$ such that $g(x) = B$. However, if $x \in B$, then $x \notin g(x) = B$, while if $x \notin B = g(A)$, then $x \in B$. Thus no such g exists. ■

Using this we can construct an infinite list of cardinalities, since $|A| < |\mathcal{P}(A)| < |\mathcal{P}(\mathcal{P}(A))| < \cdots$. We define $\{0,1\}^A = \{f : A \to \{0,1\}\}$. It is easy to verify that $|\{0,1\}^A| = |\mathcal{P}(A)|$ via the bijection $B \mapsto \chi_B$ where

$$\chi_B = \begin{cases} 0 & : x \notin B \\ 1 & : x \in B \end{cases}$$

An important argument in cardinality is the following:

1.3 Theorem. (Schroeder-Bernstein) *If $|A| \leq |B|$ and $|B| \leq |A|$ then $|A| = |B|$.*

PROOF Our general goal is to partition A into parts D and D^c so that $D^c = g(f(D)^c)$. Assuming this, we may define

$$\phi(x) = \begin{cases} f(x) & : x \in D \\ g^{-1}(x) & : x \in D^c \end{cases}$$

and it follows easily that it is a bijection.

Define $Q : \mathcal{P}(A) \to \mathcal{P}(A)$ by $Q(E) = [g(f(E)^c)]^c \subseteq A$; we want to show that Q has a fixed point. Since f and g are injective, they are inclusion preserving; thus, if $E \subseteq F$, then $Q(E) \subseteq Q(F)$.

Now let $\mathcal{D} = \{E \subseteq A : E \subseteq Q(E)\}$. Set $D = \bigcup_{E \in \mathcal{D}} E \subseteq A$ so that $E \subseteq Q(E) \subseteq Q(D)$ for any $E \in \mathcal{D}$. If $E \in \mathcal{D}$ then $E \subseteq Q(E) \subseteq Q(D)$, since $E \subseteq D$. Thus $D = \bigcup_{E \in \mathcal{D}} E \subseteq Q(D)$, so $Q(D) \subseteq Q(Q(D))$ and $Q(D) \in \mathcal{D}$ so that $D = Q(D)$. ■

Example.

1. *If $A_1 \subseteq A_2 \subseteq A_3$, and $|A_1| = |A_3|$, then $|A_1| = |A_2| = |A_3|$.* We have injections i, j

$$A_1 \hookrightarrow iA_2 \hookrightarrow jA_3$$

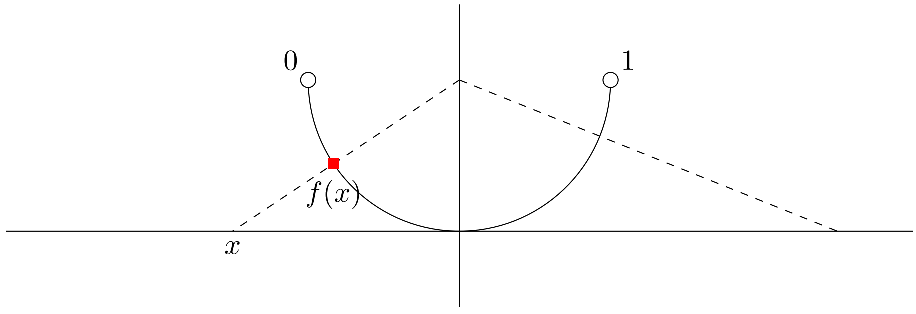
given by the embedding maps, and a bijection $k : A_3 \to A_1$. Then $k \circ j : A_2 \to A_1$ is an injection, so by Theorem 1.3, $|A_1| = |A_2|$ and $|A_2| = |A_3|$ by transitivity.

2. $|(0,1)| = |[0,1)| = |\mathbb{R}|$. It suffices to show $|(0,1)| = |\mathbb{R}|$. Consider $f(x) = \arctan x$ which is a bijection $f : \mathbb{R} \to \left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$. Thus

$$\frac{1}{\pi}\arctan x + \frac{1}{2} : \mathbb{R} \to (0,1)$$

is a bijection.

Alternatively, a proof by picture:



3. $|\mathbb{R}| = |\{0,1\}^{\mathbb{N}}|$. It suffices to show that $|[0,1)| = |\{0,1\}^{\mathbb{N}}|$. Take $r \in [0,1)$ and write as $r = .r_1 r_2 r_3 \ldots$ where $r_j \in \{0,1\}$ (binary representation of r). Define $f_r(n) = r_n$ for each $n \in \mathbb{N}$; clearly, the $r \mapsto f_r$ is injective.
   Similarly, we have an injection $\{0,1\}^{\mathbb{N}} \to [0,1)$ given
   $$f \mapsto 0.0f(1)0f(2)0f(3)\ldots \in [0,1)$$
   This is an injection because non-unique binary representation have to end with a tail of 1's or a tail of 0's. Thus by Theorem 1.3, the result follows.

1.3 CARDINAL ARITHMETIC

Definition. The equivalence classes of sets modulo bijective equivalence are called *cardinal numbers*.

We may define operations $|A|+|B| = |A \cup B|$ and $|A| \cdot |B| = |A \times B|$. It is easy to verify that these operations are commutative, associative, and distributive. Note that with n copies of A, we have $|A| + \cdots + |A| = |A \cup \cdots \cup A| = |\{1,2,\ldots,n\} \times A|$ and we write $n|A|$. Let $A^B = \{f : B \to A | f \text{ a function}\}$, so we may define $|A|^{|B|} = |A^B|$. One may verify that the usual power rules apply, namely $(|A|^{|B|})^{|C|} = |A|^{|B|\cdot|C|}$ and $|A|^n = |A^{1,\ldots,n}| = |A \times \cdots \times A|$ n times.

If S is finite, we simply write $|S| = n$ to denote the cardinal. We also write $\aleph_0 = |\mathbb{N}|$ and $\aleph_1 = |\mathbb{R}|$.

[***TODO: prove these***]

Example. 1. $|A| \geq \aleph_0$ *if and only if there is* $A \subsetneq B$ *such that* $|A| = |B|$.

2. *A is infinite if and only if* $\aleph_0|A| = |A|$ *if and only if* $|A| = n|A$ *for each* $n \in \mathbb{N}$.
   Idea: Use a Zorn argument to show A can be partitioned into infinte countable sets. Manually show that $|N \times N| = |N|$ to further partition each element of the partition into infinite countable sets.
3. *Given any two sets* A, B, *either* $|A| \leq |B|$ *or* $|B| \leq |A|$.
   Idea: Find a maximal pair (E, f) such that $E \subseteq A$ and $f : E \to B$ is injective, i.e. maximal w.r.t. $(E,f) \leq (E', f')$ if and only if $E \subseteq'$ and $f'|_E = f$. Verify that either $|E| = |A|$ or $|E| = |B|$.
   If any two cardinals are comparable, then through an ordinal-arithmetic idea called "Hartog's number", it can be proved that any set A is well-orderable. Thus it is impossible to prove (iii) without A of C.
4. *A is infinite if and only if* $n|A| = |A|$ *for each* $n \in \mathbb{N}$.
   It suffices to show for $n = 2$. Find a maximal pair (B, f), $B \subseteq A$ and $f : B \to B \times B$ bijection, same partial ordering as in 3 above. If $|B| < |A|$ then $|A \setminus B| = |A|$ (why?). There would be $B' \subset A \setminus B$ with $|B'| = |B|$ and one could construct $\tilde{f}$ for which $(B, f) \leq (B \cup B', \tilde{f})$, which violates assumptions on (B, f).
5. *A is infinite if and only if* $|\mathcal{F}(A)| = |A|$, *where* $\mathcal{F}(A) = \{F \in \mathcal{P}(A) : |F| < \aleph_0\}$.

1.4 INDUCTIVE CONSTRUCTION

[***TODO: write***]

1.5 THE AXIOM OF CHOICE

Definition. Let S be a non-empty set. A *partial ordering* is a binary relation $\leq$ on S which satisfies for $s, t, u \in S$,

(i) *(reflexivity)*: $s \leq s$

(ii) *(transitivity)*: $s \leq t$, $t \leq u$ implies $s \leq u$

(iii) *(anti-symmetry)*: $s \leq t$, $t \leq s$ implies $s = t$

We call the pair $(S, \leq)$ a *partially ordered set*. We say that $(S, \leq)$ is *totally ordered* if, given $s, t \in S$, at least one of $s \leq t$ or $t \leq s$ holds. We say that $(S, \leq)$ is *well-ordered* if given any $\emptyset \neq S_0 \subseteq S$, there is some $s_0 \in S_0$ such that $s_0 \leq s$ for $s \in S_0$. A *chain* in a poset $(S, \leq)$ is any $\emptyset \neq C \subseteq S$ such that $(S, \leq|_{C\times C})$ is totally ordered.

Example. (i) $X \neq \emptyset$, $(\mathcal{P}(X), \subseteq)$ is a partially ordered set.

(ii) $(\mathbb{R}, \leq)$ is a totally ordered set.

(iii) If $(S, \leq)$ is partially ordered and $T \subset S$, then $\leq|_{T\times T}$ is a partial order on T.

The main discussion of this section is the Axiom of Choice, and various equivalent formulations of it.

1.4 Theorem. *The following are equivalent:*

(i) (Axiom of Choice 1): For any $x \neq \emptyset$, there is a function $\gamma : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ such that $\gamma(A) \in A$ for each $A \in \mathcal{P}(X)\setminus\{\emptyset\}$.

(ii) (Axiom of Choice 2): Given any $\{A_\lambda\}_{\lambda\in\Lambda}$ where $A_\lambda \neq \emptyset$ for each λ,

$$\prod_{\lambda\in\Lambda} A_\lambda = \{(a_\lambda)_{\lambda\in\Lambda} : a_\lambda \in A_\lambda \textit{ for each } \lambda\} \neq \emptyset.$$

(iii) (Hausdorff Maximality Principle): In any partially ordered set $(S, \leq)$ there is a maximal chain, i.e. a chain M for which no $M \cup \{s\}$ is a chain for any s in $S \setminus M$.

(iv) (Zorn's Lemma): In a partially ordered set $(S, \leq)$, if each chain $C \subseteq S$ admits an upper bound in S, then $(S, \leq)$ admits a maximal element.

(v) (Well-ordering principle): Any $S \neq \emptyset$ admits a well-ordering.

PROOF $(i \Rightarrow ii)$ Let $X = \bigcup_{\lambda\in\Lambda} A_\lambda$ and let γ be a choice function on X. Then $(\gamma(X_\lambda))_{\lambda\in\Lambda} \in \prod_{\lambda\in\Lambda} A_\lambda$.

$(ii \Rightarrow i)$ If X is any non-empty set, get $x \in \prod_{A\in\mathcal{P}(X)\setminus\{\emptyset\}} A$ and define $\gamma(A) = x_A$.

$(ii \Rightarrow iii)$ We first prove the following result: let $\mathcal{F} \subset \mathcal{P}(X)$ satisfy

- $\emptyset \in \mathcal{F}$ and
- if $\mathcal{K} \subset \mathcal{F}$ is a chain (with respect to inclusion), then $\bigcup_{K\in\mathcal{K}} \in \mathcal{F}$.

Then $\mathcal{F}$ contains an element M so that $M \cup \{x\} \notin \mathcal{F}$ for any $x \in X \setminus M$.

For each $A \in \mathcal{F}$, define $A^* = \{x \in X : A \cup \{x\} \in \mathcal{F}\}$. Fix a choice function $\gamma : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ and let $\Gamma(A) = A \cup \{\gamma(A^*)\}$ if $A^* \neq \emptyset$, and $\Gamma(A) = A$ otherwise. Note that $\Gamma(A) \in \mathcal{F}$.

We define a *tower* to be any subcollection $\mathcal{T} \subset \mathcal{F}$ such that

- $\emptyset \in \mathcal{T}$,
- $A \in \mathcal{T}$ implies $\Gamma(A) \in \mathcal{T}$, and
- If $\mathcal{K} \subset \mathcal{T}$ is a chain, then $\bigcup_{K\in\mathcal{K}} \in \mathcal{T}$.

It is easy to see that

$$\mathcal{T}_0 := \bigcap_{\mathcal{T}\subset\mathcal{F} \text{ is a tower}} \mathcal{T}$$

is a tower. Note that $\emptyset \in \mathcal{T}_0$, so that $\{\gamma(\emptyset^*)\} \in \mathcal{T}_0$, $\{\gamma(\emptyset^*), \gamma(\{\gamma(\emptyset^*)\}^*)\} \in \mathcal{T}_0$, etc. We will show that $\mathcal{T}_0$ is totally ordered. We say that $C \in \mathcal{T}_0$ is *comparable* if for all $A \in \mathcal{T}_0$ either $A \subset C$ or $C \subset A$. For such a C consider

$$\mathcal{T}_C = \{A \in \mathcal{T}_0 : A \subsetneq C\} \cup \{C\} \cup \{A \in \mathcal{T}_0 : \Gamma(C) \subset A\}.$$

Note that $\emptyset \in \mathcal{T}_C$ and there is no proper inclusion $C \subsetneq A \subsetneq \Gamma(C)$ by definition of Γ. Thus since C is comparable, one may observe that $\mathcal{T}_C$ is a tower so in fact $\mathcal{T}_C = \mathcal{T}_0$. Thus if C is comparable, so is $\Gamma(C)$.

Let $\mathcal{C}$ denote the family of comparable sets in $\mathcal{T}_0$. But then the family $\mathcal{C}$ of comparable sets in $\mathcal{T}_0$ is $\Gamma-$closed, and the chain condition is immediately verified, so that $\mathcal{C}$ is itself a tower and $\mathcal{C} = \mathcal{T}_0$. Thus $\mathcal{T}_0$ is totally ordered and thus a chain in $\mathcal{F}$. In particular, $M = \bigcup_{T \in \mathcal{T}_0} T \in \mathcal{T}_0$. Suppose for contradiction $M^* \neq \emptyset$; but then $M \subsetneq \Gamma(M) \in \mathcal{T}_0$, a contradiction, so $M^* = \emptyset$ is maximal, proving the claim.

Now let $(S, \leq)$ be partially ordered and let $\mathcal{F}$ denote the set of all chains in S. Clearly $\mathcal{F}$ satisfies the requirements in the claim, and the corresponding maximal element in $(\mathcal{F}, \subseteq)$ is a maximal chain in S.

$(iii \Rightarrow iv)$ Let $(S, \leq)$ be a partially ordered set in which each chain has a maximal element. Let M be a maximal chain and m an upper bound for M, so that $M \cup \{m\}$ is a chain so $m \in M$. But now if there is any $s \in S$ with $m \leq s$, then $M \cup \{s\}$ is a chain so $s \in M$ and $s \leq m$. Thus m is maximal.

$(iv \Rightarrow v)$ Let $\mathcal{W} = \{(A, \leq_A) : A \in \mathcal{P}(S), \leq_A \text{ is a well-ordering on } A\}$. We say $(A, \leq_A) \leq (B, \leq_B)$ if A is an *initial segment* of B, i.e. if $A \subseteq B$, $\leq_B|_{A \times A} = \leq_A$, and for $a \in A$ and $b \in B$, $a \leq_B b$. Let $\mathcal{C}$ be a chain in $(\mathcal{W}, \leq)$, let $U = \bigcup_{(C, \leq_C) \in \mathcal{C}} C$, and say $s \leq_U t$ if $s \leq_C t$ for some C containing s and t. Clearly $\leq_U$ is well-defined. To see that it is a well-ordering, let $A \subseteq U$ be non-empty and get some $C \in \mathcal{C}$ so $C \cap A \neq \emptyset$ and thus admits a minimal element a_C. This choice is independent of C for if C' has $C' \cap A \neq \emptyset$ with, say, $C \subseteq C'$, then $(C, \leq_C)$ is an initial segment of $(C', \leq_{C'})$ so $a_{C'} = a_C$. Thus $(U, \leq_U)$ is an upper bound for $\mathcal{C}$.

Thus by Zorn's lemma, $\mathcal{W}$ has some maximal element $M, \leq_M$. Note that $M = S$; if not, get $s \in S \setminus M$ and extend $\leq_M$ to $M \cup \{s\}$ by $m \leq_M s$ for any $m \in M$.

$(v \Rightarrow i)$ Fix any well-ordering $\leq$ on X and for any $A \in \mathcal{P}(X) \setminus \{0\}$ let $\gamma(A)$ be a minimal element. ■

2 FUNDAMENTALS OF TOPOLOGY

2.1 BASIC DEFINITIONS AND EXAMPLES

Definition. A *topology* on a set X is a set τ of subsets of X such that
(i) $\emptyset \in \tau$ and $X \in \tau$
(ii) If $U_\alpha \in \tau$ for all $\alpha \in A$, then $\bigcup_{\alpha \in A} U_\alpha \in \tau$.
(iii) If $n \in \mathbb{N}$ and $U_i \in \tau$ for each $1 \leq i \leq n$, then $\bigcap_{i=1}^n U_i \in \tau$.
The sets $U \in \tau$ are the *open sets* in X, and sets $X \setminus U$ for some open set U are the *closed sets* in X. The pair (X, τ) is called a *topological space*.

To simplify notation, we will say that any $A \subset X$ is a *neighbourhood* of some $x \in X$ if there exists some $U \in \tau$ with $x \in U$ and $U \subset A$. On may construct an analgous set of axioms for the closed sets on X using DeMorgan's identities.

An important feature of topological spaces is the type of functions that one categorically cares about.

Definition. Let (X, τ) and (Y, σ) be topological spaces, and $f : X \to Y$. We say that f is **$(\tau - \sigma-)$continuous at x_0** in X if for any $V \in \sigma$ such that $f(x_0) \in V$, then there exists $U \in \tau$ such that $x_0 \in U$ and $f(U) \subseteq V$. We say that f is **$(\tau - \sigma-)$continuous** if it is continuous at each x_0 in X.

An easy application of definitions yields the following:

> ***2.1 Proposition.*** *Let (X, τ), (Y, σ) be topological spaces and $f : X \to Y$. Then f is continuous if and only if for any $U \in \sigma$, $f^{-1}(U) \in \tau$.*

If $f : X \to Y$ is continuous and invertible with f^{-1} also continuous, then we say that f is a *homeomorphism* and that X and Y are *homeomorphic*. If $f : X \to \operatorname{im} X$ is a homeomorphism onto its image (with the subspace topology from Y), then we say that f is an *imbedding*

We define the *interior* of some $A \subset X$ as

$$A^\circ = \bigcup_{U \subseteq A, U \in \tau} U.$$

Dually, we define the *closure* of A as

$$\overline{A} = \bigcap_{F \supseteq A, X \setminus F \in \tau} F.$$

Then the *boundary* of A is $\partial A = \overline{A} \setminus A^\circ$.

One may naturally compare topologies as follows:

Definition. Suppose σ, τ are two topologies on X. We say that σ is *finer* than τ if $\sigma \supseteq \tau$, and that σ is *coarser* than τ if $\sigma \subseteq \tau$.

One can think of a finer topology as having more ("smaller") open sets.

Definition. Let (X, τ) be a topological space. A *base* for τ is any family $\beta \subseteq \tau$ such that for any $U \in \tau$ and $x \in U$, there is $B \in \beta$ such that $x \in B \subseteq U$. A *subbase* for τ is any family $\alpha \subseteq \tau$ such that $\{\bigcap_{k=1}^n U_k : n \in \mathbb{N}, U_1, \ldots, U_n \in \alpha\}$ is a base for τ.

Note that if $\beta \subseteq \mathcal{P}(X)$ has $\bigcup_{V \in \beta} V = X$, then

$$\langle \beta \rangle = \{U \subseteq X : \text{ for each } x \in U, \text{ there are } S_i \text{ s.t. } x \in \bigcap_{k=1}^n S_k \subseteq U\}$$

which defines a topology on X; in fact, this topology is the coarsest topology containing β.

Example. Let $(S, \leq)$ be a totally ordered set. Write $x < y$ if $x \neq y$ and $x \leq y$. Then S has a natural *order topology* induced by the open sets $(a, b) = \{x \in S : a < x, x < b\}$, and if S has a minimal element a_0, we also include the sets $[a_0, b) = \{x \in S : a_0 < x, x < b\}$, and similarly if S has a maximal element b_0.

Definition. Let $X \neq \emptyset$. Suppose we are given

- a family $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in A}$ of topological spaces, and
- for each $\alpha \in A$, a function $f_\alpha : X \to X_\alpha$

Then the *initial topology* on X given this data is denoted

$$\sigma = \sigma(X, (f_\alpha)_{\alpha \in A}) = \sigma(X, (f_\alpha, \tau_\alpha)_{\alpha \in A})$$

and is the topology with base

$$\beta = \left\{ \bigcap_{k=1}^{n} f_{\alpha_k}^{-1}(U_{\alpha_k}), n \in \mathbb{N}, \alpha_1, \dots, \alpha_n \in A, \text{ each } U_{\alpha_k} \in \tau_{\alpha_k} \right\}$$

In particular, $\{f_\alpha^{-1}(U_\alpha) : U_\alpha \in \tau_\alpha, \alpha \in A\}$ is a subbase for σ.

Remark. The topology is chosen so that each $f_\alpha : X \to X_\alpha$ is $\sigma - \tau_\alpha$−continuous; in particular, it is the coarsest topology so that all the f_α are continuous.

A number of common topologies can be defined using this notation:

Example. (i) *Relative (or subspace) topology.* If (Y, τ)−topological space, $\emptyset \neq X \subseteq Y$, and $\iota : X \to Y$ is the inclusion map. Then $\tau|_X = \sigma(X, \{\iota\})$.

(ii) *Product topology.* Let $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in A}$ be a family of topological spaces. Let $X = \prod_{\alpha \in A} X_\alpha$. Let for $\alpha \in A$, $p_\alpha : X \to X_\alpha$ denote the projection map onto the component α. Then the product topology $\pi = \sigma(X, \{p_\alpha\}_{\alpha \in A})$. Hence, $V \in \mathcal{P}(X)$, then $V \in \pi$ if and only if for any $x \in V$, there is $\alpha_1, \dots, \alpha_n \in A$ and $U_{\alpha_k} \in \tau_{\alpha_k}$ such that $x_{\alpha_k} = p_{\alpha_k(x)} \in U_{\alpha_k}$ and $x \in \bigcap_{k=1}^n p_{\alpha_k}^{-1}(U_{\alpha_j}) \subseteq V$.
Note that if $X = \prod_{n=1}^\infty X_n$, each (X_n, τ_n) is a topological space, then the basic open sets look like $U_1 \times U_2 \times \cdots \times U_m \times X_{m+1} \times X_{m+2} \times \cdots$ for some $m \in \mathbb{N}$.

2.2 COMPACTNESS AND ULTRAFILTERS

Let (X, τ) be a topological space.

Definition. A subset $K \subseteq X$ is called *compact* if for any collection $\{U_\alpha\}_{\alpha \in A} \subseteq \tau$ with $\bigcup_{\alpha \in A} U_\alpha \supseteq K$, there exists some finite $U_1, \dots, U_n$ covering K. If X itself is τ−compact, we call (X, τ) a compact space.

A basic application of the definitions yields the following:

2.2 *Proposition.* *Let (X, τ) be a compact space.*
(i) If $K \subseteq X$ is closed, then K is compact.
(ii) If (Y, σ) is a topological space and $f : X \to Y$ is continuous, then $f(X)$ is compact.

Here we give an alternative way to characterize compactness in a topological space:

Definition. A family $\mathcal{F} \subseteq \mathcal{P}(X)$ has the **finite intersection property** if for any $F_1, \dots, F_n \in \mathcal{F}$, $\bigcap_{l=1}^n F_k \neq \emptyset$.

2.3 *Proposition.* *Let (X, τ) be a topological space. Then (X, τ) is compact if and only if any $\mathcal{F} \subseteq \mathcal{P}(X)$ with the finite intersection property has $\bigcap_{F \in \mathcal{F}} \overline{F} \neq \emptyset$.*

PROOF ($\Longrightarrow$) Suppose $\mathcal{F} \subset \mathcal{P}(X)$ has the finite intersection property but with $\bigcap_{F \in \mathcal{F}} \overline{F} = \emptyset$. Then $\{X \setminus \overline{F}\}_{F \in \mathcal{F}}$ is an open cover of X with no finite subcover.

($\Longleftarrow$) If $\{U_\alpha\}_{\alpha \in A}$ is an open cover of X, then $\mathcal{F} = \{X \setminus U_\alpha\}_{\alpha \in A}$ satisfies $\bigcap_{F \in \mathcal{F}} F = \emptyset$, so there is $F_1, \dots, F_n \in \mathcal{F}$ with $\bigcap_{k=1}^n F_k = \emptyset$. Then $\{X \setminus F_i\}_{i=1}^k$ is a finite subcover. ■

In this sense, compactness is very closely tied to the nature of families of sets which satisfy the finite intersection property. As a result, a nice way to study compactness (and in particular, Tychonoff's theorem), is through ultrafilters.

Definition. Let X be a non-empty set. An **ultrafilter** is a family $\mathcal{U} \subset \mathcal{P}(X)$ such that

- $\mathcal{U}$ has the finite intersection property
- If $A \in \mathcal{P}(X)$, then either $A \in \mathcal{U}$ or $X \setminus A \in \mathcal{U}$.

Example. The *Principal or trivial ultrafilter*. If $x_0 \in X$, let $\mathcal{U}_{x_0} = \{U \subseteq X : x_0 \in U\}$.

2.4 Lemma. (Ultrafilter) *If $\mathcal{F} \subseteq \mathcal{P}(X)$ is any set with the finite intersection property, then there is an ultrafilter $\mathcal{U}$ with $\mathcal{F} \subset \mathcal{U}$.*

PROOF Let $\Phi = \{\mathcal{G} \subseteq \mathcal{P}(X) : \mathcal{F} \subseteq \mathcal{G}, \mathcal{G} \text{ has f.i.p.}\}$ partially ordered by inclusion. If $\Gamma \subseteq \Phi$ is a chain, then $\mathcal{G}_\Phi = \bigcup_{\mathcal{G} \in \Gamma} \mathcal{G}$ contains $\mathcal{F}$ and has the finite intersection property. Thus by Zorn's lemma, Φ admits a maximal element $\mathcal{U}$. Let $A \in \mathcal{P}(X) \setminus \mathcal{U}$. Then $U \cup \{A\} \supsetneq \mathcal{U}$, so $\mathcal{U} \cup \{A\}$ fails the finite intersection property. Hence get $U_1, \dots, U_n$ so $A \cap \bigcap_{k=1}^n U_k = \emptyset$. Now if $V_1, \dots, V_m \in \mathcal{U}$, then

$$\left(\bigcap_{j=1}^m V_j\right) \cap \left(\bigcap_{k=1}^n U_k\right) \subseteq \bigcap_{k=1}^n U_k \subseteq X \setminus A,$$

so $(X \setminus A) \cap \bigcap_{j=1}^m V_j \neq \emptyset$. Thus $\mathcal{U} \cup \{X \setminus A\}$ has the finite intersection property, so $X \setminus A \in \mathcal{U}$ by maximality. ∎

2.5 Corollary. *If $\mathcal{U} \subseteq \mathcal{P}(X)$ is an ultrafilter, then*

(i) If $A \in \mathcal{P}(X)$, $A \in \mathcal{U}$ if and only if $A \cap U \neq \emptyset$ for each $U \in \mathcal{U}$
(ii) If $A, B \in \mathcal{P}(X)$, then $A \cup B \in \mathcal{U}$ implies at least one of A or B is in $\mathcal{U}$
(iii) If $A \in \mathcal{U}$ and $A \subseteq V$, then $V \in \mathcal{U}$

PROOF The forward implication of (i) follows since $\mathcal{U}$ has finite intersection. Conversely, $X \setminus A \notin \mathcal{U}$, so $A \in \mathcal{U}$. (ii) and (iii) follow consequently. ∎

An ultrafilter is, unsurprisingly, a special case of a filter, which we give the definition here:

Definition. A family $\mathcal{F} \subset \mathcal{P}(X) \setminus \{\emptyset\}$ is a *filtering family* if for each $F_1, F_2 \in \mathcal{F}$, there is $F_3 \in \mathcal{F}$ such that $F_3 \subseteq F_1 \cap F_2$. We say that $\mathcal{F}$ is a *filter* if for any $F \in \mathcal{F}$ and $A \supseteq F$ we have $A \in F$ as well.

The previous corollary highlights this relationship.

2.6 Corollary. *If X is an infinite set, it admits a non-principle ultrafilter.*

PROOF Let $\mathcal{F} = \{F \in \mathcal{P}(X) : X \setminus F \text{ is finite}\}$, which has the finite intersection property; apply Theorem 2.4. ∎

2.7 Proposition. *There are at least $\mathfrak{c}$ many ultrafilters in $\mathcal{P}(\mathbb{N})$.*

PROOF [***TODO: Fix proof***] We let $\mathcal{F} \subset \mathcal{P}(\mathbb{N})$ be a collection of infinite sets such that $E \neq F$ in $\mathcal{F}$ implies $|E \cap F| < \infty$, and $|\mathcal{F}| = \mathfrak{c}$. For each $F \in \mathcal{F}$, we let $\mathcal{F}_F = \mathcal{F}_0 \cup \{F\}$, which has the finite intersection property. Moreover, if $E \in \mathcal{F} \setminus \{F\}$, then $\mathcal{F}_F \cup \{E\}$ would fail f.i.p. Hence, for $F \in \mathcal{F}$, let $\mathcal{U}_F$ be any ultrafilter containing $\mathcal{F}_F$, giving $\mathfrak{c}$ many ultrafilters. ∎

Remark. It can be shown (with a lot more work) that $\mathbb{N}$ admits $2^{\mathfrak{c}}$ ultrafilters.

Example. It is possible to construct elements of $\ell_\infty^* \cong \mathcal{FA}(\mathbb{N})$ using ultrafilters.

Let $\mathcal{U} \subset \mathcal{P}(\mathbb{N})$ be a non-principal ultrafilter and define $\delta_\mathcal{U} : \mathcal{P}(\mathbb{N}) \to \{0,1\} \subset \mathbb{R}$ by

$$\delta_\mathcal{U}(A) = \begin{cases} 1 & : A \in \mathcal{U} \\ 0 & : X \setminus A \in \mathcal{U} \end{cases}$$

It is easy to verify that this function is finitely additive:

- Since $\mathbb{N} \in \mathcal{U}$, we see that $\delta_\mathcal{U}(\emptyset) = 0$.
- If $\emptyset \neq A, B \in \mathcal{P}(\mathbb{N})$ with $A \cap B = \emptyset$, then if $A \cup B \in \mathcal{U}$, then exactly one of A or B is in $\mathcal{U}$. Thus $\delta_U(A \cup B) = \delta_U(A) + \delta_U(B)$.
- If $E_1, \ldots, E_n \subseteq \mathbb{N}$ with $E_j \cap E_k = \emptyset$ for $j \neq k$, then $\sum_{k=1}^n |\delta_\mathcal{U}(E_k)| \leq 1$ so $\|\delta_\mathcal{U}\|_{\text{var}} \leq 1$.

In fact, since $\delta_\mathcal{U}(\mathbb{N}) = 1$, we have $\|\delta_\mathcal{U}\|_{\text{var}} = 1$. Let $L_\mathcal{U} \in \ell_\infty^*$ be the linear functional associated to $\delta_\mathcal{U}$, so $L_\mathcal{U}(\mathbf{1}) = 1$ and $\|L_\mathcal{U}\| = 1$. It is left as an exercise to verify the following:

(i) $L_\mathcal{U}|_{\mathbf{c_0}} = 0$, so if $x \in \ell_\infty^\mathbb{R}$, then $\liminf_{n\to\infty} x_n \leq L_\mathcal{U} \leq \limsup_{n\to\infty} x_n$

(ii) Exactly one of $2\,\mathbb{N}$ and $2\,\mathbb{N} - 1$ is in $\mathcal{U}$, so $L(\chi_{2\,\mathbb{N}}) \neq L_\mathcal{U}(\chi_{2\,\mathbb{N}-1})$, so $L_\mathcal{U}$ is not translation invariant.

(iii) Let $S \in \mathcal{B}(\ell_\infty)$ be given by $Sx = \left(\frac{x_1 + \cdots + x_n}{n}\right)_{n=1}^\infty$. Then $L_\mathcal{U} \circ S$ is a Banach limit.

We now transition to a discussion of ultrafilters in the context of topology.

Definition. If (X, τ) is a topological space, $\mathcal{U}$ an ultrafilter on X, we say that $x_0 \in X$ is a $(\tau-)$limit point for $\mathcal{U}$ if for each $U \in \tau$ with $x_0 \in U$, we have $U \in \mathcal{U}$.

2.8 Proposition. *Let (X, τ) be a topological space. Then (X, τ) is compact if and only if any ultrafilter on X admits a $\tau-$limit point.*

PROOF It suffices to show compactness in the sense of Theorem 2.3. First note that if $x \in X$ and $\mathcal{U}$ is an ultrafilter on X, then $x \in \bigcap_{V\in\mathcal{U}} \overline{V}$ if and only if x is a $\tau-$limit point of $\mathcal{U}$. If (X, τ) is compact, then $\bigcap_{V\in\mathcal{U}} \overline{V} \neq \emptyset$. Conversely, if $\mathcal{F} \subseteq \mathcal{P}(X)$ has the finite intersection property, then there exists an ultrafilter $\mathcal{U} \supseteq \mathcal{F}$, so $\bigcap_{F\in\mathcal{F}} \overline{F} \supseteq \bigcap_{V\in\mathcal{U}} \overline{V} \neq \emptyset$. ■

2.9 Theorem. (Tychonoff) *Let $\{(X_\alpha, \tau_\alpha)\}_{\alpha\in A}$ be a family of compact spaces, and $X = \prod_{\alpha\in A} X_\alpha$ with the product topology π. Then (X, π) is compact.*

PROOF [***TODO: touch up proof***] We use Theorem 2.8: let $\mathcal{U}$ be an ultrafilter on X, and it suffices to show that it admits a $\pi-$limit point. Fix some $\alpha \in A$ and let $\mathcal{U}_\alpha = \{p_\alpha(V) : V \in \mathcal{U}\}$, where p_α is the coordinate projection onto α. If $\emptyset \neq S_\alpha \subseteq X_\alpha$ is any non-empty subset, then $S_\alpha = p_\alpha^{-1}(p_\alpha^{-1}(S_\alpha))$, so $S_\alpha \in \mathcal{U}_\alpha$ if and only if $p^{-1}(S_\alpha) \in \mathcal{U}$, and since p^{-1} commutes with complementation, $\mathcal{U}_\alpha$ is an ultrafilter. The last proposition provides a $\tau_\alpha-$limit point x_α for $\mathcal{U}_\alpha$. Now let $x = (x_\alpha)_{\alpha\in A}$, where x_α is found as above. If $W \in \pi$ with $x \in W$, then there are $\alpha_1, \ldots, \alpha_n$ in A, $U_{\alpha_i} \in \tau_{\alpha_i}$ with $x \in \bigcap_{k=1}^n p_{\alpha_k}^{-1}(U_{\alpha_k}) \subseteq W$. Since each x_{α_k} is a $\tau_{\alpha_k}-$limit point of $\mathcal{U}_{\alpha_k}$, we see that each $U_{\alpha_k} \in \mathcal{U}_{\alpha_k}$, so $p_{\alpha_k}^{-1}(U_{\alpha_k}) \in \mathcal{U}$. Thus we see that $W \in \mathcal{U}$, so x is a $\pi-$limit point of $\mathcal{U}$. ■

Remark. (i) Tychonoff's theorem implies the axiom of choice. Given $\{X_\alpha\}_{\alpha\in A}$ be a family of non-empty sets. Find y which is not a member of any X_α, and let $Y_\alpha = X_\alpha \cup \{y\}$ and $\tau_\alpha = \{\emptyset, \{y\}, X_\alpha, Y_\alpha\}$, and $(Y_\alpha.\tau_\alpha)$ is compact. The constant element y is an element of Y, so by Tychonoff, (Y, π) is compact. Given $\alpha_1, \ldots, \alpha_n \in A$, then $\bigcup_{k=1}^n p_{\alpha_k}^{-1}(\{y\})$. Since $\prod_{k=1}^n X_{\alpha_k} \neq 0$, we see that $Y \subsetneq \bigcup_{k=1}^n p_{\alpha_k}^{-1}(\{y\})$. Hence by compactness, $Y \not\subseteq \bigcup_{\alpha\in A} p_\alpha^{-1}(\{y\})$. Hence $\prod_{x\in A} X_\alpha = Y \setminus \bigcup_{\alpha\in A} p_\alpha^{-1}(\{y\}) \neq 0$.

(ii) If we are given $(X_\alpha, \tau_\alpha)_{\alpha \in A}$ a family of topological spaces, $X = \prod_{\alpha \in A} X_\alpha$, we can define the *box topology*, i.e. the topology with base $\{\prod_{\alpha \in A} U_\alpha : U_\alpha \in \tau_\alpha \setminus \{\emptyset\} \text{ for each } \alpha\}$ Of course, $\pi \subseteq \tau$, and the inclusion is proper on infinite products.

Definition. We say that (X, τ) is *locally compact* if for any $x \in X$, there exists some compact neighbourhood K containing x.

For example, $\mathbb{R}$ with the usual topology is locally compact.

2.10 Proposition. *Let $\{(X_i, \tau_i)\}_{i \in I}$ be locally compact spaces. Then $\prod_{i \in I} X_i$ is locally compact if and only if all but finitely many (X_i, τ_i) are compact.*

PROOF ($\Longleftarrow$) This follows directly from Theorem 2.9.
($\Longrightarrow$) We prove by contrapositive. Note that closures of basic open sets are of the form

$$\overline{U} = \overline{U}_{i_1} \times \cdots \times \overline{U}_{i_n} \times \prod_{i \in I \setminus \{i_1, \ldots, i_n\}} X_i.$$

For at least one $i \in I \setminus \{i_1, \ldots, i_n\}$, say (X_i, τ_i) is not compact, so that $X_i = \pi_i(\overline{U})$ is not compact. Thus $\overline{U}$ cannot be compact. ■

Example. [***TODO: move somewhere better***] Suppose X is a $\mathbb{F}-$vector space which is infinite dimensional. If $\|\cdot\|$ is a norm on X, then the Riesz lemma shows that for no $\epsilon > 0$ can the closed ball be compact. Suppose $\mathcal{F} \subseteq X'$ is separating, i.e. $\bigcap_{f \in \mathcal{F}} \ker f = \{0\}$. Every $\sigma(X, \mathcal{F})-$neighbourhood U of 0 contains a subspace $Y \neq \{0\}$. Since $\mathcal{F}$ is separating, there is $f \in \mathcal{F}$ and $Y \not\subseteq \ker f$, then $f(\overline{U}) \supseteq \mathbb{F}$.

2.3 CONNECTEDNESS

[***TODO: move stuff from pmath 465 assignment 1***]

Definition. Let (X, τ) be a topological space. Say that U, V are *disconnecting sets* for X if $X = U \cup V$ and $U, V \in \tau$. Then X is *connected* if X does not have disconnecting sets.

2.11 Proposition. *If $f : X \to f(X)$ is continuous and X is connected, then $f(X)$ is connected.*

PROOF Suppose $f(X)$ has disconnecting sets U, V. Then $X = f^{-1}(U) \cup f^{-1}(V)$ is disconnected. ■

A stronger notion of connected is path-connected.

Definition. We say E is *path-connected* if for any $x, y \in E$ there exists some continuous $f : [a, b] \to E$ so that $f(a) = x$, $f(b) = y$, and $\operatorname{im} f \subseteq E$.

2.12 Proposition. *If X is path connected, then X is connected.*

PROOF Say it is not connected, so we have $E = A \cup B$ where A, B are open, disjoint, and non-empty. Then pick $x \in A, y \in B$ and let $\gamma : [0, 1] \to E$ be a path connecting. Note that $f([a, b])$ is connected since f is continuous and $[a, b]$ is connected. But then $f([a, b]) = \Im f = (f([a, b]) \cap A) \cup (f([a, b]) \cap B)$ are disjoint (since A, B are disjoint) and non-empty (since x, y are in the first and second respectively), so A,B are disconnecting sets for $f([a, b])$, a contradiction. ■

Example. The "Topologist Sine Curve": let $X = \{x, \sin 1/x : x > 0\} \cup \{(0,0)\} = E \cup \{(0,0)\} \subseteq \mathbb{R}^2$. Then X is connected. but not path-connected.

Note that $T_p = \{(x, \sin(1/x) : x \in [p, 1]\}$ is connected for any $p > 0$ since it is the continuous image of a connected set. With this in mind, suppose for contradiction that T has disconnecting sets U and V, and assume without loss of generality that $(0,0) \in U$ and $f(t) = x \in V$ for some $t \neq 0$. Since U is open, there exists some $\epsilon-$ball $B_\epsilon((0,0)) \subseteq U$.

Furthermore, the zeros of $\sin(1/x)$ form a sequence which converges to 0, so that there exists some u such that $\sin(1/u) = 0$ and $u < \epsilon$. Then fix $p = \min\{t, u\}$. We claim that U and V are disconnecting sets for T_p. U and V are certainly still open and satisfy $T_p \subseteq T \subseteq U \cup V$, and $U \cap V$ is empty (equivalent characterization given by past assignment). Furthermore, $B_\epsilon((0,0)) \cap T_p \neq \emptyset$, so $U \cap T_p \neq \emptyset$ by construction, and $V \cap T_p \neq \emptyset$ by choice of p. Thus U and V disconnect T_p, a contradiction.

However, T is not path connected. To see this, suppose it is. Then there exists a continuous function γ on $[0,1]$ with $\gamma(0) = (0,0)$ and $\gamma(1) = (1, \sin(1))$. We claim that the image of γ, Γ, must be equal to T. We certainly have $\Gamma \subseteq T$ by definition. Then let $(x, f(x)) \in \Gamma$. We know that $(0, f(0))$ and $(1, f(1))$ are in T, so suppose for contradiction again that $(a, f(a)) \notin \Gamma$ for some $a \in (0,1)$. But then any point along the line $x = a$ is not in Γ. Note that γ is continuous, so it is componentwise continuous. Write $\gamma(t) = (x(t), y(t))$. Then $x(0) = 0$ and $x(1) = 1$, so by the intermediate value theorem we must have $x(c) = a$ for some c, which is impossible by assumption.

Therefore, $T = \Gamma$. However, this is a contradiction since Γ must be compact and T is not compact.

2.4 NETS AND TOPOLOGY

Definition. A pair $(N, \leq)$ is a *preorder* on N if it is

- *(symmetric)*: $v \leq v$ for $v \in N$
- *(transitive)*: $v_1 \leq v_2$ and $v_2 \leq v_3$ implies $v_1 \leq v_3$.

This pair is *cofinal* if for any $v_1, v_2 \in N$, there is $v_3 \in N$ so $v_1 \leq v_3$ and $v_2 \leq v_3$. Then $(N, \leq)$ is a *directed set* if $\leq$ is a cofinal preorder. Given a non-empty set X, a *net* is a function $x : N \to X$.

We usually write $x_\nu = x(\nu)$ and denote the net by $(x_\nu)_{\nu \in N}$. The motivating example of a net is a sequence $a : \mathbb{N} \to X$ with the standard well-ordering.

Definition. If $(x_\nu)_{\nu \in N} \subset X$ and $A \subseteq X$ is a subset, we say that $(x_\nu)_{\nu \in N}$ is

- *eventually* in A if there is $\nu_A \in N$ so $x_\nu \in A$ whenever $\nu \geq \nu_A$
- *frequently* in A if for any $\nu \in N$, there is $\nu' \in N$ with $\nu' \geq \nu$ and $x_{\nu'} \in A$.

These definitions align with the intuitive meanings for sequences.

Definition. Let $(N, \leq)$ and $(M, \leq)$ be directed sets. A map $\varphi : M \to N$ is *eventually cofinal* if for any $\nu \in N$, there is $\mu_\nu \in N$ so that $\phi(\mu) \geq \nu$ whenever $\mu \geq \mu_\nu$. Given a net $(x_\nu)_{\nu \in N}$ and an eventually cofinal $\varphi : M \to N$, we call $(x_{\varphi(\mu)})_{\mu \in M}$ a *subnet* of N.

Remark. Some sources use directed maps in place of eventually cofinal maps: we call $\varphi : M \to N$ a *directed map* if

(i) $\mu \leq \mu'$ in M implies $\varphi(\mu) \leq \varphi(\mu')$ in N

(ii) For any $\nu \in N$, there is $\mu \in M$ with $\varphi(\mu) \geq \nu$.

Certainly, directed maps are always cofinal. In this treatment, we will use cofinal maps for subnets.

Example. (i) $(\mathbb{N}, \leq)$ is directed and subsequences are special types of subnets.
(ii) $(\mathbb{R}, \leq)$ is directed.
(iii) *Riemann sums*. Let $a < b$ in $\mathbb{R}$. We let

$$N = \{(P, P^*) : P = \{a = t_0 < t_1 < \cdots < t_n = b\}, P^* = \{t_1^*, \ldots, t_n^*\}, t_k^* \in [t_{k-1}, t_k]\}$$

and say $(P, P^*) \leq (Q, Q^*)$ if $P \subseteq Q$. Given $f : [a, b] \to \mathbb{F}$ where $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, we define the *Riemann sum*

$$f_{(P,P^*)} = \sum_{j=1}^{n} f(t_j^*)(t_j - t_{j-1}) \in \mathbb{F}$$

where P, P^* are as above. Then $(f_{(P,P^*)})_{(P,P^*)\in N}$ is a net.
(iv) *Nets from filtering families*. Let $\mathcal{F}$ be a filtering family and define

$$N_{\mathcal{F}} = \{(x, F) : x \in F, F \in \mathcal{F}\}$$

equipped with the pre-order $(x, F) \leq (x', F')$ if and only if $F \supseteq F'$. Since $\mathcal{F}$ is a filtering family, $(N_{\mathcal{F}}, \leq)$ is directed. Let $f : N_{\mathcal{F}} \to X$ be given by $f(x, F) = x$, so $(x)_{(x,F)\in N_{\mathcal{F}}}$ is the net built from $\mathcal{F}$. Note that if $F \in \mathcal{F}$, then $(x)_{(x,F)\in\mathcal{F}}$ is eventually in F.
An *ultranet* $(x_\nu)_{\nu\in N} \subset X$ is a net for which any $A \in \mathcal{P}(X)$, $(x_\nu)_{\nu\in N}$ is either eventually in A or eventually in $X \setminus A$. If $\mathcal{U}$ is an ultrafilter, then $(x)_{(x,U)\in N_{\mathcal{U}}}$ is an ultranet.

The main gain of using nets over sequences is that they give a full (albeit, to an extent, tautological) characterization of continuity and other aspects of topological spaces.

Definition. Let (X, τ) be a topological space. We say that some $x_0 \in X$ is a

- *limit point* if for any $U \in \tau$ with $x_0 \in U$, $(x_\nu)_{\nu\in N}$ is eventually in U. We denote by $\lim_{\nu\in N} x_\nu$ the set of limit points of x. Often, we will write $x_0 = \lim_{\nu\in N} x_\nu$, but this is an abuse of notation since limit points need not be unique (which may happen when (X, τ) is not Hausdorff).
- *cluster point* of $(x_\nu)_{\nu\in N}$ if for any $U \in \tau$ with $x_0 \in U$, $(x_\nu)_{\nu\in N}$ is frequently in U.

2.13 Proposition. *If $(x_\nu)_{\nu\in N}$ is a net in (X, τ) and $x_0 \in X$, then x_0 is a cluster point for $(x_\nu)_{\nu\in N}$ if and only if x_0 is a $\tau-$limit point of x_{ν_μ} for some subnet $(x_{\nu_\mu})_{\mu\in M}$ of $(x_\nu)_{\nu\in N}$.*

PROOF ($\Longrightarrow$) Suppose x_0 is a cluster point for $(x_\nu)_{\nu\in N}$. Then for each $\nu \in N$ and $U \in \tau$ with $x_0 \in U$, define

$$F_{\nu,U} = \{\nu' \in N : \nu' \geq \nu, x_{\nu'} \in U\}$$

which is non-empty since x_0 is a cluster point. Then it is easy to see that $\mathcal{F} = \{F_{\nu,U} : \nu \in N, U \in \tau, x_0 \in U\} \subset \mathcal{P}(N)$ is a filtering family. Let $N_{\mathcal{F}}$ be the corresponding net, so that the map $(\nu, F) \mapsto \nu : N_{\mathcal{F}} \to N$ is cofinal and $(x_\nu)_{(\nu,F)\in N_{\mathcal{F}}}$ is a subnet of $(x_\nu)_{\nu\in N}$. Moreover, if $U \in \tau$ has $x_0 \in U$, then for any $\nu \in N$ and $\nu' \in F$ where $F \subseteq F_{\nu,U}$, we have $x_{\nu'} \in U$ so $(x_\nu)_{(\nu,F)\in N_{\mathcal{F}}}$ is eventually in U.

($\Longleftarrow$) If for some subnet $(x_{\nu_\mu})_{\mu\in M}$ is eventually in U for any $U \in \tau$ with $x_0 \in U$, then $(x_\nu)_{\nu\in N}$ is frequently in U for such U by definition of a subnet. ■

In the following proposition, it is clear that the forward implication holds in the special case when the net is a sequence. However, as the following example shows, the converse is not necessarily true:

Example. [***TODO: example(s?) here***]

2.14 Proposition. *Let (X, τ) and (Y, σ) be topological spaces. Then $f : X \to Y$ is continuous if and only if for any $x_0 \in X$ and net $(x_\nu)_{\nu \in N}$ having x_0 as a limit, $f(x_0) = \lim_{v \in N} f(x_v)$.*

PROOF ($\Longrightarrow$) If $V \in \sigma$ with $f(x_0) \in V$, then $f^{-1}(V) \in \tau$ with $x_0 \in f^{-1}(V)$. Since $(x_\nu)_{\nu \in N}$ is eventually in $f^{-1}(V)$, so $(f(x_\nu))_{\nu \in N}$ is eventually in V.

($\Longleftarrow$) Conversely, let $\tau_{x_0} = \{U \in \tau : x_0 \in U\}$, which is filtering on X, and N_{x_0} the corresponding net; then $x_0 = \lim_{(x,U) \in N_{x_0}} x$. Now, let $V \in \sigma$ have $f(x_0) \in V$. By assumption on f, there is some $(x, W) \in N_{x_0}$ such that for all $(x', W') \geq (x, W)$, we have $f(x') \in V$. Since $(x', W) \geq (x, W)$ for any $x' \in W$, we have

$$f(W) = \bigcup_{x' \in W} \{f(x')\} \subseteq V$$

or that $W \subseteq f^{-1}(V)$. Thus f is continuous at x_0, which was chosen arbitrarily. ■

This proposition is particular when proving convergence with respect to initial topoligies. Let X be a non-empty space and $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in A}$ is a family of topological spaces such that for each α there is a function $f : X \to X_\alpha$. The intial topology σ generated by this data is the coarsest topology such that each f_α is σ-τ_α-continuous. In particular, if $(x_\nu)_{\nu \in N}$ is a net, then $\lim_{\nu \in N} x_\nu = x$ if and only $\lim_{\nu \in N} f_\alpha(x_\nu) = f_\alpha(x)$ for each $\alpha \in A$.

Remark. We get the following consequences of this result:

(i) Given topologies τ, τ' on X, $\tau' \subseteq \tau$ if and only if $\lim_{v \in N} x_v = x_0$ with respect to τ' whenever $\lim_{v \in N} x_v = x_0$ with respect to τ for any $x_0 \in X$.

(ii) *Limits in product topology.* Let $\{(x_\alpha, \tau_\alpha)\}_{\alpha \in A}$ be topological spaces and $X = \prod_{\alpha \in A} X_\alpha$ equipped with the product topology π. [***TODO: change sequential superscript to brackets everywhere***] If $(x^{(\nu)})_{\nu \in N}$ is a net in X and $x^{(0)} \in X$, then $\lim_{\nu \in N} x^{(\nu)} = x^{(0)}$ with respect to π if and only if for every $\alpha \in A$, $\lim_{v \in N} x_\alpha^{(v)} = x_\alpha^{(0)}$ with respect to τ_α.

(iii) If X is a normed space and $(f_v)_{v \in N} \subset X^*$, $f_0 \in X^*$, then $w^* - \lim_{v \in N} f_v = f_0$ if and only if $\lim_{v \in N} f_v(x) = f_0(x)$ for each $x \in X$. [***TODO: move: forward direction from previous proposition, why reverse?***]

2.5 COUNTABILITY AND SEPARATION AXIOMS

[***TODO: Hausdorff + locally compact =¿ (completely) Regular***] [***TODO: Hausdorff + locally compact =¿ Baire***] An important class of topologies which are particularly tractable are the Hausdorff spaces. Most of the important topological spaces which occur in analysis are Hausdorff spaces.

Definition. A topological space (X, τ) is *Hausdorff* if given $x \neq y$ in X, there are $U, V \in \tau$ such that $x \in U$, $y \in V$, and $U \cap V = \emptyset$.

Example. (i) A metric space is Hausdorff.

(ii) X a normed space, $w = \sigma(X, X^*)$ is Hausdorff (by Hahn-Banach and A2Q1).

(iii) If X is a normed space, then $w^* = \sigma(X^*, \hat{X})$ on X^* is Hausdorff.

(iv) $\{(X_\alpha, \tau_\alpha)\}_{\alpha \in A}$ family of topological spaces, $X = \prod_{\alpha \in A} X_\alpha$ with π the product topology. Then (X, π) is Hausdorff if and only if all (X_α, τ_α) are Hausdorff. (Straightfoward exercise).

The following proposition is left as a straightforward exercise:

2.15 Proposition. *Let (X, τ) be a Hausdorff space with $K \subseteq X$ compact. Then K is closed.*

This yields the following nice result for compact subsets of Hausdorff spaces.

2.16 Proposition. *Let (X, τ) be a compact space.*

(i) If (Y, σ) is a Hausdorff space and $\phi : X \to Y$ is continuous and bijective, then $\phi^{-1} : Y \to X$ is continuous.

(ii) If $\tau' \subseteq \tau$ are Hausdorff topologies on X, then $\tau' = \tau$.

PROOF (i) If $F \subset X$ is closed, it is compact, so that $(\phi^{-1})^{-1}\phi(F) \subset Y$ is compact, and hence closed.

(ii) $\mathrm{id} : X \to X$ is continuous, so if $U \in \tau'$, then $\mathrm{id}^{-1}(U) = U \in \tau$, so id is continuous. Hence by (1) id^{-1} is continuous so $\tau \subseteq \tau'$. ■

2.17 Proposition. *Let X be a first countable space. Then X is Hausdorff if and only if sequential limits in X are unique; that is whenever $(x_i)_{i=1}^{\infty}$ converges to x and to y, then $x = y$.*

PROOF ($\Longrightarrow$) Let X be Hausdorff, $(x_n)_{n=1}^{\infty} \to x$, and $y \neq x$. Get $U \ni x$ and $V \ni y$ with $U \cap V = \emptyset$; but then $(x_n)_{n=1}^{\infty}$ is eventually in U and hence eventually not in V.

($\Longleftarrow$) Suppose X is not Hausdorff and let x, y be points such that any neighbourhood of x intersects any neighbourhood of y. By first countability, get neighbourhood bases $\{U_i : i \in \mathbb{N}\}$ and $\{V_i : i \in \mathbb{N}\}$ for x and y respectively; by assumption, get $x_i \in U_i \cap V_i$. But then $(x_i)_{i=1}^{\infty}$ converges to x and y. ■

A straightforward modification of the above proof yields the following result:

2.18 Proposition. *Let X be a topological space. Then X is Hausdorff if and only if net limits in X are unique.*

[***TODO: discussion about how a Hausdorff space has enough continuous functions to see points (Urysohn lemma, tietze extension, etc.)***] [***TODO: In metric section, discuss briefly how these results specialize to sequential characterizations.***]

2.6 METRIZATION THEOREMS

[***TODO: Urysohn metrization, maybe Nagata-Smirnov?***]

2.7 COMPACTIFICATION

One can define the extended real line as follows: set the space $X = \mathbb{R} \cup \{-\infty, +\infty\}$. Then the topology is given by

$$G \in \tau \Leftrightarrow \begin{cases} \text{for all } x \in G \cap \mathbb{R} & \exists r > 0 \text{ s.t. } (x - r, x + r) \subset G \\ -\infty \in G & \exists b \in \mathbb{R} \text{ s.t. } (-\infty, b) \subset G \\ +\infty \in G & \exists a \in \mathbb{R} \text{ s.t. } (a, \infty) \subset G \end{cases}$$

The same can be done with a single symbol as well. In either case, the extended real line is a compact set. We also extend the general operations so that $a + \infty = \infty$ for any $a \in (0, \infty]$, and $\infty = \sup[0, \infty] = \sup[0, \infty)$, and similarly for $-\infty$.

[***TODO: Stone-Cech compactification?***] [***TODO: One-point or Alexandroff extension / compactification***] [***TODO: connectivity***] [***TODO: separation axioms (see Munkres)***]

II. Algebra

3 Group Theory

4 Commutative Algebra

[***TODO: Add a section on language / terminology from module theory***] [***TODO: section on linear algebra? maybe save that for algebra lecture notes...***] [***TODO: commutative algebra***]

5 Linear Algebra

5.1 Vector Spaces

Definition. Let $\mathfrak{k}$ be a field. Then X is a **vector space** if X is a $\mathfrak{k}-$module. If $Y \subset X$ is a $\mathfrak{k}-$vector space, we say that Y is a **subspace** of X.

Given a subset $S \subset X$, we write

$$\operatorname{span} S = \{\alpha_1 v_1 + \cdots + \alpha_n v_n : \alpha \in \mathfrak{k}, v_i \in S\}$$

to denote the smallest subspace of X containing S.

A basic property of vector spaces is that they have bases:

Definition. Let X be a vector space over $\mathfrak{k}$ A subset $S \subseteq X$ is called

- **linearly independent** if for any distinct $x_1, \ldots, x_n \in S$, the equation $0 = \alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ where $\alpha_i \in \mathfrak{k}$ implies $\alpha_1 = \cdots = \alpha_n = 0$,
- **spanning** if $\operatorname{span} S = X$, and
- a **Hamel basis** if it is both linearly independent and spanning

As a basic consequence of Zorn's lemma, we have the following result:

> ***5.1 Lemma.*** *Suppose $S \subset X$ is linearly independent. Then there exists a Hamel basis M containing S.*

Proof Let $\mathcal{L} = \{S \subset L \subseteq X : L \text{ is linearly independent}\}$. Then $(\mathcal{L}, \subseteq)$ is a poset. If $\mathcal{C} \subseteq \mathcal{L}$ is a chain, it is easy to verify that $U = \bigcup_{L \in \mathcal{C}} L \in \mathcal{L}$ and is an upper bound for $\mathcal{C}$. Thus by Zorn's lemma, $\mathcal{L}$ has some maximal element L; by maximality, M is spanning for X. ■

> ***5.2 Proposition.*** *Let X be a vector space. Then X admits a Hamel basis M. Moreover, if M, M' are distinct Hamel bases, them $|M| = |M'|$.*

Proof Apply Theorem 5.1 to $\emptyset \subset X$ to get some Hamel basis $M = \{v_i : i \in I\}$. For the second part, let $\{w_j : j \in J\}$ be any spanning set: it suffices to show that $|M| \leq |J|$. If J is finite, this is a standard exercise in (finite dimensional) linear algebra.

Now suppose J is infinite and for any $j \in J$, write $w_j = \sum_{i \in A_j \subset I} \lambda_{i,j} v_i$ where $A_j \subset I$ is finite. Thus we have an injection $\phi : \bigcup_{j \in J} \{j\} \times A_j \to M$. But then since J is infinite, $|J| = \left| \bigcup_{j \in J} A_j \right| \leq |M|$ as required. ■

The previous proposition implies that the following notion is well-defined:

Definition. The **dimension** of a vector space X is the cardinality of any Hamel basis for X.

An important feature of any category is to consider the set of maps between objects.

Definition. Given vector spaces X and Y over $\mathfrak{k}$, let $\mathcal{L}(X, Y)$ denote the set of all linear maps from X to Y. In particular, when $Y = \mathfrak{k}$, we write $\mathcal{L}(X, \mathfrak{k}) = X'$ and call this the **algebraic dual** of X'.

Note that $\mathcal{L}(X, Y)$ also has a natural vector space structure given by $(\alpha S + T)(u) = \alpha S(u) + T(u)$. We will discuss such spaces in more details later.

III. Analysis in $\mathbb{R}^n$ and Vector Calculus

[***TODO: Include a bunch of standard sequential inequalities (and proofs) - maybe see book on real analysis?***]

Some things to prove:

- l'Hôpital
- Stolz-Cesàro: application, proof that harmonic sum is asymptotic to the natural logarithm
- mean value tricks $(1/n^p - 1/(n+1)^p \; p/n^{p+1})$

6 The Real Line

6.1 Existence and Uniqueness

We begin with a rigorous construction of the real line. The standard argument through Cauchy sequences is relegated to the section on metric completions in the following section; in this section, we present another standard argument using Dedekind cuts.

6.1 Theorem. *There exists a field $\mathbb{R}$ satisfying the following axioms:*

[***TODO: Include a construction of $\mathbb{R}$: metric completion, or Dedekind cuts?***] [***TODO: How about a proof of uniqueness of $\mathbb{R}$ is a complete ordered field***] We say that $\mathbb{R}$ satisfies the *Archimedean Principle*, which states that for any $r \in \mathbb{R}$, there exists some $N \in \mathbb{N}$ so that $N \geq r$.

Definition. Let $S \subseteq \mathbb{R}$. An *upper bound* for S is some $r \in \mathbb{R}$ such that for any $x \in S$, $x \leq r$. If a set has an upper bound, then we say it is *bounded above*. If x is an upper bound such that for any upper bound y of S, $x \leq y$, then x is a *least upper bound*, and write $x = \sup S$.

We can similarly define lower bounds and greatest lower bounds, denoted by $\inf S$.

6.2 Sequences and Completeness

A fundamental notion in the theory of metric spaces is that of a Cauchy sequence. Here, we first present the definition over the real line.

Definition. We say that $(x_n)_{n=1}^{\infty}$ is *Cauchy* if for any $\epsilon > 0$, there exists some $N \in \mathbb{N}$ so that for all $n, m \geq N$, $|x_n - x_m| < \epsilon$.

Definition. The *limit inferior*, denoted $\liminf$, is defined

$$\liminf_{n \in \mathbb{N}} x_n = \lim_{n \in \mathbb{N}} \left(\inf_{k \geq n} x_k \right)$$

which converges since the sequence on the right hand side is monotonic. Similarly, we define the *limit superior*, denoted $\limsup$, as

$$\limsup_{n \in \mathbb{N}} x_n = \lim_{n \in \mathbb{N}} \left(\sup_{k \geq n} x_k \right).$$

Here, we list some basic properties of the limit inferior and limit superior:

6.2 Proposition. *(i) Every bounded sequence has a subsequence that converges to* $\limsup x_n$ *and a subsequence that converges to* $\liminf x_n$*.*

(ii) $L = \limsup x_n$ *if and only if for any* $\epsilon > 0$*,* $x_n < L + \epsilon$ *for all but finitely many* n *and* $x_n > L - \epsilon$

(iii) $\liminf(x_n) \leq \limsup(x_n)$*, with equality if and only if* $\lim_{n\to\infty} x_n = L$ *exists. In this case, equality holds at* L*.*

The proof of this proposition is easy and is left as an exercise.

6.3 Theorem. *The following are equivalent:*

(i) Every bounded, increasing sequence converges.

(ii) Every Cauchy sequence has a unique limit.

(iii) (Bolzano-Weierstrass): Every bounded sequence has a convergent subsequence.

(iv) Every bounded set has a least upper bound.

6.4 Theorem. (Bolzano-Weierstrass) *Every bounded sequence has a convergent subsequence.*

PROOF $\limsup x_n$ is the limit of a subsequence of x_n. ■

PROOF [***TODO: prove***] ■

IV. Analysis in Metric Spaces

7 METRIC TOPOLOGY

Definition. The pair (X, d) consisting of a set X along with a function $d : X \times X \to [0, \infty)$ (called a *distance function* or *metric*) satisfying

(i) *(symmetric)*: $d(x, y) = d(y, x)$

(ii) *(triangle inequality)*: $d(x, y) \leq d(x, z) + d(z, y)$ for all x, y, z

(iii) *(positive definite)*: $d(x, y) = 0$ if and only if $x = y$

is called a *metric space*.

As is commonly used, we write $B(x, r) = \{y \in X : d(x, y) < r\}$ to denote a *basic open ball* about x with radius r. We call these open balls since metric space are equipped with a natural topology induced by the basis

$$\mathcal{B} = \{B(x, r) : x \in X :, r > 0\}$$

and these balls in fact provide a basis for each point in X. As a result, one can see that this topology is first countable and Hausdorff. Equivalently, given $f_x(y) = d(x, y)$, we have $\tau_d = \sigma(X, \{f_x\}_{x \in X})$, i.e. the initial topology generated by the distance evaluation maps.

PROOF [***TODO: prove***] ■

A notable advantage of metric spaces over topological spaces is that, rather than passing to nets, it suffices to consider sequences. Here, we collect (without proof) the sequential versions of the net results proven earlier

7.1 Proposition.

1. *x_0 is a cluster point of $(x_n)_{n=1}^{\infty}$ if and only if there is a subsequence converging to x_0*

2. *f is continuous if and only if for any sequence $(x_n)_{n=1}^{\infty}$ converging to x_0, then $(f(x_n))_{n=1}^{\infty}$ converges to $f(x_0)$.*

3. *$x \in \overline{E}$ if and only if there is a sequence $(x_n)_{n=1}^{\infty} \subset E$ such that $x_n \to x$.*

Suppose (X, d_X) and (Y, d_Y) are metric spaces. Define a function d by $d((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2)$ for $(x_j, y_j) \in X \times Y$ for $j = 1, 2$. It is easy to see that d defines a metric on $X \times Y$.

7.1 COMPLETENESS

Core to the idea of a metric space, and generalizing one of the fundamental properties of the real line, is the notion of metric completeness. The approach is nearly identical to the notion of completeness in the real line.

Definition. We say that $(x_n)_{n=1}^{\infty}$ is *Cauchy* if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n, m \geq N$, $d(x_n, x_m) < \epsilon$.

Trivially, any Cauchy sequence is bounded and every convergent sequence is Cauchy.

Definition. A metric space in which every Cauchy sequence converges (to an element in X) is called *complete*.

For example, $\mathbb{R}^n$ is complete, but $\mathbb{Q}$ is not complete. Discrete metric spaces are complete because the only Cauchy sequences are sequences that are eventually constant. The following proposition is a straightforward exercise.

7.2 Proposition. *If X is complete and $E \subset X$ is closed, then E is complete.*

We also have results for products of metric spaces:

7.3 Proposition. *If X, Y are complete, then so is $X \times Y$.*

Proof Let $((x_n, y_n))_{n=1}^{\infty}$ be Cauchy in $X \times Y$; then $(x_n)_{n=1}^{\infty}$ and $(y_n)_{n=1}^{\infty}$ are Cauchy in X and Y respectively. Say $(x_n)_{n=1}^{\infty} \to x$ and $(y_n)_{n=1}^{\infty} \to y$, so $(x_n, y_n)_{n=1}^{\infty} \to (x, y)$ in $X \times Y$. ■

7.2 Compactness

In this section, we discuss equivalent characterizations of compactness in the metric space section and some useful consequences. The main notion that one need generalize is the idea of boundedness:

Definition. A finite set $\{x_1, \ldots, x_n\} \subseteq X$ is called an $\epsilon-$*net* for $A \subseteq X$ if for any $x \in A$, there exists some j so that $d(x, x_j) < \epsilon$. We then say that A is *totally bounded* if A has an $\epsilon-$net for any $\epsilon > 0$.

7.4 Proposition. *If A is totally bounded, so is $\overline{A}$.*

Proof Take an $\frac{\epsilon}{2}-$net for A and use Theorem 7.1 to verify that it works. ■

7.5 Theorem. (Cantor Intersection) *If $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$ are non-empty closed sets in a complete metric space X and $\operatorname{diam} A_n \to 0$ where $\operatorname{diam} A = \sup\{d(x, y) : x, y \in A\}$, then $\bigcap_{n=1}^{\infty} A_n$ is exactly one point.*

Proof For each $n \in \mathbb{N}$, let $x_n \in A_n$ be arbitrary. Since $\operatorname{diam} A_n \to 0$, $(x_n)_{n=1}^{\infty}$ is Cauchy in X and hence admits a limit x_0; in particular, since the A_i are nested and closed, $x_0 \in A_i$ for each i so that $x_0 \in \bigcap_{n=1}^{\infty} A_n$. If $y \in \bigcap_{n=1}^{\infty} A_n$ is any point, then $d(x_0, y) \leq \operatorname{diam} A_n$ for each $n \in \mathbb{N}$ so that $x_0 = y$. ■

7.6 Theorem. *Let X be a metric space. Then the following are equivalent:*
(i) X is compact.
(ii) Every sequence in X has a convergent subsequence.
(iii) X is complete and totally bounded.

Proof ($i \Rightarrow ii$) Let (x_n) be a sequence in X and define $S_n = \{x_k : k \geq n\}$. Then $\{\overline{S_n}\}_{n=1}^{\infty}$ is nested and has the finite intersection property, so by Theorem 2.3, there exists some $x \in \bigcap_{n=1}^{\infty} \overline{S_n}$. Construct a subsequence as follows:

- Let $x_{k_1} \in S_1$ so that $d(x, x_{k_1}) < 1$.
- Having chosen $x_{k_1} < \ldots < x_{k_n}$ so that $d(x, x_{k_j}) < 1/j$ for each $1 \leq j \leq n$, get $x_{k_{n+1}} \in S_{k_n+1}$ so that $d(x, x_{k_{n+1}}) < 1/n$.

By construction, $(x_{k_j})_{j=1}^\infty \to x$, as required.

$(ii \Rightarrow iii)$ Let $(x_n)_{n=1}^\infty \subset X$ be Cauchy and let $(x_{n_j})_{j=1}^\infty \to x$ be a convergent subsequence. Straightforward verification yields that $(x_n)_{n=1}^\infty \to x$, so that X is in fact complete.

To show that X is totally bounded, suppose not, and get some $\epsilon > 0$ so that X has no $\epsilon-$net. Thus we may inductively get $(x_k)_{k=1}^\infty$ so that $d(x_i, x_j) \geq \epsilon$ for any $i \neq j$. Such a sequence has no convergent subsequence, a contradiction.

$(iii \Rightarrow i)$ Suppose X is complete and totally bounded and, for contradiction, let $\{U_\alpha\}_{\alpha \in A}$ be an open cover with no finite subcover. Since X is totally bounded, for any $n \in \mathbb{N}$, X has a $1/2^n$ =net, say $\{x_1^{(n)}, \dots, x_{k_n}^{(n)}\}$. Write $X = \bigcup_{j=1}^{k_n} \overline{B(x_j^{(n)}, 1/2^n}$ so that, without loss of generality, $D_n := \overline{B(x_1^{(n)}, 1/2^n)}$ cannot be covered by finitely many U_α. Note that D_n is complete and totally bounded, so we may repeat the above argument with D_n in place of X inductively to get $D_1 \supseteq D_2 \supseteq \cdots$, D_i closed, and $\operatorname{diam} D_i \leq 1/n$. By Theorem 7.5, get $x_0 \in \bigcap_{n=1}^\infty D_n$; say $x_0 \in U_{\alpha_0}$. Since U_{α_0} is open, get $B(x_0, \epsilon) \subset U_{\alpha_0}$. But then if m is so that $1/2^m < \epsilon$, since $x_0 \in D_j$ for each j, $D_m \subset B(x_0, \epsilon) \subset U_{\alpha_0}$. This contradicts the fact that no D_j has a finite subcover from $\{U_\alpha\}_{\alpha \in A}$. ■

Note that $E \subset \mathbb{R}^n$ is bounded if and only if it is totally bounded, so we immediately deduce the following.

7.7 Corollary. *For $A \subseteq \mathbb{R}^n$, the following are equivalent:*
- *(i) A is compact*
- *(ii) A is closed and bounded*
- *(iii) Every sequence in A has a convergent subsequence with limit in A*

Clasically, the equivalence $(i \Leftrightarrow iii)$ is the Bolzano-Weierstrass Theorem, and the equivalence $(i \Leftrightarrow ii)$ is the Heine-Borel Theorem. [***TODO: compact Hausdorff is separable?***]

7.8 Proposition. *A compact metric space is separable.*

PROOF Consider $\{B(x, 1/n) : x \in X\}$, an open cover of X. By compactness of X, there is a finite subcover, say $B(x_1^n, 1/n), \dots, B(x_{r_n}^n, 1/n)$. Let $K_n = \{x_1^n, x_2^n, \dots, x_{r_n}^n\}$ and set $K = \bigcup_{n=1}^\infty K_n$, a countable set. Check that K is dense, so let $x \in X$. Since each K_n is an open cover, $x \in K_n$ for all n, and in particular is in some $B(x_{i(n)}^n, 1/n)$ for all n. But then by construction $(x_{i(n)}^n)_{n=1}^\infty \to x$ so K is dense. ■

7.3 BAIRE CATEGORY AND CONSEQUENCES

Definition. Let (X, d) be a metric space. A subset $F \subset X$ is *nowhere dense* if $X \setminus F$ is dense in X; equivalently, $\overline{F}$ contains no non-trivlal open subsets. We say that a subset $M \subseteq X$ is *meagre* (or *first category*) if $M = \bigcup_{n=1}^\infty F_n$ and each F_n is nowehere dense; and a set is *non-meagre* (or *second category*) otherwise.

The main relevant discussion in this section is the following theorem, which is straightforward in content but deep in consequences.

7.9 Theorem. (Baire Category I) *If (X, d) is a complete metric space and $\{U_n\}_{n=1}^\infty$ is a countable collection of dense, open subsets, then $\bigcap_{n=1}^\infty U_n$ is dense in X.*

PROOF [***TODO: strengthen proof to show density, rather than non-empty***] Let $\{A_n\}_{n=1}^\infty$ be open and dense, and show $\bigcap_{n=1}^\infty A_n \neq \emptyset$. Get $x_1 \in A_1$, so get $B(x_1, r_1) = U_1 \subseteq A_1$.

A_2 is dense, so intersects U_1, and $A_2 \cap U_1 \neq \emptyset$. Say $x_2 \in A_2 \cap U_1 \subseteq A_2 \cap A_1$. Since finite intersections are open, get $V_2 \subseteq A_2 \cap U_1$, without loss of generality $r_2 \leq r_1/2$. Set $U_2 = B(x_2, r_2/2)$. Then $\overline{U}_2 \subseteq V_2 \subseteq A_2 \cap U_1$, and $\operatorname{diam} \overline{U}_2 \leq \frac{1}{2} \operatorname{diam} \overline{U}_1$.

Proceed inductively to get x_n, open sets $U_n \ni x_n$ and $U_n \subseteq \bigcap_{k=1}^n A_k$, $\overline{U}_n \subseteq U_{n-1}$ and $\operatorname{diam} U_n \to 0$. Check (x_n) is Caucy. Let $\epsilon > 0$ and pick N such that $\operatorname{diam} U_N < \epsilon$. If $n, m \geq N$, then $x_n, x_m \in U_n$ so $d(x_n, x_m) < \operatorname{diam} U_n < \epsilon$. Thus (x_n) is Cauchy so by completeness, $\lim x_n = x \in X$. Furthermore, $x_n \in \overline{U}_n$ for all $n \geq N$, so $x \in \overline{U}_n \subseteq U_{n-1} \subseteq \bigcap_{j=1}^{n-1} A_j$ for all n. But then $x \in A_j$ for all j, so the intersection is non-empty. ■

7.10 Theorem. (Baire Category II) *Let (X, d) be a complete metric space. Then a non-empty open $U \subseteq X$ is non-meagre.*

PROOF Suppose not, so $U = \bigcup_{n=1}^\infty F_n \subseteq \bigcup_{n=1}^\infty \overline{F}_n$, each F_n (hence $\overline{F_n}$) nowhere dense. Then each $V_n = X \setminus \overline{F_n}$ is open and dense, and hence by BCT I, $G = \bigcap_{n=1}^\infty V_n$ is dense in X, and hence $U \cap G \neq \emptyset$, violating assumption ■

Theorem 7.9 can furnish proofs of many interesting results; we will spend the remainder of this section discussing some of the consequences.

Example. Say that a set S is *perfect* if S is closed and $S = \overline{S \setminus \{x\}}$ for any $x \in S$. Then Theorem 7.9 implies that S is uncountable.

Suppose E is countable and write $E = \bigcup_{n=1}^\infty \{x_n\}$. Note that each $\{x_n\}$ cannot be open in the subspace topology since x_n is an accumulation point, so that $\{x_n\}$ is closed and nowhere dense, contradicting Theorem 7.9.

Since each x_n is an accumulation point of E, any open set must contain infinitely many points of E. Consider E as a metric space, and since E is closed, E is complete. Furthermore, $\{x_n\}$ is not open in E for it not, then $\{x_n\} = V \cap E$ where V is open in X, contradicting the fact that $V \cap E$ must contain infinitely many points in E. Thus E is a countable union of closed, nowhere dense sets, a contradiction to the Baire Category Theorem.

7.11 Theorem. *No function on $[0, 1]$ can be continuous on precisely the rational points.*

PROOF We first see that the set of points at which the function $f : X \to \mathbb{R}$ is discontinuous is a countable union of closed sets. Define the set $S_k = \{x \in X : \forall \delta > 0, \exists y, z \in B(x, \delta) \text{ s.t. } |f(y) - f(x)| \geq 1/k\}$. I will show that (1) the S_k are closed and (2) $D_f = \bigcup_{k=1}^\infty S_k$ where D_f denotes the set of points of discontinuity of f.

To see that S_k is closed, let $x \in S_k^c$. Thus we have for some $\delta > 0$ for any $y, z \in B(x, \delta)$ we have $|f(y) - f(x)| < 1/k$. But then for any $y \in B(x, \delta)$, choose δ' so $B(y, \delta') \subseteq B(x, \delta)$ and for any $y, z \in B(y, \delta')$ we have $|f(y) - f(x)| < 1/k$. Therefore $B(x, \delta) \subseteq S_k^c$ so S_k^c is open.

Furthermore, we certainly have $S_k \subseteq D_f$ by definition, and if $x \in D_f$, then there exists an $\epsilon > 0$ so that for all $\delta > 0$, there exist $y, z \in B(x, \delta)$ so that $|f(y) - f(x)| \geq \epsilon$. Then let k be such that $1/k < \epsilon$ so $x \in S_k$. Thus D_f is a countable union of closed sets.

It now suffices to show that $\mathbb{R} \setminus \mathbb{Q}$ cannot be written as a countable union of closed sets. Suppose for contradiction that $\mathbb{R} \setminus \mathbb{Q} = \displaystyle\bigcup_{i=1}^\infty V_i$, and write

$$\mathbb{R} = \bigcup_{i=1}^\infty V_i \cup \bigcup_{i=1}^\infty \{r_i\}$$

where $\{r_i\}$ is an enumeration of $\mathbb{Q}$. However by BCT, $\mathbb{R}$ is second category, so some V_k must not be nowhere dense. But then since the V_i are closed, $V_k = \overline{V_k}$ has non-empty interior, a contradiction since $\mathbb{R} \setminus \mathbb{Q}$ has empty interior.

However, by (8a), D_f is a countable union of closed sets, a contradiction so $D_f \neq \mathbb{R} \setminus \mathbb{Q}$. ■

Exercise: generalize this statement to continuous functions on $C(X)$!

7.12 Theorem. *The set of functions on $C[0,1]$ which have a derivative at some point on $(0,1)$ is first category.*

PROOF Let

$$E_j = \left\{ f \in C[0,1] : \exists x \in [0, 1-1/j] \text{ s.t. } \forall h \in (0, 1/j], \left| \frac{f(x+h) - f(x)}{h} \right| \leq j \right\}$$

and define

$$E = \{f \in C[0,1]\ f \text{ has a derivative somewhere}\}$$

We now prove the theorem in three parts

1. $E \subseteq \bigcup_{j=1}^{\infty} E_j$
2. E_j has empty interior.
3. E_j is closed.

Once this is done, then notice

$$E = \bigcup_{j=1}^{\infty} (E \cap E_j)$$

so $\overline{(E \cap E_j)} \subseteq \overline{E_j} = E_j$. Then $\overline{E \cap E_j}$ has empty interior since E_j has empty interior, so $E \cap E_j$ is nowhere dense and E is first category. Let's now proceed to prove those three statements!

(1) Say $f \in E$ and f is differentiable at $x \in (0,1)$. Thus there exist some j_1 such that for all $j \geq j_1$, $x \in [0, 1-1/j]$. Since $\lim_{h \to 0} \frac{f(x+h)-f(x)}{h}$ exists, we have C such that for all $h \leq j_2$, we have

$$\left| \frac{f(x+h) - f(x)}{h} \right| \leq C \leq j_3$$

Let $j = \max\{j_1, j_2, j_3\}$, then $f \in E_j$.

(2) To show that E_j has empty interior, we will show that if $f \in E_j$ and $\epsilon > 0$, then there exists $g \in B(f, \epsilon)$ such that $g \notin E_j$. First, by Stone-Weierstrass get p such that $d(p, f) < \epsilon/2$. We then have some M such that $|p'(x)| \leq M$ for all $x \in (0,1]$. Let Q be a piecewise linear function of slope $\pm(M + j + 1)$ with $0 \leq Q(x) \leq \epsilon/2$. Set $g = p + Q$, so

that $d(f,g) < \epsilon$. Now let $x \in (0,1)$ and we have

$$\begin{aligned}\left|\frac{g(x+h)-g(x)}{h}\right| &= \left|\frac{p(x+h)-p(x)+q(x+h)-q(x)}{h}\right| \\ &\geq -\left|\frac{p(x+h)-p(x)}{h}\right| + \left|\frac{Q(x+h)-Q(x)}{h}\right| \\ &\geq -\left(M+\frac{1}{2}\right) + M + j + 1 \\ &\geq j + \frac{1}{2}\end{aligned}$$

for sufficiently small h.

(3) We finally see that E_j is closed. Suppose $f_n \in E_j$ and $f_n \to f$ uniformly. We see that $f \in E_j$. By definition of E_j, for all n there exists $x_n \in [0, 1-1/j]$ such that for any $h \in (0, 1/j)$

$$\left|\frac{f_n(x_n+h)-f_n(x_n)}{h}\right| \leq j$$

By Bolzano-Weierstrass, there exists a subsequence $x_{n_k} \to x_0 \in [0, 1-1/j]$. We will ceck that x_0 "works" for f, so $f \in E_j$. Let $\epsilon > 0$ and fix $h \in [0, 1/j]$. We have M such that $d(f_m, f) < \epsilon h/4$ for all $m \geq M$. Furthermore, f is uniformly continuous so there exists $\delta > 0$ such that $|f(x) - f(y)| \leq \epsilon h/4$ whenever $|x-y| < \delta$. We also have M_2 such that $|x_m - x_0| < \delta$ if $m \geq M_2$, and fix $M = \max\{M_1, M_2\}$. Finally

$$\begin{aligned}\frac{|f(x_0+h)-f(x_0)|}{h} &\leq \frac{|f(x_0+h)-f(x_m+h)|}{h} + \frac{|f(x_m+h)-f_M(x_m+h)|}{h} + \frac{|f_M(x_m+h)-f_M(x_M)|}{h} \\ &\quad + \frac{|f_M(x_M)-f(x_M)|}{h} + \frac{|f(x_M)-f(x_0)|}{h} \\ &\leq \frac{\frac{\epsilon h}{4}}{h} + \frac{d(f,f_M)}{h} + j + \frac{d(f_m,f)}{h} + \frac{\frac{\epsilon h}{4}}{h} \\ &\leq \epsilon + j\end{aligned}$$

Since $\epsilon > 0$ was arbitrary, we must have

$$\frac{|f(x_0+h)-f(x_0)|}{h} \leq j$$

for all $h \in (0, 1/j)$. ■

7.13 Corollary. *The set of nowhere differentiable functions is second category.*

PROOF The union of two first category sets is first category, and $C[0,1]$ is second category. ■

8 THEORY OF CONTINUOUS FUNCTION SPACES

8.1 CONTINUOUS FUNCTIONS ON METRIC SPACES

A common type of metric space that one will encounter is a space of functions equipped with a uniform norm. We approach this topic first through uniform continuity.

Definition. Let (X, d_X) and (Y, d_Y) be metric spaces. Then $f : X \to Y$ is *uniformly continuous* if for all $\epsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), f(y)) < \epsilon$ whenever $d_X(x, y) < \delta$.

8.1 Theorem. *Let K be compact and $f : K \to Y$ continuous. Then f is uniformly continuous.*

PROOF Fix $\epsilon > 0$. Since f is continuous, for all $x \in K$, there exists $\delta_x > 0$ such that if $d_X(x, y) < \delta$ then $d_Y(f(x), f(y)) < \epsilon$. Consider $B(x, \delta_x/2)$ as an open cover, and get a finite subcover, say $\{B(x_i, \delta_{x_i}/2)\}$ for $i = 1, \dots, n$ and choose $\delta = \min\{\delta_{x_i}\}$. Suppose $x, y \in K$ and $d_X(x, y) < \delta$. Then choose i such that $x \in B(x_i, \delta_{x_i}/2)$. Then

$$d(y, x_i) \leq d(y, x) + d(x, x_i) < \delta + \frac{\delta_{x_i}}{2} \leq \delta_{x_i}$$

but this implies $d(x, x_i) < \delta_{x_i}$ so that $d(f(x), f(x_i)) < \epsilon$. Furthermore, $d(x_i, y) < \delta_{x_i}$ so

$$(f(y), f(x_i)) < \epsilon \Rightarrow d(f(x), f(y)) < 2\epsilon$$

by the triangle inequality. ■

Given metric spaces X, Y, let $C(X, Y)$ denote the *space of continuous functions* $f : X \to Y$. $C(X, Y)$ is naturally a metric space equipped with the uniform distance $d(f, g) := \sup_{x \in X} d_Y(f(x), g(x))$ where d_Y is the metric on Y. Often, we will denote by $C(X) := C(X, \mathbb{R})$. We also denote $C_b(X) = \{f \in C(X) : \sup_{x \in X} |f(x)| < \infty\}$; if X is compact, certainly $C(X) = C_b(X)$.

Definition. We say that A is an *algebra* over a field $\mathfrak{k}$ if A is a vector space equipped with a multiplication $\cdot : A \times A \to \mathfrak{k}$ that is associative and distributive over vector addition.

In this sense, $C(X)$ and $C_b(X)$ are (infinite dimensional) vector spaces with pointwise addition and scalar multiplication, and algebra struture given by pointwise multiplication. In fact, it is a normed space with respect to the norm $\|f\| = d(f, 0)$. The following proposition verifying that the spaces are indeed algebras is left as an exercise:

8.2 Proposition. *If $f, g : X \to \mathbb{R}$ is continuous, then so are $f \pm g$, $f \cdot g$, f/g for $g \neq 0$. In particular, $C(X)$ is an algebra and inverse-closed.*

Let's now define two notions of convergence:

Definition. We say $(f_n)_{n=1}^{\infty}$ *converges pointwise* if for each $x \in X$, the sequence $(f_n(x))_{n=1}^{\infty}$ converges in Y. We say $(f_n)_{n=1}^{\infty} \to f$ *converges uniformly* if the convergence is with respect to the uniform norm.

Uniform limits are categorically a more correct notion of convergence to consider than pointwise convergence for continuous functions, since uniform limits preserve continuity. In fact, a sequence of functions is Cauchy if and only if it is uniformly continuous. We see this summarized in the following theorem:

8.3 Theorem. *Let X, Y be metric spaces where Y is complete, and $(f_n)_{n=1}^{\infty}$ where each $f_n : X \to Y$. Then $(f_n)_{n=1}^{\infty}$ is uniformly Cauchy if and only if it is uniformly convergent.*

PROOF ($\Longleftarrow$) This direction does not require completeness. Say $(f_n)_{n=1}^{\infty} \to f$ uniformly and get N so that for all $n \geq N$ and $x \in X$, $d(f_n(x), f(x)) < \epsilon/2$. Then if $n, m \geq N$,

$$d(f_n(x), f_m(x)) \leq d(f_n(x), f(x)) + d(f(x), f_m(x)) < \epsilon$$

which holds for any $x \in X$ so that $(f_n)_{n=1}^\infty$ is uniformly Cauchy.

($\Longrightarrow$) Conversely, suppose $(f_n)_{n=1}^\infty$ is uniformly Cauchy so that $(f_n(x))_{n=1}^\infty \subset Y$ is Cauchy for each $x \in X$. Since Y is complete, $(f_n(x))_{n=1}^\infty \to f(x)$ for some pointwise limit f; it remains to verify that the convergence is uniform.

Fix some $\epsilon > 0$. For any $x \in X$, get M_x so that for all $m \geq M_x$, $d(f_m(x), f(x)) < \epsilon/2$. Since the sequence is uniformly Cauchy, get N so that for all $n, m \geq N$ and $x \in X$, $d(f_n(x), f_m(x)) < \epsilon/2$. Let $x \in X$ be arbitrary and $M \geq \max(N, M_x)$, so that

$$d(f_n(x), f(x)) \leq d(f_n(x), f_m(x)) + d(f_m(x), f(x)) < \epsilon$$

for all $n \geq M$ and $x \in X$, so that $(f_n)_{n=1}^\infty \to f$ uniformly. ■

8.4 Corollary. *Let (X, d_X) and (Y, d_Y) be metric spaces where Y is complete. Then $C(X, Y)$ is a Banach space.*

8.5 Theorem. (Weierstrass M-test) *Let $(f_n)_{n=1}^\infty$ where each $f_n : X \to Y$, X a metric space, and Y a Banach space. Suppose there exist constants $M_n \in \mathbb{R}$ so that $\|f_n(x)\|_Y \leq M_n$ for all x, n and $\sum_{n=1}^\infty M_n < \infty$. Then $\sum_{n=1}^\infty f_n$ converges uniformly.*

PROOF Let $S_n(x) = \sum\limits_{k=1}^{n} f_n(x)$. Then for $n < m$,

$$\begin{aligned} d(S_n(x), S_m(x)) = \|S_n(x) - S_m(x)\| = \left\| \sum_{k=m+1}^{n} f_k(x) \right\| &\leq \sum_{k=1}^{n} \|f_k(x)\| \\ &\leq \sum_{k=n+1}^{n} M_k < \epsilon \end{aligned}$$

for sufficiently large n, m since $\sum_{k=1}^\infty M_k$ converges independently of x. ■

8.6 Theorem. (Dini's Theorem) *Suppose K is compact and $(f_n)_{n=1}^\infty$ with each $f_n : K \to \mathbb{R}$ converges pointwise to some f. If the f_n, f are continuous and $f_1 \leq f_2 \leq \cdots$, then $(f_n)_{n=1}^\infty \to f$ uniformly.*

PROOF Let $g_n(x) = f_n(x) - f(x)$. We want to show $g_n \to 0$ uniformly. We certainly have $g_n \to 0$ pointwise, that $g_n \geq 0$ and g_n are monotonic (from monotonicity of f_n).

Let $\epsilon > 0$. We want to show that there exists N such that for all $n \geq N$, $0 \leq g_n(x) < \epsilon$. Since $g_n \to 0$ pointwise, for all $t \in K$ there exists N_t such that $0 \leq g_{N_t}(t) < \epsilon/2$ implies $0 \leq g_n(t) < \epsilon/2$ for all $n \geq N_t$. Furthermore, since each g_{N_t} is continuous, there exists $\delta_t > 0$ so that $d(t, y) < \delta_t$ implies $|g_{N_t}(t) - g_{N_t}(y)| < \epsilon/2$. Consider $B(t, \delta_t) \subseteq K$ for each $t \in K$. These form an open cover of K, so take a finite subcover

$$B(t_1, \delta_{t-1}), \ldots, B(t_r, \delta_{t_r})$$

Now suppose $x \in B(t_i, \delta_{t_i})$ so that $d(x, t_i) < \delta_{t_i}$

$$|g_{N_{t_i}}(x)| \leq |g_{N_{t_i}}(t_i) - g_{N_{t_i}}(x)| + |g_{N_{t_i}}(t_i)| < \epsilon$$

But then fix $N = \max\{N_{t_1}, \ldots, N_{t_r}\}$. Now suppose $n \geq N$, and let $x \in K$ be arbitrary. We have some i so that $x \in B(t_i, \delta_{t_i})$. By monotonicity of g_n and the choice of N, we have

$$0 \leq g_n(x) \leq g_N(x) \leq g_{N_{t_i}}(x) < \epsilon$$

and since this holds for every $x \in K$ and $\epsilon > 0$, we have $g_n \to 0$ uniformly. ■

Example. Consider $U = \{f \in C([0,1]) : f(x) > 0 \forall x \in [0,1]\}$. Then U is open.

PROOF Take $f \in U$. By E.V.T. f has a minimum, say $f(x_1) \leq f(x)$ for all $x \in [0,1]$. Set $\epsilon = f(x_1)$. But then $B(f, \epsilon) \subseteq U$ since if $g \in B(f, \epsilon)$, then $\|f - g\| < \epsilon$. Thus for all x, $g(x) > f(x) - \epsilon = f(x) - f(x_1) \geq 0$ so that $g(x) > 0$ for all x, and $g \in U$.

We can also see that U^c is closed, where $U^c = \{f \in C([0,1]) : f(x) \leq 0 \text{ for some } x\}$. We will see that U^c is closed by checking it contains all its accumulation points. Say f is an accumulation point of U^c. Then $\exists f_n \to f$ uniformly where $f_n \in U^c$. We have some $x_n \in [0,1]$ such that $f_n(x_n) \leq 0$. Then by Bolzano-Weierstrass, there is a subsequence $x_{n_k} \to x_0$. Since f is continuous, notice $f(x_{n_k}) \to f(x_0)$ so that $f(x_{n_k}) \leq 0$, $f(x_0) \leq 0$. Thus $f \in U^c$ so we're done! ∎

9 THE FUNCTION SPACE $C(X)$

9.1 COMPACTNESS IN $C_b(X)$

Recall that $E \subseteq C_b(X)$ is compact if and only if it is complete and totally bounded, and that compactness implies closedness and boundedness (but the converse does not hold). In fact, $E \subseteq C_b(X)$ is bounded if and only if there exists N such that $B(0, N) \supseteq E$. So in our case, we can take $N = M + \|f\|$. If $g \in E$, then $d(g, f) < M$ so $d(g, 0) \leq d(g, f) + d(f, 0) < M + \|f\| = N$. Thus $E \subseteq C_b(X)$ is bounded if and only if there exists N such that $E \subseteq B(0, N)$ if and only of $\|f\| < N$ for all $f \in E$. We thus say that "E is uniformly bounded".

However, closed and bounded does not imply compact. For example, consider $E = \{x^n : n = 1, 2, \ldots\} \subseteq C([0,1])$. Then $\|x^n\| \leq 1$ for all $x^n \in E$, so E is bounded. If f is an accumulation point of E, then there exists n_k such that $f_{n_k} \to f$ uniformly. But

$$f_{n_k}(x) = x^{n_k} \to \begin{cases} 0 & : x \neq 1 \\ 1 & : x = 1 \end{cases}$$

Since $f_{n_k} \to f$ uniformly, it must also be pointwise. However, $f_{n_k} \to g$ pointwise so $g = f$, but $g \notin C[0,1]$. Since g is not continuous, we don't have $f_n \to g$ uniformly, so f cannot exist. Thus means that E has no accumulation points, so E is closed.

Furthermore, E is not compact because the sequence $(f_n(x) = x^n)$ has no convergent subsequence. So by Bolzano-Weierstrass, E is not compact.

Example. The set

$$E = \left\{ \frac{x^2}{x^2 + (1 - nx)^2} : n = 1, 2, 3, \ldots \right\} \subseteq C([0,1])$$

is closed, bounded, and not compact.

9.2 EQUICONTINUITY

Definition. Let $E \subseteq C_b(X)$. We say E is equicontinuous if for all $\epsilon > 0$, there exists $\delta > 0$ such that for all $f \in E$ and $x, y \in X$ such that $d(x, y) < \delta$ we have $|f(x) - f(y)| < \epsilon$.

This is to say that this family is uniformly continuous across the points and between functions as well. Given an ϵ, the same δ must show uniform continuity for all $f \in E$.

Example. 1. $E = \{f\}$ where f is uniformly continuous.

2. $E = \{f_1, f_2, \ldots, f_n\}$ is equicontinuous if and only if each f_i is uniformly continuous for $i = 1, \ldots, n$. Given $\delta > 0$, we take δ_i from each f_i being uniformly continuous, and take δ to be their minimum.

3. $E = \{x^n : n = 1, 2, 3, \ldots\} \subseteq C([0, 1])$ is not equicontinuous, even though each function is uniformly continuous. Not equicontinuous means that there exists $\epsilon > 0$ such that there exists $x, y \in X$ with $d(x, y) < \delta$ and some $f \in E$ with $f|(x) - f(y)| \geq \epsilon$. Take $\epsilon = 1/2$ and $\delta > 0$. Take $x = 1$, $y = 1 - \delta/2$, and choose n such that $y^n < 1/2$. But then $|f_n(x) - f_n(y)| = x^n - y^n > 1/2 = \epsilon$.

9.3 ARZELA-ASCOLI THEOREM

Our goal is to characterize compactness of $E \subseteq C_b(X)$. We have the following important theorem:

> ***9.1 Theorem.*** *Let X be compact. Assume the set $\{f_n : n = 1, 2, 3, \ldots\} \subseteq C(X)$ is a pointwise bounded, equicontinuous family. Then there exists a subsequence of $(f_n)_{n=1}^{\infty}$ which converges uniformly in $C(X)$.*

We will prove the theorem in parts.

> ***9.2 Lemma.*** *Let K be a countable set and let $\{f_n : K \to \mathbb{R}, n \in \mathbb{N}\}$ be a pointwise bonded family. Then there exists a subsequence of $(f_n)_{n=1}^{\infty}$ that converges pointwise at every $k \in K$.*

PROOF Let $K = \{x_j\}_{j=1}^{\infty}$. Start with $(f_n(x_1))_{n=1}^{\infty}$. This is a bounded sequence of real numbers since $\{f_n\}$ is pointwise bounded. Then by Bolzano-Weierstrass, there exists a convergent subsequence $(f_{n_j}(x_1))_{j=1}^{\infty}$. Rename f_{n_j} as $f_j^{(1)}$, so $(f_j^{(1)}(x_1))$ converges. Next, look at $(f_j^{(1)}(x_2))$, which as above has a convergent subsequence, say $(f_j^{(2)}(x_2))$. Repeat this to get a collection of convergent subsequences $(f_j^{(m)}(x_m))$ for $m \in \mathbb{N}$. But then by construction the sequence given by $(f_i^{(i)}(x_i))$ is eventually a subsequence of a convergent subsequence for any $x_i \in K$, so it converges at all $k \in K$. ∎

We can now prove the theorem!

PROOF Since X is compact, it is separable, so let $K \subseteq X$ be a countable dense subset. Consider $\{f_n|_K\}$, a pointwise bounded family. By Lemma 2, there is a subsequence (f_{n_k}) that converges pointwise on K. Our goal is to show that (f_{n_k}) converges uniformly on X. We will do this by showing that it is uniformly Cauchy on X. Thus let $\epsilon > 0$. Since $\{f_n\}$ are equicontinuous, there exists $\delta > 0$ such that $d(x, y), \delta$ implies that $|f_n(x) - f_n(y)| < \epsilon/3$ for all f_n. Now consider $\{B(x, \delta) : x \in K\}$. Since K is ense, for all $y \in X$, there exists $x \in K$ such that $y \in B(x, \delta)$. Thus these balls cover X. Since X is compact there is a finite subcover, say $B(x_1, \delta), B(x_r, \delta)$. We have $(f_{n_k}(x_i))_{k=1}^{\infty}$ converges for each $i \in 1, \ldots, r$. Thus for ech i, there are Caucy sequences (in $\mathbb{R}$). Hence for each i, there exists N_i such that if $n_k, n_l \geq N_i$, then

$$|f_{n_k}(x_i) - f_{n_l}(x_i)| < \frac{\epsilon}{3}$$

Let $N = \max\{N_1, \ldots, N_r\}$. Then if $n_k, n_l \geq N$, then $|f_{n_k}(x_i) - f_{n_l}(x_i)| < \epsilon/3$. Furthermore, if $y \in X$, then $y \in B(x_i, \delta)$ and $d(y, x_i) < \delta$ and equicontinuity implies $|f_{n_k}(y) - f_{n_l}(x_i)| <$

$\epsilon/3$ as well. But now with a direct application of the triangle inequality and the above statements, for any $x \in X$, we have

$$|f_{n_k}(x) - f_{n_l}(x)| \leq |f_{n_k}(x) - f_{n_k}(x_i)| + |f_{n_k}(x_i) - f_{n_l}(x_i)| + |f_{n_l}(x_i) - f_{n_l}(x)| < \epsilon$$

as desired. ■

9.3 Corollary. (Arzela-Ascoli) *Suppose X is compact. Then $E \subseteq C(X)$ is compact iff E is pointwise bounded, closed, and equicontinuous.*

PROOF ($\Rightarrow$) is independent of the theorem. Let E be compact. Then E is closed and bounded (meaning uniformly bounded and hence pointwise bounded). We check that E is equicontinuous. Suppose not, then there exists $\epsilon > 0$ such that for all $\delta = 1/n$, $n \in \mathbb{N}$, there exists $x_n, y_n \in X$ such that $d(x_n, y_n) < 1/n$ with $f_n \in E$ with $|f_n(x_n) - f_n(y_n)| \geq \epsilon$. Since E is comact by Bolzano-Weierstrass, there exists a convergent subseuence $(f_{n_k})_{k=1}^{\infty}$. By the previous proposition, $\{f_{n_k} : k = 1, 2, \ldots\}$ is equicontinuous. So for this choice of ϵ there is a δ that "works". Pick n_k large enough that $1/n_k < \delta$. Then

$$d(x_{n_k}, y_{n_k}) < \frac{1}{n_k} < \delta$$

so $|f_{n_k}(x_{n_k}) - f_{n_k}(y_{n_k})| < \epsilon$, a contradiction.

($\Leftarrow$) We will verify the Bolzano-Weierstrass characterization of compactness - that every sequence from E has a convergent subsequence with limit in E. Take a sequence $(f_n)_{n=1}^{\infty}$ in E. Since E is pointwise bounded and equicontinuous, so is $\{f_n : n \in \mathbb{N}\}$. Thus by the above theorem, there is a convergent subsequence, but since E is closed, this limit must be in E. ■

10 DENSE SUBSETS OF $C(X)$

10.1 POLYNOMIALS IN $C[0,1]$

10.1 Theorem. (Weierstrass) *Let $f : [0,1] \to \mathbb{R}$ be continuous and let $\epsilon > 0$. Then there is a polynomial p such that $\|p - f\| < \epsilon$. So $\{Poly\}$ is dense in $C[0,1]$.*

In fact, the Bernstein polynomial

$$p_n(x) = \sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$$

converge uniformly to f.

Intuitively, suppose we toss a biased coin with probability x of heads and $1 - x$ of tails. Suppose the payoff is $f(k/n)$ dollars if you get k heads and n tosses. Then the expected payoff is $\sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k} = p_n(x)$. In the long run we expected xn heads in n tosses, so the expected payoff in the long run should $f(xn/n) = f(x)$. With this in mind, let's prove the theorem

PROOF First, some technical calculations. By the Binomial theorem, we have

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

which we can view as a function of x. Differentiate with respect to x twice to get

$$n(x+y)^{n-1} = \sum_{k=0}^{n} \binom{n}{k} k x^{k-1} y^{n-k}$$

$$n(n-1)(x+y)^{n-2} = \sum_{k=0}^{n} \binom{n}{k} k(k-1) x^{k-2} y^{n-k}$$

Set $r_k(x) = \binom{n}{k} x^k (1-x)^{n-k}$ and $p_n(x) = \sum\limits_{k=0}^{n} f(k/n) r_k(x)$. At $y = 1-x$, we have $nx = \sum\limits_{k=0}^{n} k r_k(x)$ and $x^2 n(n-1) = \sum\limits_{k=0}^{n} k(k-1) r_k(x) = \sum\limits_{k=0}^{n} k^2 r_k(x) - nx$. So

$$\begin{aligned}
\sum_{k=0}^{n} (k-nx)^2 r_k(x) &= \sum_{k=0}^{n} (k^2 - 2knx + n^2x^2) r_k(x) \\
&= (nx + x^2 n(n-1)) - 2nx \cdot nx + n^2 x^2 \cdot 1 \\
&= nx - x^2 n \\
&= nx(1-x)
\end{aligned}$$

f is continuous on $[0,1]$ so there exists M such that $|f(x)| \leq M$ for all $x \in [0,1]$. Also, f is uniformly continuous so for all $\epsilon > 0$, there exists $\delta > 0$ such that $|x-y| < \delta$ implies that $|f(x) - f(y)| < \epsilon$. We want to prove that for every $\epsilon > 0$ there exists N such that $|p_n(x) - f(x)| < \epsilon$ for all $x \in [0,1]$ and $n \geq N$. Fix ϵ. Take N such that $\frac{2M}{\delta^2 N} < \epsilon$. Let $n \geq N$ and $x \in [0,1]$. Then

$$\begin{aligned}
|p_n(x) - f(x)| &= \left| \sum_{k=0}^{n} f\left(\frac{k}{n}\right) r_k(x) - f(x) \sum_{k=0}^{n} r_k(x) \right| \\
&\leq \left| \sum_{k=0}^{n} \left(f\left(\frac{k}{n}\right) - f(x) r_k(x) \right) \right| | \\
&\leq \sum_{k \in A} \left| f\left(\frac{k}{n}\right) - f(x) \right| | |r_k(x)| + \sum_{k \in B} \left| f\left(\frac{k}{n}\right) - f(x) \right| | |r_k(x)|
\end{aligned}$$

Where we break the sum into (i) $|k - nx| < n\delta$ denoted A and (ii) otherwise, denoted B For A, $|k/n - x| < \delta$ implies $f(k/n) - f(x)| < \epsilon$. Thus

$$\sum_{k \in A} \left| f\left(\frac{k}{n}\right) - f(x) \right| | |r_k(x)| \leq \sum_{k \in A} \epsilon r_k(x) < \epsilon$$

For B, we have $|k - nx| \geq \delta n$ so that

$$\begin{aligned}
\sum_{k \in B} \left| f(x) - f\left(\frac{k}{n}\right) \right| r_k(x) &\leq \sum_{k \in B} \left(|f(x)| + \left| f\left(\frac{k}{n}\right) \right| \right) r_k(x)| \\
&\leq 2M \sum_{k \in B} r_k(x) \\
&= 2m \sum_{k \in B} \frac{(k-nx)^2}{(k-nx)^2} r_k(x) \\
&\leq \frac{2M}{(n\delta)^2} \sum_{k=0}^{n} (k-nx)^2 r_k(x) \\
&\leq \frac{2M}{(n\delta)^2} nx(1-x) \\
&\leq \frac{2M}{n\delta^2} < \epsilon
\end{aligned}$$

by choice of N. ■

10.2 THE STONE-WEIERSTRASS THEOREM

Recall that an algebra $\mathcal{A}$ is a vector space V over $\mathbb{F}$ equipped with a bilinear form $\cdot : V \times V \to \mathbb{F}$. The space $C(X)$ is an algebra with pointwise multiplication.

Definition. An algebra $\mathcal{A}$ of functions separates points if, for any $x \neq y \in X$, then there exists $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

10.2 Theorem. (Stone-Weierstrass) *Let X be compact. Let $\mathcal{A} \subseteq C(x)$ be an algebra that separates points and contains the constants. Then $\mathcal{A}$ is dense in $C(X)$*

This will take a bit of time!

10.3 Lemma. *If $\mathcal{A}$ is an algebra that separates points, then so is $\overline{\mathcal{A}}$.*

PROOF If $f, g \in \overline{\mathcal{A}}$, then write $f = \lim f_n$, $g = \lim g_n$, so that for some constant c

$$\begin{aligned}
cf + g &= c \lim f_n + g \lim g_n \\
&= \lim(cf_n + g_n)
\end{aligned}$$

where the limit exists by the arithmetic properties of limits, and since $\mathcal{A}$ is an algebra, $cf_n + g_n \in \mathcal{A}$ and by closure of $\overline{\mathcal{A}}$, the limit is is in $\overline{\mathcal{A}}$. In the exact same way,

$$\begin{aligned}
fg &= (\lim f_n)(\lim g_n) \\
&= \lim f_n g_n
\end{aligned}$$

because, when the limits all exist, a product of limits is a limit of products. But then in the same way as above, $f_n g_n \in \mathcal{A}$ so by closedness $fg \in \overline{\mathcal{A}}$ as well. Since $\overline{\mathcal{A}} \supseteq \mathcal{A}$, it certainly separates points as well. ■

10.4 Lemma. *Suppose $B \subseteq \mathcal{A}$ is an algebra that separates points. Then if $f, g \in B$, then $\max\{f, g\}$ and $\min\{f, g\}$ are in B.*

PROOF Since $\max\{f, g\} = (f + g + |f - g|)/2$ and $\min\{f, g\} = (f + g - |f - g|)/2$, it suffices to show that $f \in B \Rightarrow |f| \in \overline{B}$¿ Note that there exists $(p_n(t)) \to \sqrt{t}$ uniformly on $[0, 1]$ with constant term 0. Let $c = \|f\|$ and note that $F_n := \frac{f^2}{c^2} \in \mathcal{A}$. We claim that $F_n \to |f|/c$. Let $\epsilon > 0$ be arbitrary and let N be such that $|p_n(t) - \sqrt{t}| < \epsilon/2$ for all $n \geq N$ and $t \in [0, 1]$. Then

$$\left|F_n(x) - \frac{|f|}{c}(x)\right| = \left|p_n \circ (f^2/c^2)(x) - \sqrt{\frac{f^2}{c^2}(x)}\right|$$
$$< \epsilon$$

since $f^2/c^2(x) \in [0, 1]$. Thus $|f|/c \in \overline{\mathcal{A}}$ and since $\overline{\mathcal{A}}$ is an algebra, $|f| \in \overline{\mathcal{A}}$. ■

10.5 Lemma. *Suppose $x \neq y$ are in X and $a, b \in \mathbb{R}$. Then there exists $f \in \mathcal{A}$ so that $f(x) = a$ and $f(y) = b$.*

PROOF Since $\mathcal{A}$ separates points, let g be such that $g(x) \neq g(y)$. Then define

$$f(t) = a + (b - a)\left(\frac{g(t) - g(x)}{g(y) - g(x)}\right)$$

which has the desired property. ■

10.6 Lemma. *If $f \in C(X)$, $x_0 \in X$ and $\epsilon > 0$, then there exists $g \in \overline{\mathcal{A}}$ such that $g(x_0) = f(x_0)$ and $g(x) \leq f(x) + \epsilon$ for all $x \in X$*

PROOF For each $y \in X$, there exists $h_y \in \mathcal{A}$ so that $h_y(x_0) = f(x_0)$ and $h_y(y) = f(y)$. If $y = x_0$, take $h_y = f$, and otherwise use the previous lemma. Now, for any $y \in X$, $|h_y - f| = 0$ so there exists a δ_y such that on $B(y, \delta_y)$, $h_y - f < \epsilon$. Thus the balls $B(y, \delta_y)$ form an open cover for X, so we have a finite subcover $\{B(y_1, \delta_{y_1}), \ldots, B(y_n, \delta_{y_n})\}$. Then fix $g = \min\{g_{y_i}\}$ so for any $x \in X$, $x \in B(y_j, \delta_{y_j})$ and $g(x) \leq g_{y_j}(x) < f + \epsilon$, as desired. ■

We can finally prove the main theorem!

PROOF First, we show for all $f \in C(X)$, $\epsilon > 0$, there exists $g \in \overline{\mathcal{A}}$ such that $\|g - f\| < \epsilon$. Once we have this, then we get $h \in \mathcal{A}$ such that $\|g - h\| < \epsilon$ and then $\|f - h\| \leq \|f - g\| + \|g - h\| < 2\epsilon$. This proves that $\mathcal{A}$ is dense.

Now for all $x \in X$, there exists $g_x \in \overline{\mathcal{A}}$ such that $g_x(x) = f(x)$ and $g(z) \leq f(z) + \epsilon$ for all $z \in X$. The function $f - g_x$ is continuous and 0 at x. Get $\delta_x > 0$ such that for all $y \in B(x, \delta_x)$, we have $|f(y) - g_x(y)| < \epsilon$. Look at $B(x, \delta_x)$ as $x \in X$. These cover X so we have a finite subcover $B(x_1, \delta_{x_1}), \ldots, B(x_n, \delta_{x_n})$. Put $g = \max(g_{x_1}, \ldots, g_{x_n}) \in \overline{\mathcal{A}}$ by Lemma 2. Let $y \in X$. Then there exists i such that $y \in B(x_i, \delta_{x_i})$, s $|f(y) - g_{x_i}(y)| < \epsilon$ and $g_{x_i}(y) > f(y) - \epsilon$. Thus $f(y) - \epsilon < g_{x_i}(y) \leq g(y) = g_{x_j}(y) \leq f(y) + \epsilon$ for some j by choice of g_x. But then $|g(y) - f(y)| \leq \epsilon$ for all $y \in X$, so $\|g - f\| \leq \epsilon$. ■

Definition. A function $f : X \to \mathbb{C}$ is continuous at $x \in X$ if whenever $x_n \to x$, then $f(x_n) \to f(x)$. We say $z_n \to z$ if $|z_n - z| \to 0$.

We define $\Re f : X \to \mathbb{R}$, $\Im f : X \to \mathbb{R}$ and $f = \Re f + i\Im f$. Then f is continuous if and only if $\Re f$ and $\Im f$ are continuous.

Given f, $\overline{f}(x) = \overline{f(x)} = \Re f(x) - i\Im f(x)$.

10.3 AN ALTERNATIVE PROOF OF STONE-WEIERSTRASS

10.7 Theorem. *Let X be a compact metric space. Suppose $\mathcal{A}$ is a subalgebra of $C(X)$ that separates points. If $\overline{\mathcal{A}} \neq C(X)$, prove there is a point $x_0 \in X$ such that $\overline{\mathcal{A}} = \{f \in C(X) : f(x_0) = 0\}$.*

10.8 Lemma. *If for every $x, y \in X$ there exists $g \in \overline{\mathcal{A}}$ such that $g(x) = f(x)$ and $g(y) = f(y)$, then $f \in \overline{\mathcal{A}}$.*

PROOF Recall the lemma from class that

$$f, g \in \mathcal{A} \Rightarrow \min\{f, g\}, \max\{f, g\} \in \overline{\mathcal{A}} \tag{$*$}$$

Since $\overline{\mathcal{A}}$ is an algebra, this holds for $\overline{\mathcal{A}}$ as well. Let $\epsilon > 0$ be arbitrary. Fix $x, y \in X$ and let g_{xy} be such that $g_{xy}(x) = f(x)$ and $g_{xy}(y) = f(y)$. Then define

$$U_{xy} = \{u \in X : f(u) < g_{xy}(u) + \epsilon\}$$

$$V_{xy} = \{u \in X : f(u) > g_{xy}(u) - \epsilon\}$$

Note that $x \in U_{xy}$, and that U, V are open by continuity of f and g. To see the latter statement, let $u \in U_{xy}$. Let $\gamma = (g_{xy}(u) + \epsilon) - f(u) > 0$ since the inequality is strict. By continuity of f, g_{xy} let $\delta > 0$ be such that for all $v \in B(u, \delta)$, $|f(u) - f(v)| < \gamma/2$ and $|g_{xy}(u) - g_{xy}(v)| < \gamma/2$. But then we must have

$$\begin{aligned} f(v) - g_{xy}(v) &< f(u) + \gamma/2 - g_{xy}(u) + \gamma/2 \\ &= f(u) - g_{xy}(u) + \gamma \\ &= f(u) - g_{xy}(u) + g_{xy}(u) - f(u) + \epsilon \\ &= \epsilon \end{aligned}$$

so that $v \in U_{xy}$. The identical argument works for V_{xy}.

Furthermore, $x \in U_{xy}$ so $\{U_{xy} : x \in X\}$ is an open cover for X. By compactness of X it has a subcover $\{U_{x_1 y}, \ldots, U_{x_n y}\}$, and write $g_y = \max\{g_{x_1 y}, \ldots, g_{x_n y}\}$. Now for a given y, define $V_y = \bigcap_{i=1}^{n} V_{x_i y}$. V_y are open since it is a finite intersection of open sets, so since $y \in V_y$, $\{V_y : y \in X\}$ is an open cover for X. Thus we also have a finite subcover $\{V_{y_1}, \ldots, V_{y_k}\}$. Finally, define $g = \min\{g_{y_1}, \ldots, g_{y_k}\}$. By $(*)$, $g \in \overline{\mathcal{A}}$.

I claim that $d(g, f) < \epsilon$. We have $f < g_y + \epsilon$ on X for all y since on each $U_{x_i y}$, $f < g_{x_i y} + \epsilon \leq g_y + \epsilon$. Since this holds for all y, it certainly holds for g as well, so $f < g + \epsilon$. Similarly, on each $V_{x_i y}$, we have $f > g_{x_i y} - \epsilon$, so $f > g_{x_i y} - \epsilon$ for $i = 1, \ldots, n$ on V_y. Thus $f > g_{y_i} - \epsilon$ on each V_{y_i}, and since $g \leq g_{y_i}$ for all y we have $f > g_{y_i} - \epsilon \geq g - \epsilon$. Combining these expressions, we have $f - g < \epsilon$ and $g - f < \epsilon$ so $d(g, f) < \epsilon$.

Furthermore, since ϵ was arbitrary, for any f, we can find a sequence $g_n \to f$ by choosing $d(g_n, f) < 1/n$. Then since $\overline{\mathcal{A}}$ is closed, we must have $f \in \overline{\mathcal{A}}$ as desired. ■

We are now in position to prove the main theorem!

PROOF Let $x, y \in X$ with $x \neq y$, and consider the algebra $A_{xy} = \{(f(x), f(y)) : f \in \overline{\mathcal{A}}\} \subseteq \mathbb{R}^2$ with component-wise addition, scalar multiplication, and multiplication. Thse are algebras since $\mathcal{A}$ is an algebra, and we must have

$$A_{xy} \in \{\{(0, x) : x \in \mathbb{R}\}, \{(x, 0) : x \in \mathbb{R}\}, \mathbb{R}^2\}$$

To see this, suppose $(a, b) \in A_{xy}$. If $a \neq b$ and $a, b \neq 0$, then (a, b) and (a^2, b^2) provide a basis for $\mathbb{R}^2$, so $S = \mathbb{R}^2$ by closure under multiplication of $\mathcal{A}$. If $a = b$, then $\mathcal{A}$ does not separate points, and if $a = 0, b \neq 0$ or the reverse, we have the first two cases above.

Now suppose that for all x, y, $A_{xy} = \mathbb{R}^2$. Then for any $f \in C(X)$, by Lemma 1, we must have $f \in \overline{\mathcal{A}}$, so $\overline{\mathcal{A}} = C(X)$, the contents of the Stone-Weierstrass Theorem.

Otherwise there exists some x_0, y_0 so that $A_{x_0y_0} = \{(x, 0) : x \in \mathbb{R}\}$ without loss of generality. Note that such a y_0 must satisfy $f(y_0) = 0$ for all $f \in \overline{\mathcal{A}}$. Furthermore, y_0 is unique. Suppose not, so we have y_0, y_1 such that $f(y_0) = 0$ and $f(y_1) = 0$ for all $f \in \overline{\mathcal{A}}$ and $\overline{\mathcal{A}}$ would not separate points. Finally, for any x, we have $A_{xy_0} = \{(x, 0) : x \in \mathbb{R}\}$ since $f(y_0) = 0$ for all $f \in \overline{\mathcal{A}}$. We can summarize this by saying

$$A_{xy} = \begin{cases} \mathbb{R}^2 & : x, y \neq y_0 \\ \{(u, 0) : u \in \mathbb{R}\} & : y = y_0 \\ \{(0, u) : u \in \mathbb{R}\} & : x = y_0 \end{cases}$$

We certainly have $\overline{\mathcal{A}} \subseteq \{f \in C(X) : f(y_0) = 0\}$ by the existence of y_0 above. Now to show the reverse inclusion, suppose $f(y_0) = 0$. Suppose without loss of generality that $x \neq y_0$. Since $A_{xy_0} = \{(x, 0) : x \in \mathbb{R}\}$, we have $g \in \overline{\mathcal{A}}$ so that $g(y_0) = 0 = f(y_0)$ and $g(x_0) = f(x_0)$. Similarly, for A_{xy} with $y \neq y_0$, $A_{xy} = \mathbb{R}^2$ so there exists $g \in \overline{\mathcal{A}}$ so that $g(x) = f(x)$ and $g(y) = f(y)$. Thus by Lemma 1, $f \in \overline{\mathcal{A}}$. And we're done! ■

10.4 APPLICATIONS OF STONE-WEIERSTRASS

10.9 Theorem. (Stone-Weierstrass for $\mathbb{C}$ −valued functions) *Let X be compact and suppose $\mathcal{A}$ is an algebra in $C(X, \mathcal{C}) = \{f : X \to \mathbb{C}, \text{continuous}\}$ (the scalars are from $\mathbb{C}$). that separates points, contains the constants, and is closed under conjugation. Then $\mathcal{A}$ is dense in $C(X, \mathbb{C})$.*

PROOF Let $\mathcal{A}_{\mathbb{R}}$ denote the set of real-valued functions in $\mathcal{A}$. If $f \in \mathcal{A}$, then $(f \pm \overline{f})/2 \in \mathcal{A}$ so $\Re f$ and $\Im f$ are in $\mathcal{A}$. Notice $\mathcal{A}_{\mathbb{R}}$ is an algebra over $\mathbb{R}$ in $C(x)$ and it contains the constants. Furthermore, since $\mathcal{A}$ separates points, either the real part or imaginary part separates points, so $\mathcal{A}_{\mathbb{R}}$ separates points as well. Thus the result holds by Stone-Weierstrass over $\mathbb{R}$. ■

Example. Consider the set $X = \{z \in \mathbb{C} : |z| = 1\}$, and

$$\mathcal{A} = \left\{ \sum_{n=-N}^{N} a_n z^n : a_n \in \mathbb{C} \right\}$$

The collection $\mathcal{A}$ is an algebra, separates points, contains constants, and is closed under conjugation.

10.10 Corollary. *Suppose X is a compact metric space. Then $C(X)$ is separable.*

PROOF Since X is compact, it is separable, so let $S \subseteq X$ be countable and dense. Then consider the collection of functions $\mathcal{S} = \{f_u(x) = d(x, u) : u \in S\}$. I claim that $\mathcal{S}$ separates points. Thus suppose $x, y \in X$ and fix $d(x, y) = \epsilon$ since $x \neq y$. Then since S is dense, choose x_0 so that $d(x_0, x) < \epsilon/2$. But then $f_{x_0}(x) < \epsilon/2$ and $f_{x_0}(y) > \epsilon/2$ for if not, then

$d(x,y) \leq d(x,x_0) + d(x_0,y) = f_{x_0}(x) + f_{x_0}(y) < \epsilon$, a contradiction. Thus $f_{x_0}(x) \neq f_{x_0}(y)$ so $\mathcal{S}$ separates points.

Note that the polynomials of functions in $\mathcal{S}$ (in other words, the smallest algebra containing S) is dense in $C(X)$ by Stone-Weierstrass. Thus let $\mathcal{H}$ denote the polynomials over $\mathcal{S}$ with rational coefficients. It suffices to show that any polynomial in $\mathcal{S}$ can be approximated by functions in $\mathcal{H}$. We can write an arbitrary polynomial as $p = \sum c_i \prod f_j^{c_j}$ for $f_i \in \mathcal{S}$ where the sum and all the products are finite. Note that each f_j is uniformly continuous on X, so $\prod f_j^{c_j}$ is also uniformly continuous. Thus let $\left|\prod_{j \in A_i} f_j^{c_j}\right| < M$ for all terms in p. Furthermore, suppose the summation consists of N terms. Let $\epsilon > 0$ be arbitrary and let $b_i \in \mathbb{Q}$ be such that $|b_i - c_i| < \epsilon/(NM)$. But then define $q = \sum b_i \prod f_j^{c_j}$ so that

$$\begin{aligned}
|p(x) - q(x)| &\leq \sum |b_i - a_i| \left|\prod f_j^{c_j}\right| \\
&\leq M \sum |b_i - a_i| \\
&\leq M \sum \frac{\epsilon}{Mn} \\
&< \epsilon
\end{aligned}$$

as desired. ■

10.11 Corollary. *If $f \in C[0,1]$ and $\int_0^1 f(x)x^n dx = 0$ for all $n = 0, 1, 2, \ldots$, then $f(x) = 0$ for all $x \in [0,1]$.*

PROOF $\int_0^1 f(x)x^n = 0$ for all n implies $\int_0^1 f(x)p(x)dx = 0$ for any polynomial p. By Stone-Weierstrass, let $(p_n) \to f$ uniformly. Then

$$\int_0^1 f(x)p_n(x)dx \to \int_0^1 (f(x))^2 dx$$

Assuming this, it follows that $\int_0^1 f^2 = 0$ so $f = 0$. For the claim, consider

$$|\int_0^1 f(x)(p_n(x) - f(x))dx| \leq \int_0^1 |f||p_n - f|dx \leq \|f\| \int_0^1 |p_n - f| \leq \|f\| \, \|p_n - f\| \int_0^1 dx$$

as desired. ■