PMATH 351, Real Analysis

Brian E. Forrest

August 23, 2017

Contents

Chapter 1

Axiom of Choice, Zorn's Lemma and Cardinality

1.0 Notation

We will introduce some basic material that will be used throughout the rest of the course.

We will use the following notation:

- $\mathbb{N}$ will denote the set of natural numbers $\{1, 2, 3, \ldots\}$.
- $\mathbb{Z}$ will denote the set of integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$.
- $\mathbb{Q}$ will denote the set of rational numbers $\left\{\frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N}\right\}$.
- $\mathbb{R}$ will denote the set of real numbers.

We will use the notation $A \subset B$ and $A \subseteq B$ interchangeably to mean that A is a subset of B with the possibility that $A = B$ though when we explicitly wish to emphasize that $A = B$ is a possibility, we will generally use $A \subseteq B$. When we wish to express that A is a *proper subset* of B, then we can either specify further that $A \neq B$, or we can use the notation $A \subsetneq B$.

DEFINITION **1.0.1.** *Given a set X, we let*

$$\mathcal{P}(X) = \{A \mid A \subset X\}.$$

We call $\mathcal{P}(X)$ the Power Set of X. In this case, we call X the universal set.

The *union* of A and B is the set

$$A \cup B = \{x \mid x \in A \textbf{ or } x \in B\}.$$

More generally, if for each $\alpha \in I$, $A_\alpha \subseteq X$, then

$$\bigcup_{\alpha \in I} A_\alpha = \{x \mid x \in A_\alpha \textbf{ for some } \alpha \in I\}.$$

The *intersection* of A and B is the set

$$A \cap B = \{x \mid x \in A \textbf{ and } x \in B\}.$$

More generally, if for each $\alpha \in I$, $A_\alpha \subseteq X$, then

$$\bigcap_{\alpha \in I} A_\alpha = \{x \mid x \in A_\alpha \textbf{ for all } \alpha \in I\}.$$

Problem **1.** *What would we mean by* $\bigcup_{\alpha\in I} A_\alpha$ *if* $I = \emptyset$*?*

Let $A, B \subset X$. We will let

$$B \setminus A = \{x \in B \mid x \notin A\}.$$

In the special case, when $B = X$, we also call the set $X \setminus A$ the *complement* of A in X and denote this set by A^c. Observe that for any sets $A, B \subseteq X$, we have that $(A^c)^c = A$ and also $A^c = B^c$ if and only if $A = B$.
.

Theorem **1.0.2.** ***(DeMorgan's Laws)***

1. $(\bigcup_{\alpha\in I} A_\alpha)^c = \bigcap_{\alpha\in I} (A_\alpha^c)$.
2. $(\bigcap_{\alpha\in I} A_\alpha)^c = \bigcup_{\alpha\in I} (A_\alpha^c)$.

Proof.

1. This follows since $x \in (\bigcup_{\alpha\in I} A_\alpha)^c$ if and only if $x \notin \bigcup_{\alpha\in I} A_\alpha$. In turn this happens if and only if $x \in A_\alpha^c$ for each $\alpha \in I$, and hence if and only if $x \in \bigcap_{\alpha\in I} (A_\alpha^c)$.
2. This can be seen to be exactly the same statement as in 1. if we simply replace A_α by A_α^c and apply the complementation operation to both sides of the equality.

∎

1.1 Products and the Axiom of Choice

Definition **1.1.1.** *Given two sets* X, Y*, define the product of* X *and* Y *by*

$$X \times Y = \{(x, y) \mid x \in X \text{ and } y \in Y\}.$$

In this case, x *is called the x-coordinate of* (x, y) *and* y *is called the y-coordinate of* (x, y)*.*

Given n *sets* $\{X_1, X_2, X_3, \cdots, X_n\}$*, define the product of* $\{X_1, X_2, X_3, \cdots, X_n\}$ *by*

$$X_1 \times X_2 \times \cdots \times X_n = \prod_{i=1}^{n} X_i = \{(x_1, x_2, x_3, \cdots, x_n) \mid x_i \in X_i\}.$$

$(x_1, x_2, \cdots, x_n) \in \prod_{i=1}^{n} X_i$ *is called an n-tuple and* x_i *is called the ith coordinate.*

If $X_i = X$ *for all* i*, we write* X^n *for* $\prod_{i=1}^{n} X_i$*.*

Notation: Given a finite set X, let $\mid X \mid$ be the number of elements in X. $\mid X \mid$ is called the cardinality of X.

The following theorem is clear.

Theorem **1.1.2.** *Let* $\{X_1, X_2, \cdots, X_n\}$ *be a finite collection of finite sets. Then*

$$\mid \prod_{i=1}^{n} X_i \mid = \prod_{i=1}^{n} \mid X_i \mid .$$

In particular, if $X_i = X$ *for all* i*, then*

$$\mid \prod_{i=1}^{n} X_i \mid = \mid X \mid^n .$$

PROBLEM **2.** *How do we define the product of an arbitrary collection of sets?*

Note: Each $(x_1, x_2, \cdots, x_n) \in \prod_{i=1}^{n} X_i$ determines a function

$$f_{(x_1,x_2,\cdots,x_n)}(i) : \{1, 2, \cdots, n\} \to \bigcup_{i=1}^{n} X_i$$

by

$$f_{(x_1,x_2,\cdots,x_n)}(i) = x_i.$$

Given $f : \{1, 2, \cdots, n\} \to \bigcup_{i=1}^{n} X_i$ with $f(i) \in X_i$, define $(x_1, x_2, \cdots, x_n) \in \prod_{i=1}^{n} X_i$, by

$$x_i = f(i).$$

Fact:

$$\prod_{i=1}^{n} X_i \cong \{f : \{1, 2, \cdots, n\} \to \bigcup_{i=1}^{n} X_i \mid f(i) \in X_i\}.$$

f is called a *choice function*.

DEFINITION **1.1.3.** *Given a collection $\{X_\alpha\}_{\alpha \in I}$ of sets define*

$$\prod_{\alpha \in I} X_\alpha = \{f : I \to \bigcup_{\alpha \in I} X_\alpha \mid f(\alpha) \in X_\alpha\}.$$

If $X_\alpha = X$ for all $\alpha \in I$, $\prod_{\alpha \in I} X_\alpha$ is written as X^I.

Fundamental Problem:

Given a non-empty collection $\{X_\alpha\}_{\alpha \in I}$ of non-empty sets is

$$\prod_{\alpha \in I} X_\alpha \neq \emptyset?$$

AXIOM **1.1.4** [ZERMELO'S AXIOM OF CHOICE]**.** *Given a non-empty collection $\{X_\alpha\}_{\alpha \in I}$ of non-empty sets*

$$\prod_{\alpha \in I} X_\alpha \neq \emptyset.$$

Note: The following statement is equivalent to the Axiom of Choice:

AXIOM **1.1.5** [AXIOM OF CHOICE]**.** *Given a non-empty set X there exists a function*

$$f : \mathcal{P}(X) \setminus \{\emptyset\} \to X$$

such that for every $A \subset X$ with $A \neq \emptyset$, we have $f(A) \in A$.

1.2 Relations and Zorn's Lemma

DEFINITION **1.2.1.** *A relation is a subset $\mathcal{R}$ of $X \times Y$. We generally will write $x\mathcal{R}y$ if $(x, y) \in \mathcal{R}$. R is often called the graph of the relation.*

In the case, that $X = Y$ we say that $\mathcal{R}$ determines a relation on X.

EXAMPLE **1.1.** *A function $f : X \to Y$ can be viewed as a relation $\mathcal{R}$ with the property that for each $x \in X$ there exists exactly one $y \in Y$ such that $x\mathcal{R}y$. In this case, we have:*

$$f(x) = y \text{ if and only if } \ x\mathcal{R}y.$$

DEFINITION **1.2.2.** *A relation $\mathcal{R}$ on X is*

1. *reflexive if $x\mathcal{R}x$ for every $x \in X$.*
2. *symmetric if $x\mathcal{R}y \Rightarrow y\mathcal{R}x$.*
3. *anti-symmetric if $x\mathcal{R}y$ and $y\mathcal{R}x$ implies $x = y$.*
4. *transitive if $x\mathcal{R}y$ and $y\mathcal{R}z$ implies $x\mathcal{R}z$.*

EXAMPLE **1.2.** *Let $X = \mathbb{R}$. Then we can define a relation $\mathcal{R}$ on $\mathbb{R}$ by*

$$x\mathcal{R}y \text{ if and only if } \ x \leq y.$$

In this case, the relation is easily seen to be refelxive, transitive, and anti-symmetric.

EXAMPLE **1.3.** *Let X be any set. Then we can define a relation $\mathcal{R}$ on $\mathcal{P}(X)$ by*

$$A\mathcal{R}B \text{ if and only if } \ A \subseteq B.$$

This relation is easily seen to be reflexive, transitive, and anti-symmetric. In this case, we say that $\subseteq$ orders $\mathcal{P}(X)$ by inclusion.

We can define a second relation $\mathcal{R}^*$ on $\mathcal{P}(X)$ by

$$A\mathcal{R}^*B \text{ if and only if } \ B \subseteq A \text{ or } A \supseteq B.$$

Again, this relation is easily seen to be reflexive, transitive, and anti-symmetric. In this case, we say that $\supseteq$ orders $\mathcal{P}(X)$ by containment.

REMARK **1.2.3.** *Note that when we use the relation $\leq$ to order the elements of $\mathbb{R}$, then for any two elements $x, y \in \mathbb{R}$ we have either $x \leq y$ or $y \leq x$. As such we say that $\leq$ totally orders $\mathbb{R}$.*

In general, for a set X, if we use $\subseteq$ to order $\mathcal{P}(X)$ it is not possible to compare every pair of sets A and B. As such we say that $\subseteq$ partially orders $\mathcal{P}(X)$.

DEFINITION **1.2.4.** *A relation $\mathcal{R}$ on a set X is called a partial order if it is*

1. *reflexive*
2. *anti-symmetric*
3. *transitive*

We call the pair $(X, \mathcal{R})$ a partially ordered set or a poset for short.

If for any $x, y \in X$, either $x\mathcal{R}y$ or $y\mathcal{R}x$, we call $\mathcal{R}$ a total order on X. In this case, we call $(X, \mathcal{R})$ a totally ordered set or a chain for short.

Because $(\mathbb{R}, \leq)$ is the fundamental example of a partially ordered set we will often use the symbol $\leq$ or the stylized $\preceq$ to denote the relation on on a parially ordered set.

Definition **1.2.5.** *Let $(X, \leq)$ be a partially ordered set. Let $A \subseteq X$.*

We say that $x \in X$ is an upper bound for A if $y \leq x$ for every $y \in A$. We say that A is bounded above if it has an upper bound.

We say that $x \in X$ is the least upper bound (or the supremum) for A if

1. *x is an upper bound of A.*

2. *if y is an upper bound of A, then $x \leq y$.*

If A has a least upper bound, we denote it by $\operatorname{lub}(A)$ *or* $\sup(A)$. *If $x = \operatorname{lub}(A)$ and $x \in A$, then we call x the maximum of A and denote this by* $\max(A)$.

We say that $x \in X$ is a lower bound for A if $x \leq y$ for every $y \in A$. We say that A is bounded below if it has a lower bound.

We say that $x \in X$ is the greatest lower bound (or the infimum) for A if

1. *x is a lower bound of A.*

2. *if y is a lower bound of A, then $y \leq x$.*

If A has a greatest lower bound, we denote it by $\operatorname{glb}(A)$ *or* $\inf(A)$. *If $x = \operatorname{glb}(A)$ and $x \in A$, then we call x the minimum of A and denote this by* $\min(A)$.

Example **1.4.**

1. ***[Least Upper Bound Axiom for** $\mathbb{R}$**]** Consider $\mathbb{R}$ with the usual order. Let A be a non-empty subset if $\mathbb{R}$. If A is bounded above, then A has a least upper bound.*

2. *Consider $(\mathcal{P}(X), \subseteq)$. Then if $\{A_\alpha\}_{\alpha\in I}$ be any non=empty collection of subsets of X. Then X is an upper bound for $\{A_\alpha\}_{\alpha\in I}$ and $\emptyset$ is a lower bound for $\{A_\alpha\}_{\alpha\in I}$. Moreover,*

$$\bigcup_{\alpha\in I} A_\alpha = \operatorname{lub}(\{A_\alpha\}_{\alpha\in I})$$

   and

$$\bigcap_{\alpha\in I} A_\alpha = \operatorname{glb}(\{A_\alpha\}_{\alpha\in I}).$$

3. *Consider $(\mathcal{P}(X), \supseteq)$. Then if $\{A_\alpha\}_{\alpha\in I}$ be any non-empty collection of subsets of X. Then $\emptyset$ is an upper bound for $\{A_\alpha\}_{\alpha\in I}$ and X is a lower bound for $\{A_\alpha\}_{\alpha\in I}$. Moreover,*

$$\bigcap_{\alpha\in I} A_\alpha = \operatorname{lub}(\{A_\alpha\}_{\alpha\in I})$$

   and

$$\bigcup_{\alpha\in I} A_\alpha = \operatorname{glb}(\{A_\alpha\}_{\alpha\in I}).$$

Definition **1.2.6.** *Let $(X, \leq)$ be a partially ordered set. A element $x \in X$ is said to be maximal if $x \leq y$ implies that $x = y$.*

EXAMPLE **1.5.** *1. Consider $\mathbb{R}$ with the usual order. Then R has no maximal element.*

2. Consider $(\mathcal{P}(X), \subseteq)$. Then X is maximal.

3. Consider $(\mathcal{P}(X), \supseteq)$. Then $\emptyset$ is maximal.

PROPOSITION **1.2.7.** *Every finite, non-empty poset $(X, \leq)$ has a maximal element.*

The proof of the previous Proposition can be obtained by induction on the number of elements in X and is is left as an exercise.

We have already seen that there are posets without maximal elements. The next result is a fundamental tool in much of mathematics.

AXIOM **1.2.8** [ZORN'S LEMMA]**.** *Let $(X, \leq)$ be a non-empty partially ordered set. If every totally ordered subset $\mathcal{C}$ of X has an upper bound, then $(X, \leq)$ has a maximal element.*

REMARK **1.2.9.** ***Fundamental Fact:*** *Zorn's Lemma is logically equivalent to the Axiom of Choice.*

DEFINITION **1.2.10.** *Let V be a non-zero vector space. Let $\mathcal{L} = \{A \subset V \mid A \text{ is linearly independent}\}$. We can order $\mathcal{L}$ by inclusion $\subseteq$. A basis B for V is a maximal element in $(\mathcal{L}, \subseteq)$.*

THEOREM **1.2.11.** *Every non-zero vector space has a basis.*

Proof. Order $\mathcal{L}$ by inclusion. Let $\mathcal{C} = \{A_\alpha \mid \alpha \in I\}$ be a chain in $\mathcal{L}$. Let

$$A = \bigcup_{\alpha \in I} A_\alpha.$$

We claim that A is linearly independent. To see why assume that $\{x_1, x_2, \ldots, x_n\} \subset A$ and that $\{\beta_1, \beta_2, \ldots, \beta_n\} \subset \mathbb{R}$ are such that

$$\beta_1 x_1 + \beta_2 x_2 + \cdots \beta_n x_n = 0.$$

For each $i = 1, 2, \ldots, n$, we can find an $A_{\alpha_i} \in \mathcal{C}$ such that $x_i \in A_{\alpha_i}$. We can also assume that

$$A_{\alpha_1} \subseteq A_{\alpha_2} \subseteq \cdots \subseteq A_{\alpha_n}$$

since $\mathcal{C}$ is a chain. Consequently, $\{x_1, x_2, \ldots, x_n\} \subset A_{\alpha_n}$. But since A_{α_n} is linearly independent, we have $\beta_i = 0$ for each $i = 1, 2, \ldots, n$.

We have seen that every chain has an upper bound. It now follows from Zorn's Lemma that $\mathcal{L}$ has a maximal element.

∎

REMARK **1.2.12.** *Observe that the above proof can be modified to show that given any linearly independent set $A \subset V$, there exists a basis B for V with $A \subseteq B$.*

DEFINITION **1.2.13.** *A poset $(X, \leq)$ is said to be well ordered if every non-empty subset A has a least element.*

EXAMPLE **1.6.** *1. Consider $(\mathbb{N}, \leq)$, the natural numbers with the usual order. It follows from the Principle of Mathematical Induction that $(\mathbb{N}, \leq)$ is well ordered.*

2. *Let $X = \mathbb{Q} = \left\{\frac{n}{m} : n \in \mathbb{Z}, m \in \mathbb{N}, \gcd(n,m) = 1\right\}$. Then $\mathbb{Q}$ is not well ordered with respect to the usual order. However, we can still construct a well order on $\mathbb{Q}$.*

   Define $\phi : \mathbb{Q} \to \mathbb{N}$ by
$$\phi(\frac{n}{m}) := \begin{cases} 2^n 3^m & \text{if } n > 0, \\ 1 & \text{if } n = 0 \\ 5^{-n} 7^m & \text{if } n < 0. \end{cases}$$
   Then ϕ is $1-1$. We can use this to define an order $\preceq$ on $\mathbb{Q}$ by
$$\frac{n}{m} \preceq \frac{p}{q} \text{ if and only if } \phi(\frac{n}{m}) \leq \phi(\frac{p}{q}).$$

Axiom **1.2.14** [Well Ordering Principle]. *Given any set X there exists a partial order $\leq$ such that $(X, \leq)$ is well ordered.*

Theorem **1.2.15.** *The following are equivalent:*

1. *The Axiom of Choice*
2. *Zorn's Lemma*
3. *Well Ordering Principle*

Proof. 3) $\Rightarrow$ 1) : Assume that X is non-empty. By the Well ordering Principle there exists and order $\leq$ on X that makes $(X, \leq)$ a well ordered set. Define $f : \mathcal{P}(X) \setminus \{\emptyset\} \to X$ by
$$f(A) = \text{ the least element of } A.$$
∎

1.3 Equivalence Relations and Cardinality

Definition **1.3.1.** *A relation $\sim$ on a set X is called an equivalnce relation if it is:*

1. *reflexive*
2. *symmetric*
3. *transitive*

Given $x \in X$ we let
$$[x] = \{y \in X : x \sim y\}.$$
$[x]$ is called the equivalence class of x.

The following imprortant observation is easily proved.

Proposition **1.3.2.** *Let $\sim$ be an equivalnce relation on X.*

1. *For each $x \in X$ we have $x \in [x]$ so $[x] \neq \emptyset$.*
2. *For each $x, y \in X$ either $[x] = [y]$ or $[x] \cap [y] = \emptyset$.*
3. *$X = \bigcup_{x \in X} [x]$.*

Definition **1.3.3.** *Given a set X, a partition of X is a collection*

$$\mathcal{P} = \{A_\alpha \subseteq X : \alpha \in I\}$$

such that:

1. *for each $\alpha \in I$ we have $A_\alpha \neq \emptyset$.*
2. *if $\alpha, \beta \in I$ with $\alpha \neq \beta$ we have $A_\alpha \cap A_\beta = \emptyset$.*
3. *$X = \bigcup\limits_{\alpha \in I} A_\alpha$.*

Remark **1.3.4.** *Observe that the previous proposition implies that every equivalence relation $\sim$ on a set X induces a partition on X consisting of the distinct equivalence classes of $\sim$. Conversely, given a partition $\mathcal{P} = \{A_\alpha \subseteq X : \alpha \in I\}$ there is an equivalence realtion $\sim$ on X such that $\mathcal{P}$ is precisely the set of equivalence classes of $\sim$. In particular, we say that $x \sim y$ if and only if there exists an $\alpha \in I$ such that $x, y \in A_\alpha$.*

Example **1.7.** *Let X be any set. Define a relation $\sim$ on $\mathcal{P}(X)$ by $A \sim B$ if and only if there exists a $1-1$ and onto function $f : A \to B$. Then we claim that $\sim$ is an equivalence relation on X.*

1. *Reflexivity: For each $A \in \mathcal{P}(X)$ define $id_A : A \to A$ by*

   $$id_A(x) = x$$

   for all $x \in A$. (In case, $A = \emptyset, id_A$ is the empty function.)
2. *Symmetry: Assume that $A \sim B$ and that $f : A \to B$ is $1-1$ and onto. Then f is invertible with inverse g. Since $g : B \to A$ is $1-1$ and onto, we have $B \sim A$.*
3. *Transitivity: Assume that $A \sim B$ and $B \sim C$ with $f : A \to B$ and $g : B \to C$ being $1-1$ and onto. Then $g \circ f : A \to C$ is also $1-1$ and onto. Hence $A \sim C$.*

In essence, we can view two subsets A and B of X to be *equivalent* as above if they have the *same number of elements*, in the sense that there is a $1-1$ correspondence bewteen the elements of the two sets.

Definition **1.3.5.** *We say that two sets X and Y are equivalent if there exists a $1-1$ and onto function $f : X \to Y$. In this case, we write $X \sim Y$. In this case, we also say that the two sets have the same cardinality and write $\mid X \mid = \mid Y \mid$.*

We say that a set X is finite if $X = \emptyset$ or if $X \sim \{1, 2, 3, \dots, n\}$ for some $n \in \mathbb{N}$. In this case, we say that X has cardinality n and write $\mid X \mid = n$. Otherwise, we say that X is infinite.

Problem **3.** *Can a set X be equivalent to both $\{1, 2, 3, \dots, n\}$ and $\{1, 2, 3, \dots, m\}$ if $n \neq m$?*

If such a set exits then we can show that $\{1, 2, 3, \dots, n\} \sim \{1, 2, 3, \dots, m\}$. We may also assume WLOG that $n < m$ so that $\{1, 2, 3, \dots, n\} \subsetneq \{1, 2, 3, \dots, m\}$. As such we can consider the following related question: Can $\{1, 2, 3, \dots, m\}$ be equivalent to a proper subset of itself?

Proposition **1.3.6.** *The set $\{1, 2, 3, \dots, m\}$ is not equivalent to any proper subset of itself.*

Proof. We will prove this by induction on m. This is clear if $m = 1$ as the only proper subset is the empty set.

Assume that the statement holds for the set $\{1, 2, 3, \dots, k\}$. Assume also that there is a proper subset S of $\{1, 2, 3, \dots, k, k+1\}$ such that there exists a $1-1$ and onto function $f : \{1, 2, 3, \dots, k, k+1\} \to S$. If $k+1 \notin S$, then the restriction of f to $\{1,2,3,\ldots,k\}$ defines a one to one function onto $S\setminus\{f(k+1)\} \subsetneq \{1,2,3,\ldots,k\}$, which is impossible by assumption.

Suppose that $k+1 \in S$ and let $S' = S \setminus \{k+1\}$. Then $S' \subsetneq \{1,2,3,\ldots,k\}$. If $f(k+1) = k+1$ then the restriction of f to $\{1,2,3,\ldots,k\}$ defines a one to one function onto S' which is again impossible. So we may assume that $f(j) = k+1$ for some $j \in \{1,2,3,\ldots,k\}$. From here we define a new function $f^* : \{1,2,3,\ldots,k,k+1\} \to S$ by

$$f^*(i) := \begin{cases} k+1 & \text{if } i = k+1, \\ f(k+1) & \text{if } i = j \\ f(i) & \text{if } i \neq k+1, j. \end{cases}$$

But then f^* is a $1-1$ function from $\{1,2,3,\ldots,k,k+1\}$ onto S that maps $k+1$ to $k+1$, something we already know to be impossible.

∎

COROLLARY **1.3.7.** *If a set X is finite, then X is not equivalent to any proper subset of itself.*

Proof. We can clearly assume that X is not empty. As such $X \sim \{1,2,3,\ldots,n\}$ for some $n \in \mathbb{N}$. Therefore, there exists a $1-1$ and onto function $f : X \to \{1,2,3,\ldots,n\}$.

Assume that there is a proper subset $S \subsetneq X$ with $X \sim S$. Then it folows that $\{1,2,3,\ldots,n\} \sim S$ as well. Hence there exits a $1-1$ and onto function $g : \{1,2,3,\ldots,n\} \to S$. From this it follow that $h = f \circ g$ defines a $1-1$ function from $\{1,2,3,\ldots,n\}$ into $\{1,2,3,\ldots,n\}$ with proper range $f(S)$. This shows that $\{1,2,3,\ldots,n\}$ is equivalent to $f(S) \subsetneq \{1,2,3,\ldots,n\}$ which we know is impossible.

∎

REMARK **1.3.8.** *The proposition above is a formal realisation of the Pigeonhole Principle: If $m > n$ and m objects are placed in n containers, then at least one of the containers must contain more than one object.*

EXAMPLE **1.8.** *The function $f(n) = n+1$ defines a $1-1$ function from $\mathbb{N}$ onto the proper subset $\{2,3,4,\ldots\}$.*

DEFINITION **1.3.9.** *A set X is countable if it is either finite or if $X \sim \mathbb{N}$. If X is not countable, we say it is uncountable.*

We say that X is countably infinite if $X \sim \mathbb{N}$. In this case we write

$$\mid X \mid = \mid \mathbb{N} \mid = \aleph_0.$$

($\aleph_0$ = aleph naught).

PROPOSITION **1.3.10.** *Every infinite set countains a countably infinite subset.*

Proof. The Axiom of Choice allows us to define a function

$$f : \mathcal{P}(X) \setminus \{\emptyset\} \to X$$

such that for every $A \subset X$ with $A \neq \emptyset$, we have $f(A) \in A$.

Let $x_1 = f(X)$. Now define a function recursively by

$$x_{n+1} = f(X \setminus \{x_1, x_2, x_3, \ldots, x_n\}).$$

Then $A = \{x_1, x_2, \ldots, x_n, \ldots\}$ is a countably infinite subset of X.

∎

COROLLARY **1.3.11.** *A set X is infinite if and only if it is equivalent to a proper subset of itself.*

Proof. Assume that X is not infinite. Then we know it cannot be equivalent to a proper subset of itself.

Assume that X is infinite. Then X contains a countably infinite subset $\{x_1, x_2, \ldots, x_n, \ldots\}$. We can define $f : X \to X$ by

$$f(x) := \begin{cases} x_{n+1} & \text{if } x = x_n \text{ for some } n, \\ x & \text{if } x \neq x_n \text{ for any } n. \end{cases}$$

Then $f(x)$ is $1-1$ and has proper range. As such $X \sim f(X)$.

∎

Assume that there exists a $1-1$ function $f : X \to Y$. Then $X \sim f(X) \subseteq Y$. This should suggest that Y has at least as many elements as X. With this in mind we introduce the following notation:

INFORMAL DEFINITION **1.** *If there exists a $1-1$ function $f : X \to Y$ we write $\mid X \mid \preceq \mid Y \mid$.*

EXAMPLE **1.9.** *If we let $\mathbb{Q} = \left\{\frac{n}{m} : n \in \mathbb{Z}, m \in \mathbb{N}, \gcd(n,m) = 1\right\}$. If we define $\phi : \mathbb{Q} \to \mathbb{N}$ by*

$$\phi(\frac{n}{m}) := \begin{cases} 2^n 3^m & \text{if } n > 0, \\ 1 & \text{if } n = 0, \\ 5^{-n} 7^m & \text{if } n < 0, \end{cases}$$

then ϕ is $1-1$. Hence we have that $\mid \mathbb{Q} \mid \preceq \mid \mathbb{N} \mid$. However, the map $\psi : \mathbb{N} \to \mathbb{Q}$ given by $\psi(n) = \frac{n}{1}$ is also clealry $1-1$, so we have $\mid \mathbb{N} \mid \preceq \mid \mathbb{Q} \mid$. This leads us to ask: Is $\mathbb{N} \sim \mathbb{Q}$?

The question that arises from the previous example can be generalized as follows:

PROBLEM **4.** *Assume that $A_1 \subseteq A$ and that $B_1 \subseteq B$. If $A \sim B_1$ and $B \sim A_1$, is $A \sim B$?*

While it would seem reasonable that the answer to the above question is Yes, it turns out that the proof is not so straight forward. It can however, be deduced from the following important result:

THEOREM **1.3.12** [CANTOR-SCHROEDER-BERNSTEIN THEOREM]**.** *Assume that $A_2 \subseteq A_1 \subseteq A_0 = A$. If $A_0 \sim A_2$, then $A_0 \sim A_1$.*

Proof. Assume that $\phi : A_0 \to A_2$ is $1-1$ and onto. We can recursively define

$$A_{n+2} = \phi(A_n).$$

We get a sequence $\{A_n\}$ of subsets of A_0 such that .....

We can write

$$A_0 = (A_0 \setminus A_1) \cup (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup (A_3 \setminus A_4) \cup \cdots \cup A_\infty$$

where

$$A_\infty = \bigcap_{n=0}^{\infty} A_n = \bigcap_{n=1}^{\infty} A_n.$$

Similarly

$$A_1 = (A_1 \setminus A_2) \cup (A_2 \setminus A_3) \cup (A_3 \setminus A_4) \cup (A_4 \setminus A_5) \cup \cdots \cup A_\infty.$$

Now becase ϕ is $1-1$ and the range is A_2, we get that for each $k = 0, 1, 2, \ldots$ we have $(A_k \setminus A_{k+1}) \sim (A_{k+2} \setminus A_{k+3})$. In particular,

$$(A_0 \setminus A_1) \sim (A_2 \setminus A_3), (A_2 \setminus A_3) \sim (A_4 \setminus A_5), (A_4 \setminus A_5) \sim (A_6 \setminus A_7), \ldots.$$

We can now define $f : A_0 \to A_1$ by

$$f(x) := \begin{cases} \phi(x) & \text{if } x \in A_{2k} \setminus A_{2k+1} \text{ for some } k = 0, 1, 2, \ldots, \\ x & \text{if } x \in A_{2k+1} \setminus A_{2k+2} \text{ for some } k = 0, 1, 2, \ldots, \\ x & \text{if } x \in A_\infty. \end{cases}$$

It follows that f is both $1-1$ and onto, and as such that $A_0 \sim A_1$.

∎

COROLLARY **1.3.13** [CANTOR-SCHROEDER-BERNSTEIN THEOREM]. *Assume that $A_1 \subseteq A$ and that $B_1 \subseteq B$. If $A \sim B_1$ and $B \sim A_1$, then $A \sim B$.*

Proof. Let $f : A \to B_1$ be 1-1 and onto. Let $g : B \to A_1$ be $1-1$ and onto. Let

$$A_2 = g \circ f(A) = g(B_1).$$

Then since both f and g are $1-1$, we have $A \sim A_2$. But since $A_2 \subseteq A_1 \subseteq A$ the Cantor-Schroter-Bernstein Theorem shows that $A \sim A_1$. However, $A_1 \sim B$ so indeed $A \sim B$. ∎

COROLLARY **1.3.14.** *An infinite set X is countably infinite if and only if there exists a $1-1$ function $f : X \to \mathbb{N}$. In particular, $\mathbb{Q} \sim \mathbb{N}$.*

Proof. Assume that X is countably infinite, then be definition there is a $1-1$ and onto function $f : X \to \mathbb{N}$.

Assume that there exists a $1-1$ function $f : X \to \mathbb{N}$. It follows that $\mid X \mid \preceq \mid \mathbb{N} \mid$. Conversely, since X is infinite, there exists a countably infinite subset $\{x_1, x_2, x_3, \ldots\}$. The function $g : \mathbb{N} \to X$ defined by $g(n) = x_n$ is $1-1$. As such $\mid \mathbb{N} \mid \preceq \mid X \mid$. The CSB Theorem shows that $\mid \mathbb{N} \mid = \mid X \mid$. ∎

REMARK **1.3.15.** *We have seen that if there exists a $1-1$ function $f : X \to Y$, then X is equivalent to a subset of Y. It follows that we should view Y as containing at least as many elements as X. But what happens if there exists an onto function $g : X \to Y$? In this case, it would seem that there are enough points in X to cover all of Y. As such intuitively we would expect that $\mid Y \mid \preceq \mid X \mid$.*

PROPOSITION **1.3.16.** *Assume that there exists an onto function $g : X \to Y$. Then there exists a $1-1$ function $f : Y \to X$. That is $\mid Y \mid \preceq \mid X \mid$.*

Proof. Assume that $g : X \to Y$ is onto. For each $y \in Y$, the set

$$g^{-1}(\{y\}) \neq \emptyset.$$

By the Axiom of Choice there is a choice function

$$h : \mathcal{P}(X) \setminus \{\emptyset\} \to X$$

such that for every $A \subset X$ with $A \neq \emptyset$, we have $h(A) \in A$. Then if we define $f : Y \to X$ by

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

then f is $1-1$.

∎

COROLLARY **1.3.17.** *Given two sets X and Y. The following are equivalent.*

1. *There exists a $1-1$ function $f : X \to Y$.*
2. *There exists an onto function $g : Y \to X$.*

3. $\mid X \mid \preceq \mid Y \mid$.

We have already seen that if X is infinite, then $\mid \mathbb{N} \mid \preceq \mid X \mid$. It is therefore natural to ask if every pair of infinite sets are equivalent to one another.

Theorem **1.3.18.** $[0,1]$ *is uncountable.*

Proof. Assume that $[0,1]$ is countable. Then we can list the elements of $[0,1]$ in a sequence $\{a_1, a_2, a_3, \ldots\}$. Now each a_i has a decimal expansion which is actually unique if we do not allow for an infinite string of consequetive 9's. We can write

$$\begin{array}{rcl} a_1 & = & .a_{1,1}a_{1,2}a_{1,3}\cdots a_{1,n}\cdots \\ a_2 & = & .a_{2,1}a_{2,2}a_{2,3}\cdots a_{2,n}\cdots \\ a_3 & = & .a_{3,1}a_{3,2}a_{3,3}\cdots a_{3,n}\cdots \\ & \vdots & \\ a_n & = & .a_{n,1}a_{n,2}a_{n,3}\cdots a_{n,n}\cdots \\ & \vdots & \end{array}$$

We can construct $b_0 \in [0,1]$ as follows:

$$b_0 = .b_1 b_2 b_3 \cdots b_n \cdots$$

where

$$b_n := \begin{cases} 7 & \text{if } a_{n,n} \neq 7, \\ 3 & \text{if } a_{n,n} = 7. \end{cases}$$

Then it is clear that $b_0 \in [0,1]$ but that $b_0 \neq a_n$ for any n. ∎

Corollary **1.3.19.** $\mathbb{R}$ *is uncountable.*

In fact, it is easy to see that $(0,1) \sim \mathbb{R}$. The map

$$x \to \tan(\pi x - \frac{\pi}{2})$$

establishes a $1-1$ correspondence between $(0,1)$ and $\mathbb{R}$.

Notation: We write $\mid \mathbb{R} \mid = c$.

Problem **5.** *Given two sets X and Y is it always the case that either $\mid X \mid \preceq \mid Y \mid$ or $\mid Y \mid \preceq \mid X \mid$? That is, can we always compare the size of two sets.*

Theorem **1.3.20.** *([Comparability Theorem for Cardinals])*
Given two sets X and Y, either $\mid X \mid \preceq \mid Y \mid$ or $\mid Y \mid \preceq \mid X \mid$.

Proof. Let

$$\mathcal{S} = \{(A, B, f) \mid A \subseteq X, B \subseteq Y, f : A \to B \text{ is 1-1 and onto}\}.$$

We can order $\mathcal{S}$ by $(A_1, B_1, f) \leq (A_2, B_2, g)$ if and only if $A_1 \subseteq A_2$, $B_1 \subseteq B_2$, and $g_{|_A} = f$.
Let $\mathcal{C} = \{(A_\alpha, B_\alpha, f_\alpha)\}_{\alpha \in I}$ be a chain in $\mathcal{S}$. Let

$$A = \bigcup_{\alpha \in I} A_\alpha$$

and

$$B = \bigcup_{\alpha \in I} B_\alpha.$$

Define $f : A \to B$ by

$$f(x) = f_\alpha(x)$$

if $x \in A_\alpha$.

Note that f is well defined because if $x \in A_\alpha$ and $x \in A_\beta$ where $A_\alpha \subseteq A_\beta$, then

$$f_\alpha = (f_\beta)_{|A_\alpha}$$

so $f_\alpha(x) = f_\beta(x)$.

We claim that f is both one to one and onto.

To see that it is 1-1 let $x, y \in A$ be such that $x \neq y$. Then we can find α and β with $x \in A_\alpha$ and $y \in A_\beta$. Moreover we can assume that $A_\alpha \subseteq A_\beta$ and hence that $x, y \in A_\beta$. But then

$$f(x) = f_\beta(x) \neq f_\beta(y) = f(y).$$

To see that f is onto choose $z \in B$. Then $z \in B_\alpha$ for some $\alpha \in I$. But f_α is onto so there exists $x \in A_\alpha$ such that $f_\alpha(x) = z$. Then cleary $f(x) = z$.

It follows that (A, B, f) is an upper bound for $\mathcal{C}$. By Zorn's Lemma $\mathcal{S}$ has a maximal element (A_0, B_0, f_0).

If $A_0 = X$, we are done since then $\mid X \mid \preceq \mid Y \mid$.

Assume that $A_0 \neq X$. Assume also that $B_0 \neq Y$. Then there exists $x_0 \in X \setminus A_0$ and $z_0 \in Y \setminus B_0$. We can know define a function $f_1 : A_0 \cup \{x_0\} \to B_0 \cup \{z_0\}$ by

$$f_1(x) := \begin{cases} f_0(x) & \text{if } x \in A_0, \\ z_0 & \text{if } x = x_0. \end{cases}$$

Since f is also both 1-1 and onto, we have $(A_0, B_0, f_0) < (A_1, B_1, f_1)$ contraditing the maximality of (A_0, B_0, f_0). It follows that if $A_0 \neq X$, then $B_0 = Y$ and we get $\mid Y \mid \preceq \mid X \mid$.

∎

1.4 Cardinal Arithmetic

1.4.1 Sums of Cardinals

REMARK **1.4.1.** *Suppose that two finite sets X and Y are disjoint with $\mid X \mid = n$ and $\mid Y \mid = m$. Then it is easy to see that*

$$\mid X \cup Y \mid = n + m = \mid X \mid + \mid Y \mid .$$

This leads us to the following definition:

DEFINITION **1.4.2** [SUM OF CARDINALS]**.** *Let X and Y be disjoint sets. We define*

$$\mid X \mid + \mid Y \mid \overset{def}{=} \mid X \cup Y \mid .$$

(It is actually easy to see that this is well defined.)

EXAMPLE **1.10.** *Let $X = \{1, 3, 5, \ldots\}$ and $Y = \{2, 4, 6, \ldots\}$. Then it is clear that $X \sim Y \sim \mathbb{N}$, and since $X \cup Y = \mathbb{N}$ we get that*

$$\aleph_0 + \aleph_0 = \aleph_0.$$

REMARK **1.4.3.** *We have just seen that $\aleph_0 + \aleph_0 = \aleph_0$ and since $(0, 1) \sim \mathbb{R} \sim (2, 3)$ we get that $c + c = c$ as well. We call such cardinal numbers idempotent.*

The next result shows that every infinite cardinal number is in fact idempotent

THEOREM **1.4.4.** *If X is infinite, then*

$$| X | + | Y |= \max\{| X |, | Y |\}.$$

In particular

$$| X | + | X |=| X | .$$

Proof. **(Sketch):** We first show that $| X | + | X |=| X |$. To do so use Zorn's Lemma to show that every infinite set X can be written as the disjoint union of a family $\{A_\alpha\}_{\alpha\in I}$ of countably infinite sets and then note that each such set can be split into two disjoint countably infinite subsets.

Finally, it is easy to see that

$$\max\{| X |, | Y |\} \leq| X | + | Y |\leq \max\{| X |, | Y |\} + \max\{| X |, | Y |\} = \max\{| X |, | Y |\}.$$

From here the CBS Theorem shows that

$$| X | + | Y |= \max\{| X |, | Y |\}.$$

∎

PROBLEM **6.** *The previous proposition can be extended to show that the union of finitely many countable sets is countable. What can we say about a countable union of countable sets?*

THEOREM **1.4.5.** *Assume that $\{X_i\}_{i=1}^{\infty}$ is a countable collection of countable sets. Then*

$$X = \bigcup_{i=1}^{\infty} X_i$$

is countable.

Proof. First we note that we may assume that $X_i \cap X_j = \emptyset$ if $i \neq j$. If not we can define a new collection $\{E_i\}_{i=1}^{\infty}$ by $E_1 = X_1$, $E_2 = X_2 \setminus X_1$,

$$E_{n+1} = X_{n+1} \setminus \bigcup_{i=1}^{n} X_i.$$

Assuming that the collection $\{X_i\}_{i=1}^{\infty}$ is pairwise disjoint, for each $i = 1, 2, 3, \ldots$, if $X_i \neq \emptyset$ let

$$X_i = \{x_{i,j}\}$$

where the index j may run over a finite set or over $\mathbb{N}$. Now define $f : X = \bigcup\limits_{i=1}^{\infty} X_i \to \mathbb{N}$ by

$$f(x_{i,j}) = 2^i 3^j.$$

Since f is 1-1, we see that X is countable.

∎

1.4.2 Products of Cardinals

REMARK **1.4.6.** *Suppose that two finite sets X and Y are disjoint with $| X |= n$ and $| Y |= m$. Then it is easy to see that*

$$| X \times Y |= n \cdot m =| X | \cdot | Y | .$$

This leads us to the following definition:

DEFINITION **1.4.7** [PRODUCT OF CARDINALS]. *Let X and Y be two sets. We define*

$$| X | \cdot | Y | \overset{def}{=} | X \times Y | .$$

(It is again easy to see that this is well defined.)

EXAMPLE **1.11.** *Show that $\aleph_0 \cdot \aleph_0 = \aleph_0$.*

It suffices to show that $\mathbb{N} \times \mathbb{N}$ is countable. To do so we define $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by

$$f((n, m)) = 2^n 3^m.$$

As before, f is 1-1, and hence $\mathbb{N} \times \mathbb{N}$ is countable.

REMARK **1.4.8.** *We have already seen that $\aleph_0 \cdot \aleph_0 = \aleph_0$. The next result shows that every infinite has the same property.*

THEOREM **1.4.9.** *If X is infinite and Y is nonempty, then*

$$| X | \cdot | Y |= \max\{| X | | Y |\}.$$

In particular

$$| X | \cdot | X |=| X | .$$

The proof of the previous theorem relies on a clever use of Zorn's Lemma. We omit it here for now.

1.4.3 Exponentiation of Cardinals

REMARK **1.4.10.** *Recall that given a collection $\{Y_x\}_{x \in X}$ of sets we defined*

$$\prod_{x \in X} Y_x = \{f : X \to \bigcup_{x \in X} Y_x \mid f(x) \in Y_x\}.$$

If we modify our notation so that for each $x \in X$ we have $Y_x = Y$ for some fixed Y, then we get

$$Y^X = \prod_{x \in X} Y_\alpha = \prod_{x \in X} Y = \{f : X \to Y\}.$$

Now if $X = \{1, 2, 3, \ldots, n\}$ and $Y = \{1, 2, 3, \cdots, m\}$, then

$$| Y^X |=| \{f : \{1, 2, 3, \ldots, n\} \to \{1, 2, 3, \ldots, m\}\} |= m^n =| Y |^{|X|} .$$

DEFINITION **1.4.11.** *Let X and Y be non-empty sets. Then we define*

$$| Y |^{|X|} \overset{def}{=} | Y^X |$$

The following theorem shows that our familiar laws for exponentiation do hold.

THEOREM **1.4.12.** *Let X, Y and Z be non-empty sets. Then*

1) $| Y |^{|X|} \cdot | Y |^{|Z|}=| Y |^{|X|+|Z|}$

2) $(| Y |^{|X|})^{|Z|} =| Y |^{(|X| \cdot |Z|)}$

Example **1.12.** *Show that* $2^{\aleph_0} = c$.

Observe that $2^{\aleph_0}$ *represents the cardinality of the set* $\{0,1\}^{\mathbb{N}} = \{\{a_1, a_2, \ldots, a_n, \ldots\} \mid a_i = 0, 1\}$. *Define a function* $f : \{0,1\}^{\mathbb{N}} \to [0,1]$ *by*

$$f(\{a_1, a_2, \ldots, a_n, \ldots\}) = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$$

Since f *is 1-1 we get that* $2^{\aleph_0} \preceq c$.

For each $\alpha \in [0,1]$ *choose a sequence* $\{a_n\} \in \{0,1\}^{\mathbb{N}}$ *such that*

$$\alpha = \sum_{n=1}^{\infty} \frac{a_n}{2^n}.$$

Then the function given by $g(\alpha) = \{a_n\}$ *is a 1-1 functiopn from* $[0,1]$ *into* $\{0,1\}^{\mathbb{N}}$. *It follows that* $c \preceq 2^{\aleph_0}$ *and as such by the Cantor Bernstein Theorem that* $2^{\aleph_0} = c$.

Remark **1.4.13.** *Let* X *be any set. Let* $A \subseteq X$. *The characteristic function of* A *is the function*

$$\chi_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases}$$

Then $\chi_A \in \{0,1\}^X$.

Conversely, if $f \in \{0,1\}^X$, *and* $A = \{x \in X \mid f(x) = 1\}$, *then* $f = \chi_A$.

It follows that the map $\Gamma : \mathcal{P}(X) \to \{0,1\}^X$ *given by*

$$\Gamma(A) = \chi_A$$

is 1-1 and onto. Consequently, we have that $\mid \mathcal{P}(X) \mid = 2^{|X|}$.

The next theorem shows that the Power set of a set X is always strictly larger than X.

Theorem **1.4.14** [Russel's Paradox]. *For any set* X, *we have that* $\mid X \mid \prec \mid \mathcal{P}(X) \mid$.

Proof. Given that we know that any two cardinal numbers are comparable, it suffices to show there is no onto function $f : X \to \mathcal{P}(X)$.

Assume to the contrary that there is an onto function $f : X \to \mathcal{P}(X)$. Let

$$A = \{x \in X \mid x \notin f(x)\}.$$

Since $A \subseteq X$ and f is onto, there exists an $x_0 \in X$ such that $f(x_0) = A$.

First we assume that $x_0 \in A$. Then by definition we have $x_0 \notin f(x_0) = A$, which is impossible.

Next we assume that $x_0 \notin A$. Then we have $x_0 \notin f(x_0)$ and as such by the definition of A we have $x_0 \in A$ which is again impossible.

It follows that no such function f could exist and hence that $\mid X \mid \prec \mid \mathcal{P}(X) \mid$. ∎

Finally, we note that $\aleph_0 \prec c$. We can ask; Is there a set with cardinality strictly between $\aleph_0$ and c. The answer to this question is actually not derivable from the standard rules of set theory. THis leads us to adopt the following additional axiom which is both consistent with and independent of the other typical axioms of set theory:

Axiom **1.4.15.** ***[Continuum Hypothesis]*** *Assume that X is such that $\aleph_0 \preceq \mid X \mid \preceq c$. Then either $\mid X \mid = \aleph_0$ or $\mid X \mid = c$.*

Next we observe that $c = 2^{\aleph_0}$ and that in general $\mid X \mid \prec 2^{\mid X \mid}$. We can extend the Continuum Hypothesis further by adopting:

Axiom **1.4.16.** ***[Generalized Continuum Hypothesis]*** *Assume that X and Y are such that $\mid X \mid \preceq \mid Y \mid \preceq 2^{\mid X \mid}$. Then either $\mid Y \mid = \mid X \mid$ or $\mid Y \mid = 2^{\mid X \mid}$.*

Chapter 2

Metric Spaces

2.0 Basic Concepts and Examples

DEFINITION **2.0.1.** *Let X be a non-empty set. A metric on X is a function $d : X \times X \to \mathbb{R}$ such that*

M1) $d(x, y) \geq 0$ *for all* $x, y \in X$ *and* $d(x, y) = 0$ *if and only if* $x = y$.

M2) $d(x, y) = d(y, x)$ *for all* $x, y \in X$.

M3) $d(x, y) \leq d(x, z) + d(z, y)$ *for all* $x, y, z \in X$. *[Triangle Inequality]*

The pair (X, d) *is called a metric space.*

REMARK **2.0.2.** *A metric is an abstract distance function. The first example below is the motivating example.*

EXAMPLE **2.1.** *1) Let* $X = \mathbb{R}$ *and let* $d(x, y) =| x - y |$.

Here the first two conditions are clearly satisfied by the definition of the absolute value function, and the triangle inequality is the usual triangle inequality property of the absoulte value function.

2) Let X *be any set. Define*

$$d(x, y) := \begin{cases} 1 & \text{if } x \neq y, \\ 0 & \text{if } x = y. \end{cases}$$

It is easy to verify that this is a metric. It is called the discrete metric.

3) Let $X = \mathbb{R}^n$. *Let* $\boldsymbol{x} = \{x_1, x_2, x_3 \ldots, x_n\}$ *and* $\boldsymbol{y} = \{y_1, y_2, y_3 \ldots, y_n\}$. *Define*

$$d_2(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}.$$

d_2 *is called the Euclidean metric.*

This metric carries very little information about the underlying set, though it will prove to be an important example.

It clearly satisfies $M1)$ *and* $M2)$. *Heuristically, the triangle inequality is the familiar result that the legth of any one side of a triangle is less than or equal to the sum of the legths of the other two side. We will however prove later that it does satisfy the triangle inequality.*

REMARK **2.0.3.** *Many of the most important examples of metric spaces are vector spaces with an abstract length function or norm.*

DEFINITION **2.0.4.** *Let V be a vector space. A norm on V is a function $\|\cdot\| : V \to \mathbb{R}$ such that*

N1) $\|x\| \geq 0$ *for all* $x \in V$ *and* $\|x\| = 0$ *if and only if* $x = 0$.

N2) $\|\alpha x\| = \mid \alpha \mid \|x\|$ *for all* $x \in V$, $\alpha \in \mathbb{R}$.

N3) $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in V$. *[Triangle Inequality]*

The pair $(V, \|\cdot\|)$ is called a normed linear space.

REMARK **2.0.5.** *Given a normed linear space $(V, \|\cdot\|)$ we have a natural metric on V induced by $\|\cdot\|$ defined as follows:*

$$d_{\|\cdot\|}(x, y) = \|x - y\|.$$

To see that this is indeed a metric observe:

M1) $d_{\|\cdot\|}(x, y) = \|x - y\| \geq 0$ *for all* $x, y \in V$ *and* $d_{\|\cdot\|}(x, y) = 0$ *if and only if* $x - y = 0 \Leftrightarrow x = y$.

M2) $d_{\|\cdot\|}(x, y) = \|x - y\| = \|y - x\| = d(y, x)$ *for all* $x, y \in X$.

M3) $d_{\|\cdot\|}(x, y) = \|x - y\| \leq \|x - z\| + \|z - y\| = d_{\|\cdot\|}(x, z) + d_{\|\cdot\|}(z, y)$ *for all* $x, y, z \in V$.

EXAMPLE **2.2.** *1) Let $X = \mathbb{R}^n$. Let $\boldsymbol{x} = \{x_1, x_2, x_3 \ldots, x_n\}$. Define*

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

Then we will see that $\|\cdot\|_2$ is a norm. We call $\|\cdot\|_2$, the 2-norm or the Euclidean norm.

2) Let $X = \mathbb{R}^n$. Let $\boldsymbol{x} = \{x_1, x_2, x_3 \ldots, x_n\}$. Define

$$\|x\|_1 = \mid x_1 \mid + \mid x_2 \mid + \cdots + \mid x_n \mid = \sum_{i=1}^{n} \mid x_i \mid .$$

Then it is actually easy to see that $\|\cdot\|_1$ is a norm. We call $\|\cdot\|_1$, the 1-norm .

3) Let $X = \mathbb{R}^n$. Let $\boldsymbol{x} = \{x_1, x_2, x_3 \ldots, x_n\}$. Define

$$\|x\|_\infty = max\{\mid x_1 \mid, \mid x_2 \mid, \ldots, \mid x_n \mid\}.$$

Then it is again easy to see that $\|\cdot\|_\infty$ is a norm. We call $\|\cdot\|_\infty$, the ∞-norm or the sup norm .

4) Let $X = \mathbb{R}^n$. Let $\boldsymbol{x} = \{x_1, x_2, x_3 \ldots, x_n\}$. Let $1 < p < \infty$. Define

$$\|x\|_p = (\mid x_1 \mid^p + \mid x_2 \mid^p + \cdots + \mid x_n \mid^p)^{\frac{1}{p}} = (\sum_{i=1}^{n} \mid x_i \mid^p)^{\frac{1}{p}}.$$

REMARK **2.0.6.** *To show that $\|\cdot\|_p$ determines a norm on $\mathbb{R}^n$, we begin with the following lemma. Before we state and prove the result we note that if $1 < p < \infty$ and if $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$1 + \frac{p}{q} = p$$

so

$$p - 1 = \frac{p}{q} \Rightarrow \frac{1}{p-1} = \frac{q}{p} = q - 1$$

and

$$(p-1)q = p.$$

LEMMA **2.0.7.** *Let $\alpha, \beta > 0$, Let $1 < p < \infty$. Then if $\frac{1}{p} + \frac{1}{q} = 1$,*

$$\alpha\beta \leq \frac{\alpha^p}{p} + \frac{\beta^q}{q}.$$

Proof. Let $u = t^{p-1}$ and as such $t = u^{\frac{1}{p-1}} = u^{q-1}$. Then

$$\alpha\beta \leq \int_0^\alpha t^{p-1}\, dt + \int_0^\beta u^{q-1}\, dt = \frac{\alpha^p}{p} + \frac{\beta^q}{q}.$$

∎

THEOREM **2.0.8** [HÖLDERS INEQUALITY I]. *Let $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ and $(b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$. Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\sum_{i=1}^n \mid a_i b_i \mid \leq (\sum_{i=1}^n \mid a_i \mid^p))^{\frac{1}{p}} (\sum_{i=1}^n \mid b_i \mid^q))^{\frac{1}{q}}. \qquad (*)$$

Proof. First observe that we can assume that

$$(a_1, a_2, \ldots, a_n) \neq (0, 0, \ldots 0) \neq (b_1, b_2, \ldots, b_n),$$

otherwise the inequality is trivially and equality holds. Moreover since for any $\alpha, \beta > 0$ we have that

$$\sum_{i=1}^n \mid (\alpha a_i)(\beta b_i) \mid = \alpha\beta \sum_{i=1}^n \mid a_i b_i \mid,$$

$$(\sum_{i=1}^n \mid \alpha a_i \mid^p)^{\frac{1}{p}} = \alpha (\sum_{i=1}^n \mid a_i \mid^p)^{\frac{1}{p}},$$

and

$$(\sum_{i=1}^n \mid \beta b_i \mid^q))^{\frac{1}{q}} = \beta (\sum_{i=1}^n \mid b_i \mid^q))^{\frac{1}{q}},$$

it follows that $(*)$ holds for $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ and $(b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$ if and only if it holds for $(\alpha a_1, \alpha a_2, \ldots, \alpha a_n) \in \mathbb{R}^n$ and $(\beta b_1, \beta b_2, \ldots, \beta b_n) \in \mathbb{R}^n$ for some $\alpha, \beta > 0$. As such, by scaling if necessary, we may assume that

$$\sum_{i=1}^n \mid a_i \mid^p = 1 = \sum_{i=1}^n \mid b_i \mid^q .$$

Now for each $i = 1, 2, \ldots, n$, we have

$$\mid a_i b_i \mid \leq \frac{\mid a_i \mid^p}{p} + \frac{\mid b_i \mid^q}{q}.$$

As such

$$\begin{aligned}
\sum_{i=1}^n \mid a_i b_i \mid &\leq \frac{\sum_{i=1}^n \mid a_i \mid^p}{p} + \frac{\sum_{i=1}^n \mid b_i \mid^q}{q} \\
&= \frac{1}{p} + \frac{1}{q} \\
&= 1 \\
&= (\sum_{i=1}^n \mid a_i \mid^p))^{\frac{1}{p}} (\sum_{i=1}^n \mid b_i \mid^q))^{\frac{1}{q}}.
\end{aligned}$$

∎

Theorem **2.0.9** [Minkowski's Inequality]. *Let* $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ *and* $(b_1, b_2, \ldots, b_n) \in \mathbb{R}^n$. *Let* $1 < p < \infty$. *Then*

$$(\sum_{i=1}^{n} \mid a_i + b_i \mid^p)^{\frac{1}{p}} \leq (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}} + (\sum_{i=1}^{n} \mid b_i \mid^p)^{\frac{1}{p}}$$

Proof. First let $\frac{1}{p} + \frac{1}{q} = 1$.

We can again assume that

$$(a_1, a_2, \ldots, a_n) \neq (0, 0, \ldots 0) \neq (b_1, b_2, \ldots, b_n),$$

otherwise the inequality is trivially true.

Now

$$\begin{aligned} \sum_{i=1}^{n} \mid a_i + b_i \mid^p &= \sum_{i=1}^{n} \mid a_i + b_i \mid \cdot \mid a_i + b_i \mid^{p-1} \\ &\leq \sum_{i=1}^{n} \mid a_i \mid \cdot \mid a_i + b_i \mid^{p-1} + \sum_{i=1}^{n} \mid b_i \mid \cdot \mid a_i + b_i \mid^{p-1} \end{aligned}$$

By Hölders Inequality

$$\begin{aligned} \sum_{i=1}^{n} \mid a_i \mid \cdot \mid a_i + b_i \mid^{p-1} &\leq (\sum_{i=1}^{n} \mid a_i \mid^p))^{\frac{1}{p}} \cdot (\sum_{i=1}^{n} \mid a_i + b_i \mid^{(p-1)q})^{\frac{1}{q}} \\ &= (\sum_{i=1}^{n} \mid a_i \mid^p))^{\frac{1}{p}} \cdot (\sum_{i=1}^{n} \mid a_i + b_i \mid^p)^{\frac{1}{q}} \end{aligned}$$

Similarly we get that

$$\sum_{i=1}^{n} \mid b_i \mid \cdot \mid a_i + b_i \mid^{p-1} \leq (\sum_{i=1}^{n} \mid b_i \mid^p))^{\frac{1}{p}} \cdot (\sum_{i=1}^{n} \mid a_i + b_i \mid^p)^{\frac{1}{q}}.$$

Putting everything together we get

$$\begin{aligned} \sum_{i=1}^{n} \mid a_i + b_i \mid^p &\leq \sum_{i=1}^{n} \mid a_i \mid \cdot \mid a_i + b_i \mid^{p-1} + \sum_{i=1}^{n} \mid b_i \mid \cdot \mid a_i + b_i \mid^{p-1} \\ &\leq [(\sum_{i=1}^{n} \mid a_i \mid^p))^{\frac{1}{p}} + (\sum_{i=1}^{n} \mid b_i \mid^p))^{\frac{1}{p}}] \cdot (\sum_{i=1}^{n} \mid a_i + b_i \mid^p)^{\frac{1}{q}} \end{aligned}$$

Therefore, dividing both sides of the above inequality by $(\sum\limits_{i=1}^{n} \mid a_i + b_i \mid^p)^{\frac{1}{q}}$, we get,

$$(\sum_{i=1}^{n} \mid a_i + b_i \mid^p)^{1-\frac{1}{q}} = (\sum_{i=1}^{n} \mid a_i + b_i \mid^p)^{\frac{1}{p}} \leq (\sum_{i=1}^{n} \mid a_i \mid^p)^{\frac{1}{p}} + (\sum_{i=1}^{n} \mid b_i \mid^p)^{\frac{1}{p}}$$

∎

Example **2.3** [Sequence Spaces]. *1) Let* $l_1 = \{\{x_i\} \mid \sum\limits_{i=1}^{\infty} \mid x_i \mid < \infty\}$. *Define*

$$\|\{x_i\}\|_1 = \sum_{i=1}^{\infty} \mid x_i \mid .$$

Observe that if $\{x_i\}$ *and* $\{y_i\}$ *are in* l_1*, then for any* $k \in \mathbb{N}$*,*

$$\begin{aligned}\sum_{i=1}^{k} \mid x_i + y_i \mid &\leq \sum_{i=1}^{k} \mid x_i \mid + \sum_{i=1}^{k} \mid y_i \mid \\ &\leq \sum_{i=1}^{\infty} \mid x_i \mid + \sum_{i=1}^{\infty} \mid y_i \mid \\ &= \|\{x_i\}\|_1 + \|\{y_i\}\|_1\end{aligned}$$

It follows from the Monotone Convergence Theorem that $\sum_{i=1}^{\infty} \mid x_i + y_i \mid$ *converges and hence that* $\{x_i + y_i\} \in l_1$*. Moreover it also shows us that*

$$\|\{x_i + y_i\}\|_1 \leq \|\{x_i\}\|_1 + \|\{y_i\}\|_1.$$

A similar arguement shows that if $\{x_i\} \in l_1$ *and* $\alpha \in \mathbb{R}$*, then* $\{\alpha x_i\} \in l_1$ *and*

$$\|\{\alpha x_i\}\|_1 = \mid \alpha \mid \|\{x_i\}\|_1.$$

It follows that $(l_1, \|\{\cdot\}\|_1)$ *is a normed linear space.*

2) Let $1 < p < \infty$*. Let* $l_p = \{\{x_i\} \mid \sum_{i=1}^{\infty} \mid x_i \mid^p < \infty\}$*. Define*

$$\|\{x_i\}\|_p = (\sum_{i=1}^{\infty} \mid x_i \mid^p)^{\frac{1}{p}}.$$

Observe that if $\{x_i\}$ *and* $\{y_i\}$ *are in* l_p*, then for any* $k \in \mathbb{N}$*,*

$$\begin{aligned}(\sum_{i=1}^{k} (\mid x_i + y_i \mid)^p)^{\frac{1}{p}} &\leq (\sum_{i=1}^{k} \mid x_i \mid^p)^{\frac{1}{p}} + (\sum_{i=1}^{k} \mid y_i \mid^p)^{\frac{1}{p}} \\ &\leq (\sum_{i=1}^{\infty} \mid x_i \mid^p)^{\frac{1}{p}} + (\sum_{i=1}^{\infty} \mid y_i \mid^p)^{\frac{1}{p}} \\ &= \|\{x_i\}\|_p + \|\{y_i\}\|_p\end{aligned}$$

As before, this shows that $\{x_i + y_i\} \in l_p$ *and that*

$$\|\{x_i + y_i\}\|_p \leq \|\{x_i\}\|_p + \|\{y_i\}\|_p.$$

A similar arguement shows that if $\{x_i\} \in l_p$ *and* $\alpha \in \mathbb{R}$*, then* $\{\alpha x_i\} \in l_p$ *and*

$$\|\{\alpha x_i\}\|_p = \mid \alpha \mid \|\{x_i\}\|_p.$$

3) Let $l_\infty = \{\{x_i\} \mid sup\{\mid x_i \mid\} < \infty\}$*. Define*

$$\|x\|_\infty = sup\{\mid x_i \mid\}.$$

Then it is again easy to see that $\|\cdot\|_\infty$ *is a norm.*

EXAMPLE **2.4.** *The space* $(C([a,b]), \|\cdot\|_\infty)$.

Let $C([a,b]) = \{f : [a,b] \to \mathbb{R} \text{ such that } f(x) \text{ is continuous on } [a,b]\}$. *Let*

$$\|f\|_\infty = max\{| f(x) || x \in [a,b]\}.$$

If $f, g \in C([a,b])$, *then for each* $x \in [a,b]$,*we have*

$$\begin{aligned} |(f+g)(x)| &= |f(x)+g(x)| \\ &\leq |f(x)| + |g(x)| \\ &\leq \|f\|_\infty + \|g\|_\infty. \end{aligned}$$

From this it follows immediately that

$$\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty.$$

It is then a straight forward task to show that $\|\cdot\|_\infty$ *defines a norm on* $C([a,b])$.
This space will play a fundamental role in much of the later portion of this course.

EXAMPLE **2.5.** *The space* $(C([a,b]), \|\cdot\|_1)$.

Given $f \in C([a,b])$, *define*

$$\|f\|_1 = \int_a^b | f(t) | \, dt \qquad (*).$$

It follows from the linearity of integration and the usual triangle inequality for $\mathbb{R}$ *that* $\|\cdot\|_1$ *defines a norm on* $C([a,b])$.

EXAMPLE **2.6.** *The space* $(C([a,b]), \|\cdot\|_p)$.

Given $f \in C([a,b])$, *we claim that*

$$\|f\|_p = (\int_a^b | f(t) |^p \, dt)^{\frac{1}{p}}$$

defines a norm on $C([a,b])$. *To see this we observe that all of the properties of a norm hold trivially, with the exception of the triangle inequality. To establish the triangle inequality, we will first need to prove analogues of Hölders Inequality and Minkowski's Inequality.*

THEOREM **2.0.10** [HÖLDERS INEQUALITY II]**.** *Let* $1 < p < \infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$. *Then for each* $f, g \in C([a,b]$, *we have*

$$\int_a^b | f(t)g(t) | \, dt \leq (\int_a^b | f(t) |^p \, dt)^{\frac{1}{p}} (\int_a^b | g(t) |^q \, dt)^{\frac{1}{q}}.$$

Proof. Clearly the result holds if either $f(x) = 0$ for all $x \in [a,b]$ or if $g(x) = 0$ for all $x \in [a,b]$. As such we may assume that this is not the case. Since for any $\alpha, \beta >= 0$ we have

$$\int_a^b | (\alpha f(t))(\beta g(t)) | \, dt = \alpha\beta \int_a^b | f(t)g(t) | \, dt,$$

$$(\int_a^b | \alpha f(t) |^p \, dt)^{\frac{1}{p}} = \alpha(\int_a^b | f(t) | \, dt)^{\frac{1}{p}},$$

and

$$(\int_a^b \mid \beta g(t) \mid^q dt)^{\frac{1}{q}} = \beta(\int_a^b \mid g(t) \mid^q dt)^{\frac{1}{q}},$$

using the same trick as in the previous proof of Hölders Inequality for vectors of $\mathbb{R}^n$ we see that we only need to prove $(*)$ under the additional assumption that

$$\int_a^b \mid f(t) \mid^p dt = 1 = \int_a^b \mid g(t) \mid^q dt,$$

Again, we have that for each $t \in [a, b]$,

$$\mid f(t)g(t) \mid \leq \frac{\mid f(t) \mid^p}{p} + \frac{\mid g(t) \mid^q}{q}.$$

Integrating, we get

$$\begin{aligned}
\int_a^b \mid f(t)g(t) \mid dt &\leq \int_a^b (\frac{\mid f(t) \mid^p}{p} + \frac{\mid g(t) \mid^q}{q})dt \\
&= \frac{1}{p} + \frac{1}{q} \\
&= 1 \\
&= (\int_a^b \mid f(t) \mid^p dt)^{\frac{1}{p}} (\int_a^b \mid g(t) \mid^q dt)^{\frac{1}{q}}.
\end{aligned}$$

∎

THEOREM **2.0.11** [MINKOWSKI'S INEQUALITY II]. *Let $f, g \in C([a, b])$. Let $1 < p < \infty$. Then*

$$(\int_a^b \mid (f+g)(t)) \mid^p dt)^{\frac{1}{p}} \leq (\int_a^b \mid f(t) \mid^p dt)^{\frac{1}{p}} + (\int_a^b \mid g(t) \mid^p dt)^{\frac{1}{p}}.$$

Proof. This result can be obtained from Hölders Inequality II in nearly the exact same way as Minkowski's Inequality I was deduced from Hölders Inequality I. Indeed we may assume that

$$\int_a^b \mid f(t) \mid^p dt \neq 0 \neq \int_a^b \mid g(t) \mid^p dt.$$

Then

$$\begin{aligned}
\int_a^b \mid (f+g)(t) \mid^p dt &= \int_a^b \mid f(t) + g(t) \mid \cdot \mid (f+g)(t) \mid^{p-1} dt \\
&\leq \int_a^b \mid f(t) \mid \cdot \mid (f+g)(t) \mid^{p-1} dt + \int_a^b \mid g(t) \mid \cdot \mid (f+g)(t) \mid^{p-1} dt
\end{aligned}$$

By Hölders Inequality

$$\begin{aligned}
\int_a^b \mid f(t) \mid \cdot \mid (f+g)(t) \mid^{p-1} dt &\leq (\int_a^b \mid f(t) \mid^p)dt)^{\frac{1}{p}} \cdot (\int_a^b \mid (f+g)(t) \mid^{(p-1)q})^{\frac{1}{q}} \\
&= (\int_a^b \mid f(t) \mid^p)dt)^{\frac{1}{p}} \cdot (\int_a^b \mid (f+g)(t) \mid^p)^{\frac{1}{q}}
\end{aligned}$$

Similarly we get that

$$\int_a^b | g(t) | \cdot | (f+g)(t) |^{p-1} \, dt \leq (\int_a^b | g(t) |^p) dt)^{\frac{1}{p}} \cdot (\int_a^b | (f+g)(t) |^p)^{\frac{1}{q}}.$$

Putting everything together we get

$$\begin{aligned} \int_a^b | (f+g)(t) |^p \, dt &\leq \int_a^b | f(t) | \cdot | (f+g)(t) |^{p-1} \, dt + \int_a^b | g(t) | \cdot | (f+g)(t) |^{p-1} \, dt \\ &\leq [(\int_a^b | f(t) |^p \, dt)^{\frac{1}{p}} + (\int_a^b | g(t) |^p \, dt)^{\frac{1}{p}}] \cdot (\int_a^b | (f+g)(t) |^p \, dt)^{\frac{1}{q}}). \end{aligned}$$

Finally, dividing both sides of the above inequality by $(\int_a^b | (f+g)(t) |^p \, dt)^{\frac{1}{q}}$ gives

$$\begin{aligned} (\int_a^b | (f+g)(t)) |^p \, dt)^{\frac{1}{p}} &= (\int_a^b | (f+g)(t)) |^p \, dt)^{1-\frac{1}{q}} \\ &\leq (\int_a^b | f(t) |^p \, dt)^{\frac{1}{p}} + (\int_a^b | g(t) |^p \, dt)^{\frac{1}{p}}. \end{aligned}$$

∎

DEFINITION **2.0.12.** *(Bounded Operator)*

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces. Let $T : X \to Y$ be linear. We define

$$\|T\| = sup\{\|T(x)\|_Y \mid x \in X, \|x\|_X \leq 1\}.$$

We say that T is bounded if $\|T\| < \infty$. We let $B(X,Y) = \{T : X - Y \mid T \text{ is linear and bounded}\}$.

REMARK **2.0.13.** *We claim that $B(X,Y)$ is a vector space and that $\|T\|$ is a norm on $B(X,Y)$. Assume that $T, S \in B(X,Y)$. Let $x \in X$ with $\|x\|_X \leq 1$. Then*

$$\begin{aligned} \|(T+S)(x)\|_Y &= \|T(x) + S(x))\|_Y \\ &\leq \|T(x)\|_Y + \|S(x)\|_Y \\ &\leq \|T\| + \|S\|. \end{aligned}$$

This shows that $T + S \in B(X,Y)$ and that $\|T + S\| \leq \|T\| + \|S\|$.

Now let $\alpha \in \mathbb{R}$ and $T \in B(X,Y)$. Then since

$$\begin{aligned} sup\{\|(\alpha T)(x)\|_Y \mid x \in X, \|x\|_X \leq 1\} &= sup\{\|T(\alpha x)\|_Y \mid x \in X, \|x\|_X \leq 1\} \\ &= sup\{| \alpha | \, \|T(x)\|_Y \mid x \in X, \|x\|_X \leq 1\} \\ &= | \alpha | \, sup\{\|T(x)\|_Y \mid x \in X, \|x\|_X \leq 1\} \\ &= | \alpha | \, \|T\|, \end{aligned}$$

we have that $\alpha T \in B(X,Y)$ and that $\|\alpha T\| =| \alpha | \, \|T\|$.

Finally, it is esay to see that $\|T\| = 0$ if and only if $T = 0$. Hence $\|\cdot\|$ does indeed define a norm on X.

2.1 Topology of Metric Spaces

In this section we will introduce some of the basic topological concepts associated with metric spaces.

DEFINITION **2.1.1.** *Let (X, d) be a metric space. Let $x_0 \in X$ and let $\epsilon > 0$. The open ball of radius ϵ centered at x_0 is the set*

$$B(x_0, \epsilon) = \{x \in X \mid d(x, x_0) < \epsilon\}$$

The closed ball of radius ϵ centered at x_0 is the set

$$B[x_0, \epsilon] = \{x \in X \mid d(x, x_0) \leq \epsilon\}$$

A subset $U \subseteq X$ is said to be open if for evey $x_0 \in U$ there exists an $\epsilon > 0$ such that $B(x_0, \epsilon) \subseteq U$.

A subset $F \subseteq X$ is said to be closed if F^c is open.

PROPOSITION **2.1.2.** *Let (X, d) be a metric space. Then*

1) X and $\emptyset$ are both open.

2) If $\{U_\alpha\}_{\alpha \in I}$ is any collection of open sets, then $U = \bigcup_{\alpha \in I} U_\alpha$ is also open.

3) If $\{U_1, U_2, \ldots, U_n\}$ is any finite collection of open sets, then $U = \bigcap_{i=1}^{n} U_i$ is also open.

Proof. 1) That X is open follows immediately since for any $x_0 \in X$ clearly $B(x_1, 1) \subseteq X$. That $\emptyset$ is open follows vacuously from the definition.

2) Let $\{U_\alpha\}_{\alpha \in I}$ be any collection of open setsand let $x_0 \in U = \bigcup_{\alpha \in I} U_\alpha$. Then $x_0 \in U_{\alpha_0}$ for some $\alpha_0 \in I$. But as U_{α_0} is open there exists an $\epsilon > 0$ such that $B(x_0, \epsilon) \subseteq U_{\alpha_0} \subseteq U$. Consequently U is also open.

3) If $\{U_1, U_2, \ldots, U_n\}$ is any finite collection of open sets and $x_0 \in U = \bigcap_{i=1}^{n} U_i$. It follows that for each i we can find an $\epsilon_i > 0$ such that $B(x_0, \epsilon_i) \subseteq U_i$. Now if $\epsilon_0 = min\{\epsilon_1, \epsilon_2, \ldots, \epsilon_n\}$, then for each i we have $B(x_0, \epsilon_0) \subseteq B(x_0, \epsilon_i) \subseteq U_i$. Hence $B(x_0, \epsilon_0) \subseteq \bigcap_{i=1}^{n} U_i = U$ and U is open.

∎

The following proposition follows immediately from the previous proposition and from DeMorgan's Laws. The proof is left as an exercise.

PROPOSITION **2.1.3.** *Let (X, d) be a metric space. Then*

1) X and $\emptyset$ are both closed.

2) If $\{F_\alpha\}_{\alpha \in I}$ is any collection of closed sets, then $F = \bigcap_{\alpha \in I} F_\alpha$ is also closed.

3) If $\{F_1, F_2, \ldots, F_n\}$ is any finite collection of closed sets, then $F = \bigcup_{i=1}^{n} U_i$ is also closed.

DEFINITION **2.1.4.** *Given a set X, a topology on X is a collection of sets $\tau \subseteq \mathcal{P}(X)$ such that*

1) X and $\emptyset$ are in τ.

2) If $\{U_\alpha\}_{\alpha \in I} \subseteq \tau$, then $U = \bigcup_{\alpha \in I} U_\alpha \in \tau$.

3) If $\{U_1, U_2, \ldots, U_n\} \subseteq \tau$, then $U = \bigcap_{i=1}^{n} U_i \in \tau$.

In this case, we call elements of τ τ-open sets, or open sets for short. We call the pair (X, τ) a topological space.

If (X, d) is a metric space, then we let τ_d denote the topology consisting of those subsets of X that are open with respect to the metric d.

The next proposition, which is essentially a consequence of the triangle inequality, identifies some basic open sets and basic closed sets in a metric space (X, d).

PROPOSITION **2.1.5.** *Let (X, d) be a metric space. Then*

1) For any $x_0 \in X$ and $\epsilon > 0$, we have that $B(x_0, \epsilon)$ is open.

2) A set $U \subseteq X$ is open if and only if it is the union of open balls.

3) For any $x_0 \in X$ and $\epsilon \geq 0$, we have that $B[x_0, \epsilon]$ is closed.

4) For any x_0, the set $\{x_0\}$ is closed. In particular, every finite subset of a metric space is closed.

Proof. 1) Let $z \in B(x_0, \epsilon)$ and let $d = d(x_0, z)$. Since $d < \epsilon$, if we let $r = \epsilon - d$, then $r > 0$. Now let $w \in B(z, r)$. Then by the triangle inequality

$$d(w, x_0) \leq d(w, z) + d(z, x_0) < r + d = \epsilon.$$

Consequently, $B(z, r) \subseteq B(x_0, \epsilon)$ so that $B(x_0, \epsilon)$ is open.

2) Since evey open ball is an open set, if U is the union of open balls it is an open set.

Assume that U is open. For each $x \in U$ we can find an $\epsilon_x > 0$ such that $B(x, \epsilon) \subseteq U$. From this it follows that

$$U = \bigcup_{x \in U} B(x, \epsilon_x)$$

3) Let $z \in B[x_0, \epsilon]^c$. Then $d(z, x_0) = d > \epsilon$. This time $r = d - epsilon > 0$. Now let $w \in B(z, r)$. If $w \in B[x_0, \epsilon]$, then we would have

$$d(z, x_0) \leq d(z, w) + d(w, x_0) < \epsilon + r = d$$

which is clearly a contradiction. Consequently, we have $w \in B[x_0, \epsilon]^c$. It follows that $B(z, r) \subseteq B[x_0, \epsilon]^c$ and hence that $B[x_0, \epsilon]^c$ is open. This shows that $B[x_0, \epsilon]$ is closed.

4) Let $y \in \{x_0\}^c$. Let $\epsilon = d(y, x_))$. Then $B(x_0, \epsilon) \subset \{x_0\}^c$ which shows that $\{x_0\}^c$ is open.

∎

EXAMPLE **2.7.** ***Open sets in $\mathbb{R}$.***

Recall that a subset I of $\mathbb{R}$ is an interval if whenever $x, y \in I$ and $x < z < y$, then we must have that $z \in I$.

There are three fundamental types of finite intervals:

1. Open (finite) intervals: $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$ where $a, b \in \mathbb{R}$.

2. Closed (finite) intervals: $[a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$ where $a, b \in \mathbb{R}$.

3. Half-open intervals: $[a, b) = \{x \in \mathbb{R} \mid a \leq x < b\}$ or$(a, b] = \{x \in \mathbb{R} \mid a < x \leq b\}$ where $a, b \in \mathbb{R}$.

There are also two aditional types of infinite intervals or rays.

1. Open rays: $(a, \infty) = \{x \in \mathbb{R} \mid a < x\}$ and $(infty, b) = \{x \in \mathbb{R} \mid x < b\}$ where $a, b \in \mathbb{R}$.

2. *Closed rays:* $[a, \infty) = \{x \in \mathbb{R} \mid a \leq x\}$ *and* $(infty, b] = \{x \in \mathbb{R} \mid x \leq b\}$ *where* $a, b \in \mathbb{R}$.

In addition, both $\mathbb{R}$ *and* $\emptyset$ *are also intervals.*

It is a straight forward exercise to show that every open interval or open ray is actually an open subset of $\mathbb{R}$ *with the ususal metric. It is also clear that* $\mathbb{R}$ *and* $\emptyset$ *are also open. Together we call these the open intervals in* $\mathbb{R}$*. The following theorem tells us exactly how the structure of an arbitrary open set depends upon these key open sets:*

Theorem **2.1.6.** *Let* $U \subseteq \mathbb{R}$ *be open. Then there is a countable collection* $\{I_n\}$ *of pairwise disjoint open intervals such that*

$$U = \bigcup_n I_n.$$

It is also easy to see that every closed interval is closed, as is every closed ray. However, the interval $(0, 1]$ *is neither open or closed.*

Problem **7.** *We know that every open set is the countable union of open intervals. Is every closed set the countable union of closed intervals?*

Example **2.8.** ***Cantor set*** *Let* $P_0 = [0, 1]$*. Construct* P_1 *for* P_0 *by removing the open middle* $\frac{1}{3}$ *of the interval. That is*

$$P_1 = [0, 1] \setminus (\frac{1}{3}, \frac{2}{3}) = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1].$$

Next construct P_2 *from* P_1 *by removing the open interval of length* $\frac{1}{9}$ *from the middle* $\frac{1}{3}$ *each of the two closed subintervals in* P_1*.*

$$P_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1].$$

Continue recursively to construct P_n *from* P_{n-1} *by removing the open interval of length* $\frac{1}{3^n}$ *from the middle* $\frac{1}{3}$ *each of the* 2^{n-1} *closed subintervals in* P_{n-1}*.*

Properties of P_n*:*

1. *Each* P_n *is the union of* 2^n *closed intervals of length* $\frac{1}{3^n}$ *and hence* P_n *is closed.*
2. P_n *contains no interval of length greater than* $\frac{1}{3^n}$*.*

Let

$$P = \bigcap_{n=1}^{\infty} P_n.$$

P *is callled the Cantor ternary set or simply the Cantor set. As the intersection of closed sets* P *is also closed.*

Properties of P*:*

1. $x \in P$ *if and only if we can express* x *as a series with* $x = \sum_{n=1}^{\infty} \frac{a_n}{3^n}$ *where* $a_n = 0, 2$.
2. P *is uncountable.*
3. P *contains no interval of positive length.*

Example **2.9.** ***Open sets and the Discrete Metric.***

Let (X, d) *be a metric space, where* d *is the discrete metric. That is* $d(x, y) = 1$ *if* $x \neq y$*. Then each singleton* $\{x\}$ *is open because*

$$\{x\} = B(x, 1).$$

If $U \subseteq X$, then

$$U = \bigcup_{x \in U} \{x\},$$

so U is an open set. Moreover, since every set is open it is also true that every set is closed.

PROBLEM **8.** *We know that in $\mathbb{R}$ with the usual metric, that $\mathbb{R}$ and $\emptyset$ are both open and closed. Are there any other such subsets in $\mathbb{R}$? The answer to this question is essentiallly the Intermediate Value Theorm.*

2.2 Boundaries, Interiors and Closures of a set.

DEFINITION **2.2.1.** *Let (X, d) be a metric space. Let $A \subseteq X$.*

1) We define the closure of A to be the set

$$\bar{A} = \bigcap \{F \mid A \subseteq F \text{ and } F \text{ is closed in } X\}.$$

That is, $\bar{A}$ is the smallest closed set that contains A.

2) We define the interior of A to be the set

$$int(A) = \bigcup \{U \mid U \subseteq A \text{ and } U \text{ is open in } X\}.$$

That is, $int(A)$ is the largest open set that contained in A. The elements of $int(A)$ are called interior points of A.

3) Let $x \in X$. A subset $N \subseteq X$ is said to be a neighbourhood of x, if $x \in int(N)$.

4) Given a set $A \subseteq X$, an element $x \in X$ is said to be a boundary point of A if for every neighbourhood N of x we have $N \cap A \neq \emptyset$ and $N \cap A^c \neq \emptyset$. (Note: It is easy to see that this is equivalent to the statement that for every $\epsilon > 0$ we have $B(x, \epsilon) \cap A \neq \emptyset$ and $B(x, \epsilon) \cap A^c \neq \emptyset$.)

We denote the set of all boundary points of A by $bdy(A)$.

The following is a useful proposition that tells us when a set is closed in terms of the nature of its boundary points.

PROPOSITION **2.2.2.** *Let $A \subseteq (X, d)$.*

1) A is closed if and only if $bdy(A) \subseteq A$.

2) $\bar{A} = A \cup bdy(A)$.

Proof. 1) First assume that A is closed. Then A^c is open. Let $x \in A^c$. Then there exists an $\epsilon > 0$ such that $B(x, \epsilon) \subseteq A^c$. In particular, $B(x, \epsilon) \cap A = \emptyset$ and $x \notin bdy(A)$. Thus $bdy(A) \subseteq A$.

Conversely, assume that $bdy(A) \subseteq A$. Let $x \in A^c$, then $x \notin bdy(A)$ and since $x \in A^c$, this means that there exists $\epsilon > 0$ such that $B(x, \epsilon) \subseteq A^c$. This shows the A^c is open.

2) We first show that $bdy(A) \subseteq \bar{A}$. Suppose that $x \notin \bar{A}$ Since $\bar{A}$ is closed there exists an $\epsilon > 0$ such that $B(x, \epsilon) \subseteq (\bar{A})^c$. In particular, $B(x, \epsilon) \cap A = \emptyset$ and $x \notin bdy(A)$. Thus $bdy(A) \subseteq \bar{A}$ and as such $A \cup bdy(A) \subseteq \bar{A}$.

Next let $F = A \cup bdy(A)$. We will show that F is closed. To see why this is so let $x \in F^c$. Then $x \notin bdy(A)$ and since $x \notin A$, this means that there exists an $\epsilon > 0$ such that $B(x, \epsilon) \cap A = \emptyset$. Suppose the $z \in bdy(A) \cap B(x, \epsilon)$. Then since $B(x, \epsilon)$ is a neighbourhood of z by definition of a boundary point we must also have $B(x, \epsilon) \cap A \neq \emptyset$ which is a contradiction. It follows that $B(x, \epsilon) \subseteq F^c$ and hence that F is closed.

Since $F = A \cup bdy(A)$ is closed set containing A it follows that $\bar{A} \subseteq A \cup bdy(A)$ completing the proof. ∎

DEFINITION **2.2.3.** *Let $A \subseteq X$. We say that x is a limit point of A if every neighbourhood N of x is such that $N \cap (A \setminus \{x\}) \neq \emptyset$ or equivalently if for every $\epsilon > 0$ the set $B(x, \epsilon)$ contains at least one point in A different from x.*

We denote the set of limit points of A by $Lim(A)$.

Note: *Limit points are also often called cluster points.*

We have the following analogue of our previous result concerning boundary points. The proof is virtually identical.

PROPOSITION **2.2.4.** *Let $A \subseteq (X, d)$.*

1) *A is closed if and only if $Lim(A) \subseteq A$.*
2) *$\bar{A} = A \cup Lim(A)$.*

Proof. 1) First assume that A is closed. Then A^c is open. Let $x \in A^c$. Then there exists an $\epsilon > 0$ such that $B(x, \epsilon) \subseteq A^c$. In particular, $B(x, \epsilon) \cap A = \emptyset$ and $x \notin Lim(A)$. Thus $Lim(A) \subseteq A$.

Conversely, assume that $Lim(A) \subseteq A$. Let $x \in A^c$, then $x \notin Lim(A)$ and since $x \in A^c$, this means that there exists $\epsilon > 0$ such that $B(x, \epsilon) \subseteq A^c$. This shows the A^c is open.

2) We first show that $Lim(A) \subseteq \bar{A}$. Suppose that $x \notin \bar{A}$ Since $\bar{A}$ is closed there exists an $\epsilon > 0$ such that $B(x, \epsilon) \subseteq (\bar{A})^c$. In particular, $B(x, \epsilon) \cap A = \emptyset$ and hence $x \notin Lim(A)$. Thus $Lim(A) \subseteq \bar{A}$ and as such $A \cup Lim(A) \subseteq \bar{A}$.

Next let $F = A \cup Lim(A)$. We will show that F is closed. To see why this is so let $x \in F^c$. Then $x \notin Lim(A)$ and since $x \notin A$, this means that there exists an $\epsilon > 0$ such that $B(x, \epsilon) \cap A = \emptyset$. Suppose the $z \in Lim(A) \cap B(x, \epsilon)$. Then since $B(x, \epsilon)$ is a neighbourhood of z by definition of a limit point we must also have $B(x, \epsilon) \cap (A \setminus \{z\}) \neq \emptyset$ which is a contradiction. It follows that $B(x, \epsilon) \subseteq F^c$ and hence that F is closed.

Since $F = A \cup Lim(A)$ is a closed set containing A it follows that $\bar{A} \subseteq A \cup Lim(A)$ completing the proof.

∎

The proof of the following proposition is straight forward and is left to the reader.

PROPOSITION **2.2.5.** *Let $A \subseteq B \subseteq X$.*

1) *$\bar{A} \subseteq \bar{B}$.*
2) *$int(A) \subseteq int(B)$.*
3) *$(\bar{A})^c = int(A^c)$.*
4) *$int(A) = A \setminus bdy(A)$.*

PROPOSITION **2.2.6.** *Let $A, B \subseteq X$.*

1) *$\overline{A \cup B} = \bar{A} \cup \bar{B}$.*
2) *$int(A \cap B) = int(A) \cap int(B)$.*

Proof. 1) Since $A \cup B \subseteq \bar{A} \cup \bar{B}$ and $\bar{A} \cup \bar{B}$ is closed, it is clear that $\overline{A \cup B} \subseteq \bar{A} \cup \bar{B}$. Similarly, since $A \subseteq \overline{A \cup B}$ and $B \subseteq \overline{A \cup B}$, we get $\bar{A} \subseteq \overline{A \cup B}$ and $\bar{B} \subseteq \overline{A \cup B}$.

2) Since $int(A) \cap int(B) \subseteq A \cap B$ and $int(A) \cap int(B)$ is open we have $int(A) \cap int(B) \subseteq int(A \cap B)$. Similarly, since $int(A \cap B) \subseteq A$ and $int(A \cap B) \subseteq B$, we get $int(A \cap B) \subseteq int(A)$ and $int(A \cap B) \subseteq int(B)$.

■

PROBLEM **9.** *Let (X, d) be any metric space. Let $x \in X$ and $\epsilon > 0$. Then we know that $B(x, \epsilon) \subseteq B[x, \epsilon]$, and that $B[x, \epsilon]$ is a closed set. Is $B[x, \epsilon] = \overline{B(x, \epsilon)}$?*

DEFINITION **2.2.7.** *Given a set $A \subset X$, we say that A is dense in X if $\overline{A} = X$.*

We say that a metric space (X, d) is separable if X has a dense subset A which is countable. Otherwise we say that (X, d) is non-separable.

EXAMPLE **2.10.** *1) $\mathbb{R}$ is separable since $\mathbb{Q}$ is countable and dense.*

2) $\mathbb{R}^n$ is separable since $\mathbb{Q}^n$ is countable and dense.

3) l_1 is separable. (Exercise).

4) l_∞ is nonseparable. (Exercise).

PROBLEM **10.** *Is $(C[a, b], \|\cdot\|_\infty)$ separable?*

2.3 Convergence of Sequences and Topology in a Metric Space

.

DEFINITION **2.3.1.** *Let (X, d) be a metric space. Let $\{x_n\} \subseteq X$ be a sequence in X. We say that the sequence converges to a point x_0 in X if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n \geq N$, then $d(x_n, x_0) < \epsilon$. In this case, we call x_0 the limit of the sequence $\{x_n\}$ and write*

$$x_0 = \lim_{n \to \infty} x_n.$$

Equivalently, we have that $x_0 = \lim\limits_{n \to \infty} x_n$ if and only if $0 = \lim\limits_{n \to \infty} d(x_n, x_0)$ when $\{d(x_n, x_0)\}$ is viewed as a sequence of real numbers.

We will often use the notation $x_n \to x_0$ to mean $x_0 = \lim\limits_{n \to \infty} x_n$.

We say that a sequence is convergent if it has a limit. Otherwise we say it is divergent.

REMARK **2.3.2.** *In the previous definition the language we use suggests that a limit if it exists must be unigue. In the case of the real line, this was a consequence of the Triangle Inequality. We will see next that this fact also carries over to an abstract metric space.*

PROPOSITION **2.3.3.** ***[Uniqueness of Limits]*** *Let (X, d) be a metric space. Let $\{x_n\} \subseteq X$ be a sequence in X. Assume that*

$$x_0 = \lim_{n \to \infty} x_n = y_0.$$

Then $x_0 = y_0$.

Proof. Assume that $x_0 = \lim\limits_{n \to \infty} x_n = y_0$. Assume also that $x_0 \neq y_0$ with $d(x_0, y_0) = \epsilon > 0$.

Since $x_0 = \lim_{n\to\infty} x_n = y_0$, we can find an $N_0 \in \mathbb{N}$ such that if $n \geq N_0$, then $d(x_n, x_0) < \frac{\epsilon}{2}$ and $d(x_n, y_0) < \frac{\epsilon}{2}$. From this we deduce using the Triangle Inequality that

$$d(x_0, y_0) \leq d(x_{N_0}, x_0) + d(x_{N_0}, y_0) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

which is clearly impossible. ∎

We have seen that a convergent sequence can have only one limit. However, it is possible for a sequence to contain subsequences convergeing to different values. For example the real sequence $\{x_n\} = \{1, -1, 1, -1, 1, -1, \ldots\} = \{(-1)^{n+1}\}$ is such that $x_{2k-1} \to 1$ and $x_{2k} \to -1$.

Definition **2.3.4.** *We say that a point x_0 is a limit point of the sequence $\{x_n\}$ if there is a subsequence $\{x_{n_k}\}$ of $\{x_n\}$ with $x_{n_k} \to x_0$.*

Note: *The as the example below shows set of limit points of the sequence $\{x_n\}$ can be different than the limit point of the collection of elements in the sequence viewed simply as a subset of the metric space (X, d). For this reason we use the notation $\lim^*(\{x_n\})$ to denote the set of limit points of $\{x_n\}$.*

In general $Lim(\{x_n\}) \subseteq \lim^*(\{x_n\})$. The proof of this claim is left as an exercise.

Example **2.11.** *Let $x_n = 1$ for each $n \in \mathbb{N}$. Then the constant sequence $\{x_n\} = \{1, 1, 1, \ldots\}$ clearly converges to 1. As such $\lim^*(\{x_n\}) = \{1\}$. However, when the elements of the sequence are viewed as a subset of $\mathbb{R}$ all we have is the singleton $\{1\}$. And in this case $Lim(\{1\}) = \emptyset$.*

We can now state a very important result which essentially shows that the topology of a metric space is determined by its convergent sequences.

Proposition **2.3.5.** *Let $A \subseteq X$.*

1) $x_0 \in Lim(A)$ if and only if there exists a sequence $\{x_n\} \subseteq A$ with $x_n \neq x_0$ and $x_n \to x_0$.

2) $x_0 \in bdy(A)$ if and only if there exists sequences $\{x_n\} \subseteq A$ and $\{y_n\} \subseteq A^c$ with $x_n \to x_0$ and $y_n \to x_0$.

3) A set $A \subseteq X$ is closed if and only if

(*) whenever $\{x_n\}$ is a sequence in A with $x_n \to x_0$, we must have $x_0 \in A$.

Proof. 1) Assume that $x_0 \in Lim(A)$. For each $n \in \mathbb{N}$, there exists an $x_n \in A \cap (B(x_o, \frac{1}{n}) \setminus \{x_0\})$. It follows that the sequence $\{x_n\}$ is such that $\{x_n\} \subseteq A$ with $x_n \neq x_0$ and $x_n \to x_0$.

Assume that there exists a sequence $\{x_n\} \subseteq A$ with $x_n \neq x_0$ and $x_n \to x_0$. Let $\epsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $x_n \in B(x_0, \epsilon)$ for every $n \geq N$. But then $x_N \in B(x_o, \epsilon) \setminus \{x_0\}$ which shows that $x_0 \in Lim(A)$.

2) Assume that $x_0 \in bdy(A)$. For each $n \in \mathbb{N}$, there exists an $x_n \in A \cap (B(x_o, \frac{1}{n})$ and a $y_n \in A^c \cap (B(x_o, \frac{1}{n})$. It follows that there exists sequences $\{x_n\} \subseteq A$ and $\{y_n\} \subseteq A^c$ with $x_n \to x_0$ and $y_n \to x_0$.

Suppose that there exists sequences $\{x_n\} \subseteq A$ and $\{y_n\} \subseteq A^c$ with $x_n \to x_0$ and $y_n \to x_0$. Let $\epsilon > 0$. Then there exists an $N \in \mathbb{N}$ such that $x_n \in B(x_0, \epsilon)$ and $y_n \in B(x_0, \epsilon)$ for every $n \geq N$. But then $x_N, y_N \in B(x_o, \epsilon)$. This shows that $x_0 \in bdy(A)$).

3) Assume that A is closed and $\{x_n\}$ is a sequence in A with $x_n \to x_0$. Assume also that $x_0 \in A^c$. Since A^c is open we can find an $\epsilon > 0$ such that $B(x_0, \epsilon) \subseteq A^c$. But this is impossible since $x_n \in B(x_0, \epsilon)$ for large enough $n's$. Hence $x_0 \in A$.

Conversely, assume that A is not closed. Then there exists a point $x_0 \in bdy(A) \setminus A$. By 2) above there is a sequence $\{x_n\} \subseteq A$ with $x_n \to x_0$. This shows that $(*)$ fails. Consequently, if $(*)$ holds, A must be closed.

∎

We close this section with a brief remark about convergent sequences in a discrete metric space.

REMARK **2.3.6.** *Let (X, d) be any set with the discrete metric. Assume that $x_n \to x_0$. Then there exists an $N \in \mathbb{N}$ such that if $n \geq N$, then $x_n \in B(x_0, 1/2)$. But $B(x_0, 1/2) = \{x_0\}$ and hence $x_n = x_0$ for all $n \geq N$. That is, the only convergent sequences are those that are eventually constant.*

2.4 Induced Metric and the Relative Topology

DEFINITION **2.4.1.** *Let (X, d) be a metric space and let $A \subseteq X$. We can define a function d_A on $A \times A$ as follows:*

$$d_A(x, y) \stackrel{def}{=} d(x, y)$$

for every $x, y \in A$. It is easy to see that d_A is a metric on A which we call the induced metric.

If we let

$$\tau_A = \{W \subset A \mid W = U \cap A \text{ for some open set } U \subseteq X\}$$

then it is easy to see that τ is a topology on A, which we call the relative topology on A inherited from τ_d on X. .

We will now show that the relative topology τ_A is the natural topology obtained from the induced metric d_A.

THEOREM **2.4.2.** *Let (X, d) be a metric space and let $A \subseteq X$. Let τ_A and τ_{d_A} be the relative topology and the metric topology on A respectively. Then*

$$\tau_A = \tau_{d_A}.$$

Proof. Let $W \subset A$ be in τ_A and let $x \in W$. We know there exists an open set U in X so that $W = A \cap U$. But since $x \in U$ we can find an $\epsilon > 0$ such that

$$B_d(x, \epsilon) = \{y \in X \mid d(x, y) < \epsilon\} \subset U.$$

But then

$$B_{d_A}(x, \epsilon) = \{y \in A \mid d(x, y) < \epsilon\} \subset W$$

which shows that $W \in \tau_{d_A}$.

Now let $W \subset A$ be in τ_{d_A}. Then for each $x \in W$ we can find an $\epsilon_x > 0$ so that

$$W = \bigcup_{x \in W} B_{d_A}(x, \epsilon_x).$$

But then if

$$U = \bigcup_{x \in W} B_d(x, \epsilon_x),$$

we have U is open in X and $W = U \cap A$. Hence $W \in \tau_A$.

∎

2.5 Continuity

DEFINITION **2.5.1.** *Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$.*

We say that $f(x)$ is continuous at $x_0 \in X$ if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $x \in X$ with $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \epsilon$.

Otherwise, we say that $f(x)$ is discontinuous at x_0.

We say that $f(x)$ is continuous if it is continuous at each $x \in X$.

REMARK **2.5.2.** *Recall that every function $f : X \to Y$ induces a function $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ given by*

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$$

for each $B \subseteq Y$. The map f^{-1} is called the pullback of f.

The three theorems establish continuity of a function in terms of the nature of the pull back and in terms of the way the function acts on convergent sequences.

THEOREM **2.5.3.** *Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. Then the following are equivalent:*

1) $f(x)$ is continuous at x_0.

2) If W is a neighborhood of $y_0 = f(x_0)$, then $V = f^{-1}(W)$ is a neighborhood of x_0.

Proof. 1) $\Rightarrow$ 2) Assume that $f(x)$ is continuous at x_0. Since W is a neighbourhood of y_0, there exists an $\epsilon_0 > 0$ such that $B(y_0, \epsilon_0) \subseteq W$. Now by definition of continuity at x_0, there exists a $\delta_0 > 0$ such that if $x \in B(x_0, \delta_0)$, then $f(x) \in B(y_0, \epsilon_0) \subseteq W$. Hence $B(x_0, \delta_0) \subseteq V$.

2) $\Rightarrow$ 1) Assume that $\epsilon > 0$. Then $W = B(y_0, \epsilon)$ is a neigbourhood of y_0. Let $V = f^{-1}(W)$. Then V is a neighborhood of x_0. In particular, There is a $\delta > 0$ such that $B(x_0, \delta) \subseteq V$. This means that if $x \in X$ with $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \epsilon$. ∎

THEOREM **2.5.4.** ***[Sequential Characterization of Continuity]:*** *Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. Then the following are equivalent:*

1) $f(x)$ is continuous at x_0.

2) If $\{x_n\}$ is a sequence in X with $x_n \to x_0$, then $f(x_n) \to f(x_0)$.

Proof. 1) $\Rightarrow$ 2) Assume that $f(x)$ is continuous at x_0 and that $x_n \to x_0$. Let$\epsilon_0 > 0$ and let $\delta > 0$ be such that if $x \in X$ with $d_X(x, x_0) < \delta$, then $d_Y(f(x), f(x_0)) < \epsilon$. Since $x_n \to x_0$ there exists an $N \in \mathbb{N}$ such that if $n \geq N$, we have $d_X(x_n, x_0) < \delta$. It folows that if $n \geq N$, we have $d_Y(f(x_n), f(x_0)) < \epsilon$. That is, $f(x_n) \to f(x_0)$.

2) $\Rightarrow$ 1) Assume that $f(x)$ is not continuous at x_0. Then there exists an $\epsilon_0 > 0$ such that in every ball $B(x_0, \delta)$ there is a point x_δ with $d_Y(f(x_\delta), f(x_0)) \geq \epsilon_0$. In particular, for each $n \in \mathbb{N}$, there exists an x_n with $d_X(x_n, x_0) < \frac{1}{n}$ but $d_Y(f(x_n), f(x_0)) \geq \epsilon_0$. It follows that $x_n \to x_0$ but $f(x_n) \not\to f(x_0)$. Consequently, 2) fails.

∎

THEOREM **2.5.5.** *Let (X, d_X) and (Y, d_Y) be metric spaces and $f : X \to Y$. Then the following are equivalent:*

1) $f(x)$ is continuous.

2) If W is an open set in Y, then $V = f^{-1}(W)$ is open in X.

3) If $\{x_n\}$ is a sequence in X with $x_n \to x_0$ where $x_0 \in X$, then $f(x_n) \to f(x_0)$ in Y .

Proof. 1) $\Rightarrow$ 2) Assume that $f(x)$ is continuous. Let W be open in Y and let $V = f^{-1}(W)$. Assume that $x_0 \in V$. Then W is an open set containing $y_0 = f(x_0)$. It then follows that V is a neighbourhood of x_0. That is $x_0 \in int(V)$. But since $V = int(V)$, V is open in X.

2) $\Rightarrow$ 3) Assume that $x_n \to x_0$. Let $y_0 = f(x_0)$ and let $\epsilon > 0$. If $W = B(y_0, \epsilon)$, then since W is open so is $V = f^{-1}(W)$. But $x_0 \in V$ so there exists a $\delta > 0$ for which $B(x_0, \delta) \subseteq V$. And since Since $x_n \to x_0$ there exists an $N \in \mathbb{N}$ such that if $n \geq N$, we have $d_X(x_n, x_0) < \delta$. It folows that if $n \geq N$, we have

$d_Y(f(x_n), f(x_0)) < \epsilon$. That is, $f(x_n) \to f(x_0)$.

3) $\Rightarrow$ 1) This follows immediately from the sequential characterization of continuity.

■

EXAMPLE **2.12.** *Let (x, d) be any set together with the discrete topology. Since every subset of (X, d) is open it follows that for any function $f : (X, d) \to (Y, d_Y)$, $f(x)$ is automatcally continuous.*

DEFINITION **2.5.6.** *Let (X, d_X) and (Y, d_Y) be two metric spaces. A function $\phi : X \to Y$ is said to be a homeomorphism if ϕ is 1-1 and onto and if both ϕ and ϕ^{-1} are continuous. We then say that (X, d_X) and (Y, d_Y) are homeomorphic.*

Two metric spaces (X, d_X) and (Y, d_Y) are said to be equivalent if there exists a 1-1 onto map $\phi : X \to Y$ and two constants $c_1, c_2 > 0$ such that

$$c_1 d_X(x_1, x_2) \leq d_Y(\phi(x_1), \phi(x_2)) \leq c_2 d_X(x_1, x_2)$$

for each $x_1, x_2 \in X$

REMARK **2.5.7.**

1) *If $\phi : X \to Y$ is a homeomorphism, then $\phi(U)$ is open in Y if and only if U is open in X. As such, homeomomorphic spaces can be viewed as identical topologically.*
2) *If (X, d_X) and (Y, d_Y) are equivalent, then they are homoeomorphic. But the converse need not hold. (Exercise)*

DEFINITION **2.5.8.** ***[Continuity on a Set]*** *Given a function $f : X \to Y$ and a subset $A \subseteq X$. The restriction of $f(x)$ to A is the function $f_{|A} : A \to Y$ given by*

$$f_{|A}(x) \stackrel{def}{=} f(x)$$

for all $x \in A$.

Given a function $f : X \to Y$ and a subset $A \subseteq X$, we say that $f(x)$ is continuous on A if the restriction $f_{|A}$ is continuous on the metric space (A, d_A).

REMARK **2.5.9.** *It follows immediately from the sequential characterization of continuity that $f(x)$ is continuous on A if and only if whenever $\{x_n\}$ is a sequence in A with $x_n \to x_0$ for some $x_0 \in A$, then we have $f(x_n) \to f(x_0)$.*

2.6 Complete Metric Spaces: Cauchy Sequences

.

Recall that a sequence $\{x_n\}$ converges to a point x_0 in X if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n \geq N$, then $d(x_n, x_0) < \epsilon$. At this point, if we want to test to see if a sequence converges it seems we need to have a possible limit in mind. This leads to the following question:

PROBLEM **11.** *Is there an intrinsic test to see if a sequnece $\{x_n\}$ converges?*

We beign with the following strategy:

Strategy: Assume that $x_n \to x_0$. Let $\epsilon > 0$ .Then there exists an $N \in \mathbb{N}$ such that if $n \geq N$, then $d(x_n, x_0) < \frac{\epsilon}{2}$. In particular, if $n, m \geq N$, then

$$d(x_n, x_m) \leq d(x_n, x_0) + d(x_0, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

This leads us to the following definition:

Definition **2.6.1.** *A sequence $\{x_n\}$ is said to be Cauchy in (X,d) if for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that if $n, m \geq N$, then $d(x_n, x_m) < \epsilon$.*

Theorem **2.6.2.** *Let $\{x_n\}$ be a convergent sequence in (X,d). Then $\{x_n\}$ is Cauchy.*

This takes us to the fundamental question:

Problem **12.** *Does every Cauchy sequence converge?*

It turns out that the answer to the above question can be no. For example, let $X = (0,1)$ with the usual metric. Then $\{\frac{1}{n}\}$ is Cauchy but does not converge. In contrast this sequence, and any Cauchy sequence for that matter, does converge in $\mathbb{R}$ with the usual metric.

Definition **2.6.3.** *A metric space (X,d) is said to be complete if every Cauchy sequence converges.*

We will show that several of the usual examples of metric spaces we have are complete. In this respect we will begin with $\mathbb{R}$ and the usual metric. However to do so we need to first observe to important properties of Cauchy sequences.

Definition **2.6.4.** *A set $A \subset (X,d)$ is said to be bounded if there exists an $x_0 \in X$ and an $M > 0$ such that $A \subseteq B[x_0, M]$.*

Proposition **2.6.5.** *Every Cauchy sequence is bounded.*

Proof. Let $\{x_n\}$ be Cauchy. There exists an $N \in \mathbb{N}$ such that if $n, m \geq N$, then $d(x_n, x_m) < 1$. In particular, $d(x_N, x_m) < 1$ for all $m \geq N$. Now let

$$M = max\{d(x_1, x_N), d(x_2, x_N), \ldots, d(x_{N-1}, x_N), 1\}.$$

Then $\{x_n\} \subset B[x_N, M]$. ∎

We know that the sequence $\{1, -1, 1, -1, \ldots\}$ has a convergent subsequence but it does not converge. We will now see that for a Cauchy sequence this is not possible.

Proposition **2.6.6.** *Assume that $\{x_n\}$ is a Cauchy sequence with a subsequence $x_{n_k} \to x_0$. Then $x_n \to x_0$.*

Proof. Let $\epsilon > 0$. There exists an $N \in \mathbb{N}$ such that if m, $n \geq N$, then

$$d(x_n, x_m) < \frac{\epsilon}{2}.$$

Since $\lim\limits_{k\to\infty} x_{n_k} = x_0$, we can find a $K \in \mathbb{N}$ so that $n_K \geq N$ and

$$d(x_{n_K}, x_0) < \frac{\epsilon}{2}.$$

But now we have that for all $n \geq N$,

$$d(x_n, x_0) \leq d(x_n, x_{n_K}) + d(x_{n_K}, x_0) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

showing that $\{x_n\}$ converges and $\lim\limits_{n\to\infty} x_n = x_0$. ∎

2.7 Completeness of $\mathbb{R}$, $\mathbb{R}^n$ and l_p

To show that $\mathbb{R}$ with the usual metric is complete we need the following important theorem with its proof left as an excercise:

THEOREM **2.7.1.** *(Bolzano-Weierstrass Theorem)*
Every bounded sequence $\{x_n\} \subset \mathbb{R}$ has a convergent subsequence.

We are now prepared to establish the completeness of $\mathbb{R}$.

THEOREM **2.7.2.** *(Completeness Theorem for $\mathbb{R}$)*
Every Cauchy sequence $\{x_n\} \subset \mathbb{R}$ converges.

Proof. Let $\{x_n\} \subset \mathbb{R}$ be Cauchy. Then we know that $\{x_n\}$ is also bounded. By the Bolzano-Weierstrass Theorem $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$. Form here it follows that $\{x_n\}$ itself converges. ∎

Next we will show that $(\mathbb{R}^n, \|\cdot\|_2)$ is also complete. To do so we need the following useful observation.

PROPOSITION **2.7.3.** *Let $\{\boldsymbol{x}_k\} = \{(x_{k,1}, x_{k,2}, \ldots, x_{k,n})\}$ be a sequence in $(\mathbb{R}^n, \|\cdot\|_2)$. Then $\boldsymbol{x}_k \to \boldsymbol{x}_0 = (x_{0,1}, x_{0,2}, \ldots, x_{o,n})$ if and only if $x_{k,i} \to x_{0,i}$ for each $i = 1, 2, \ldots, n$.*

THEOREM **2.7.4.** *(Completeness Theorem for $(\mathbb{R}^n, \|\cdot\|_2)$)*
Every Cauchy sequence $\{\boldsymbol{x}_k\} \subset (\mathbb{R}^n, \|\cdot\|_2)$ converges.

It is in fact possible to modify our previous argument to show that for any $1 \leq p \leq \infty$ that a sequence $\{\mathbf{x}_k\} \subset (\mathbb{R}^n, \|\cdot\|_p)$ converges in $\|\cdot\|_p$ if and only if the component sequence $\{x_{k,i}\}$ converges for each $i = 1, 2, \ldots, n$. Consequently we have the following:

THEOREM **2.7.5.** *(Completeness Theorem for $(\mathbb{R}^n, \|\cdot\|_p)$)*
Let $1 \leq p \leq \infty$. Every Cauchy sequence $\{\boldsymbol{x}_k\} \subset (\mathbb{R}^n, \|\cdot\|_p)$ converges.

We will now show that our completeness theorems extend to sequence spaces.

LEMMA **2.7.6.** *Let $1 \leq p \leq \infty$. Let $\{\boldsymbol{x}_k\} = \{(x_{k,1}, x_{k,2}, x_{k,3}, \ldots)\}$ be a Cauchy sequence in $(l_p, \|\cdot\|_p)$. Then for each $i \in \mathbb{N}$, the component sequence $\{x_{k,i}\}$ is Cauchy in $\mathbb{R}$.*

Proof. As before, this follows because for each $i \in \mathbb{N}$,

$$\mid x_{k,i} - x_{m,i} \mid \leq \|\mathbf{x}_k - \mathbf{x}_m\|_p.$$

∎

THEOREM **2.7.7.** *(Completeness Theorem for $(l_p, \|\cdot\|_p)$)*
Let $1 \leq p \leq \infty$. Every Cauchy sequence $\{\boldsymbol{x}_k\} \subset (l_p, \|\cdot\|_p)$ converges.

Proof. **Case 1)** $(l_\infty, \|\cdot\|_\infty)$
Assume that $\{\mathbf{x}_k\} \subset (l_\infty, \|\cdot\|_\infty)$ is Cauchy. Since each component sequence $\{x_{k,i}\}$ is also Cauchy we can define

$$x_{0,i} = \lim_{k\to\infty} x_{k,i}.$$

We claim that the sequence $\mathbf{x}_0 = \{x_{0,i}\} \in l_\infty$ and that $\mathbf{x}_k \to \mathbf{x}_0$.

To do so first fix an $\epsilon > 0$. Then choose an N_0 such that if $k, m > N_0$ we have

$$\|\mathbf{x}_k - \mathbf{x}_m\|_\infty < \frac{\epsilon}{2}.$$

Let $k \geq N_0$. Then for each $i \in \mathbb{N}$, we have

$$\mid x_{k,i} - x_{m,i} \mid \leq \|\mathbf{x}_k - \mathbf{x}_m\|_\infty < \frac{\epsilon}{2}$$

for all $m \geq N_0$. It follows that for each $i \in \mathbb{N}$,

$$\mid x_{k,i} - x_{0,i} \mid = \lim_{m\to\infty} \mid x_{k,i} - x_{m,i} \mid \leq \frac{\epsilon}{2} < \epsilon \qquad (*)$$

It follows that the sequence $\{x_{k,i} - x_{0,i}\}_{i=1}^{\infty}$ is in l_∞ and hence that $\mathbf{x}_0 = \{x_{0,i}\}$ is also in l_∞. Moreover, $(*)$ also shows that if $k \geq N_0$, then

$$\|\mathbf{x}_k - \mathbf{x}_0\|_\infty \leq \frac{\epsilon}{2} < \epsilon.$$

This shows that $\mathbf{x}_k \to \mathbf{x}_0$.

Case 2) $(l_1, \|\cdot\|_1)$

Assume that $\{\mathbf{x}_k\} \subset (l_1, \|\cdot\|_1)$ is Cauchy. Since each component sequence $\{x_{k,i}\}$ is also Cauchy we can define

$$x_{0,i} = \lim_{k\to\infty} x_{k,i}.$$

We claim that the sequence $\mathbf{x}_0 = \{x_{0,i}\} \in l_1$ and that $\mathbf{x}_k \to \mathbf{x}_0$.

To do so first fix an $\epsilon > 0$. Then choose an N_0 such that if $k, m > N_0$ we have

$$\|\mathbf{x}_k - \mathbf{x}_m\|_1 < \frac{\epsilon}{2}.$$

Let $k \geq N_0$. Then for each $j \in \mathbb{N}$, we have

$$\sum_{i=1}^{j} \mid x_{k,i} - x_{m,i} \mid \leq \|\mathbf{x}_k - \mathbf{x}_m\|_\infty < \frac{\epsilon}{2}$$

for all $m \geq N_0$. It follows that for each $j \in \mathbb{N}$,

$$\sum_{i=1}^{j} \mid x_{k,i} - x_{0,i} \mid = \lim_{m\to\infty} \sum_{i=1}^{j} \mid x_{k,i} - x_{m,i} \mid \leq \frac{\epsilon}{2} < \epsilon \qquad (*)$$

Since $j \in \mathbb{N}$ was arbitrary, we get

$$\sum_{i=1}^{\infty} \mid x_{k,i} - x_{0,i} \mid \leq \frac{\epsilon}{2} < \epsilon \qquad (**)$$

It follows that the sequence $\{x_{k,i} - x_{0,i}\}_{i=1}^{\infty}$ is in l_1 and hence that $\mathbf{x}_0 = \{x_{0,1}\}$ is also in l_1. Moreover, $(**)$ also shows that if $k \geq N_0$, then

$$\|\mathbf{x}_k - \mathbf{x}_0\|_1 \leq \frac{\epsilon}{2} < \epsilon.$$

This shows that $\mathbf{x}_k \to \mathbf{x}_0$.

Case 3) $(l_p, \|\cdot\|_p)$ where $1 < p < \infty$

Assume that $\{\mathbf{x}_k\} \subset (l_p, \|\cdot\|_p)$ is Cauchy. Since each component sequence $\{x_{k,i}\}$ is also Cauchy we can define

$$x_{0,i} = \lim_{k\to\infty} x_{k,i}.$$

We claim that the sequence $\mathbf{x}_0 = \{x_{0,i}\} \in l_p$ and that $\mathbf{x}_k \to \mathbf{x}_0$.

To do so first fix an $\epsilon > 0$. Then choose an N_0 such that if $k, m > N_0$ we have

$$\|\mathbf{x}_k - \mathbf{x}_m\|_p < \frac{\epsilon}{2}.$$

Let $k \geq N_0$. Then for each $j \in \mathbb{N}$, we have

$$(\sum_{i=1}^{j} \mid x_{k,i} - x_{m,i} \mid^p)^{\frac{1}{p}} \leq \|\mathbf{x}_k - \mathbf{x}_m\|_\infty < \frac{\epsilon}{2}$$

for all $m \geq N_0$. It follows that for each $j \in \mathbb{N}$,

$$(\sum_{i=1}^{j} \mid x_{k,i} - x_{0,i} \mid^p)^{\frac{1}{p}} = \lim_{m\to\infty} (\sum_{i=1}^{j} \mid x_{k,i} - x_{m,i} \mid^p)^{\frac{1}{p}} \leq \frac{\epsilon}{2} < \epsilon \qquad (*)$$

Since $j \in \mathbb{N}$ was arbitrary, we get

$$(\sum_{i=1}^{\infty} \mid x_{k,i} - x_{0,i} \mid^p)^{\frac{1}{p}} \leq \frac{\epsilon}{2} < \epsilon \qquad (**)$$

It follows that the sequence $\{x_{k,i} - x_{0,i}\}_{i=1}^{\infty}$ is in l_p and hence that $\mathbf{x}_0 = \{x_{0,1}\}$ is also in l_p. Moreover, $(**)$ also shows that if $k \geq N_0$, then

$$\|\mathbf{x}_k - \mathbf{x}_0\|_p \leq \frac{\epsilon}{2} < \epsilon.$$

This shows that $\mathbf{x}_k \to \mathbf{x}_0$.

∎

All of the examples in this section are normed linear spaces which are complete in the induced metric. This leads us to the following definition.

Definition **2.7.8.** *A normed linear space $(X, \|\cdot\|)$ which is complete under the metric space induced by the norm is called a Banach space.*

2.8 Completeness of $(C_b(X), \|\cdot\|_\infty)$

.

Definition **2.8.1.** *Let (X, d_X), (Y, d_Y) be metric spaces. Let $\{f_n\}$ be a sequnce of functions from X to Y. We say that a the sequence $\{f_n\}$ converges pointwise on X to $f_0(x)$ if*

$$\lim_{n\to\infty} f_n(x_0) = f_0(x_0)$$

for every $x_0 \in X$.

We say thay $\{f_n\}$ converges to $f(x)$ uniformly on X if for every $\epsilon > 0$ there exists an $N_0 \in \mathbb{N}$ such that if $n \geq N_0$, then $d_Y(f_n(x), f_0(x)) < \epsilon$ for all $x \in X$.

Remark **2.8.2.** *It should be clear that if $f_n \to f_0$ uniformly, then $f_n \to f_0$ pointwise but the converse does not necessarily hold. For example if $f_n(x) = x^n$ on $[0, 1]$, then $f_n \to f_0$ pointwise where*

$$f_0(x) := \begin{cases} 0 & \text{if } x \in [0, 1), \\ 1 & \text{if } x = 1 \end{cases}$$

However, each f_n is continuous on $[0, 1]$ but f_0 is not. As we will soon see the convergence is not uniform.

The next theorem shows that unlike pointwsie convergence, uniform convergence preserves continuity. The proof uses an important technique known as a *three ϵ argument*.

THEOREM **2.8.3.** *Let (X, d_X), (Y, d_Y) be metric spaces. Let $\{f_n : X \to Y\}$ be a sequence of functions that converges uniformly on X to f_0. If each $f_n(x)$ is continuous at x_0, then so is $f_0(x)$. In particular, if each f_n is continuous, so is f_0.*

Proof. Assme that $f_n \to f_0$ uniformly and that each $f_n(x)$ is continuous at x_0. Let $\epsilon > o$. By uniform convergence, we can find an $N_0 \in \mathbb{N}$ such that if $n \geq N_0$, then $d_Y(f_n(x), f_0(x)) < \frac{\epsilon}{3}$ for all $x \in X$. Since f_{N_0} is continuous at x_0, there exists a $\delta > 0$ such that if $x \in B(x_0, \delta)$, then $d_Y(f_{N_0}(x), f_{N_0}(x_0)) < \frac{\epsilon}{3}$.So assume that $x \in B(x_0, \delta)$. Then

$$\begin{aligned} d_Y(f_0(x), f_0(x_0)) &\leq d_Y(f_0(x), f_{N_0}(x)) + d_Y(f_{N_0}(x), f_{N_0}(x_0)) + d_Y(f_{N_0}(x_0), f_0(x_0)) \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon \end{aligned}$$

∎

DEFINITION **2.8.4.** *Let (X, d) be a metric space. Let*

$$C_b(X) = \{f : X \to \mathbb{R} \mid f(x) \text{is continuous on } X \text{ and } f(X) \text{ is bounded in } \mathbb{R}\}.$$

Define the norm $\|\cdot\|_\infty$ on $C_b(X)$ by

$$\|f\|_\infty = sup\{\mid f(x) \mid \mid x \in X\}.$$

Then $(C_b(X), \|\cdot\|_\infty)$ is a normed linear space.

REMARK **2.8.5.** *Observe that it follows immediately from the definition of uniform convergence that $f_n \to f_0$ in $(C_b(X), \|\cdot\|_\infty)$ if and only if $f_n \to f_0$ uniformly on X.*

We are now in a position to show that $(C_b(X), \|\cdot\|_\infty)$ is complete.

THEOREM **2.8.6.** *[Completeness Theorem for $(C_b(X), \|\cdot\|_\infty)$]*
$(C_b(X), \|\cdot\|_\infty)$ is complete.

Proof. Assume that $\{f_n\}$ is Cauchy in $(C_b(X), \|\cdot\|_\infty)$. Let $x_0 \in X$. Since

$$\mid f_n(x_0) - f_m(x_0) \mid \leq \|f_n - f_m\|_\infty,$$

it follows that $\{f_n(x_0)\}$ is also Cauchy in $\mathbb{R}$. Now define

$$f_0(x) \stackrel{def}{=} \lim_{n \to \infty} f_n(x)$$

for each $x \in X$.

Claim: $f_n \to f_0$ uniformly on X.

To see why this is so let $e > 0$. Choose an N_0 so that if $n, m \geq N_0$, then $\|f_n - f_m\|_\infty < \frac{\epsilon}{2}$. Now let $n \geq N_0$ and $x \in X$. Then

$$\begin{aligned} \mid f_n(x) - f_0(x) \mid &= \lim_{n\to\infty} \mid f_n(x) - f_m(x) \mid \\ &\leq \frac{\epsilon}{2} \\ &< \epsilon \end{aligned}$$

and $f_n \to f_0$ uniformly.

The Since each f_n is continuous, it follows that f_0 is also continuous on X. To complete the proof we need only show that f_0 is also bounded. However, since f_n is Cauchy, it is bounded. As such there exists an M so that $\|f_n\|_\infty \leq M$ for each $n \in \mathbb{N}$. By uniform convergence, we can find an n_0 such that if $x \in X$, then $\mid f_0(x) - f_n(x) \mid \leq 1$ for all $x \in X$. Hence

$$\mid f_0(x) \mid \leq \mid f_0(x) - f_n(x) \mid + \mid f_n(x) \mid \leq 1 + M$$

for all $x \in X$. Consequently $f_0 \in C_b(X)$ and $f_n \to f_0$.

∎

Remark **2.8.7.** *Observe that if we let $X = \mathbb{N}$ and give $\mathbb{N}$ the discrete metric d, then in this case $(C_b(\mathbb{N}), \|\cdot\|_\infty)$ is exactly $(l_\infty, \|\cdot\|_\infty)$. In particular, or previous result is a generalization of the proof that $(l_\infty, \|\cdot\|_\infty)$ is complete.*

Following what we did above, given any set X we can give X the discrete topology. In this case, we define

$$(l_\infty(X), \|\cdot\|_\infty) \stackrel{def}{=} (C_b(X), \|\cdot\|_\infty).$$

Problem **13.** *In the previous remark, we saw how we could define the analog of $(l_\infty(X), \|\cdot\|_\infty)$ of $(l_\infty, \|\cdot\|_\infty)$ for any set X. Is there a way to define the analogue $(l_p(X), \|\cdot\|_p)$ of $(l_p, \|\cdot\|_p)$ for any $1 \leq p < \infty$?*

2.9 Characterizations of Complete Metric Spaces

.

In this section we will give several useful characterizations of completeness for metric and normed linear spaces.

Remark **2.9.1.** *Recall that the Nested Interval Theorem states the following: Suppose that $\{[a_n, b_n]\}$ is a sequence of closed intervals in $\mathbb{R}$ with $[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n]$, then $\bigcap\limits_{n=1}^{\infty} [a_n, b_n] \neq \emptyset$.*

The simplest proof of the Nested Interval Theoerm uses the Monotone Convergence Theorem, which in turn follows from the Least Upper Bound Property. In fact, it turns out that all three statements are logically equivalent and are actually different variants of the Completeness Property for $\mathbb{R}$. As such it makes sense to ask if there is an analog of the Nested Interval Theorem for general metric spaces that serves as a characterization of completeness.

One might ask if the following may be true:

Conjecture: *A metric space is complete if and only if whenever $\{F_n\}$ is a sequence of non-empty closed subsets of X with $F_{n+1} \subseteq F_n$, then $\bigcap\limits_{n=1}^{\infty} F_n \neq \emptyset$.*

Unfortunately, this result is false even for $\mathbb{R}$. In particular, if $F_n = [n, \infty)$, then F_n is a closed interval but $\bigcap\limits_{n=1}^{\infty} F_n = \emptyset$. However, it does turn out that an appropriate analogue to the Nested Interval Theorem does in fact exist. To formulate this result we first need the notion of the diameter of a set.

Definition **2.9.2.** *Let $A \subseteq (X,d)$. We let*

$$diam(A) \stackrel{def}{=} sup\{d(x,y) \mid x,y \in A\}.$$

In this case, $diam(A)$ is called the diameter of A.

The following prroposition will be needed later on.

Proposition **2.9.3.** *Let $A \subseteq B \subseteq X$. Then*

1) $diam(A) \leq diam(B)$

2) $diam(\overline{A}) = diam(A)$

Proof. Statement 1) is obvious so we will only prove 2).

We know from 1) that $diam(A) \leq diam(\overline{A})$. From this it follows that the statement is true if $diam(A) = \infty$. Hence we may assume that $diam(A) = d < \infty$.

Let $x, y \in \overline{A}$ and let $\epsilon > 0$. We can find $x_0, y_0 \in A$ so that $d(x,x_0) \leq \frac{\epsilon}{2}$ and $d(y,y_0) \leq \frac{\epsilon}{2}$. It follows that

$$d(x,y) \leq d(x,x_0) + d(x_0,y_0) + d(y_0,y) \leq \frac{\epsilon}{2} + d + \frac{\epsilon}{2} = d + \epsilon.$$

From this it follows that $diam(A) \leq diam(\overline{A}) \leq diam(A) + \epsilon$. But ϵ was arbitrary so indeed $diam(\overline{A}) = diam(A)$.

∎

We are now in a position to state and prove our analogue of the Netsed Interval Theorem.

Theorem **2.9.4.** *[Cantor's Intersection Theorem]*

Let (X,d) be a metric space. Then the following are equivalent:

1) (X,d) is complete.

2) (X,d) satisfies the following property:

$(*)$ *If $\{F_n\}$ is a sequence of non-empty closed sets such that $F_{n+1} \subseteq F_n$ for each $n \in \mathbb{N}$ and $\lim\limits_{n\to\infty} diam(F_n) = 0$, then $\bigcap\limits_{n=1}^{\infty} F_n \neq \emptyset$.*

Proof. 1) $\Rightarrow$ 2) Assume that $\{F_n\}$ is as above. For each $n \in \mathbb{N}$, choose an $x_n \in F_n$ and let $\epsilon > 0$.

Since $diam(F_n) \to 0$, we can find an N_0 so that $diam(F_{N_0}) < \epsilon$. If $n, m \geq N_0$, then because $\{F_n\}$ is nested $x_n, x_m \in F_{N_0}$. It follows that $d(x_n, x_m) \leq diam(F_{N_0}) < \epsilon$, and hence that $\{x_n\}$ is Cauchy. Because (X,d) is complete, we have $x_n \to x_0$ for some $x_0 \in X$.

Now for each $n \in \mathbb{N}$, again because $\{F_n\}$ is nested, we have $\{x_n, x_{n+1}, x_{n+2}, \ldots\} \subseteq F_n$. But the sequence $\{x_n, x_{n+1}, x_{n+2}, \ldots\}$ also converges to x_0. Since F_n is closed, this means that $x_0 \in F_n$. Since n was arbitrary, $x_0 \in \bigcap\limits_{n=1}^{\infty} F_n$.

2) $\Rightarrow$ 1) Let $\{x_n\}$ be a Cauchy sequence in X. For each $n \in \mathbb{N}$ let $A_n = \{x_n, x_{n+1}, x_{n+2}, \ldots\}$. The fact that $\{x_n\}$ is Cauchy means that $diam(A_n) \to 0$.

Next, let $F_n = \overline{A_n}$. Then clearly, $F_n \neq \emptyset$, $F_{n+1} \subseteq F_n$ and since $diam(A_n) = diam(F_n)$, we have $diam(F_n) \to 0$. By $(*)$, we can conclude that there exists an $x_0 \in \bigcap\limits_{n=1}^{\infty} F_n$.

Let $\epsilon > 0$. Choose $N_0 \in \mathbb{N}$ so that $diam(F_{N_0}) < \epsilon$. Then $F_{N_0} \subset B(x_0, \epsilon)$. In particular, is $n \geq N_0$, then $d(x_n, x_0) < \epsilon$.

∎

Our next result, which characterises completeness for a normed linear space, is an analogue of the familiar fact that if a series $\sum\limits_{n=1}^{\infty} a_n$ converges absolutely, then it converges.

Definition **2.9.5.** *Let $(X, \|\cdot\|)$ be a normed linear space. Let $\{x_n\} \subset X$. A series with terms $\{x_n\}$ is a formal sum*

$$\sum_{n=1}^{\infty} x_n = x_1 + x_2 + x_3 + \cdots .$$

For each $k \in \mathbb{N}$, we define the k-th partial sum of $\sum\limits_{n=1}^{\infty} x_n$ by

$$S_k = \sum_{n=1}^{k} x_n \in X$$

As we do in $\mathbb{R}$, we say that the series $\sum\limits_{n=1}^{\infty} a_n$ convegres if the sequence $\{S_k\}$ converges. Otherwise we say that the series diverges.

Theorem **2.9.6.** *[Generalized Weierstrass M-Test]*

Let $(X, \|\cdot\|)$ be a normed linear space. Then the following are equivalent:

1) $(X, \|\cdot\|)$ is a Banach space.

2) $(X, \|\cdot\|)$ satisfies the following property:

$(*)$ *Let $\{x_n\}$ be a sequence in $(X, \|\cdot\|)$. If $\sum\limits_{n=1}^{\infty} \|x_n\|$ converges in $\mathbb{R}$, then $\sum\limits_{n=1}^{\infty} x_n$ converges in $(X, \|\cdot\|)$.*

Proof. 1) $\Rightarrow$ 2) $\sum\limits_{n=1}^{\infty} \|x_n\|$ convegres in $\mathbb{R}$. For each $k \in \mathbb{N}$, let $T_k = \sum\limits_{n=1}^{k} \|x_n\|$. Then $\{T_k\}$ is a Cauchy sequence. Hence, given $\epsilon > 0$ we can find an N_0 so that in $k > m > N_0$, then

$$\sum_{n=m+1}^{k} \|x_n\| = \mid T_k - T_m \mid < \epsilon.$$

If we let $S_k = \sum\limits_{n=1}^{k} x_n$ and $k > m > N_0$, then

$$\|S_k - S_m\| = \| \sum_{n=m+1}^{k} x_n\| \le \sum_{n=m+1}^{k} \|x_n\| < \epsilon.$$

This shows that $\{S_k\}$ is Cauchy and hence convergent.

2) $\Rightarrow$ 1) Assume that $(*)$ holds and that $\{x_n\}$ is Cauchy in $(X, \|\cdot\|)$. We can choose an n_1 such that if $i, j \ge n_1$, then $\|x_i - x_j\| < \frac{1}{2}$. Then we can choose $n_2 > n_1$ so that if $i, j \ge n_2$, then $\|x_i - x_j\| < \frac{1}{2^2}$. Next we choose $n_3 > n_2 > n_1$ so that $i, j \ge n_1$, then $\|x_i - x_j\| < \frac{1}{2^3}$. Proceeding inductively we construct a striclty increasing sequence of natural numbers $\{n_k\}$ such that for each $k \in \mathbb{N}$ if $i, j \ge n_k$, then $\|x_i - x_j\| < \frac{1}{2^k}$. In particular, for each $k \in \mathbb{N}$ we have $\|x_{n_k} - x_{n_{k+1}}\| < \frac{1}{2^k}$.

For each $k \in \mathbb{N}$, let $g_k = x_{n_k} - x_{n_{k+1}}$. Then

$$\sum_{k=1}^{\infty} \|g_k\| = \sum_{k=1}^{\infty} \|x_{n_k} - x_{n_{k+1}}\| < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

It now follows from $(*)$ that the sequence $\{S_j\} = \{\sum_{k=1}^{j} x_{n_k} - x_{n_{k+1}}\}$ also converges. But

$$S_j = \sum_{k=1}^{j} x_{n_k} - x_{n_{k+1}} = (x_{n_1} - x_{n_2}) + (x_{n_2} - x_{n_3}) + \cdots + (x_{n_j} - x_{n_{j+1}}) = x_{n_1} - x_{n_{j+1}}$$

as the series telescopes. It follows that

$$x_{n_{j+1}} \overset{j\to\infty}{\to} x_{n_1} - \sum_{k=1}^{\infty} g_k.$$

Finally, because the subsequence $\{x_{n_{j+1}}\}$ converges, so does $\{x_n\}$. ∎

EXAMPLE **2.13.** *[A continuous nowhere differentiable function]*

Let

$$\varphi(x) = \begin{cases} x & \text{if } x \in [0,1] \\ 2-x & \text{if } x \in [1,2] \end{cases} .$$

and then extend φ to all of $\mathbb{R}$ by letting $\varphi(x+2) = \varphi(x)$.

Let

$$f(x) = \sum_{n=0}^{\infty} (\frac{3}{4})^n \varphi(4^n x).$$

Then by the Weierstrass M-Test, $f(x)$ is continuous on $\mathbb{R}$. We claim that $f(x)$ is nowhere diffferentiable.

Fix $x \in \mathbb{R}$. For each $m \in \mathbb{N}$, there exists a $k \in \mathbb{Z}$ such that

$$k \leq 4^m x \leq k+1$$

or equivalently that

$$\frac{k}{4^m} \leq x \leq \frac{k+1}{4^m}.$$

Let

$$p_m = \frac{k}{4^m}, q_m = \frac{k+1}{4^m}$$

For $n \in \mathbb{N}$, let

$$\alpha = 4^n p_m = 4^{n-m} k$$

and

$$\beta = 4^n q_m = 4^{n-m}(k+1).$$

- *If $n > m$, then α and β differ by an even integer so that $\mid \varphi(\alpha) - \varphi(\beta) \mid = 0$.*
- *If $n = m$, then $\alpha = k$ and $\beta = k+1$ so that $\mid \varphi(\alpha) - \varphi(\beta) \mid = 1$.*
- *If $n < m$, then there is no integer strictly between α and β so that $\mid \varphi(\alpha) - \varphi(\beta) \mid = \mid \alpha - \beta \mid = 4^{n-m}$.*

It follows that

$$\begin{aligned}
\mid f(p_m) - f(q_m) \mid &= \mid \sum_{n=0}^{m} (\frac{3}{4})^n [\varphi(4^n p_m) - \varphi(4^n q_m)] \mid \\
&\geq (\frac{3}{4})^m - \sum_{n=0}^{m-1} (\frac{3}{4})^n \mid \varphi(4^n p_m) - \varphi(4^n q_m) \mid \\
&= (\frac{3}{4})^m - \sum_{n=0}^{m-1} (\frac{3}{4})^n 4^{n-m} \\
&= (\frac{3}{4})^m - \frac{1}{4^m} \sum_{n=0}^{m-1} 3^n \\
&= \frac{3^m}{4^m} - \frac{1}{4^m} (\frac{3^m - 1}{3 - 1}) \\
&= \frac{2 \cdot 3^m}{2 \cdot 4^m} - \frac{3^m - 1}{2 \cdot 4^m} \\
&= \frac{3^m + 1}{2 \cdot 4^m} \\
&> \frac{1}{2} (\frac{3}{4})^m
\end{aligned}$$

But $\mid p_m - q_m \mid = \frac{1}{4^m}$ *so that*

$$\frac{\mid f(p_m) - f(q_m) \mid}{\mid p_m - q_m \mid} > \frac{1}{2} 3^m$$

If $x = p_m$, *then we have*

$$\frac{\mid f(x) - f(q_m) \mid}{\mid x - q_m \mid} > \frac{1}{2} 3^m.$$

If $x = q_m$, *then we have*

$$\frac{\mid f(p_m) - f(x) \mid}{\mid p_m - x \mid} > \frac{1}{2} 3^m.$$

If $x \neq p_m$, $x \neq q_m$, *then since* $p_m < x < q_m$, *we get*

$$\begin{aligned}
\frac{1}{2} 3^m &< \frac{\mid f(p_m) - f(q_m) \mid}{\mid p_m - q_m \mid} \\
&\leq \frac{\mid f(p_m) - f(x) \mid}{\mid p_m - q_m \mid} + \frac{\mid f(x) - f(q_m) \mid}{\mid p_m - q_m \mid} \\
&\leq \frac{\mid f(p_m) - f(x) \mid}{\mid p_m - x \mid} + \frac{\mid f(x) - f(q_m) \mid}{\mid x - q_m \mid}
\end{aligned}$$

As such either

$$\frac{\mid f(p_m) - f(x) \mid}{\mid p_m - x \mid} > \frac{1}{4} 3^m$$

or

$$\frac{\mid f(x) - f(q_m) \mid}{\mid x - q_m \mid} > \frac{1}{4} 3^m.$$

All together this shows that there exists a sequence $\{y_m\} \to x$ *such that for each* $m \in \mathbb{N}$

$$\frac{\mid f(y_m) - f(x) \mid}{\mid y_m - x \mid} > \frac{1}{4} 3^m.$$

We can conclude that f *is not differentiable at* x.

2.10 Completions of Metric Spaces

We know that $\mathbb{R}$ can be constructed from $\mathbb{Q}$ by essentially *completing* $\mathbb{Q}$. In this section, we will see that every metric space may be viewed as a dense subset of a complete metric space.

We begin with the following simple, but very useful result.

PROPOSITION **2.10.1.** *Let (X, d) be a complete metric space. Let $A \subseteq X$. Then A is complete with respect to the induced metric if and only if A is closed in X.*

Proof. Assume that A is closed in X and let $\{x_n\}$ be Cauchy in (A, d_A). Then it is clear that $\{x_n\}$ is also Cauchy in (X, d). As such there exists an $x_0 \in X$ with $x_n \to x_0$. Since A is closed, $x_o \in A$ and A is complete.

For the converse, assume that A is not closed. Then there exists a point $x_0 \in bdy(A) \setminus A$. It follows that there exists a sequence $\{x_n\} \subseteq A$ with $x_n \to x_0$ in (X, d). But then $\{x_n\}$ is Cauchy in (A, d_A) but it cannot have a limit in A. That is A is not complete.

∎

DEFINITION **2.10.2.** *Let (X, d_X) and (Y, d_Y) be two metric spaces. A map $\phi : X \to Y$ is said to be an isometry if $d_Y(\phi(x_1), \phi(x_2)) = d_X(x_1, x_2)$ for every $x_1, x_2 \in X$.*

Note that it is clear that isometries are $1-1$. If ϕ is onto we say that (X, d_X) and (Y, d_Y) are isometric metric spaces.

A completion of (X, d_X) is a pair $((Y, d_Y), \phi)$ where (Y, d_Y) is a complete metric space, $\phi : X \to Y$ is an isometry and $\phi(X)$ is dense in Y.

It is not obvious that every metric space (X, d) can be completed. However, we will now show that this is infact the case.

THEOREM **2.10.3.** *Given a metric space (X, d), there exists and isometry $\phi : X \to (C_b(X), \|\cdot\|_\infty)$.*

Proof. Fix a point $a \in X$. For each $u \in X$ define $\phi(u) = f_u : X \to \mathbb{R}$ by

$$f_u(x) = d(u, x) - d(x, a).$$

Then f_u is continuous on X.

The triangle inequality shows that for any $x \in X$ that

$$\mid f_u(x) \mid = \mid d(u, x) - d(x, a) \mid \leq d(u, a)$$

so $f_u \in C_b(X)$.

Next we observe that for any $u, v \in X$, we have

$$\begin{aligned} \|f_u - f_v\|_\infty &= sup\{\mid f_u(x) - f_v(x) \mid \mid x \in X\} \\ &= \mid d(u, x) - d(v, x) \mid \\ &\leq d(u, v). \end{aligned}$$

On the other hand, $(f_u - f_v)(v) = (d(u, v) - d(v, a)) - (d(v, v) - d(v, a)) = d(u, v)$ so in fact $\|f_u - f_v\|_\infty = d(u, v)$.

∎

COROLLARY **2.10.4.** *Every metric space has a completion.*

Proof. Let $\phi : X \to (C_b(X), \|\cdot\|_\infty)$ be as in the previous theorem. Let $Y = \overline{\phi(X)}$. Since Y is a closed subset of a complete metric space it is also complete.

2.11 Banach Contractive Mapping Theorem

To motivate the main result of this section we begin with the following problem.

PROBLEM **14.** *Does there exists a function $f \in \mathbf{C}[0,1]$ so that*

$$f(x) = e^x + \int_0^x \frac{\sin(t)}{2} f(t)\,dt. \tag{2.1}$$

The strategy we will follow in answering this question is to first define a map $\Gamma : C[0,1] \to C[0,1]$ defined by

$$\Gamma(g) = e^x + \int_0^x \frac{\sin(t)}{2} g(t)\,dt$$

for each $g \in \mathbf{C}[a,b]$. Note that the fact that $\Gamma(g) \in C[0,1]$ follows from the Fundamental Theorem of Calculus. Next we will show that in fact there exists a unique function $f \in C[0,1]$ such that Γ fixes f. That is $\Gamma(f) = f$ and hence

$$f(x) = \Gamma(f) = e^x + \int_0^x \frac{\sin(t)}{2} f(t)\,dt.$$

as desired.

DEFINITION **2.11.1.** *Let $\Gamma : X \to X$. We call $x_0 \in X$ a fixed point of Γ if $\Gamma(x_0) = x_0$.*

If (X,d) is a metric space and $\Gamma : X \to X$, we say that Γ is Lipschitz if there exists constant $0 \leq \alpha$ such that

$$d(\Gamma(x), \Gamma(y)) \leq \alpha d(x,y)$$

for every $x, y \in X$.

We say that Γ is a contraction if there exists a constant $0 \leq k < 1$ such that

$$d(\Gamma(x), \Gamma(y)) \leq k d(x,y)$$

for every $x, y \in X$.

The following theorem, which is also know as the Banach Fixed Point Theorem, shows that every contraction map on a complete metric space has a unique fixed point.

THEOREM **2.11.2.** *[Banach Contractive Mapping Theorem]*
Let (X,d) be a complete metric space. Let $\Gamma : X \to X$ be a contraction. Then Γ has a unique fixed point $x_0 \in X$.

Proof. Let $x_1 \in X$. Then let $x_2 = \Gamma(x_1)$, $x_3 = \Gamma(x_2)$, and proceed recursively by defining

$$x_{n+1} = \Gamma(x_n).$$

Note that

$$d(x_3, x_2) = d(\Gamma(x_2), \Gamma(x_1)) \leq k d(x_2, x_1).$$

Similarly,

$$d(x_4, x_3) = d(\Gamma(x_3), \Gamma(x_2)) \leq k d(x_3, x_2) \leq k^2 d(x_2, x_1).$$

In fact, we can proceed inductively to show that

$$d(x_{n+1}, x_n) = \leq k^{n-1} d(x_2, x_1).$$

From this it follows that if $m > n$, we have

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n) \\ &\leq k^{m-2} d(x_2, x_1) + k^{m-1} d(x_2, x_1) + \cdots + k^{n-1} d(x_2, x_1) \\ &= k^{n-1} d(x_2, x_1)[k^{m-n-1} + k^{m-n-2} + \cdots + k + 1] \\ &\leq \frac{k^{n-1} d(x_2, x_1)}{1-k} \end{aligned}$$

Since $k^n \to 0$, it follows that $\{x_n\}$ is Cauchy. As (X, d) is complete $\{x_n\}$ converges to some $x_0 \in X$.

Now, It is clear that Γ is continuous. As such we have that $\Gamma(x_n) \to \Gamma(x_0)$. But $\Gamma(x_n) = x_{n+1} \to x_0$, so it follows that

$$\Gamma(x_0) = x_0.$$

Finally assume that y_0 also satisfies $\Gamma(y_0) = y_0$. Then

$$d(x_0, y_0) = d(\Gamma(x_0), \Gamma(y_0)) \leq k d(x_0, y_0).$$

As $0 < k < 1$, this implies that $d(x_0, y_0) = 0$ and hence that $x_0 = y_0$. ∎

REMARK **2.11.3.** *One might be tempted to ask if we could replace the condition $d(\Gamma(x), \Gamma(y)) \leq kd(x, y)$ with the weaker condition $d(\Gamma(x), \Gamma(y)) < d(x, y)$ and still obtain a unique fixed point? Unfortunately, this is not the case as the example $f : [1, \infty) \to [1, \infty)$ where $f(x) = x + \frac{1}{x}$ shows. (The details are left as an exercise.)*

The Banach Contraction Mapping Theorem has many applications to the theory of both integral and differential equations. We illustrate one such application by solving the problem with which we began this section.

EXAMPLE **2.14.** *Show that there exists a unique $f \in C[0, 1]$ so that*

$$f(x) = e^x + \int_0^x \frac{\sin(t)}{2} f(t)\, dt. \tag{2.2}$$

Let $\Gamma : C[0, 1] \to C[0, 1]$ be defined by

$$\Gamma(g) = e^x + \int_0^x \frac{\sin(t)}{2} g(t)\, dt$$

for each $g \in \mathbf{C}[0, 1]$. We note that for any $x \in [0, 1]$, and $f,\ g \in C[0, 1]$,

$$\begin{aligned} |\Gamma(g)(x) - \Gamma(f)(x)| &= \left| \left[e^x + \int_0^x \frac{\sin(t)}{2} g(t)\, dt \right] - \right. \\ &\qquad \left. \left[e^x + \int_0^x \frac{\sin(t)}{2} f(t)\, dt \right] \right| \\ &= \left| \int_0^x \frac{\sin(t)}{2} (g(t) - f(t))\, dt \right| \\ &\leq \int_0^x \left| \frac{\sin(t)}{2} \right| |g(t) - f(t)|\, dt \\ &\leq \int_0^x \left| \frac{\sin(t)}{2} \right| \|g - f\|_\infty\, dt, \end{aligned}$$

whence by linearity of definite integrals.

$$\begin{aligned}
\int_0^x \left|\frac{\sin(t)}{2}\right| \|g-f\|_\infty \, dt &= \|g-f\|_\infty \int_0^x \left|\frac{\sin(t)}{2}\right| dt \\
&\leq \|g-f\|_\infty \int_0^1 \left|\frac{\sin(t)}{2}\right| dt \\
&\leq \|g-f\|_\infty \int_0^1 \frac{1}{2}\, dt \\
&= \frac{1}{2}\|g-f\|_\infty.
\end{aligned}$$

This shows that $\|\Gamma(g)-\Gamma(f)\|_\infty \leq \frac{1}{2}\|g-f\|_\infty$; Γ *is a contraction. By the Banach Contraction Mapping Theorem, there exists a unique function* $f_0 \in C[0,1]$ *such that* $\Gamma(f_0) = f_0$. *But a function* f *satisfies the integral equation 2.2 if and only if* $\Gamma(f) = f$. *Hence* f_0 *is the unique solution to the integral equation 2.2.*

Note that not only does the Banach Contraction Mapping Theorem guarantee uniqueness and existence of fixed points, it also provides a constructive method to find the fixed point; namely, start with any function $f_0 \in C[0,1]$ and iteratively apply the contractive map Γ to it. The limit of this iteration will be the desired fixed point of Γ.

EXAMPLE **2.15.** *Show that there exists a unique function* $f_0(x) \in C[0,1]$ *such that*

$$f_0(x) = x + \int_0^x t^2 f_0(t)\, dt. \tag{2.3}$$

Find a power series representation for this function on $[0,1]$.

Let $\Gamma : C[0,1] \to C[0,1]$ *be defined by*

$$\Gamma(g)(x) = x + \int_0^x t^2 g(t)\, dt.$$

Note that f *is a solution to* (2.3) *if and only if* $\Gamma(f) = f$. *Observe that for any* $x \in [0,1]$, *and* $f,\, g \in C[0,1]$, *we have that*

$$\begin{aligned}
|\Gamma(g)(x) - \Gamma(f)(x)| &:= \left|\left(x + \int_0^x t^2 g(t)\, dt\right) - \left(x + \int_0^x t^2 f(t)\, dt\right)\right| \\
&= \left|\int_0^x t^2 (g(t) - f(t))\, dt\right| \\
&\leq \int_0^x |t^2| \cdot |g(t) - f(t)|\, dt \\
&\leq \int_0^1 |t^2| \cdot |g(t) - f(t)|\, dt \\
&\leq \int_0^1 t^2 \cdot \|g-f\|_\infty\, dt \\
&= \|g-f\|_\infty \cdot \int_0^1 t^2\, dt \\
&= \|g-f\|_\infty \cdot \left[\frac{t^3}{3}\right]_0^1 \\
&= \frac{1}{3}\|g-f\|_\infty.
\end{aligned}$$

This shows that Γ *is contractive. By the Banach Contraction Mapping Theorem,* Γ *has a unique fixed point* f_0 *with* $\Gamma(f_0) = f_0$, *so* f_0 *is the unique solution to* (2.3).

To find the series representation of $f_0(x)$, we begin with $f_1 = 0$, and $f_{n+1} = \Gamma(f_n)$. So

$$\begin{aligned}
f_2 = \Gamma(f_1) &= x + \int_0^x t^2 \cdot 0\, dt = x, \\
f_3 = \Gamma(f_2) &= x + \int_0^x t^2 \cdot t\, dt = x + \left.\frac{t^4}{4}\right|_0^x = x + \frac{x^4}{4}, \\
f_4 = \Gamma(f_3) &= x + \int_0^x t^2 \cdot \left(t + \frac{t^4}{4}\right) dt = x + \int_0^x \left(t^3 + \frac{t^6}{4}\right) dt \\
&= x + \frac{x^4}{4} + \frac{x^7}{4 \cdot 7}, \\
f_5 = \Gamma(f_4) &= x + \int_0^2 t^2 \cdot \left(t + \frac{t^4}{4} + \frac{t^7}{4 \cdot 7}\right) dt \\
&= x + \int_0^x \left(t^3 + \frac{t^6}{4} + \frac{t^9}{4 \cdot 7}\right) dt \\
&= x + \frac{x^4}{4} + \frac{x^7}{4 \cdot 7} + \frac{x^{10}}{4 \cdot 7 \cdot 10},
\end{aligned}$$

and so on. We can easily show using induction that for $n \geq 2$,

$$f_n(x) = \sum_{i=0}^{n-2} \frac{x^{3i+1}}{1 \cdot 4 \cdot 7 \cdot \cdots \cdot (3i+1)}.$$

We know from the Banach contractive mapping theorem that if $f_{n+1} = \Gamma(f_n)$ for all $n \in \mathbb{N}$, then $f_n \to f_0$ uniformly. This shows that

$$f_0(x) = \sum_{n=0}^{\infty} \frac{x^{3n+1}}{1 \cdot 4 \cdot 7 \cdot \cdots \cdot (3n+1)}$$

is the required power series representation.

■

Perhaps the most significant application of the Banach Contraction Mapping Theorem is the Picard-Lindelöf Theorem.

Theorem **2.11.4.** *[Picard-Lindelöf Theorem]*
Let $f : [0, b] \times \mathbb{R} \to \mathbb{R}$ be continuous and Lipschitz in y. That is there exists an $\alpha \geq 0$ such that

$$| f(t, y) - f(t, z) | \leq \alpha \mid y - z \mid$$

for all $y, z \in \mathbb{R}$. Let $y_0 \in \mathbb{R}$. Then there exists a unique function $y(t) \in C[0, b]$ such that

$$\begin{cases} y'(t) = f(t, y(t)) & \text{for all } t \in (0, b), \\ y(0) = y_0 \end{cases}$$

2.12 Baire's Category Theorem

We begin this section with the following interesting example of a function $f : \mathbb{R} \to \mathbb{R}$ which is continous every rational point in $\mathbb{R}$, but discontinuous at each irrational.

Example **2.16.** *[A Strange Function]*

Let

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}, \\ \frac{1}{n} & \text{if } x = \frac{m}{n},\ m \in \mathbb{Z},\ n \in \mathbb{N},\ \gcd(m,n) = 1,\ m \neq 0, \\ 1 & \text{if } x = 0. \end{cases}$$

For any $\alpha \in \mathbb{R}$, there exists a sequence $\{x_n\} \subset \mathbb{R} \setminus \mathbb{Q}$ with $\lim\limits_{n \to \infty} x_n = \alpha$ due to the density of irrationals. Since $f(x_n) = 0$ for every $n \in \mathbb{N}$, the sequential characterization of continuity shows that if $f(x)$ is continuous at $x = \alpha$, then $f(\alpha) = 0$. This shows that $f(x)$ is discontinuous at $x = r$ for all $r \in \mathbb{Q}$.

On the other hand, assume the $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Let $\epsilon > 0$ and choose $N_0 \in \mathbb{N}$ so that $\frac{1}{N_0} < \epsilon$. We note that in the interval $[\alpha - 1, \alpha + 1]$ there are only finitely any rationals of the form $r = \frac{m}{n}$ where $n < N_0$. As such, and because α is irrational, we can find a $\delta > 0$ so that if $r = \frac{m}{n} \in (\alpha - \delta, \alpha + \delta)$, then it must be that $n \geq N_0$. It follows that if $\mid x - \alpha \mid < \delta$, then

$$\mid f(x) - f(\alpha) \mid = f(x) < \epsilon$$

and hence that $f(x)$ is continuous at α.

In the previous example, we saw a function that is continous every irrational point in $\mathbb{R}$, but discontinuous at each rational. This leads us to the following natural question:

PROBLEM **15.** *Does there exist a function $f : \mathbb{R} \to \mathbb{R}$ such that $f(x)$ is continuous at each rational but discontinuous otherwise?*

To answer this question we will first show that the set of discontinuities of a function $f : (X, d) \to \mathbb{R}$, has a particular topological nature. Before we do so we will present a series of definitions that we will need going forward.

DEFINITION **2.12.1.** *Let (X, d) be a metric space. A set $A \subseteq X$ is said to be an F_σ set if*

$$A = \bigcup_{n=1}^{\infty} F_n$$

where $\{F_n\}$ is a sequence of closed subsets of X.

A set $A \subseteq X$ is said to be a G_δ set if

$$A = \bigcap_{n=1}^{\infty} U_n$$

where $\{U_n\}$ is a sequence of open subsets of X.

A set $A \subseteq X$ is said to be nowhere dense if $int(\overline{A}) = \emptyset$. A is said to be of 1st category in X if

$$A = \bigcup_{n=1}^{\infty} A_n$$

where $\{A_n\}$ is a sequence of nowhere dense subsets of X. Otherwise, we say that A is of second category in X.

We say that A is residual in X if A^c is of 1st category in X.

REMARK **2.12.2.** *1) It follows from DeMorgan's Laws that a set A is F_σ in X if and only if A^c is G_δ.*

2) The set $[0, 1)$ is both F_σ and G_δ in $\mathbb{R}$ even though it is neither open or closed. ($[0, 1) = \bigcup\limits_{n=1}^{\infty} [0, 1 - \frac{1}{n}]$ and $[0, 1) = \bigcap\limits_{n=1}^{\infty} (-\frac{1}{n}, 1)$.)

3) *Every closed set $F \subset X$ is also G_δ, and hence every open set U is F_σ. (Exercise)*

4) *Nowhere dense subsets, and indeed first category subsets of a metric space are thought to be "topologically thin" while second category sets, and more so residual sets are seen to be "topogically fat".*

5) *The Cantor set is nowhere dense in $\mathbb{R}$ but has cardinality c.*

6) *A closed set F is nowhere dense if and only if $U = F^c$ is dense.*

Definition **2.12.3.** *Let $f : (X, d_X) \to (Y, d_Y)$. Let $D(f)$ denote the set of all points in X at which f is not continuous.*

For each $n \in \mathbb{N}$, let

$$D_n(f) = \{x \in X \mid \text{for every } \delta > 0 \text{ there exists } y, z \in B(x, \delta) \text{ with } D_Y(f(y), f(z)) \geq \frac{1}{n}\}$$

The next theorem show that the set of discontinuities for a function between metric spaces must be an F_σ set. The proof is left as an excercise.

Theorem **2.12.4.** *Let $f : (X, d_X) \to (Y, d_Y)$. Then for each $n \in \mathbb{N}$, $D_n(f)$ is closed in X. Moreover,*

$$D(f) = \bigcup_{n=1}^{\infty} D_n(f).$$

In particular, $D(f)$ is F_σ.

Theorem **2.12.5.** *[Baire Category Theorem I]*

Let (X, d) be complete metric space. Let $\{U_n\}$ be a sequence of open dense sets. Then $\bigcap\limits_{n=1}^{\infty} U_n$ is dense in X.

Proof. Let W be open and non-empty. Then there exists an $x_1 \in X$ and $0 < r_1 \leq 1$ such that

$$B(x_1, r_1) \subseteq B[x_1, r_1] \subseteq W \cap U_1.$$

Next we can find $x_2 \in X$ and $0 < r_2 < \frac{1}{2}$ such that

$$B(x_2, r_2) \subseteq B[x_2, r_2] \subseteq B(x_1, r_1) \cap U_2.$$

We can then proceed recursively to find sequences $\{x_n\} \subseteq X$ and $\{r_n\} \subset \mathbb{R}$ with $0 < r_n \leq \frac{1}{n}$, and

$$B(x_{n+1}, r_{n+1}) \subseteq B[x_{n+1}, r_{n+1}] \subseteq B(x_n, r_n) \cap U_{n+1}.$$

Since $r_n \to 0$ and $B[x_{n+1}, r_{n+1}] \subseteq B[x_n, r_n]$, Cantor's Intersection Theorem implies that there exists an

$$x_0 \in \bigcap_{n=1}^{\infty} B[x_n, r_n.]$$

But then $x_0 \in B[x_1, r_1] \subseteq W$ and $x_0 \in B[x_n, r_n] \subseteq U_n$ for each $n \in \mathbb{N}$. This shows that

$$x_0 \in W \cap (\bigcap_{n=1}^{\infty} U_n).$$

∎

REMARK **2.12.6.** *The Baire Category Theorem shows that if $\{U_n\}$ is a sequence of open dense sets, then $\bigcap_{n=1}^{\infty} U_n$ is dense in X. We also know that $\bigcap_{n=1}^{\infty} U_n$ is a G_δ. These dense G_δ subsets of a complete metric space are always residual, as we will see below, and as such are topogically fat.*

Our next corollary shows the connection between the Baire Category Theorem and our notion of *category*.

COROLLARY **2.12.7.** *[Baire Category Theorem II]*
Every complete metric space (X, d) is of second category in itself.

Proof. Assume that X is of first category. Then there exists a sequence A_n of nowhere dense sets so that

$$X = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \overline{A_n}.$$

Now Let $U_n = (\overline{A_n})^c$. Then U_n is open and dense. But

$$\bigcap_{n=1}^{\infty} U_n = \emptyset$$

which is impossible. ∎

COROLLARY **2.12.8.** *$\mathbb{Q}$ is not a G_δ subset of $\mathbb{R}$.*

Proof. Suppose the

$$\mathbb{Q} = \bigcap_{n=1}^{\infty} U_n$$

where each U_n is open. Let $F_n = (U_n)^c$. Since $\mathbb{Q} \subseteq U_n$, it follows that U_n is dense, and hence that F_n is nowhere dense. Next Let $\{r_1, r_2, r_3, \ldots\}$ be an enumeration of $\mathbb{Q}$. Let $S_n = F_n \cup \{r_n\}$. Then S_n is closed and nowhere dense. However,

$$\mathbb{R} = \bigcup_{n=1}^{\infty} S_n$$

contradicting the Baire Category Theorem II.

∎

COROLLARY **2.12.9.** *There is no function $f : \mathbb{R} \to \mathbb{R}$ for which $D(f) = \mathbb{R} \setminus \mathbb{Q}$.*

We are now able to show that for a sequence $\{f_n\} \subset C([a, b])$ that converges pointwise, the limit function must be continuous at each point on a residual set. To do this we first introduce a new form of convergence which lies between pointwise and uniform convergence.

DEFINITION **2.12.10.** *Let (X, d_X), (Y, d_Y) be metric spaces. Let $\{f_n : X \to Y\}$ be a sequence of functions that converges pointwise on X to f_0. We say that $\{f_n\}$ converges uniformly at $x_0 \in X$ if for every $\epsilon > 0$ there exists a $\delta > 0$ and an $N_0 \in \mathbb{N}$ such that if $n, m \geq N_0$ and if $d_X(x, x_0) < \delta$, then $d_Y(f_n(x), f_m(x)) < \epsilon$.*

The following theorem shows that if $f_n \to f_0$ uniformly at a point x_0, then this is enough to preserve continuity. The proof is left as a exercise.

THEOREM **2.12.11.** *Let (X, d_X), (Y, d_Y) be metric spaces. Let $\{f_n : X \to Y\}$ be a sequence of functions that converges pointwise on X to f_0. Assume also that $\{f_n\}$ converges uniformly at $x_0 \in X$. If each f_n is continuous at x_0, then so is f_0.*

THEOREM **2.12.12.** *Let $f_n : (a,b) \to \mathbb{R}$ be a sequence of continuous functions that converges pointwise to $f(x)$. Then there exists an $x_0 \in (a,b)$ such that $f_n \to f$ uniformly at x_0.*

Proof. We first show:

(*) There exists a closed interval $[\alpha_1, \beta_1] \subset (a,b)$ with $\alpha_1 < \beta_1$, and an $N_1 \in \mathbb{N}$ so that if $n, m \geq N_1$ and $x \in [\alpha_1, \beta_1]$, then

$$| f_n(x) - f_m(x) | \leq 1.$$

Suppose that (*) fails and no such interval and N_1 exist. Then pick $a < t_1 < b$ and $n_1, m_1 \in \mathbb{N}$ with

$$| f_{n_1}(t_1) - f_{m_1}(t_1) | > 1.$$

Since $f_{n_1}(x)$ and $f_{m_1}(x)$ are continuous, there exists an open interval I_1 with $\overline{I_1} \subset (a,b)$ and

$$| f_{n_1}(x) - f_{m_1}(x) | > 1$$

for every $x \in I_1$. But then there must exist a $t_2 \in I_1$ and $n_2, m_2 > \max\{n_1, m_1\}$ such that

$$| f_{n_2}(t_2) - f_{m_2}(t_2) | > 1.$$

As before there exists an open interval I_2 with $\overline{I_2} \subset I_1$ and

$$| f_{n_2}(x) - f_{m_2}(x) | > 1$$

for every $x \in I_2$.

By induction we get a sequence $\{I_k\}$ of open intervals such that $(a,b) \supset \overline{I_1} \supset I_1 \supset \overline{I_2} \supset I_2 \supset \overline{I_3} \supset \cdots$ and two sequences $\{m_k\}, \{n_k\} \subseteq \mathbb{N}$ with $n_{k+1}, m_{k+1} > \max\{n_k, m_k\}$ and

$$| f_{n_k}(x) - f_{m_k}(x) | > 1$$

for every $x \in I_k$.

Now by the Nested Interval Theorem there exists

$$t_0 \in \bigcap_{k=1}^{\infty} \overline{I_k}.$$

It follows that

$$| f_{n_k}(t_0) - f_{m_k}(t_0) | > 1$$

for all $k \in \mathbb{N}$. This contradicts the fact that $\{f_n(t_0)\}$ is Cauchy.

From here we can proceed inductively to construct a sequence $\{[\alpha_k, \beta_k]\}$ of closed intervals with $(a,b) \supset [\alpha_1, \beta_1] \supset (\alpha_1, \beta_1) \supset [\alpha_2, \beta_2] \supset (\alpha_2, \beta_2) \supset [\alpha_3, \beta_3] \supset \cdots$ and a sequence $N_1 < N_2 < N_3 < \cdots$ so that if $n, m \geq N_k$ and $x \in [\alpha_k, \beta_k]$, then

$$| f_n(x) - f_m(x) | < \frac{1}{k}.$$

Finally, let

$$x_0 \in \bigcap_{k=1}^{\infty} [\alpha_k, \beta_k] = \bigcap_{k=1}^{\infty} (\alpha_k, \beta_k)$$

Now given $\epsilon > 0$, if $\frac{1}{k} < \epsilon$, then if $n, m \geq N_k$ and $x \in (\alpha_k, \beta_k)$, then

$$| f_n(x) - f_m(x) | < \epsilon.$$

But since $x_0 \in (\alpha_k, \beta_k)$ we need only choose a $\delta > 0$ so that $(x_0 - \delta, x_0 + \delta) \subseteq (\alpha_k, \beta_k)$ to conclude that $f_n \to f$ uniformly at x_0.

∎

Corollary **2.12.13.** *Let $\{f_n\} \subset C([a,b])$ be such that $f_n \to f_0$ pointwise on $[a,b]$. Then there exists a residual set $A \subset [a,b]$ such that $f_0(x)$ is continuous at each $x \in A$.*

Proof. It follows from the previous Theorem that the set A on which $f_0(x)$ is continuous is dense in $[a,b]$. However, we also know that $D(f_0)$ is an F_σ set so that A is a dense G_δ set. ∎

Corollary **2.12.14.** *Assume that $f(x)$ is differentiable on $\mathbb{R}$. Then $f'(x)$ is continuous for every point in a dense G_δ-subset of $\mathbb{R}$.*

Proof. Observe that $f'(x)$ is the pointwise limit of the sequence of continuous functions $\{\frac{f(x+\frac{1}{n})-f(x)}{\frac{1}{n}}\}$. ∎

2.13 Compactness

In this section we will discuss three important properties of a topological space, namely compactness, sequential compactness and the Bolzano-Weierstrass Property. In fact we will show that for metric spaces, these three properties are equivalent. We begin by defining these three properties.

Definition **2.13.1.** *[Compactness]*
Let (X,d) be a metric space. An (open) cover for X is a colection $\{U_\alpha\}_{\alpha\in I}$ for which

$$X = \bigcup_{\alpha\in I} U_\alpha.$$

Given a cover $\{U_\alpha\}_{\alpha\in I}$ a subcover is a collection $\{U_\alpha\}_{\alpha\in J}$ where $J \subseteq I$ and

$$X = \bigcup_{\alpha\in J} U_\alpha.$$

$\{U_\alpha\}_{\alpha\in J}$ is a finite subcover if J is finite.

We say that X is compact if every cover $\{U_\alpha\}_{\alpha\in I}$ has a finite subcover.

A subset $A \subseteq X$ is said to be compact, if for every cover of A, that is a collection $\{U_\alpha\}_{\alpha\in I}$ of open sets with

$$A \subseteq \bigcup_{\alpha\in I} U_\alpha,$$

there is a finite subcover $\{U_{\alpha_1}, U_{\alpha_2}, U_{\alpha_3}, \ldots, U_{\alpha_n}\}$ or equivalently, if (A, d_A) is a compact metric space.}

Definition **2.13.2.** *[Sequential Compactness]*
Let (X,d) be a metric space. We say that X is sequentially compact, if every sequence $\{x_n\} \subseteq X$ has a convergent subsequence.

If $A \subseteq X$, we say A is sequentially compact if every sequence $\{x_n\} \subseteq A$ has a subsequence converging to a point in A.

Definition **2.13.3.** *[Bolzano-Weierstrass Property]*
Let (X,d) be a metric space. We say that X has the Bolzano-Weierstrass Property (BWP) if every infinite subsets in X has a limit point.

We will begin by showing that a metric space (X,d) is sequentially compact if and only if it has the BWP.

Theorem **2.13.4.** *Let (X,d) be a metric space. Then the following are equivalent:*

1) *X is sequentially compact.*

2) *X has the Bolzano-Weierstrass Property.*

Proof. 1) $\Rightarrow$ 2) Assume that X is sequentially compact and let $S \subseteq X$ be infinite. Then we know that we can extract from S a sequence $\{x_n\}$ consisting of distinct points. But then by sequential compactness, this sequence has a convergent subsequence $\{x_{n_k}\}$ with $x_{n_k} \to x_0$. But this means that if $\epsilon > 0$, then $B(x_0,\epsilon)$ contains infinitely many terms in the sequence x_{n_k} and consequently $x_0 \in Lim(S)$.

2) $\Rightarrow$ 1) Assume that X has the BWP. Assume also that $\{x_n\}$ is a sequence in X. If there is an element $x_0 \in X$ which appears infinitely many times in $\{x_n\}$, then $\{x_n\}$ has a subsequence which is constant and as such convergent. So we may assume without loss of generality that no such x_0 exists. This means that when viewed as a subset of X, $\{x_n\}$ is infinite. Moreover, by passing to a subsequence if necessary we can assume that the terms of $\{x_n\}$ are distinct, which we do going forward.

Given that $\{x_n\}$ is an infinite set, there exists $x_0 \in Lim(\{x_n\})$. We can thus find an $n_1 \in \mathbb{N}$ so that $d(x_{n_1},x_0) < 1$. We can then find an $n_2 > n_1$ so that $d(x_{n_2},x_0) < \frac{1}{2}$. Indeed if we have chosen $n_1 < n_2 < n_3 < \cdots < n_k$ with $d(x_{n_1},x_0) < \frac{1}{k}$, then we choose recursively $n_{k+1} > n_k$ with $d(x_{n_{k+1}},x_0) < \frac{1}{k+1}$. This gives us a subsequence $\{x_{n_k}\}$ with $d(x_{n_k},x_0) < \frac{1}{k}$. In particular, $x_{n_k} \to x_0$.

∎

The next result is useful in establishing some significant restrictions on the nature of compact and sequentially compact metric spaces and subsets of metric spaces.

Proposition **2.13.5.** *Let (X,d) be a metric space and $A \subset X$.*

1) *If A is compact then A is closed and bounded.*

2) *If A is closed and X is compact, then so is A*

3) *If A is sequentially compact then A is closed and bounded.*

4) *If A is closed and X is sequentially compact, then so is A.*

5) *If X is sequentially compact, then X is complete.*

Proof. We may of course assume that $X \neq \emptyset$ and that $A \neq \emptyset$ in each case.

1) To see that A is bounded choose an $x_0 \in X$ and then observe that $\{B(x_0,n)\}_{n=1}^{\infty}$ is an open cover of A. If A is compact it has a finite subcover $\{B(x_0,n_1), B(x_0,n_2),\ldots,B(x_0,n_k)\}$ where we may also assume that $n_1 < n_2 < n_3 < \cdots < n_k$. It follows that $A \subset B(x_0,n_k)$ and hence that A is bounded.

   Assume that A is not closed. Then there exists $x_0 \in Lim(A) \setminus A$. For each $n \in \mathbb{N}$, let $U_n = B[x_0,\frac{1}{n}]^c$. Then $\{U_n\}_{n=1}^{\infty}$ is an open cover of A with no finite subcover.

2) Assume that A is closed and X is compact. Let $\{U_\alpha\}_{\alpha\in I}$ be a cover of A. Then $\{U_\alpha\}_{\alpha\in I} \cup \{A^c\}$ is a cover of X. Hence there exists $\{U_{\alpha_1},U_{\alpha_2},U_{\alpha_3},\ldots,U_{\alpha_n}\}$ such that $\{U_{\alpha_1},U_{\alpha_2},U_{\alpha_3},\ldots,U_{\alpha_n}\} \cup \{A^c\}$ is a cover of X. It follows that $\{U_{\alpha_1},U_{\alpha_2},U_{\alpha_3},\ldots,U_{\alpha_n}\}$ is a finite subcover of a.

3) Assume that A is not bounded. Choose $x_1 \in A$. We can then choose $x_2 \in A$ with $d(x_1,x_2) > 1$. As A is unbounded we can then choose $x_3 \in A$ with $d(x_i,x_3) > 1$ for $i = 1,2$. We can then proceed recursively to define a sequence $\{x_n\} \subset A$ with $d(x_n,x_m) > 1$ if $n \neq m$. It is then clear that the sequence $\{x_n\}$ has no convergent subsequences, and as such that A is not sequentially compact.

   Assume that A is not closed. Then there exists a sequence $\{x_n\} \subset A$ with $x_n \to x_0$ and $x_0 \notin A$. Clearly $\{x_n\}$ has no subsequence converging in A and so A is not sequentially compact.

4) Assume that A is closed and X is sequentially compact. Let $\{x_n\} \subset A$. Then there is a subsequence $\{x_{n_k}\}$ with $x_{n_k} \to x_0$ in X. But since A is closed $x_0 \in A$.

5) Assume that X is sequentially compact and that $\{x_n\}$ is Cauchy. Then $\{x_n\}$ has a convergent subsequence by sequential compactness, and hence $\{x_n\}$ converges.

∎

Example **2.17.** *[Compactness and Sequential Compactness in $\mathbb{R}^n$]*

Assume the A is either compact or sequentially compact in $\mathbb{R}^n$. Then A is closed and bounded.

We know that in $\mathbb{R}$ a bounded sequence has a convergent subsequence, and as such for $\mathbb{R}$ sequential compactness is equivalent to being closed and bounded. By recognizing that a sequence in $\mathbb{R}^n$ converges if and only if each of its component sequences converge, it is easy to see that a simple inductive process will allow us to extend this result to $\mathbb{R}^n$.

The next theorem establishes the equivalence of compactnes and closed and boundedness for $\mathbb{R}^n$. Before we state the result we need the following definition.

Definition **2.13.6.** *A closed cell in $\mathbb{R}^n$ is a set*

$$J = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

Theorem **2.13.7.** *[Heine-Borel] A set $A \subseteq \mathbb{R}^n$ is compact if and only if A is closed and bounded.*

Proof. We need only show that if A is closed and bounded then it is compact.

Assume that A is closed and bounded. Assume also that there exists a cover $\{U_\alpha\}_{\alpha \in I}$ of A with no finite subcover. Let $F_1 = A$ and let J_1 be a closed cell containing A. We can divide J_1 into 2^n closed subcells $\{J_{1,1}, J_{1,2}, J_{1,3}, \ldots, J_{1,2^n}\}$ by bisecting each of the intervals $[a_i, b_i]$. One of these subcells, call it J_2, must be such that $F_2 = J_2 \cap A$ cannot be covered by finitely many members of $\{U_\alpha\}_{\alpha \in I}$. Observe also that $diam(J_2) = \frac{1}{2} diam(J_1)$.

We can know proceed to construct a sequence of closed cells $\{J_n\}$ such that

1) $J_{n+1} \subset J_n$ for each $n \in \mathbb{N}$.

2) $diam(J_{n+1}) = \frac{1}{2} diam(J_n)$ for each $n \in \mathbb{N}$.

3) $F_n = J_n \cap A$ cannot be covered by finitely many members of $\{U_\alpha\}_{\alpha \in I}$.

Then $\{F_n\}$ is a nested sequence of non-empty closed sets with $diam(F_n) \to 0$. By Cantor's Intersection Theorem,

$$\bigcap_{n=1}^{\infty} F_n = \{x_0\}.$$

Since $x_0 \in A$, we have $x_0 \in U_{\alpha_0}$ for some α_0. But then there exists an $\epsilon > 0$ such that $B(x_0, \epsilon) \subseteq U_{\alpha_0}$. However if n_0 is large enough so that $diam(F_{n_0}) < \epsilon$, then

$$F_{n_0} \subset B(x_0, \epsilon) \subseteq U_{\alpha_0}$$

which is impossible.

∎

The Heine-Borel Theorem and what we know so far about compactness and sequential compactness suggests the following problem:

Problem **16.** *Is (X, d) compact if and only if it is sequentially compact? Moreover, is a subset A of X compact if and only if it is closed and bounded?*

Remark **2.13.8.** *In fact, it is easy to see that the second statement above is false. Indeed, in every metric space (X, d) is closed in itself, so in reality all we need is a bounded metric space which is not compact. We can simply choose any infinite set X and give it the discrete metric d. Clearly X is bounded and closed. But $\{x\}_{x \in X}$ is an open cover with no finite subcover.*

On the other hand we will eventually be able to show that for metric spaces compactness and sequential compactness are in fact equivalent.

The next result can be viewed as an upgrade of the Cantor Intersection Theorem for compact metric spaces.

Definition **2.13.9.** *Let X be any set. A collection $\{A_\alpha\}_{\alpha \in I}$ is said to have the Finite Intersection Property (FIP) if whenever $\{A_{\alpha_1}, A_{\alpha_2}, A_{\alpha_3}, \ldots A_{\alpha_n}\}$ is any finite subcollection of $\{A_\alpha\}_{\alpha \in I}$, we have that*

$$\bigcap_{i=1}^{n} A_{\alpha_i} \neq \emptyset.$$

Theorem **2.13.10.** *Let (X, d) be a metric space. Then the following are equivalent:*

1) X is compact.

2) If $\{F_\alpha\}_{\alpha \in I}$ is a collection of closed subsets of X with the FIP, then

$$\bigcap_{\alpha \in I} F_\alpha \neq \emptyset.$$

Proof. 1) $\Rightarrow$ 2) Assume that X is compact and that $\{F_\alpha\}_{\alpha \in I}$ has the FIP. Let $U_\alpha = F_\alpha^c$ for each $\alpha \in I$.

Assume that

$$\bigcap_{\alpha \in I} F_\alpha = \emptyset.$$

Then

$$\bigcup_{\alpha \in I} U_\alpha = X.$$

But then by compactness we can find a finite subcover $\{U_{\alpha_1}, U_{\alpha_2}, U_{\alpha_3}, \ldots, U_{\alpha_n}\}$. It then follows that

$$\bigcap_{i=1}^{n} F_{\alpha_i} = \emptyset$$

contradicting the assumption that $\{F_\alpha\}_{\alpha \in I}$ has the FIP.

2) $\Rightarrow$ 1) Assume 2) holds. Let $\{U_\alpha\}_{\alpha \in I}$ be a cover for X with no finite subcover. Let $F_\alpha = U_\alpha^c$ for each $\alpha \in I$. Then

$$\bigcap_{\alpha \in I} F_\alpha = \emptyset.$$

Now since $\{U_{\alpha_1}, U_{\alpha_2}, U_{\alpha_3}, \ldots, U_{\alpha_n}\}$ is not a cover for any finite collection, we get that for any choices of finitely many elements $\{F_{\alpha_1}, F_{\alpha_2}, F_{\alpha_3}, \ldots, F_{\alpha_n}\}$

$$\bigcap_{i=1}^{n} F_{\alpha_i} \neq \emptyset.$$

This contradicts the fact that $\{F_\alpha\}_{\alpha \in I}$ has the FIP.

∎

The first corollary to the previous theorem may be viewed as a generalization of the Nested Interval Theorem for $\mathbb{R}$.

COROLLARY **2.13.11.** *Assume that (X,d) is a compact metric space. Let $\{F_n\}$ be a sequence of non-empty closed substes of X such that $F_{n+1} \subseteq F_n$ for each $n \in \mathbb{N}$, Then*

$$\bigcap_{n=1}^{\infty} F_n \neq \emptyset.$$

COROLLARY **2.13.12.** *Assume that (X,d) is compact. Then X has the BWP. In particular, X is sequentially compact.*

Proof. Assume that X is compact and that S is infinite. Then we can find a sequence $\{x_n\} \subseteq S$ consisting of distinct points. Let $F_n = \overline{\{x_n, x_{n+1}, x_{n+2}, \ldots\}}$. Then $\{F_n\}$ clearly has the FIP. As such there exists

$$x_0 \in \bigcap_{n=1}^{\infty} F_n.$$

But then for every $\epsilon > 0$, $B(x_0, \epsilon) \cap \{x_n, x_{n+1}, x_{n+2}, \ldots\} \neq \emptyset$ and hence $B(x_0, \epsilon) \cap S$ is infinite. ∎

We know that in $\mathbb{R}$ compactness is equivalent to a set being closed and bounded. We have already seen that in a general metric space that this equivalence fails. We do know that the metric space must be complete. We will now present an upgrade on boundedness that will eventually allow us to complete our characterization of compact sets.

DEFINITION **2.13.13.** *Let (X,d) be a metric space. Let $\epsilon > 0$ A collection $\{x_\alpha\}_{\alpha \in I} \subseteq X$ is said to be an ϵ-net for X if*

$$X = \bigcup_{\alpha \in I} B(x_\alpha, \epsilon).$$

We say that (X,d) is totally bounded if for each $\epsilon > 0$, X has a finite ϵ-net.

Given a subset $A \subseteq X$ we say that A is totally bounded if it is totally bounded in the induced metric. This is equivalent to saying that for every $\epsilon > 0$, there exists finitely many points $\{x_1, x_2, x_3, \ldots, x_n\} \subseteq A$ so that

$$A \subseteq \bigcup_{i=1}^{n} B(x_i, \epsilon).$$

PROPOSITION **2.13.14.** *Let (X,d) be sequentially compact. Then X is totally bounded.*

Proof. Suppose that X is not totally bounded. Thene there exists an $\epsilon_0 > 0$ with no finite ϵ_0-net. But then we can constrauct a sequence $\{x_n\} \subset X$ so that $x_i \notin B(x_j, \epsilon_0)$ if $i \neq j$. Consequently, $d(x_i, x_j) \geq e_0$ if $i \neq j$. Such a sequence cannot have a convergent subsequence. ∎

REMARK **2.13.15.**

1) *Let $(\mathbb{N}, d)$ be the natural numbers with the discrete metric. Then N is bounded, but it is not totally bounded because it has no finite $\frac{1}{2}$-net.*

2) *If $A \subseteq X$ is totally bounded, then so is $\overline{A}$. In fact, if $\{x_1, x_2, x_3, \ldots, x_n\}$ is an $\frac{\epsilon}{2}$-net for A, then it is also an ϵ-net for $\overline{A}$.*

We will need the following result, which is of independent interest:

THEOREM **2.13.16.** *Let $f : (X, d_X) \to (Y, d_Y)$ be continuous. If (X, d_X) is sequentially compact, then so is $(f(X), d_Y)$.*

Proof. Let $\{y_n\} \subseteq f(X)$. Then $y_n = f(x_n)$ for some $x_n \in X$. It follows that the sequence $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ which converges to some $x_0 \in X$. But then by continuity $y_{n_k} = f(x_{n_k}) \to f(x_0) = y_0$. ∎

COROLLARY **2.13.17.** *[Extreme Value Theorem]*
Let $f : (X, d_X) \to \mathbb{R}$ be continous. If (X, d_X) is sequentially compact, then there exists $c, d \in X$ so that

$$f(c) \leq f(x) \leq f(d)$$

for all $x \in X$.

Proof. It follows from the previous theorem that $f(X)$ is sequentially compact in $\mathbb{R}$. This means that $f(X)$ is both closed and bounded, and in particular that $glb(f(X)) \in f(X)$ and $lub(f(X)) \in f(X)$. Choose $c, d \in X$ so that $f(c) = glb(f(X))$ and $f(d) = lub(f(X))$.

∎

THEOREM **2.13.18.** *[Lebesgue]*
Let (X, d) be a sequentially compact metric space and let $\{U_\alpha\}_{\alpha \in I}$ be a cover of X. Then there exists an $\epsilon > 0$ such that for every $x \in X$ and $0 < \delta < \epsilon$ there is an $\alpha_0 \in I$ with

$$B(x, \delta) \subseteq U_{\alpha_0}$$

In this case we call ϵ a Lebesgue number for the cover $\{U_\alpha\}_{\alpha \in I}$.

Proof. If $X = U_\alpha$ for some α, then any $\epsilon > 0$ will work. So we assume that $X \neq U_\alpha$ for any α.

For each $x \in X$, let

$$\phi(x) = sup\{r \in \mathbb{R} \mid B(x, r) \subseteq U_{\alpha_0} \text{ for some } \alpha_0 \in I\}.$$

Then it is clear that $\phi(x) > 0$. Moreover, $\phi(x) < \infty$ for otherwise because X is bounded, we would have $X = U_\alpha$ for some α.

Now if $x, y \in X$ then the Triangle Inequality shows that

$$\phi(x) \leq \phi(y) + d(x, y)$$

and hence that

$$\mid \phi(x) - \phi(y) \mid \leq d(x, y).$$

From this we conclude that $\phi : X \to \mathbb{R}$ is continuous. But then by the Extreme Value Theorem there exists an $\epsilon > 0$ such that $\phi(x) \geq \epsilon$ for all $x \in X$.

∎

THEOREM **2.13.19.** *[Borel-Lebesgue] Let (X, d) be a metric space. Then the following are equivalent:*

1) X is compact.

2) X has the BWP.

3) X is sequentially compact.

Proof. We need only show that 3) $\Rightarrow$ 1). The other implications have already been established. Assume that X is sequentially compact and that $\{U_\alpha\}_{\alpha \in I}$ is an open cover for X. Then $\{U_\alpha\}_{\alpha \in I}$ has a Lebesgue number $\epsilon > 0$.

Now choose $0 < \delta < \epsilon$. Since X is also totally bounded, we can find a finite subset $\{x_1, x_2, \ldots, x_n\}$ of X so that $\{B(x_i, \delta)\}_{i=1}^n\}$ is also a cover of X. However, for each $i = 1, 2, \ldots, n$ we can also find an $\alpha_i \in I$ with

$$B(x_i, \delta) \subseteq U_{\alpha_i}.$$

It follows that $\{U_{\alpha_i}\}_{i=1}^n$ is a finite subcover of $\{U_\alpha\}_{\alpha \in I}$ and hence that (X, d) is compact.

∎

We are now in a position to establish the proper generalization of the Heine-Borel Theorem for general metric spaces.

THEOREM **2.13.20.** *Let (X, d) be a metric space. Then the following are equivalent:*

1) X is compact.

2) X is complete and totally bounded.

Proof. We have already established that 1) $\Rightarrow$ 2) As such we need only prove that 2) $\Rightarrow$ 1). Moreover, as a consequence of the Borel-Lebesgue Theorem, we must only show that 2) implies that X is sequentially compact.

Assume that 2) holds and that $\{x_n\}$ is a sequence in X. Since X is totally bounded X can be covered by finitely many open balls of radius 1. It follows that one such ball $S_1 = B(y_1, 1)$ contains infinitely many terms in $\{x_n\}$.

Next we cover X with finitely many open balls of radius $\frac{1}{2}$. We then choose one such ball $S_2 = B(y_2, \frac{1}{2})$ which contains infinitely many of the terms in $\{x_n\}$ which also lie in S_1.

From here we proceed inductively to construct a sequence of open balls $\{S_k = B(y_k, \frac{1}{k})\}$ with the property that each S_{k+1} contains infinitely many terms of the sequence $\{x_n\}$ which are also in $S_1 \cap S_2 \cap \cdots \cap S_k$. In particular, we can choose a sequence $n_1 < n_2 < n_3 < \cdots$ such that

$$x_{n_k} \in S_1 \cap S_2 \cap \cdots \cap S_k.$$

Since $diam(S_k) \to 0$ and since if $k, m > N$ then $x_{n_k}, x_{n_m} \in S_N$, it follows that $\{x_{n_k}\}$ is Cauchy. From the completeness of X we conclude that $\{x_{n_k}\}$ converges and therefore that X is sequentially compact.

∎

2.14 Compactness and Continuity

Recall that we have already shown that continous functions preserve sequential compactness. From this we immediately deduce that following tow results.

THEOREM **2.14.1.** *Let $f : (X, d_X) \to (Y, d_Y)$ be continuous. If (X, d_X) is compact, then so is $(f(X), d_Y)$.*

COROLLARY **2.14.2.** *[Extreme Value Theorem]*
Let $f : (X, d_X) \to \mathbb{R}$ be continous. If (X, d_X) is compact, then there exists $c, d \in X$ so that

$$f(c) \leq f(x) \leq f(d)$$

for all $x \in X$.

We will soon show that if (X, d) is compact, then contiuous functions have the following important property.

DEFINITION **2.14.3.** *Let $f : (X, d_X) \to (Y, d_Y)$. Then $f(x)$ is said to be uniformly continuous if for every $\epsilon > 0$ there exists a $\delta > 0$ so that if $d_X(x, z) < \delta$, then $d_Y(f(x), f(z)) < \epsilon$.*

REMARK **2.14.4.** *1) It is clear that if $f : (X, d_X) \to (Y, d_Y)$ is uniformly continuous, then $f(x)$ is continuous.*

2) Uniformly continuous functions have the special property that if $\{x_n\}$ is Cauchy in X, then $\{f(x_n)\}$ is Cauchy in Y. (Exercise)

We can now state our sequential characterzation of uniform continuity.

THEOREM **2.14.5.** *[Sequential Characterization of Uniform Continuity]*
Suppose that $f : (X, d_X) \to (Y, d_Y)$. Then the following are equivalent.

1) $f(x)$ is uniformly continuous on X.

2) $(*)$ If $\{x_n\}$, $\{z_n\}$ are sequences in X with $\lim_{n\to\infty} d_X(x_n, z_n) = 0$, then $\lim_{n\to\infty} d_Y(f(x_n), f(z_n)) = 0$.

Proof. 1) $\Rightarrow$ 2) Suppose 1) holds, and $\{x_n\}$, $\{z_n\}$ are sequences in X with $\lim_{n\to\infty} d_X(x_n, z_n) = 0$. Let $\epsilon > 0$. There exists $\delta > 0$ such that if $d_X(x, z) < \delta$, then $d_Y(f(x_1), f(x_2))$. Since $\lim_{n\to\infty} d_X(x_n, z_n) = 0$, we can find an $N \in \mathbb{N}$ such that if $n \geq N$, then $d_X(x_n, z_n) < \delta$. Hence for all $n \geq N$, $d_Y(f(x_n), f(z_n)) < \epsilon$.

2) $\Rightarrow$ 1) We will proceed by contradiction. Suppose to the contrary that hypothesis 2) holds but condition 1) fails. Then there exists an $\epsilon_0 > 0$ such that for every $\delta > 0$, we can find x_δ, $z_\delta \in X$ with $d_X(x_\delta, z_\delta) < \delta$ and yet $d_Y(f(x_\delta), f(z_\delta)) \geq \epsilon_0$. In particular, for $\delta_n = \frac{1}{n}$, we can find x_n, $z_n \in X$ with $d_X(x_n, z_n) < \delta_n = \frac{1}{n}$ and $d_Y(f(x_n) - f(z_n)) \geq \epsilon_0$, for all $n \in \mathbb{N}$. We now have a pair of sequences $\{x_n\}$, $\{z_n\} \subseteq X$ that satisfies $\lim_{n\to\infty} d_X(x_n, z_n) = 0$ but $\lim_{n\to\infty} d_Y(f(x_n), f(z_n)) \neq 0$, directly contradicting hypothesis 2). ∎

THEOREM **2.14.6.** *Let $f : (X, d_X) \to (Y, d_Y)$ be continuous. If (X, d_X) is compact, then $f(x)$ is uniformly continuous.*

Proof. Assume that $f(x)$ is not uniformly continuous. Then there exists an $\epsilon_0 > 0$ and two sequences $\{x_n\}, \{z_n\} \subseteq X$ with $\lim_{n\to\infty} d_X(x_n, z_n) = 0$ but $d_Y(f(x_n), f(z_n)) \geq \epsilon_0$ for each $n \in \mathbb{N}$.

By compactness $\{x_n\}$ has a convergent subsequence $\{x_{n_k}\}$ with $x_{n_k} \to x_0$ for some $x_0 \in X$. But since $\lim_{n\to\infty} d_X(x_n, z_n) = 0$, we have $z_{n_k} \to x_0$ as well. By continuity, $f(x_{n_k}) \to f(x_0)$ and $f(z_{n_k}) \to f(x_0)$. But this is impossible since $d_Y(f(x_{n_k}), f(z_{n_k})) \geq \epsilon_0$ for each $k \in \mathbb{N}$.

∎

DEFINITION **2.14.7.** *Let $(X.d_X)$ and (Y, d_Y) be metric spaces. A function $\phi : X \to Y$ is called a homoemorphism if ϕ is $1-1$, onto, continuous and $\phi^{-1} : Y \to X$ is also continuous.*

We say that two metirc spaces $(X.d_X)$ and (Y, d_Y) are homeomorphic if there is a homeomorphism $\phi : X \to Y$

REMARK **2.14.8.** *If $(X.d_X)$ and (Y, d_Y) are homeomorphic, then the are essentailly the same topologically in the sense that U is open in X if and only if $\phi(U)$ is open in Y.*

The final result of this section shows that compact metric spaces have rather rigid topologies.

THEOREM **2.14.9.** *Let $(X.d_X)$ and (Y, d_Y) be metric spaces with X compact. Let $\phi : X \to Y$ be $1-1$, onto and continuous. Then ϕ^{-1} is also continuous.*

Proof. Since $(\phi^{-1})^{-1} = \phi$, we need only show that if $U \subset X$ is open, then $\phi(U)$ is open in Y. But if $U \subset X$ is open, then $F = U^c$ is closed, and hence compact. It follows that $\phi(F)$ is compact in Y and as such is closed. But then $\phi(F)^c = \phi(U)$ is open.

∎

2.15 Finite Dimensional Normed Linear Space

Recall that a vector space V over $\mathbb{R}$ is n-dimensional if it has a basis $\{v_1, v_2, \ldots, v_n\}$ consisting of exactly n elements. In the case of $\mathbb{R}^n$ we will denote the standard basis by $\{e_1, e_2, \ldots, e_n\}$ where $e_i =$

$(0, 0, \ldots, 0, 1, 0, \ldots, 0)$ is the n-tuple with 1 in the i-th component and 0 in each other component.

All n-dimensional vectors spaces are isomprphic as vector space with $\mathbb{R}^n$ via the mapping $\Gamma_n : \mathbb{R}^n \to V$ given by

$$\Gamma_n((a_1, a_2, \ldots, a_n)) = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

In this section we will show that all $n-$dimensional normed linear spaces are fundamentally the same, in the sense that they have the same topological structure and are all complete.

We begin by discussing continuity for linear maps between normed linear spaces. $(\mathbb{R}^3, ||| \cdot |||)$

Definition **2.15.1.** *Let $(V, \| \cdot \|_V)$ and $(W, \| \cdot \|_W)$ be normed linear spaces. Let $T : V \to W$ be linear. We say that T is bounded is*

$$\sup_{\|x\|_V \leq 1} \{\| T(x) \|_W\} < \infty.$$

In this case, we write

$$\| T \| = \sup_{\|x\|_V \leq 1} \{\| T(x) \|_W\}.$$

Otherwise, we say that T is unbounded.

The next result establishes the fundamental criterion for when a linear map between normed linear spaces is continuous. It's proof is left as an exercise.

Theorem **2.15.2.** *Let $(V, \| \cdot \|_V)$ and $(W, \| \cdot \|_W)$ be normed linear spaces. Let $T : V \to W$ be linear. Then the following are equvalent.*

1) T is continuous.

2) T is bounded.

Remark **2.15.3.** *Let $(V, \| \cdot \|_V)$ and $(W, \| \cdot \|_W)$ be normed linear spaces. Let $T : V \to W$ be linear. The following statements are easy to deduce from the p[revious theorem.*

1) If T is bounded, then T is uniformly continuous.

2) T is continuous on V if and only if T is coninuous at $0 \in V$.

3) Recall that if $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then T can be represented by an $m \times n$ matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \\ \vdots \\ \vec{a}_m \end{bmatrix}$$

where $\vec{a}_i = (a_{i,1}, a_{i,2}, a_{i,3}, \cdots, a_{i,n})$. If $\vec{x} = (x_1, x_2, x_3, \ldots, x_n) \in \mathbb{R}^n)$, then

$$T(\vec{x}) = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ a_{3,1} & a_{3,2} & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & a_{m,3} & \cdots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \vec{a}_1 \cdot \vec{x} \\ \vec{a}_2 \cdot \vec{x} \\ \vec{a}_3 \cdot \vec{x} \\ \vdots \\ \vec{a}_m \cdot \vec{x} \end{bmatrix}$$

Let $M = \max_{i=1,2,\ldots,m}\{\| \vec{a}_i \|_2\}$, *then it follows from the Cauchy-Schwartz Inequality that if* $\| \vec{x} \|_2 \leq 1$ *then*

$$\begin{aligned} \|T(\vec{x})\|_2 &= \sqrt{\sum_{i=1}^{m} |\vec{a}_i \cdot \vec{x}|^2} \\ &\leq \sqrt{\sum_{i=1}^{m} M^2} \\ &= \sqrt{m} \cdot M \end{aligned}$$

The last remark implies the following:

PROPOSITION **2.15.4.** *Let* $T : \mathbb{R}^n \to \mathbb{R}^m$ *is linear, then* T *is bounded, and hence continuous.*

The next theorem will be our key tool to establish the link between all n-dimensional normed linear spaces.

THEOREM **2.15.5.** *Let* $(V, \| \cdot \|_V)$ *be an* n*-dimensional normed linear space with basis* $\{v_1, v_2, \ldots, v_n\}$. *Let* $\Gamma_n : \mathbb{R}^n \to V$ *be given by*

$$\Gamma_n((a_1, a_2, \ldots, a_n)) = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

Then both Γ_n *and* Γ_n^{-1} *are bounded.*

Proof. Let $\vec{x} = (a_1, a_2, \ldots, a_n)$ be such that $\| \vec{x} \|_2 \leq 1$. Then $| a_i | \leq 1$ for each $i = 1, 2, \ldots, n$.It follows from the Triangle Inequality that

$$\begin{aligned} \|\Gamma_n(\vec{x})\|_V &= \|a_1 v_1 + a_2 v_2 + \cdots + a_n v_n\|_V \\ &\leq \sum_{i=1}^{n} \|v_i\| \end{aligned}$$

This shows that $\|\Gamma_n\| \leq \sum_{i=1}^{n} \|v_i\|$ and hence that Γ_n is bounded.

To see that Γ_n^{-1} is continuous we begin with the following observation. Let

$$S = \{\vec{x} \in (\mathbb{R}^n, \| \cdot \|_2) \mid \| \vec{x} \|_2 = 1\}.$$

Then by the Heine-Borel Theorem S is compact in $(\mathbb{R}^n, \| \cdot \|_2)$. Since Γ_n is continuous, $\Gamma_n(S)$ is compact in $(V, \| \cdot \|_V)$. Since the mapping $v \to \|v\|_v$ is continuous. It follows from the Extreme Value Theorem that

$$\min\{\|\Gamma_n(\vec{x})\|_V \mid \vec{x} \in S\} = \alpha > 0$$

It follows that if $\|v\|_V \leq \alpha$, then $\|\Gamma_n^{-1}(v)\|_2 \leq 1$. From this we can conclude that $\|\Gamma_n^{-1}\| = \frac{1}{\alpha}$. ∎

THEOREM **2.15.6.** *Let* $(V, \| \cdot \|_V)$ *be an* n*-dimensional normed linear space and let* $(W, \| \cdot \|_W)$ *be an* m*-dimensional normed linear space. Let* $T : V \to W$ *be linear. Then* T *is continuous.*

Proof. Let $S : (\mathbb{R}^n, \| \cdot \|_2) \to (\mathbb{R}^m, \| \cdot \|_2)$ be defines by $S = \Gamma_m^{-1} \circ T \circ \Gamma_n$ as in the diagram bellow:

$$\begin{array}{ccc} (V, \| \cdot \|_V) & \xrightarrow{T} & (W, \| \cdot \|_W) \\ \uparrow \Gamma_n & & \downarrow \Gamma_m^{-1} \\ (\mathbb{R}^n, \| \cdot \|_2) & \xrightarrow{S} & (\mathbb{R}^m, \| \cdot \|_2) \end{array}$$

Then S is continuous. But $T = \Gamma_m \circ S \circ \Gamma_n^{-1}$ so T is also continuous.

$$\begin{array}{ccc} (V, \|\cdot\|_V) & \xrightarrow{T} & (W, \|\cdot\|_W) \\ \Big\downarrow \Gamma_n^{-1} & & \Big\uparrow \Gamma_m \\ (\mathbb{R}^n, \|\cdot\|_2) & \xrightarrow{S} & (\mathbb{R}^m, \|\cdot\|_2) \end{array}$$

∎

COROLLARY **2.15.7.** *Let $(V, \|\cdot\|_V)$ be an n-dimensional normed linear space and let $(W, \|\cdot\|_W)$ be any normed linear space. Let $T : V \to W$. be linear. Then T is continuous.*

Proof. Let $W' = T(V) \subseteq W$. Then $(W', \|\cdot\|_W)$ is a finite dimensional normed linear space. Hence $T : V \to W'$ is continuous and as such $T : V \to W$ is continuous.

∎

REMARK **2.15.8.** *Let $(W, \|\cdot\|_W)$ be an n-dimensional normed linear space. Then $\Gamma_n : (\mathbb{R}^n, \|\cdot\|_2) \to (W, \|\cdot\|_W)$ is a homoeomorphism. Moreover, if $v \in V$ then $v = \Gamma_n(\Gamma_n^{-1}(v))$ so*

$$\|v\| \leq \|\Gamma_n\| \|\Gamma_n^{-1}(v)\|_2.$$

We also know that

$$\|\Gamma_n^{-1}(v)\|_2 \leq \|\Gamma_n^{-1}\| \|v\|_V.$$

It follows that if $\alpha = \frac{1}{\|\Gamma_n^{-1}}\|$ and $\beta = \|\Gamma_n\|$, then

$$(*) \quad \alpha \|\Gamma_n^{-1}(v)\|_2 \leq \|v\|_V \leq \beta \|\Gamma_n^{-1}(v)\|_2$$

for every $v \in V$.

We can now deduce the following:

1) *A set $A \subset V$ is closed and bounded in V if and only if $\Gamma_n^{-1}(A)$ is closed and bounded in $(\mathbb{R}^n, \|\cdot\|_2)$. In particular, A is compact if and only if A is closed and bounded.*

2) *A sequence $\{v_n\}$ converges to v_0 in $(V, \|\cdot\|_V)$ if and only if $\{\Gamma_n^{-1}(v_n)\}$ converges to $\Gamma_n^{-1}(v_0)$ in $(\mathbb{R}^n, \|\cdot\|_2)$.*

3) *A sequence $\{v_n\}$ is Cauchy in $(V, \|\cdot\|_V)$ is and only if $\{\Gamma_n^{-1}(v_n)\}$ is Cauchy in $(\mathbb{R}^n, \|\cdot\|_2)$.*

Items 2) and 3) above give us the following important theorem.

THEOREM **2.15.9.** *Let $(V, \|\cdot\|_V)$ be a finite dimensional normed linear space. Then $(V, \|\cdot\|_V)$ is complete. In particular, if $(W, \|\cdot\|_W)$ is any normed linear space, and V is a finite dimensional subspace of W, the V is closed in W.*

The next corollary follws from the previous theorem and the Baire Category Theorem. It's proof is left as an exercise.

COROLLARY **2.15.10.** *Let $(V, \|\cdot\|_V)$ be an infinite dimensional Banach space. Let S be a basis for V. Then S must be uncountable.*

Remark **2.15.11.** *We know that if two metric spaces* (X, d_X) *and* (Y, d_Y) *are homeomorphic then they are essentially the same topological space. It would then make sense to conjecture that if* (X, d_X) *is complete so too would be* (Y, d_Y)*. Unfortunately this is FALSE.*

Let $X = \mathbb{N}$ *and* $Y = \{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\}$ *both with the metric they inherit from* $\mathbb{R}$*. In this case both sets are discrete as metric spaces in the sense that every subset is open in the relative topology. As such* $\phi : X \to Y$ *is a homeomorphism, but* X *is complete while* Y *is not.*

Chapter 3

The Space $(C(X), \|\cdot\|_\infty)$

Throughout this chapter unless otherwise stated (X, d) will be a compact metric space. In this case, the Extreme Value Theorem every continuous function $f : X \to \mathbb{R}$ is bounded. As such we will denote $(C_b(X), \|\cdot\|_\infty)$ simply by $(C(X), \|\cdot\|_\infty)$. Moreover, unless otherwise stated we will always assume that when speaking of $C(X)$ the norm $\|\cdot\|_\infty$ will always be implied.

3.1 Weierstrass Approximation Theorem

We begin with the following problem:

PROBLEM **17.** *Let $f(x)$ be continuous on $[a, b]$. Let*

$$M_0 = \int_0^1 f(t)\, dt$$

and for $n \in \mathbb{N}$ let

$$M_n = \int_0^1 f(t)\, t^n dt.$$

The sequence $\{M_n\}$ are called the moments of $f(x)$.

Do the moments of the function $f(x)$ completely determine $f(x)$ in the sense that if $f(x)$ and $g(x)$ have the same moment sequence, then $f(x) = g(x)$?

In this section we will see that the answer to the previous problem can be deduced from the solution to the following question concerning polynomial approximations of continuous functions. .

PROBLEM **18.** *Given a function $h \in (C([a, b]), \|\cdot\|_\infty)$ and an $\epsilon > 0$, can we find a polynomial $p(x) \in (C([a, b]), \|\cdot\|_\infty)$ so that*

$$\|h(x) - p(x)\|_\infty < \epsilon?$$

REMARK **3.1.1.** *Assume that $f, g \in C([0.1])$ and that $\|f - g\|_\infty < \epsilon$. Then if $\phi : [a, b] \to [0, 1]$ is given by $\phi(x) = \frac{x-a}{b-a}$, we have that $f \circ \phi, g \circ \phi \in C([a, b])$ and*

$$\|f \circ \phi - g \circ \phi\|_\infty = \|f - g\|_\infty.$$

It is actually easy to show that the mapping $\Gamma : (C([0, 1]), \|\cdot\|_\infty) \to (C([a, b]), \|\cdot\|_\infty)$, $\Gamma(f) = f \circ \phi$, is an isometric isomorphism with inverse given by

$$\Gamma^{-1}(h) = h \circ \phi^{-1}$$

for each $h \in C([a,b])$, where $\phi^{-1}(x) = (b-a)x+a$. Moreover it is also easy to see that $\Gamma(p(x))$ is a polynomial if and only if $p(x)$ is as well. It follows that we can aproximate each $f \in C[0,1]$ by a polynomial with error at most $\epsilon > 0$ if and only if we can approximate every $h \in C[a,b]$ by a polynomial with error at most $\epsilon > 0$.

Next observe that if $f \in C([0,1])$ and we can approximate

$$g(x) = f(x) - ([f(1) - f(0)] \cdot x + f(0))$$

uniformly to within $\epsilon > 0$ with a polynomial, then we can do so for $f(x)$ as well.

Before we state our main theorem we will need th following lemma:

LEMMA **3.1.2.** *Let $n \in \mathbb{N}$. Then*

$$(1-x^2)^n \geq 1 - nx^2$$

for all $x \in [0,1]$.

Proof. Let $f(x) = (1-x^2)^n - (1-nx^2)$. Then $f(0) = 0$. Moreover, $f'(x) = 2nx(1-(1-x^2)^{n-1}) > 0$ on $(0,1)$. The result now follows from the Mean Value Theorem.

∎

THEOREM **3.1.3.** *[Weierstrass Approximation Theorem]*

Let $f \in C[a,b]$. Then there exists a sequence $p_n(x)$ of polynomials such that $p_n(x) \to f(x)$ uniformly on $[a,b]$

Proof. First we note that the previous remark shows that without loss of generality we can assume that $[a,b] = [0,1]$ and that $f(0) = 0 = f(1)$.

As such we may extend $f(x)$ to a uniformly continuous function on $\mathbb{R}$ by defining $f(x) = 0$ if $x \in (-\infty, 0] \cup [1, \infty)$.

Now let $Q_n(x) = c_n(1-x^2)^n$ where c_n is chosen so that

$$\int_{-1}^{1} Q_n(x)dx = 1.$$

We have that

$$(1-x^2)^n \geq 1 - nx^2$$

for all $x \in [0,1]$. As such

$$\begin{aligned}
\int_{-1}^{1} (1-x^2)^n dx &= 2\int_0^1 (1-x^2)^n dx \\
&\geq 2\int_0^{\frac{1}{\sqrt{n}}} 1 - nx^2 dx \\
&= \frac{4}{3\sqrt{n}} \\
&> \frac{1}{\sqrt{n}}
\end{aligned}$$

and hence we have

$$c_n < \sqrt{n}.$$

Now if $0 < \delta < 1$, then for each $x \in [-1, \delta] \cup [\delta, 1]$ we have

$$c_n(1-x^2)^n \leq \sqrt{n}(1-\delta^2)^n.$$

Let

$$\begin{aligned} p_n(x) &= \int_{-1}^{1} f(x+t)Q_n(t)dt \\ &= \int_{-x}^{1-x} f(x+t)Q_n(t)dt \\ &= \int_{0}^{1} f(u)Q_n(u-x)du \end{aligned}$$

From Leibniz's rule we have that

$$\frac{d^{2n+1}}{dx^{2n+1}}(p_n(x)) = \int_0^1 f(u)\frac{\partial^{2n+1}}{\partial x^{2n+1}}Q_n(u-x)du = 0.$$

It follows that p_n is a polynomial of degree $2n+1$ or less.

Let $\epsilon > 0$. Let $M = \| f \|_\infty$. Choose $0 < \delta < 1$ so that if $| x - y | < \delta$, then $| f(x) - f(y) | < \frac{\epsilon}{2}$. Now

$$\int_{-1}^{1} Q_n(t)dt = 1 \Rightarrow f(x) = \int_{-1}^{1} f(x)Q_n(t)dt.$$

Moreover, if $x \in [0,1]$,

$$\begin{aligned} | p_n(x) - f(x) | &= | \int_{-1}^{1} [f(x+t) - f(x)]Q_n(t)dt | \\ &\leq \int_{-1}^{1} | f(x+t) - f(x) | Q_n(t)dt \\ &= \int_{-1}^{-\delta} | f(x+t) - f(x) | Q_n(t)dt + \int_{-\delta}^{\delta} | f(x+t) - f(x) | Q_n(t)dt + \int_{\delta}^{1} | f(x+t) - f(x) | Q_n(t)dt \\ &\leq 2M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} + 2M\sqrt{n}(1-\delta^2)^n \\ &= 4M\sqrt{n}(1-\delta^2)^n + \frac{\epsilon}{2} \end{aligned}$$

Hence if we choose n large enough so that $4M\sqrt{n}(1-\delta^2)^n < \frac{\epsilon}{2}$, then

$$\| p_n - f \|_\infty < \epsilon.$$

∎

Corollary **3.1.4.** *Let $f(x) \in C([0,1])$ be such that*

$$\int_0^1 f(t)\, dt = 0$$

and for $n \in \mathbb{N}$ let

$$\int_0^1 f(t)\, t^n dt = 0.$$

Then $f(x) = 0$ for all $x \in [0,1]$

Proof. The proof will be left as an exercise.

∎

Corollary **3.1.5.** *$(C([a,b]), \|\cdot\|_\infty)$ is separable.*

Proof. For each $n \in \mathbb{N}$ let
$$\mathcal{P}_n = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n | a_i \in \mathbb{R}\}.$$
Let
$$\mathcal{Q}_n = \{r_0 + r_1 x + r_2 x^2 + \cdots + r_n x^n | a_i \in \mathbb{Q}\}.$$
Then $\overline{\mathcal{Q}_n} = \mathcal{P}_n$. And since by the Weierstrass Approximation Theorem $\bigcup\limits_{n=1}^{\infty} \mathcal{P}_n$ is dense, so is the countable set $\bigcup\limits_{n=1}^{\infty} \mathcal{Q}_n$.

∎

We will now show that the set of continuous nowhere differentiable functions is residual in $(C([a,b]), \|\cdot\|_\infty)$.

LEMMA **3.1.6.** *Let $n \in \mathbb{N}$. Define*

$\mathcal{F}_n = \{f \in C([0,1]) |$ there exists $x_0 \in [0, 1 - \dfrac{1}{n}]$ such that $|f(x_0 + h) - f(x_0)| \le nh$ for all $0 < h \le 1 - x_0\}$.

Then $\mathcal{F}_n$ is closed in $(C([0.1]), \|\cdot\|_\infty)$ and nowhere dense.

Proof. Assume that $\{f_k\} \subseteq \mathcal{F}_n$ and that $f_k \to f$. For each $k \in \mathbb{N}$ let $x_k \in [0, 1 - \frac{1}{n}]$ be such that
$$(*) \qquad |f(x_k + h) - f(x_k)| \le nh \text{ for all } 0 < h \le 1 - x_k.$$
By replacing $\{x_k\}$ by a subsequence if necessary we can assume without loss of generality that $x_k \to x_0 \in [0, 1 - \frac{1}{n}]$.

Let $0 < h < 1 - x_0$. Then since $x_k \to x_0$, we can find an $N_0 \in \mathbb{N}$ so that if $k \ge N_0$, then $0 < h < 1 - x_k$. Now if $\epsilon > 0$, we can also choose N_0 to be large enough so that if $k \ge N_0$, then

1) $|f(x_0 + h) - f(x_k + h)| < \frac{\epsilon}{4}$ (continuity of f)
2) $|f(x_0) - f(x_k)| < \frac{\epsilon}{4}$ (continuity of f)
3) $\|f_k - f\|_\infty < \frac{\epsilon}{4}$ (uniform convergence)

Then
$$\begin{aligned} |f(x_0+h) - f(x_0)| &\le |f(x_0+h) - f(x_k+h)| + |f(x_k+h) - f_k(x_k+h)| + |f_k(x_k+h) - f_k(x_k)| \\ &\quad + |f_k(x_k) - f(x_k)| + |f(x_k) - f(x_0)| \\ &\le \frac{\epsilon}{4} + \frac{\epsilon}{4} + nh + \frac{\epsilon}{4} + \frac{\epsilon}{4} \\ &= nh + \epsilon \end{aligned}$$

Since $\epsilon > 0$ was arbitary, we have
$$|f(x_0 + h) - f(x_0)| \le nh$$
for all $0 < h < 1 - x_0$ and $f \in \mathcal{F}_n$.

To see that $\mathcal{F}_n$ is nowhere dense let $f \in (C([0.1]), \|\cdot\|_\infty)$ and let $\epsilon > 0$. We know that we can find a polynomial $p(x)$ so that $\|f - p\|_\infty < \frac{\epsilon}{2}$.

Let
$$\varphi(x) = \begin{cases} x & \text{if } x \in [0,1] \\ 2 - x & \text{if } x \in [1,2] \end{cases} .$$
and then extend φ to all of $\mathbb{R}$ by letting $\varphi(x+2) = \varphi(x)$.

Let
$$g(x) = \sum_{n=0}^{\infty} (\frac{3}{4})^n \varphi(4^n x).$$

Let $F = g_{|_{[0,1]}}$. Pick $\alpha > 0$ so that $\|\alpha F\|_\infty < \frac{\epsilon}{2}$. Then $p(x) + \alpha F(x) \in \mathcal{F}_n^c$ for each $n \in \mathbb{N}$, and $\|f - (p(x) + \alpha F(x))\|_\infty < \epsilon$.

∎

THEOREM **3.1.7.** *[Banach-Mazurkiewicz Theorem]*
The set $\mathcal{ND}([0,1])$ of continuous nowhere differentiable functions is residual in $(C([0.1]), \|\cdot\|_\infty)$

Proof. This follows immediately from the Baire Category Theorem and the observation that

$$\mathcal{ND}([0,1]) \subseteq \bigcap_{n=1}^{\infty} \mathcal{F}_n.$$

∎

REMARK **3.1.8.** *In the Banach-Mazurkiewicz Theorem there is nothing special about $[0,1]$. In particular, the same result holds for $[a,b]$.*

3.2 Stone-Weierstrass Theorem

Throughout this section, unless stated otherwise, (X,d) will be a compact metric space.

In the previous section we saw that the collection

$$\mathcal{P} = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n | a_i \in \mathbb{R}, n \in \mathbb{N} \cup \{0\}\}.$$

of all polynomials is dense in $(C([a,b]), \|\cdot\|_\infty)$. In this section we will identify other classes of functions that can be shown to be dense in $(C(X), \|\cdot\|_\infty)$ when (X,d) is a compact metric space.

DEFINITION **3.2.1.** *Let (X,d) be a compact metric space. Let $\Phi \subseteq C(X)$. We say that Φ is point separating if whenever $x, y \in X$ with $x \neq y$, there exists an $f \in \Phi$ such that $f(x) \neq f(y)$.*

REMARK **3.2.2.** *1) (X,d) is a compact metric space. Let $a, b \in X$, with $a \neq b$. Then the function $f(x) = d(x,a)$ is such that $f \in C(X)$ and $f(a) \neq f(b)$. As such $C(X)$ is point separating*

2) Assume that (X,d) is a compact metric space with at least two points $x \neq y$. If $\Phi \subseteq C(X)$ is such that $f(x) = f(y)$ for every $f \in \Phi$. then $g(x) = g(y)$ for every $g \in \overline{\Phi}$. This shows that if Φ is dense in $C(X)$ it must be point separating.

DEFINITION **3.2.3.** *A linear subspace $\Phi \subseteq C(X)$ is called a lattice if for every $f, g \in \Phi$ we have $f \vee g \in \Phi$ and $f \wedge g \in \Phi$ where*

$$f \vee g(x) = \max\{f(x), g(x)\}$$

and

$$f \wedge g(x) = \min\{f(x), g(x)\}.$$

REMARK **3.2.4.** *1) Let $f, g \in C(X)$. Then*

$$f \vee g(x) = \frac{(f(x) + g(x)) + |f(x) - g(x)|}{2}$$

and

$$f \wedge g(x) = -(-f \vee -g)(x).$$

It follows that both $f \vee g$ and $f \wedge g$ are in $C(X)$. That is $C(X)$ is a lattice. Moreover, if $\Phi \subseteq C(X)$ is a linear subspace, then Φ is a lattice if $f \vee g \in \Phi$ for every $f, g \in \Phi$.

EXAMPLE **3.1.** *A function $f : [a,b] \to \mathbb{R}$ is said to be piecewise linear if there is a partition $\mathcal{P} = \{a = t_0 < t_1 < t_2 < \cdots < t_n = b\}$ of $[a,b]$ such that*

$$f_{|_{[t_{i-1},t_i]}} = m_i x + d_i.$$

It is piecewise polynomial if

$$f_{|_{[t_{i-1},t_i]}} = c_{0,i} + c_{1,i}x + c_{2,i}x^2 + \cdots + c_{n_i,i}x^{n_i}.$$

Let

$$\Phi_1 = \{f \in C([a,b]) | f \text{ is piecewise linear}\}$$

and

$$\Phi_2 = \{f \in C([a,b]) | f \text{ is piecewise polynomial}\}.$$

Both Φ_1 and Φ_2 are lattices.

Note: *If f is piecewise polynomial, then the order of f is the highest degree of any of the individual polynomials. f is called a spline of order k if it has order k as a piecewise polynomial, and if it is $k-1$-times differentiable at each t_i for $i = 1, 2, \ldots, n$.*

We have just seen that the collection of either piecewise linear or piecewise polynomial continuous functions are lattices in $C([a,b])$. We will now see that they are both dense in $(C([a,b]), \|\cdot\|_\infty)$.

THEOREM **3.2.5.** *[Stone-Weierstrass Theorem: Lattice Version]*

Let (X,d) be a compact metric space and let Φ be a linear subspace of $(C(X), \|\cdot\|_\infty)$ such that

1) The constant function $1 \in \Phi$.

2) Φ separates points.

3) If $f, g \in \Phi$, then so is $f \vee g$.

Then Φ is dense in $(C(X), \|\cdot\|_\infty)$.

Proof. We begin with the following observation: Let $\alpha, \beta \in \mathbb{R}$ and let $x \neq y \in X$. If $\phi(t)$ is a function in Φ with $\phi(x) \neq \phi(y)$ then the function

$$g(t) = \alpha + (\beta - \alpha)\frac{\phi(t) - \phi(x)}{\phi(y) - \phi(x)}$$

is such that $g \in \Phi$, $g(x) = \alpha$ and $g(y) = \beta$.

Let $\epsilon > 0$ and let $f \in C(X)$.

Step 1: Fix $x \in X$. For each $y \in X$ there exists $\phi_{x,y}(t) \in \Phi$ such that $\phi_{x,y}(x) = f(x)$ and $\phi_{x,y}(y) = f(y)$. Now for any $y \in X$, since $\phi_{x,y}(y) - f(y) = 0$, we can find a $\delta_y > 0$ such that for every $t \in B(y, \delta_y)$ we have

$$-\epsilon < \phi_{x,y}(t) - f(t) < \epsilon.$$

By compactness, we can find finitely many points $\{y_1, y_2, \ldots, y_n\}$ so that $\{B(y_i, \delta_{y_i})\}$ is a cover of X. Now let

$$\phi_x(t) = \phi_{x,y_1} \vee \phi_{x,y_2} \vee \cdots \vee \phi_{x,y_n} \in \Phi.$$

Now if $z \in X$, then $z \in B(y_i, \delta_{y_i})$ for some i and as such

$$f(z) - \epsilon < \phi_{x,y_i}(z) \leq \phi_x(z).$$

Step 2: For each $x \in X$ we have

$$\phi_x(x) - f(x) = 0.$$

As before we can find for each $x \in X$ a $\delta_x > 0$ so that if $t \in B(y, \delta_x)$ we have

$$-\epsilon < \phi_x(t) - f(t) < \epsilon.$$

And as before we can find $\{x_1, x_2, \ldots, x_k\}$ so that $\{B(x_j, \delta_{x_j})\}$ is a cover of X. This time if we let

$$\phi(t) = \phi_{x_1} \wedge \phi_{x_2} \wedge \cdots \wedge \phi_{x_k} \in \Phi,$$

then for any $z \in X$ we have

$$f(z) - \epsilon < \phi(z) < f(z) + \epsilon.$$

∎

Corollary **3.2.6.** *Let*

$$\Phi_1 = \{f \in C([a,b]) | f \text{ is piecewise linear}\}$$

and

$$\Phi_2 = \{f \in C([a,b]) | f \text{ is piecewise polynomial}\}.$$

Then both Φ_1 and Φ_2 are dense in $(C([a,b]), \|\cdot\|_\infty)$.

Proof. Both Φ_1 and Φ_2 are lattices containing the constant function 1. Since both also contain the function $f(x) = x$, both satisfy all three conditions for the Stone-Weierstrass Theorem.

∎

The second version of the Stone-Weierstrass Theorem can be viewed as a generalization of the Weirstrass Approximation Theorem.

Definition **3.2.7.** *A subspace $\Phi \subseteq C(X)$ is said to be a subalgebra if $f \cdot g \in \Phi$ for every $f, g \in \Phi$.*

Remark **3.2.8.** *1) Let $\mathcal{P} = \{c_0 + c_x + c_2x^2 + \cdots + c_nx^n |\ c_i \in \mathbb{R}, n = 0, 1, 2, \ldots\}$ be the collection of all polyomials. Then $\mathcal{P}$ is a sublagebra of $C([a,b])$.*

2) Assume that $\Phi \subseteq C(X)$ is a subalgebra. Then so is $\bar{\Phi}$. To see this let $\{f_n\}, \{g_n\} \subseteq \Phi$ with $f_n \to f$ and $g_n \to g$. It is clear that $f_n + g_n \to f + g$ and $\alpha f_n \to \alpha f$. Moreover, it is also clear that $f \cdot g \in C(X)$. Note that since $\{g_n\}$ is bounded we have that

$$\begin{aligned} \|f_n g_n - fg\|_\infty &= \|(f_n g_n - f g_n) + (f g_n - fg)\|_\infty \\ &\leq \|f_n g_n - f g_n\|_\infty + \|f g_n - fg\|_\infty \\ &= \|g_n\|_\infty \|f_n - f\|_\infty + \|f\|_\infty \|g_n - g\|_\infty \to 0. \end{aligned}$$

Theorem **3.2.9.** *[Stone-Weierstrass Theorem: Subalgebra Version]*
Let (X, d) be a compact metric space and let Φ be a linear subspace of $(C(X), \|\cdot\|_\infty)$ such that

1) The constant function $1 \in \Phi$.

2) Φ separates points.

3) If $f, g \in \Phi$, then so is fg.

Then Φ is dense in $(C(X), \|\cdot\|_\infty)$.

Proof. We first observe that since $\bar{\Phi}$ also satisfies the three conditions above we may assume without loss of generality that Φ is closed.

Let $f \in \Phi$ and let $\epsilon > 0$.

Choose $M > 0$ so that $|f(x)| \leq M$ for each $x \in X$. Then we know from the Weierstrass Approximation theorem that there is a polynomial

$$p(t) = c_0 + c_1 t + c_2 t^2 + \cdots + c_n t^n$$

so that for each $t \in [-M, M]$, we have

$$||t| - p(t)| < \epsilon.$$

Next observe that

$$p \circ f = c_0 1 + c_1 f + c_2 f^2 + \cdots + c_n f^n \in \Phi$$

and that for each $x \in X$, we have

$$||f(x)| - p(f(x))| < \epsilon.$$

Since ϵ is arbitrary it follows that $|f| \in \bar{\Phi} = \Phi$.

Finally since if $f, g \in \Phi$ then

$$f \vee g(x) = \frac{(f(x) + g(x)) + |f(x) - g(x)|}{2}$$

we see that Φ is also a Lattice and hence by the Lattice version of the Stone-Weierstrass Theorem, $\Phi = C(x)$. ∎

So far we have focused entirely on real valued functions. However we could also consider

$$C(X, \mathbb{C}) = \{f : X \to \mathbb{C}|\ f(x) \text{ is continuous on } X\}$$

with the norm

$$\|f\|_\infty = sup\{|f(x)||x \in X\}.$$

Remark **3.2.10.** *We say that a subspace $\Phi \subseteq C(X, \mathbb{C})$ is self-adjoint if $f \in \Phi$ implies $\bar{f} \in \Phi$. We also note that if Φ is self adjoint then for each $f \in \Phi$*

$$Re(f) = \frac{f + \bar{f}}{2} \in \Phi$$

and

$$Im(f) = \frac{f - \bar{f}}{2i} \in \Phi.$$

This observation allows one to deduce the following complex version of the Stone-Weierstrass Theorem:

Theorem **3.2.11.** *[Stone-Weierstrass Theorem: Complex Version]*

Let (X, d) be a compact metric space and let Φ be a self-adjoint linear subspace of $(C(X, \mathbb{C}), \|\cdot\|_\infty)$ such that

1) The constant function $1 \in \Phi$.

2) Φ separates points.

3) If $f, g \in \Phi$, then so is fg.

Then Φ is dense in $(C(X, \mathbb{C}), \|\cdot\|_\infty)$.

EXAMPLE **3.2.** *1) Let $\Pi = \{\lambda \in \mathbb{C} | \ |\lambda| = 1\}$. Define a map $\psi : \Pi \to [0, 2\pi)$ by*

$$\psi(e^{i\theta}) = \theta$$

Define the metric on $[0, 2\pi)$ to be

$$d_*(\theta_1, \theta_2) \overset{def}{=} \text{the shortest arc length between } e^{i\theta_1} \text{ and } e^{i\theta_2}.$$

Then ψ is a homeomorphism between Π with the usual metric and $([0, 2\pi), d_*)$. Moreover, $([0, 2\pi), d_*)$ is compact with respect to this metric. We also have

$$C(\Pi) \cong C([0, 2\pi)) = \{f \in C([0, 2\pi]) | f(0) = f(2\pi)\}$$

A trigonometric polynomial is an element of

$$Trig([0, 2\pi)) = span\{1, sin(nx), cos(mx) | n, m \in \mathbb{N}\}$$

Then $Trig([0, 2\pi))$ is a point separating subalgebra of $C([0, 2\pi))$. In particular, $Trig([0, 2\pi))$ is dense in $C([0, 2\pi))$.

In the complex world we have

$$Trig_{\mathbb{C}}([0, 2\pi)) = span\{f(\theta) = e^{in\theta} | n \in \mathbb{Z}\}.$$

In this case $Trig_{\mathbb{C}}([0, 2\pi))$ is a self-adjoint, point separating subalgebra of $C([0, 2\pi), \mathbb{C})$. In particular, $Trig_{\mathbb{C}}([0, 2\pi))$ is dense in $C([0, 2\pi), \mathbb{C}) \cong C(\Pi, \mathbb{C})$.

2) Let

$$\Psi = \{F(x, y) \in C([0, 1] \times [0, 1] \mid F(x, y) = \sum_{i=1}^{k} f_i(x) g_i(y)\}$$

where in the sum above the functions f_i and g_i are continuous on $[0, 1]$. Then Ψ is dense in $C([0, 1] \times [0, 1])$. (Exercise)

3.3 Compactness in $(C(X), \|\cdot\|_\infty)$ and the Ascoli-Arzela Theorem

This section has not been proof readed for typos.

The central problem of this section is the following:

PROBLEM **19.** *Can we characterise the compact subsets of $(C(X), \|\cdot\|_\infty)$?*

In fact we will look not at the compact subsets of $(C(X), \|\cdot\|_\infty)$, but rather at those subsets with compact closure. This leads us to the following definition:

DEFINITION **3.3.1.** *Let (X, d) be a metric space. We say that $A \subset X$ is relatively compact if $\bar{A}$ is compact.*

REMARK **3.3.2.** *Assume that (X, d) is complete. Then since the closure of a totally bounded set A is also totally bounded, we have that $A \subset X$ is relatively compact if and only if A is totally bounded.*

DEFINITION **3.3.3.** *Let (X, d) be a metric space. Let $\mathcal{F} \subseteq C_b(X)$. Let $x_0 \in X$. We say that $\mathcal{F}$ is equicontinuous at x_0 if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $d(x, x_0) < \delta$, then*

$$|f(x) - f(x_0)| < \epsilon$$

for every $f \in \mathcal{F}$.

We say that $\mathcal{F}$ is equicontinuous if $\mathcal{F}$ is equicontinuous at each $x_0 \in X$.

We say that $\mathcal{F}$ is uniformly equicontinuous if for every $\epsilon > 0$ there exists a $\delta > 0$ such that if $d(x, y) < \delta$, then

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

for every $f \in \mathcal{F}$.

REMARK **3.3.4.** *Let $\mathcal{F} \subseteq C(X)$ be finite. Then it is clear that $\mathcal{F}$ is equicontinuous.*

We know that know that if (X, d) is compact and if $f \in C(X)$, then f is uniformly continuous. We can now show that this uniform behaviour extends to an equicontinuous family. The proof of this result also provides an alternative way of showing that continous functions on compact sets are uniformly continuous.

PROPOSITION **3.3.5.** *Let (X, d) be a compact metric space and let $\mathcal{F} \subseteq C(X)$ be equicontinuous. Then $\mathcal{F}$ is uniformly equicontinuous.*

Proof. Let $\epsilon > 0$. For each $x_0 \in X$ there exists $\delta_{x_0} > 0$ such that if $d(x, x_0) < \delta$, then

$$|f(x) - f(x_0)| < \frac{\epsilon}{2}$$

for every $f \in \mathcal{F}$.

Since $\{B(x_0, \delta_{x_0})\}_{x_0 \in X}$ is a cover of X, there is a $\delta_0 > 0$ such that for any $y \in X$ we can find a point $x_0 \in X$ so that $B(y, \delta_0) \subseteq B(x_0, \delta_{x_0})$. In particular, if $z \in B(y, \delta_0)$, then

$$\begin{array}{rcl} |f(y) - f(z)| & \leq & |f(y) - f(x_0)| + |f(x_0) - f(z)| \\ & < & \dfrac{\epsilon}{2} + \dfrac{\epsilon}{2} \\ & = & \epsilon. \end{array}$$

∎

DEFINITION **3.3.6.** *Let $\mathcal{F} \subseteq C(X)$. We say that $\mathcal{F}$ is pointwise bounded if $\{f(x_0) | f \in \mathcal{F}\} \subseteq \mathbb{R}$ is bounded for each $x_0 \in X$.*

PROPOSITION **3.3.7.** *Let (X, d) be a compact metric space. Let $\mathcal{F} \subseteq C(X)$ be equicontinuous and pointwise bounded. Then $\mathcal{F}$ is uniformly bounded.*

Proof. We know that $\mathcal{F}$ is uniformly equicontinuous. As such we can find a $\delta > 0$ such that if $d(x, y) < \delta$, then

$$|f(x) - f(y)| < 1$$

for all $f \in \mathcal{F}$. Now let $\{x_1, x_2, \ldots, x_n\}$ be a δ-net for X and assume that

$$|f(x_i)| \leq M_i$$

for each $f \in \mathcal{F}$. Let $M_0 = max\{M_1, M_2, \ldots, M_n\}$. If $x \in X$, then there exists an x_i with $d(x, x_i) < \delta$ so that

$$|f(x)| \leq |f(x) - f(x_i)| + |f(x_i)| \leq 1 + M_0.$$

∎

We are now ready to give our characterization of relatively compact subsets of $(C(X), \|\cdot\|_\infty)$.

Theorem **3.3.8.** *[Arzelá-Ascoli]*

Let (X,d) be a compact metric space and let $\mathcal{F} \subseteq (C(X), \|\cdot\|_\infty)$. Then the following are equivalent:

1) $\mathcal{F}$ is relatively compact.

2) $\mathcal{F}$ is equicontinuous and pointwise bounded.

Proof. 1) $\Rightarrow$ 2). If $\mathcal{F}$ is relatively compact then it is bounded. As such it is clearly pointwise bounded.

Next let $\epsilon > 0$. Since $\mathcal{F}$ is relatively compact, it is totally bounded. As such there exists a finite $\frac{\epsilon}{3}$-net $\{f_1, f_2, \ldots, f_n\} \subseteq \mathcal{F}$. Since $\{f_1, f_2, \ldots, f_n\}$ is uniformly equicontinuous, we can find a $\delta > 0$ such that if $d(x,y) < \delta$, then

$$|f_i(x) - f_i(y)| < \frac{\epsilon}{3}$$

for all $i = 1, 2, \ldots, n$.

Now assume that $d(x,y) < \delta$ and $f \in \mathcal{F}$. Then we can find an $i_o \in \{1, 2, \ldots, n\}$ so that $\|f - f_{i_0}\|_\infty < \frac{\epsilon}{3}$. Hence

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_{i_0}(x)| + |f_{i_0}(x) - f_{i_0}(y)||f_{i_0}(y) - f(y)| \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon. \end{aligned}$$

2) $\Rightarrow$ 1). Since (X,d) is compact, we have that $\mathcal{F}$ is uniformly equicontinous and uniformly bounded. That is there is an $M > 0$ so that $f(x) \in [-M, M]$ for each $f \in \mathcal{F}$ and $x \in X$.

Let $\epsilon > 0$. Let $\mathcal{P} = \{-M = y_0 < y_1 < \cdots < y_m = M\}$ be a partition of $[-M, M]$ with

$$\|\mathcal{P}\| = \max_{j=1,2,\ldots,m} \{y_j - y_{j-1}\} < \frac{\epsilon}{3}.$$

Now since $\mathcal{F}$ is uniformly equicontinuous we can find a $\delta > 0$ such that if $d(x,z) < \delta$, then

$$|f(x) - f(z)| < \frac{\epsilon}{3}$$

for all $f \in \mathcal{F}$. Let $\{x_1, x_2, \ldots, x_n\}$ be a δ-net for X.

Let

$$\Phi = \{\sigma | \sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, m\}\}.$$

Then

$$\Phi = \{\sigma_1, \sigma_2, \ldots, \sigma_l\}$$

where $l = m^n$. In particular, Φ is finite.

Next, for each $k = 1, 2, \ldots, l$, let

$$\mathcal{F}_k = \{f \in \mathcal{F} | f(x_i) \in [y_{\sigma_k(i)-1}, y_{\sigma_k(i)}] \ \text{ for all } \ i = 1, 2, \ldots, n\}.$$

Note that some $\mathcal{F}_k$'s may be empty but

$$\mathcal{F} = \bigcup_{k=1}^{l} \mathcal{F}_k.$$

For each of the non-empty sets choose $f_k \in \mathcal{F}_k$. We claim that $\{f_k\}$ is a finite ϵ-net for $\mathcal{F}$. In fact, let $f \in \mathcal{F}$ and let $w \in X$. Then $f \in \mathcal{F}_k$ for some k and $w \in B(x_i, \delta)$ for some i. It follows that

$$\begin{aligned} |f(w) - f_k(w)| &\leq |f(w) - f(x_i)| + |f(x_i) - f_k(x_i)| + |f_k(x_i) - f_k(w)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon. \end{aligned}$$

In particular, $\|f - f_k\|_\infty < \epsilon$.

■

Definition **3.3.9.** *[Compact Operators]*

We say that a linear map $\Gamma : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ is compact if $\Gamma(B_X[0,1])$ is relatively compact in Y.

Example **3.3.** *Consider $(X, \|\cdot\|_X) = (C([a,b]), \|\cdot\|_\infty) = (Y, \|\cdot\|_Y)$. Let $K : [a,b] \times [a,b] \to [a,b]$ be continuous. Define for each $f \in C([a,b]$.*

$$\Gamma(f)(x) = \int_a^b K(x,y)f(y)dy.$$

We claim that $\Gamma(f) \in C([a,b])$. This is clear if $f = 0$. Otherwise, we observe that since K is uniformly continuous, given $\epsilon > 0$ we can find a $\delta > 0$ such that if $\|(x_1,y_1) - (x_2,y_2)\|_2 < \delta$ then

$$|K(x_1,y_1) - K(x_2,y_2)| < \frac{\epsilon}{(b-a)\|f\|_\infty}.$$

Hence if if $|x - z| < \delta$, then

$$\begin{aligned}
|\Gamma(f)(x) - \Gamma(f)(z)| &= |\int_a^b [K(x,y) - K(z,y)]f(y)dy| \\
&\le \int_a^b |K(x,y) - K(z,y)|\|f\|_\infty dy \\
&\le \int_a^b \frac{\epsilon\|f\|_\infty}{(b-a)\|f\|_\infty} dy \\
&= \epsilon
\end{aligned}$$

It is also clear that $\Gamma : C([a,b]) \to C([a,b])$ is a linear map.

Moreover, if we let $\epsilon > 0$, we can chose $\delta_1 > 0$ such that if $|x - z| < \delta_1$, then

$$|K(x,y) - K(z,y)| < \frac{\epsilon}{b-a}$$

for every $y \in [a,b]$. Now let $|x - z| < \delta_1$. Then for any $f \in C([a,b])$ with $\|f\|_\infty \le 1$, we have

$$\begin{aligned}
|\Gamma(f)(x) - \Gamma(f)(z)| &= |\int_a^b [K(x,y) - K(z,y)]f(y)dy| \\
&\le \int_a^b |K(x,y) - K(z,y)|\|f\|_\infty dy \\
&\le \int_a^b \frac{\epsilon}{b-a} dy \\
&= \epsilon
\end{aligned}$$

This shows that $\Gamma(B_X[0,1])$ is uniformly equicontinuous.

Finally, if $f \in C([a,b])$ with $\|f\|_\infty \le 1$, then

$$|\Gamma(f)(x) = |\int_a^b K(x,y)f(y)dy| \le \int_a^b M dy = M(b-a)$$

where M is chosen so that $|K(x,y)| \le M$ for each $(x,y) \in [a,b] \times [a,b]$. In particualr, $\Gamma(B_X[0,1])$ is also uniformly bounded. By the Arzelá-Ascoli Theorem Γ is compact.

The next theorem establishes one of the fundamental applications of the Arzelá-Ascoli Theorem.

THEOREM **3.3.10.** *[Peano's Theorem]*
Let f be continuous on an open subset D of $\mathbb{R}^2$ and let $(x_0, y_0) \in D$. Then the differential equation

$$y' = f(x, y)$$

has a local solution passing through the point (x_0, y_0).

Proof. Let R be a closed rectangle contained in D with (x_0, y_0) in its interior. Since $f(x, y)$ is continuous on the compact set R, there exists an $M \geq 1$ such that $\mid f(x, y) \mid \leq M$ on R. Let

$$W = \{(x, y) \in R : \mid y - y_0 \mid \leq M \mid x - x_0 \mid\}.$$

Let $I = [a, b] = \{x : (x, y) \in W \text{ for some } y\}$.

By uniform continuity, given an $\epsilon > 0$, there exists a $0 < \delta < 1$ such that if $(x_1, y_1), (x_2, y_2) \in W$ with $\mid x_1 - x_2 \mid < \delta$ and $\mid y_1 - y_2 \mid < \delta$, then $\mid f(x_1, y_1) - f(x_2, y_2) \mid < \epsilon$. Choose $x_0 < x_1 < \cdots < x_n = b$ with

$$\mid x_i - x_{i-1} \mid < \frac{\delta}{M}.$$

Define a function k_ϵ on $[x_0, b]$ by

i) $k_\epsilon(x_0) = y_0$ and on $[x_0, x_1]$, $k_\epsilon(x)$ is linear with slope $f(x_0, y_0)$;
ii) on $[x_1, x_2], k_\epsilon(x)$ is linear with slope $f(x_1, k_\epsilon(x_1))$;

and proceed in this manner to define the piecewise linear function $k_\epsilon(x)$ on $[x_0, b]$. Thus the graph of $k_\epsilon(x)$ is a polygonal arc that is contained in W with

$$(*) \quad \mid k_\epsilon(x) - k_\epsilon(\overline{x}) \mid \leq M \mid x - \overline{x} \mid$$

for each $x, \overline{x} \in [x_0, b]$. Let $x \in [x_0, b]$ with $x \neq x_i$ for any i. Then there exists a j such that $x_{j-1} < x < x_j$. However since $\mid x_j - x_{j-1} \mid < \frac{\delta}{M}$,

$$\mid k_\epsilon(x) - k_\epsilon(x_{j-1}) \mid \leq M \mid x - x_{j-1} \mid < \delta.$$

And hence that

$$\mid f(x_{j-1}, k_\epsilon(x_{j-1})) - f(x, k_\epsilon(x)) \mid < \epsilon.$$

Therefore since $k'_\epsilon(x_{j-1}) = f(x_{j-1}, k_\epsilon(x_{j-1}))$

$$(**) \quad \mid k'_\epsilon(x_{j-1}) - f(x, k_\epsilon(x)) \mid < \epsilon.$$

We have that the inequality $(**)$ holds but for at most finitely many points in $[x_0, b]$.

Let $\mathcal{K} = \{k_\epsilon(x) : \epsilon > 0\}$. Then $\mathcal{K}$ is pointwise bounded since each function has a graph contained in W. Moreover, by $(*)$ $\mathcal{K}$ is also equicontinuous. It follows that $\overline{\mathcal{K}}$ is compact in $C[x_0, b]$. Furthermore for each $x \in [x_0, b]$

$$\begin{aligned} (***) \quad k_\epsilon(x) &= y_0 + \int_{x_0}^{x} k'_\epsilon(t)dt \\ &= y_0 + \int_{x_0}^{x} (f(t, k_\epsilon(t)) + (k'_\epsilon(t) - f(t, k_\epsilon(t)))dt. \end{aligned}$$

By compactness, the sequence $k_{\frac{1}{n}}(x)$ has a subsequence $k_{\frac{1}{n_k}}(x)$ that converges uniformly on $[x_0, b]$ to some $k(x)$. Since f is uniformly continuous on W, $\{f(t, k_{\frac{1}{n_k}}(t))\}$ converges uniformly to $f(t, k(t))$ on $[x_0, b]$. Finally, $(***)$ shows that

$$k(x) = y_0 + \int_{x_0}^{x} f(t, k(t))dt$$

for all $x \in [x_0, b]$ and hence that $k(x)$ is a solution to the DE on $[x_0, b]$. A similar argument can be used to get a solution $k^*(x)$ on $[a, x_0]$. ∎