REAL MATHEMATICAL ANALYSIS

Lectures by Niranjan Balachandran, IIT Bombay.

Contents

Preface

Real Analysis is all about formalizing and making precise, a good deal of the intuition that resulted in the basic results in Calculus. As it turns out, the intuition is spot on, in several instances, but in some cases (and this is really why Real Analysis is important at all), our sense of intuition is so far from reality, that one needs some kind of guarantee, or validation to our heuristic arguments.

I thank all my students of this course who very actively and enthusiastically acted as scribes for the lectures for this course.

1 Preliminaries: The Real Line

As mentioned in the preface, in order to formalize the results one has studied in a first Calculus course, one needs to start at the very beginning, in order to ensure that there is no inconsistency in our settings. And we shall begin at the very beginning - with the Natural numbers.

It may be a bit difficult to formally define what the natural numbers actually are. But, we shall avoid doing this, by making our approach axiomatic. The more important thing is to ensure that the axioms are no contradictory. As it turns out, very basic axioms about the natural numbers are sufficient to set up our understanding of the natural numbers in a wholesome manner.

1.1 Relations

Definition 1 *Suppose S is a set,then the **Relation** R on S is defined as a subset of*

$$S \times S := \{(s_1, s_2)|s_1, s_2 \in S\}$$

If $(s_1, s_2) \in R$ we denote it as $s_1 \sim s_2$.

Definition 2 1. *A relation R is **Reflexive** if $a \sim a\ ,\forall\ a \in S$*

2. *A relation R is **Symmetric** if $a \sim b\ \Rightarrow a \sim b\ ,\forall\ a, b \in S$*

3. *A relation R is **Transitive** if $a \sim b\ , b \sim c\ \Rightarrow a \sim c\ ,\forall\ a, b, c \in S$*

Definition 3 ***(Equivalence Relation)***
A relation is an Equivalence relation iff it is Reflexive, Symmetric and Transitive.

Proposition 1 *Suppose $\sim$ is an equivalence relation defined on a set S, then $\sim$ induces a partition on S. Conversely, for any partition $\Pi = \{S_\alpha\}_{\alpha\in\lambda}$ of the set S, there is an equivalence relation $\sim_\Pi$ which induces partition Π on S.*

Proof: Given $\sim$, an equivalence relation, for each $a \in S$,define

$$S_a = \{b \in S | b \sim a\}$$

Claim 1 *$\{S_a\}_{a \in S}$ is a partition of S.*

Let $a \neq b$, where $a, b \in S$. Consider $S_a \cap S_b$. Suppose $S_a \cap S_b \neq \Phi$. We shall show that $S_a \subseteq (S_a \cap S_b)$. Similarly, $S_b \subseteq (S_a \cap S_b)$ and these will prove that $S_a = S_b = (S_a \cap S_b)$.

Since $S_a \cap S_b \neq \Phi$, there exists some $c \in (S_a \cap S_b)$. Let $x \in S_a$. Then we have $x \sim a$. However, since $c \in S_a$, we have $c \sim a \Rightarrow a \sim c$, by symmetry. Therefore, by transitivity, $x \sim c$. And since $c \sim b$, we also get $x \sim b \Rightarrow x \in S_b$. Now that $x \in S_a$ and $x \in S_b$, $x \in (S_a \cap S_b)$. Since this is true for any $x \in S_a$, we have $S_a \subseteq (S_a \cap S_b)$. Similarly, we can show that $S_b \subseteq (S_a \cap S_b)$. Thus, $S_a = S_b = (S_a \cap S_b)$ which is a partition of S induced by $\sim$.

Converse: Given $\Pi = \{S_\alpha\}_{\alpha \in \lambda}$, a partition of S, we define a relation $\sim_\Pi$ as follows:

$$a \sim_\Pi b \; if \; \exists \; \alpha \in \lambda \; such \; that \; a, b \in S_\alpha$$

Now, we prove that it is an equivalence relation on S.

Reflexivity: $(a \sim_\Pi a) \; \forall a \in S$ as it is in the same partition as itself.
Symmetry: $a \sim_\Pi b \Rightarrow a$ is in the same partition as $b \Rightarrow b$ is in the same partition as $a \Rightarrow b \sim_\Pi a$
Transitivity: $a \sim_\Pi b, \Rightarrow a, b \in S_\alpha$ and $b \sim_\Pi c, \Rightarrow b, c \in S_\beta$. But since S has been partitioned by Π and b belongs to both S_α and S_β, we must have $S_\alpha = S_\beta$. Hence a and c belong to S_α and are related under $\sim_\Pi$, thus proving the transitivity.
These three properties put together make $\sim_\Pi$ an equivalence relation defined on S. ♣

1.2 Natural Numbers

1.2.1 Axioms for Natural Numbers $\mathbb{N}$

1. $0 \in \mathbb{N}$.
2. For each $n \in \mathbb{N}$ there is a unique successor for n, denoted by $n + 1$.
3. If $S \subseteq \mathbb{N}$,satisfying
   - $0 \in \mathbb{N}$.
   - $n \in \mathbb{N} \Rightarrow n + 1 \in \mathbb{N}$.

   then $S = \mathbb{N}$. This is referred to as **The Principle of Mathematical Induction**.
4. For each $n \in \mathbb{N} \setminus \{0\}$, there exists a unique $m \in \mathbb{N}$ such that $m + 1 = n$.

1.2.2 Addition and Multiplication

We can define two operations on $\mathbb{N}$ called 'addition, multiplication' that satisfies the following. We skip the proofs.

Theorem 1 *Given $m, n \in \mathbb{N}$,there exists a binary operation '+' on $\mathbb{N}$ satisfying:*

1. $0 + n = n$, $\forall\ n \in \mathbb{N}$.
2. $m + n = n + m$, $\forall m, n \in \mathbb{N}$. *(Commutative Property)*
3. $m + (n + p) = (m + n) + p$, $\forall m, n, p \in \mathbb{N}$. *(Associative Property)*

♣

Theorem 2 *Given $m, n, p \in \mathbb{N}$,there exists a binary operation '.' on $\mathbb{N}$ satisfying:*

1. $1.n = n\ \forall n \in \mathbb{N}$.
2. $m.n = n.m$, $\forall m, n \in \mathbb{N}$. *(Commutative Property)*
3. $m.(n.p) = (m.n).p$, $\forall m, n, p \in \mathbb{N}$. *(Associative Property)*
4. $m.(n + p) = m.n + m.p$, $\forall m, n, p \in \mathbb{N}$. *(Distributive Property)*

♣

Example 1 ***Prove that 2+2=4.***

Proof:

$$\begin{aligned} 2 + 2 =& 2 + (1 + 1) \ldots (since\ 2\ is\ successor\ of\ 1), \\ =& (2 + 1) + 1 \ldots (Associative\ property), \\ =& 3 + 1 \ldots (3\ is\ successor\ of\ 2), \\ =& 4 \ldots (4\ is\ successor\ of\ 3). \end{aligned}$$

♣

1.2.3 Order on $\mathbb{N}$

Definition 4 *We say $a > b$, if*

1. $a = b + c$ *for some* $c \in \mathbb{N}$,
2. $a \neq b$.

The following theorem follows from the definition of order. The proof is again skipped.

Theorem 3 *$' >'$ on $\mathbb{N}$ satisfies*

1. $a > b \Rightarrow a + c > b + c \;\forall\; c \in \mathbb{Z}$.
2. $a > b$ *and* $c > 0 \Rightarrow ac > bc$.
3. $a > b$, $b > c \Rightarrow a > c$. *(Transitivity)*
4. *Given* $a, b \in \mathbb{N}$, *precisely one of* $a > b$ *or* $b > a$ *or* $a = b$ *is satisfied.*
5. $0 < 1 < 2 < 3 < \cdots$

♣

Lemma 1 *If $c, d \in \mathbb{N}$, and $c + d = 0$, then $c = d = 0$.*

Proof: Suppose $c \neq 0$. Then,

$$\begin{aligned} c =& c' + 1 \\ c + d =& (c' + 1) + d \\ =& (c' + d) + 1 \\ =& 0 \end{aligned}$$

So, it follows that 0 is the successor of $c' + d$, which is a contradiction to the fact that 0 has no predecessor in $\mathbb{N}$. Hence proved. ♣

Remark 1 *If $x = x + y$, then $y = 0$.*

Note 1 *Suppose $a, b \in \mathbb{N}$ and suppose $a > b$, can we have $b > a$ too?*

If yes, then

$$a > b \Rightarrow a = b + c$$

and

$$b > a \Rightarrow b = a + d$$

where $c, d \in \mathbb{N}$ and $c, d \neq 0 \Rightarrow (a+b) = (a+b) + (c+d)$. From earlier remark, we have $c = d = 0$ which shows that only one of the above two is possible, else they are equal.

Remark 2 *We have already seen that $0 < 1 < 2 < 3 \ldots$. In particular, we have: For any $a \in \mathbb{N}$, $a < (a+1) + b$ for any $b \in \mathbb{N}, b > 0$.*

1.2.4 The Well Ordering Principle, and the Euclidean Algorithm

We start with an equivalent formulation to the principle of Induction, known as the **Principle of Complete Induction**.

1. $0 \in S$.
2. $\{0, 1, 2, ..., n\} \subseteq S \Rightarrow n+1 \in S$.

Then $S = \mathbb{N}$. The proof is a simple consequence of the principle of Mathematical Induction. We skip the proof.

Proposition 2 ***The Well Ordering Principle (WOP)****: Every non empty subset $\boldsymbol{S} \subset \mathbb{N}$ contains a least element, i.e., there exists $s \in \boldsymbol{S}$ such that $s < s'$ for all $s' \in \boldsymbol{S}$, $s' \neq s$.*

Proof: For natural numbers $a < b$, we shall denote by $[a, b]$ the set $\{a, a+1, \ldots, b-1, b\}$.

Let $\emptyset \subsetneq \mathbf{S} \subset \mathbb{N}$. We need to show that $\mathbf{S}$ has a minimal element.
Suppose $\mathbf{S}$ has no minimal element. Let $P(n)$ be the propositional function: $n \notin \mathbf{S}$. We have two cases:

Case 1: $0 \in \mathbf{S}$. Since 0 is the least element of $\mathbb{N}$, it is also the minimal element of $\mathbf{S}$ which is a contradiction.

Case 2: $0 \notin \mathbf{S}$ so $P(0)$ holds. Suppose $P(j)$ holds for $0 \leq j \leq k$, i.e., suppose for all $j \in [0, k] : j \notin \mathbf{S}$.

If $k+1 \in \mathbf{S}$ then $k+1$ would be the minimal element of $\mathbf{S}$. So $k+1 \notin \mathbf{S}$ and so $P(k+1)$ also holds. Thus we have proven the following.

1. $P(0)$ holds.
2. For all $j \in [0, k] : P(j)$ holds) $\Rightarrow P(k+1)$ holds.

So by the principle of complete induction $P(n)$ holds for all $n \in \mathbb{N}$. But this means $\mathbf{S}$ is empty which is a contradiction. ♣

A simple consequence is the following:

Theorem 4 ***Euclidean division and the Euclidean algorithm:*** *Given positive integers m, n there exist unique non-negative integers q, r such that $m = qn+r$, $0 \leq r < n$. We describe this by saying that the Euclidean algorithm when applied to the ordered pair (m, n) gives a quotient of q and remainder r.*

Proof: Define $S = \{m - kn \mid k \in \mathbb{N}_0, m - kn > 0\}$. Now, $S \subset \mathbb{N}_0$ and $S \neq \Phi$ as $m - 0.n = m \in S$. By WOP, S has a least element; call it $r = m - qn$. We claim

1. $0 \leq r < n$.
2. q, r as determined above, are unique.

To prove this claim, note that by definition of S, $r \geq 0$. Indeed, suppose otherwise. Then $m - (q+1)n = m - qn - n = r - n \geq 0$, and this implies $m - (q+1)n \in S$. Also $m - (q+1)n = r - n < r$ as $n > 0$. Hence, r is not the least element of S, and this is a contradiction. This proves the first part.

To prove the second, again, suppose otherwise. Let $m = q_1 n + r_1$ and $m = q_2 n + r_2$ be two such representations and WLOG, let $r_2 > r_1$.
Equating the RHS of the above equations and simplifying,
$(q_1 - q_2)n = r_2 - r_1$,
$\Rightarrow r_2 - r_1$ is a multiple of n. But since, $r_1, r_2 < n, r_2 - r_1 < n$, the only possibility is $r_2 - r_1 = 0$. which implies that $r_1 = r_2$ and it follows that $q_1 = q_2$ and this proves the second claim as well.

Thus, r, q are unique integers satisfying $r \in [0, n)$ and $m = qn + r$. ♣

A very important consequence of the Euclidean algorithm is the following. For natural numbers a, b we say that n divides m if the corresponding value of r in the Euclidean algorithm above equals 0. For natural numbers a, b we say that d is the **greatest common divisor** (and denoted (a, b)) of a, b if d divides a, b and for any d' that divides both a, b we also have d divides d'. A very useful consequence of the Euclidean algorithm is this: Given m, n, r as in the theorem, $(m, n) = (r, n)$.

1.2.5 Prime Numbers

Definition 5 *Suppose $n \in \mathbb{N}$, $n > 1$. We say that n is* ***prime*** *if $n = ab$,
$\Rightarrow a = 1$ or $b = 1$. An equivalent definition is that for every $1 \leq a < p$ we must have $(a, p) = 1$.*

Theorem 5 ***(Euclid)*** *The set of Prime Numbers has* ***no largest*** *element.*

Proof: Suppose that there are only $N \in \mathbb{N}$ prime numbers. Let them be $p_1, p_2, \ldots p_N$. Consider the number

$$n = p_1 p_2 \ldots p_N + 1$$

Now, the number n is none of the prime numbers listed above and so it can be another prime number not in the listed N primes,which then contradicts our assumption. If n

is not a prime number, then n can be written as $n = ab$ for $a, b \in \mathbb{N}$ such that $a > 1$ and $b > 1$ (If either a or b is 1, then n will be a prime). Now we note that $p_1, p_2, \ldots p_N$ do not divide either a or b, which contradicts the Fundamental Theorem of Arithmetic (given below), i.e., there must exist atleast another prime number apart from $p_1, p_2, \ldots p_N$ which can divide n. This is due to our incorrect assumption that there exists only N prime numbers. Hence, there is no largest prime number. ♣

1.3 Integers

informally, an integer is a 'number' that can be represented as $(a - b)$ where $a, b \in \mathbb{N}$. But this definition is clearly a deficient one. We shall see how to make sense of this as follows.

Definition 6 *Describe a relation '$\diagdown$' on $\mathbb{N} \times \mathbb{N}$ as follows:*

$$(a, b) \diagdown (c, d)$$

if and only if

$$a + d = b + c$$

Lemma 2 ***(Cancellation Law for*** $\mathbb{N}$***)*** *If $x + y = z + y$,then $x = z$, $\forall\ x, y, z \in \mathbb{N}$.*

Proof: We prove it by induction on y. For $y = 0$, $x + 0 = z + 0 \Rightarrow x = z$.
Suppose it is true for y, i.e, if $x + y = z + y$, then $x = z$.
Now we need to prove it for $y + 1$. If $x + (y + 1) = z + (y + 1)$, then by associativity, we have, $(x + y) + 1 = (y + z) + 1$.
Since predecessors in $\mathbb{N} \setminus \{0\}$ are unique, $x + y = y + z$. By induction, it follows that $x = z$. Hence, the lemma holds good. ♣

Proposition 3 *'$\diagdown$' on $\mathbb{N} \times \mathbb{N}$ is an Equivalence relation.*

Proof: We have $(a, b) \diagdown (a, b)$ if $a + b = b + a$. Since addition is commutative ,this is satisfied and '$\diagdown$' is reflexive.
Now consider $(a, b) \diagdown (c, d)$. This gives $a + d = b + c$. Similarly $(c, d) \diagdown (a, b)$ this gives $c + b = a + d$. Since addition on natural numbers is commutative, this also proves the symmetry.
Now considering, $(a, b) \diagdown (c, d)$ i.e.,

$$a + d = b + c$$

and $(c, d) \diagdown (e, f)$ i.e.,

$$c + f = d + e$$

Adding the above gives

$$a + d + c + f = b + c + d + e$$

which is

$$(a + f) + (c + d) = (b + e) + (c + d)$$

Using the cancellation law, we get $a + f = b + e$, thus proving transitivity.
Since $'\diagdown'$ is Reflexive, Symmetric and Transitive, it is an Equivalence relation.
♣

Since the relation $\diagdown$ defines an Equivalence relation, it partitions $\mathbb{N} \times \mathbb{N}$ and these equivalence classes are called **Integers**, and the set of integers is denoted by $\mathbb{Z}$.

1.3.1 Addition on $\mathbb{Z}$

Consider the integers (a, b) and (c, d). We define addition in the following manner:

$$(a, b) + (c, d) := (a + c, b + d)$$

Claim 2 *(+) is well defined.*

Proof: Let (a, b) and (a', b') belong to one equivalence class and (c, d) and (c', d') belong to another. By definition,

$$(a, b) + (c, d) = (a + c, b + d)$$

and

$$(a', b') + (c', d') = (a' + c', b' + d')$$

Also,

$$a + b' = a' + b \dots (1)$$

$$c + d' = c' + d \dots (2)$$

Adding (1) and (2), we get

$$a + b' + c + d' = a' + b + c' + d$$

Using Associativity, we can rewrite it as

$$a + c + b' + d' = a' + c' + b + d$$

Thus, $(a + c, b + d)$ and $(a' + c', b' + d')$ belong to the same equivalence class, proving that (+) is well defined. ♣

1.3.2 Multiplication on $\mathbb{Z}$

Definition 7 *Consider the integers $(a, b), (c, d)$. Then multiplication is defined as*

$$(a, b).(c, d) := (ac + bd, ad + bc)$$

Claim 3 *(.) is well defined.*

Proof: Let (a, b) and (a', b') belong to one equivalence class and (c, d) and (c', d') belong to another. By definition,

$$(a, b).(c, d) := (ac + bd, ad + bc)$$

and

$$(a', b').(c', d') := (a'c' + b'd', a'd' + b'c')$$

Also,

$$a - b = a' - b' \ldots (1)$$

$$c - d = c' - d' \ldots (2)$$

Multiplying (1) and (2),

$$(a - b).(c - d) = ac + bd - ad - bc \ldots (3)$$

$$(a' - b').(c' - d') = a'c' + b'd' - a'd' - b'c' \ldots (4)$$

Since LHS of (3) and (4) are equal, we have

$$ac + bd - ad - bc = a'c' + b'd' - a'd' - b'c'$$

$$(ac + bd) + (a'd' + b'c') = (ad + bc) + (a'c' + b'd')$$

Thus, by definition, it follows that $(ac + bd, ad + bc)$ and $(a'c' + b'd', a'd' + b'c')$ belong to the same equivalence class, proving that (.) is well defined. ♣

1.3.3 Subtraction

If $(a, b) \in \mathbb{Z}$ then we define negation on $\mathbb{Z}$ as follows

$$-(a, b) := (b, a).$$

Definition 8 *Subtraction on integers is defined as*

$$(a, b) - (c, d) := (a, b) + (-(c, d)) \text{ i.e.},$$

$$(a, b) - (c, d) := (a, b) + (d, c)$$

It can be checked that (-) is also well defined, by using the 'well-defined'ness of (+) and the fact that $(-)$ can be represented in terms of $(+)$.

Claim 4 $\mathbb{N} \subseteq \mathbb{Z}$.

Proof: Define a set $N := \{(a, 0) \in \mathbb{Z}\}$. Now, consider the following map.

$$f : \mathbb{N} \longmapsto N$$

$$i.e., f : a \longmapsto (a, 0)$$

We observe that $f(a + b) = (a + b, 0)$ and $f(a) = (a, 0)$, $f(b) = (b, 0) \Rightarrow f(a + b) = f(a) + f(b)$ $\forall a, b \in \mathbb{N}$. This map identifies Natural Numbers sitting inside the Integers. ♣

Proposition 4 *Addition and Multiplication on $\mathbb{Z}$ satisfy the following:*

1. *They are Commutative, Associative and Addition distributes over Multiplication*
2. *O:=(0,0) satisfies (a,b)+O=(a,b) $\forall (a, b) \in \mathbb{Z}$*
3. *Given $m \in \mathbb{Z}$ there exists a unique $n \in \mathbb{Z}$ such that $m + n = 0$*
4. *If $m + x = n + x$, then $m = n$, $\forall\ x, m, n \in \mathbb{Z}$*

Proof: 3. Let us consider $m = (a, b) \in \mathbb{Z}$. Let $n = (c, d) \in \mathbb{Z}$ such that $m + n = 0$.

$$\begin{aligned}
(a, b) + (c, d) &= (0, 0) \\
(a + c, b + d) &= (0, 0) \\
(a + c) - (b + d) &= 0 - 0 \\
(a + c) - (b + d) &= 0 \\
a + c &= b + d \\
a - b &= d - c \\
-(a - b) &= c - d \\
(b, a) &= (c, d)
\end{aligned}$$

Hence $(c, d) = (b, a)$ and this shows existence of additive inverse in $\mathbb{Z}$.

Now, we prove the uniqueness of additive inverse. Let $n_1 = (c, d) \in \mathbb{Z}$ and $n_2 = (e, f) \in \mathbb{Z}$

both be additive inverses for m.

$$\begin{aligned}
(a,b)+(c,d) &= (0,0)\ldots(1)\\
(a,b)+(e,f) &= (0,0)\ldots(2)\\
(a+c,b+d) &= (0,0)\ldots(using\ (1))\\
(a+e,b+f) &= (0,0)\ldots(using\ (2))\\
(a+c,b+d) &= (a+e,b+f)\ldots(from\ above\ two)\\
(a+c)+(b+f) &= (b+d)+(a+e)\ldots(from\ definition)\\
a+b+c+f &= a+b+d+e\ldots(using\ associativity)\\
c+f &= e+d\ldots(cancellation\ law)\\
c-d &= e-f\\
(c,d) &= (e,f)
\end{aligned}$$

Hence (c,d) and (e,f) represent the same integer and hence, we have proved uniqueness of the inverse. ♣

Proof: 4. Let us consider $m=(a,b)$ and $n=(c,d)\in\mathbb{Z}$. Let $x=(x_1,x_2)\in\mathbb{Z}$ such that $m+x=n+x$. For $x=(0,0)$, we have $(a,b)+(0,0)=(c,d)+(0,0)\Rightarrow(a,b)=(c,d)$. Now, we induct on x_1.
Assume that $m+x=n+x\Rightarrow m=n$ is true for $x=(x_1,x_2)$. We now prove this to be true for $x=(x_1+1,x_2)$.

$$\begin{gathered}
(a+x_1,b+x_2)=(c+x_1,d+x_2)\Rightarrow(a,b)=(c,d)\ldots(given)\\
(a+x_1)-(b+x_2)=(c+x_1)-(d+x_2)\Rightarrow(a,b)=(c,d)\\
(a-b)+(x_1-x_2)=(c-d)+(x_1-x_2)\Rightarrow(a,b)=(c,d)\\
(a-b)+(x_1-x_2)+1=(c-d)+(x_1-x_2)+1\Rightarrow(a,b)=(c,d)\\
(a-b)+((x_1+1)-x_2)=(c-d)+((x_1+1)-x_2)\Rightarrow(a,b)=(c,d)
\end{gathered}$$

Hence, we have shown that the statement holds for $x=(x_1+1,x_2)$, thus completing the induction. ♣

1.3.4 Order on $\mathbb{Z}$

Let $m,n\in\mathbb{Z}$ We say $m>n$ if

1. $m=n+p$, for some $p\in\mathbb{Z}$
2. $m\neq n$.

The following theorem is in the same spirit as the corresponding one for the natural numbers.

Theorem 6 *$' >'$ on $\mathbb{Z}$ satisfies*

1. $a > b \Leftrightarrow a - b > 0.$

2. $a > b \Rightarrow a + c > b + c\ \forall\ c \in \mathbb{Z}.$

3. $a > b$ *and* $c > 0 \Rightarrow ac > bc.$

4. $a > b \Rightarrow -b > -a.$

5. $a > b, b > c \Rightarrow a > c.$

6. *Given* $a, b \in \mathbb{Z}$ *precisely one of* $a > b$ *or* $b > a$ *or* $a = b$ *is satisfied.*

We return to the notion of divisibility that was introduced for the natural numbers. We extend the same to the integers in the following manner. We say that $a|b$ (for integers a, b if there exists an integer c such that $b = ac$. We also define the greatest common divisor in exactly the same manner as in the case of the natural numbers.

Theorem 7 *If $a, b \in \mathbb{Z}$ and suppose $(a, b) = 1$. Then there exist integers m, n such that $1 = am + bn$.*

Proof: We will sketch the proof. Consider the set $S := \{am + bn | m, n\mathbb{Z}\}$. The following can be proved in a straightforward manner.

1. $0 \in S$.

2. $x, y \in S \Rightarrow a + b \in S$. Similarly, $x \in S \Rightarrow -x \in S$.

3. $x \in S, \lambda \in \mathbb{Z} \Rightarrow \lambda x \in S$.

In particular consider $S_0 := S \cap \mathbb{N}$. This is clearly non-empty by the above observations, and the fact that $a, b \in S$. Now, let d be the least element of S_0 which exists by the WOP. Consider the remainder r by applying the Euclidean algorithm to $(\pm a, d)$ (i.e. pick the appropriate non-negative integer). It is easy to check that $r \in S$ as well since both $a, d \in S$. By the minimality of d, it follows that $r = 0$, i.e. $d|a$. By the same reason, $d|b$. Hence $d|(a, b) = 1$, so we must have $d = 1$. ♣

1.4 Rational Numbers

Again, informally, by the rational numbers, we denote 'numbers of the form $\frac{p}{q}$, where $p, q \in \mathbb{Z}$ and $q \neq 0$'. But as we have seen before with the Integers, the path to formalizing this involves setting up the right kind of equivalence relation on pairs of integers.

Definition 9 *We define a Relation $\sim$ on $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ as*

$$a//b \sim c//d$$

if $ad = bc$, where $a, b, c, d \in \mathbb{Z}$ and $b, d \neq 0$.

Remark 3 *The integers satisfy the property that there are* ***no zero divisors****, i.e. if $a, b \in \mathbb{Z}$ and $ab = 0$,then either $a = 0$ or $b = 0$ or both $a, b = 0$. Indeed, if $a, b > 0$ then we have seen from the properties of the natural numbers that $ab > 0$. If say $a < 0$, then $(-a)b > 0$ and this also implies that $ab < 0$, so in particular, it is not zero. The same argument works in the other cases as well.*

Proposition 5 *$\sim$ defined on $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ is an Equivalence Relation.*

Proof: We prove the three properties of equivalence relations.

1. **Reflexivity**
   We have
   $$a//b \sim a//b$$
   $\Rightarrow ab = ba$ and this is true.

2. **Symmetry**
   We have , if
   $$a//b \sim c//d$$
   $\Rightarrow ad = bc$, then
   $$c//d \sim a//b$$
   $\Rightarrow bc = ad$, which is true, since $ad = bc$.

3. **Transitivity**
   If
   $$a//b \sim c//d$$
   $\Rightarrow ad = bc$, and
   $$c//d \sim e//f$$

$\Rightarrow fc = ed$, we get

$$\begin{aligned}(ad)(fc) &= (bc)(ed)\\ (af)(cd) &= (eb)(cd)........(Using\ Associativity)\\ (af-eb)(cd) &= 0\end{aligned}$$

Let $x = cd$ .We now use the property of *no zero divisors* for integers. If $x \neq 0$ we are through, since it implies that $af = eb$ and hence $\Rightarrow a//b \sim e//f$.
If x = 0, then $cd = 0, \Rightarrow c = 0$, since $d \neq 0$ by definition. But then, if $c = 0$, then $a = 0$ and $e = 0$ (by definition of the relation) and thus $\Rightarrow a//b \sim e//f$.
Since, the relation $\sim$ is Reflexive , Symmetric and Transitive, it is an Equivalence Relation.

♣

Remark 4 *The set of Rational Numbers, denoted by $\mathbb{Q}$, is the set of equivalence classes of $\mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ with respect to the equivalence relation $\sim$.*

1.4.1 Addition and Multiplication on $\mathbb{Q}$

We define Addition on $\mathbb{Q}$ in the following manner:

$$(a//b) + (c//d) := (ad + bc)//bd$$

We define Multiplication on $\mathbb{Q}$ in the following manner:

$$(a//b).(c//d) := (ac)//bd$$

We can check that both Addition(+) and Multiplication(.) on $\mathbb{Q}$ are well defined as we have done with the integers.

Claim 5 *(.) is well defined for $\mathbb{Q}$.*

Proof: Let $(a//b)$ and $(a'//b')$ belong to one equivalence class and let $(c//d)$ and $(c'//d')$ belong to another equivalence class.Then by definition,

$$(a//b).(c//d) := (ac)//bd$$

$$(a'//b').(c'//d') := (a'c')//b'd'$$

Also, since $(a//b)$ and $(a'//b')$ belong to the same equivalence class,we have $ab' = ba'$. Similarly, $cd' = dc'$ $.\Rightarrow (ab')(cd') = (ba')(dc')$. By associativity,we have $(ac)(b'd') = (bd)(a'c')$. Thus, by definition,

$$(ac//bd) \sim (a'c'//b'd')$$

and hence (.) is well defined for $\mathbb{Q}$ ♣

Claim 6 *(+) is well defined for* $\mathbb{Q}$

Proof: Let $(a//b)$ and $(a'//b')$ belong to one equivalence class and let $(c//d)$ and $(c'//d')$ belong to another equivalence class.Then by definition,

$$(a//b) + (c//d) = (ad + bc)//bd$$

$$(a'//b').(c'//d') = (a'd' + b'c')//b'd'$$

Also, since $(a//b)$ and $(a'//b')$ belong to the same equivalence class,we have

$$ab' = ba' \ldots (1)$$

Similarly,

$$cd' = dc' \ldots (2)$$

Now, multiplying (1) on both sides by dd' and (2) by bb', we get

$$ab'dd' = ba'dd' \ldots (3)$$

$$cd'bb' = dc'bb' \ldots (4)$$

By associativity,(3) and (4) can be written as

$$adb'd' = a'd'bd \ldots (5)$$

$$b'd'bc = b'c'bd \ldots (6)$$

Adding (5) and (6), we get

$$adb'd' + b'd'bc = a'd'bd + b'c'bd$$

Thus,by definition

$$\Rightarrow (ad + bc)//bd \sim (a'd' + b'c')//b'd'$$

and hence (+) is well defined for $\mathbb{Q}$. ♣

Theorem 8 *(Q* ***is a FIELD****)*
The set of Rationals , along with the binary operations (.) and (+) defined on it satisfy the following:

1. *Addition and Multiplication on* $\mathbb{Q}$ *are* ***commutative , associative*** *and* $(+)$ ***distributes*** *over* $(.)$.

2. $\mathbb{Z}$ ***is a subset of*** $\mathbb{Q}$ *.i.e.,* $\mathbb{Z} \longmapsto \mathbb{Q}$ *and this can be obtained by* $n \longmapsto (n//1)$*. Therefore, the map identifies* $\mathbb{Z}$ *inside* $\mathbb{Q}$ *and all the operations are compatible.*

3. *For* $x \in \mathbb{Q}$ *, we have*

$$x + 0 = x \ldots \{\textbf{\textit{Additive Identity}}\}$$
$$x.1 = x \ldots \{\textbf{\textit{Multiplicative Identity}}\}.$$

4. *For* $x = a//b$ *, there is a unique rational* $y = -a//b$ *such that* $x+y = 0$*. Therefore,* $\mathbb{Q}$ *has a unique* ***Additive Inverse****.*

5. *If* $x = a//b$ *where* $a \neq 0$ *, then there exists unique* $x^{-1} := b//a$ *satisfying*

$$x.x^{-1} = 1 \ldots \{\textbf{\textit{Multiplicative Inverse}}\}.$$

♣

Note 2 *For* $x = (a//b)$ *we define* $-x := (-1//1).(a//b)$*.i.e.,* $-x = -1.x$*.*

1.4.2 Order on $\mathbb{Q}$

For $x \in \mathbb{Q}$, we say that

1. $x > 0$ if , for $x = a//b$, $a, b \in \mathbb{N} - \{0\}$
2. $x < 0$ if $-1.x \ > 0$.
3. $x = 0$ if $a = 0$ and $b \neq 0$.

In general , for $x, y \in Q$ we say that $x > y$, if $\exists\ r \in \mathbb{Q}$ and $r > 0$ such that $x = y + r$.

Theorem 9 *(**Total Order**) The Rational numbers* $\mathbb{Q}$ *and* $>, <$ *defined earlier on* $\mathbb{Q}$ *satisfy the following:*

1. *For* $x, y \in \mathbb{Q}$*, exactly one of* $x > y$ *,* $x < y$ *,* $x = y$ *is satisfied.(**Trichotomy**)*
2. *For* $x, y, z \in \mathbb{Q}$*,* $x < y \Rightarrow x + z < y + z$*.*
3. *For* $x, y, z \in \mathbb{Q}$ *and* $z > 0$*,* $x < y \Rightarrow xz < yz$*.*
4. *For* $x < y$*,* $-x > -y$*.*
5. *For* $x, y, z \in \mathbb{Q}$ *, if* $x < y$ *and* $y < z$ *, then* $x < z$*.*

Thus, $(\mathbb{Q}, (+), (.), <, >)$ *is a* ***Totally Ordered Field****.* ♣

1.4.3 $\mathbb{Q}$ misses some 'numbers'

Before we move to the next important theorem we need a lemma. Recall the definition of primes (in the set $\mathbb{N}$).

Lemma 3 *For a prime number p , if $p|ab$ for natural numbers a, b, then $p|a$ or $p|b$ or both.*

Proof: First, using the Euclidean algorithm we may assume WLOG that $a, b \leq p$. Since p is prime, if p does not divide a, then we must have $(a, p) = 1$. By a previous theorem, there exist integers λ, μ such that $a\lambda + p\mu = 1$. Multiplying by b on both sides we have $b = ab\lambda + bp\mu$. Now, both the terms on the right hand side are divisible by p by assumption. Hence we must necessarily have $p|b$. ♣

Theorem 10 ***Fundamental Theorem of Arithmetic***
*Every Natural number $n > 1$ can be expressed as a product of primes in a **unique** way, except for permuting the factors.*

Proof:

1. (**Existence of the product of primes**)
We use induction to prove that all natural numbers $n > 1$ can be expressed as a product of primes. For $n = 2$, it is true since 2 is itself a prime number. Assume that for all $n > 2$ and less than N, there is a way to express the numbers as a product of primes. Now, consider $n = N$. If N is prime, we are through. If N is not a prime, then N can be written as $N = ab$, where $a, b \in \mathbb{N}$ and $0 < a, b < N$. According to induction, a and b can be expressed as a product of primes. Thus, $N = ab$ can also be expressed as a product of primes.

2. **(Uniqueness)** Suppose that the number N can be expressed in two ways as

$$N = p_1^{a_1} p_2^{a_2} \dots p_m^{a_m} = q_1^{b_1} q_2^{b_2} \dots q_n^{b_n}$$

Here all the $p_i\,'s$ and $q_j\,'s$ are distinct primes and all $a_i\,'s$ and $b_j\,'s$ are the number of times each prime occurs in the product. We want to show that $m = n$ and each $p_i^{a_i} = q_j^{b_j}$ for some j ; i.e., $p_i = q_j$ and $a_i = b_j$. Now, consider p_1.We have

$$p_1^{a_1} p_2^{a_2} \dots p_m^{a_m} = q_1^{b_1} q_2^{b_2} \dots q_n^{b_n} \quad (1)$$

p_1 divides LHS of (1). Hence p_1 divides even RHS of (1).$\Rightarrow$ $p_1|q_j^{b_j}$ for some j. $\Rightarrow p_1|q_j$. But since p_1 and q_j are primes , we must have $p_1 = q_j$. Let us renumber the index j as 1 and vice-versa. We get,

$$p_1^{a_1} p_2^{a_2} \dots p_m^{a_m} = p_1^{b_1} q_2^{b_2} \dots q_n^{b_n}$$

Now, if $a_1 > b_1$ then divide both sides by $p_1^{b_1}$ to get

$$p_1^{a_1-b_1} p_2^{a_2} \dots p_m^{a_m} = q_2^{b_2} \dots q_n^{b_n}$$

We notice that the LHS is now divisible by p_1 while the RHS is not.This is not possible. Similarly we can show that $a_1 < b_1$ is also not possible. Hence , the only possibility is that $a_1 = b_1$.The equation now becomes

$$p_2^{a_2} \dots p_m^{a_m} = p_1^{b_1} q_2^{b_2} \dots q_n^{b_n}$$

Without loss of generality, we will assume $m < n$. Proceeding similarly for $p_2, p_3 \dots p_m$ we can cancel terms on both sides until we are left with

$$1 = q_{m+1}^{b_{m+1}} q_{m+2}^{b_{m+2}} \dots q_n^{b_n}$$

However since the RHS is supposed to be a product of primes which is equal to 1, it implies that all the terms on the RHS are 1's and do not contribute to the product of (1). Hence, we will have each of the prime in LHS paired up with a prime in RHS with equal powers, thus proving the uniqueness of the product expression except for the order in which they occur.

♣

From the fundamental theorem of arithmetic it follows that every rational x can be expressed as$\frac{p}{q}$ where p, q are integers that have no prime factors in common. This representation of the rationals will be of use to us.

Firstly, the rationals are a pretty 'large' set. Indeed, we have the following:

Fact 1 *If $a, b \in \mathbb{Q}$ and $a < b$, then there exists $c \in \mathbb{Q}$ such that $a < c < b$.*

Indeed, it is easy do check that if $a < b \in \mathbb{Q}$ then $c = \frac{1}{2}(a + b)$ is also in $\mathbb{Q}$ and we have $a < c < b$.

Proposition 6 *There is no Rational Number $\boldsymbol{x}$ such that $\boldsymbol{x}^2 = 2$.*

Proof: Let us assume to the contrary that there exists $x \in \mathbb{Q}$ such that $x^2 = 2$. Writing x as $x = \frac{p}{q}$ where $p, q \in \mathbb{Z}$,$q \neq 0$ and p, q have no common factor other than 1.Then we have,

$$(p/q)^2 = 2 \tag{1.1}$$
$$\Rightarrow p^2 = 2q^2 \tag{1.2}$$
$$\Rightarrow 2|p^2 \tag{1.3}$$
$$\Rightarrow 2|p \tag{1.4}$$
$$\Rightarrow p = 2k \ for \ some \ k \in \mathbb{N} \tag{1.5}$$

Substituting p in (2),we get

$$\Rightarrow 4k^2 = 2q^2 \tag{1.6}$$
$$\Rightarrow 2k^2 = q^2 \tag{1.7}$$
$$\Rightarrow 2|q^2 \tag{1.8}$$
$$\Rightarrow 2|q \tag{1.9}$$

From (4) and (9) , we get $2|p$ and $2|q$ which is a contradiction to the assumption that p, q have no common factor other than 1.
Therefore there is no rational number such that $x^2 = 2$.

♣

1.5 Real Numbers

Definition 10 *(**Dedekind Cut**) A Dedekind Cut is a partition $\boldsymbol{A} \cup \boldsymbol{B}$ of $\mathbb{Q}$ satisfying*

1. $\boldsymbol{A} \neq \emptyset$, $\boldsymbol{B} \neq \emptyset$.
2. *For $a \in \boldsymbol{A}$, $b \in \boldsymbol{B}$, we have $a < b$.*
3. ***A** has no largest element*

Definition 11 *A Dedekind Cut $\boldsymbol{A} \wr \boldsymbol{B}$ is defined to be a **Real number**.i.e.,*

$$\mathbb{R} = \{A \wr B \mid A \wr B \textit{ is a Dedekind cut of } \mathbb{Q}\}$$

Proposition 7 $\mathbb{Q} \subset \mathbb{R}$*, i.e., $x \in \mathbb{Q} \Rightarrow x \in \mathbb{R}$.*

Proof: For every $x \in \mathbb{Q}$, we define the partition

$$\mathbf{L} = \{y \in \mathbb{Q} \,|y < x\},$$

$$\mathbf{R} = \{y \in \mathbb{Q} \,|y \geq x\},$$

By this partition, we can associate every $x \in \mathbb{Q}$ to a corresponding partition $\mathbf{L} \wr \mathbf{R}$ which is a Real Number in $\mathbb{R}$. ♣

Definition 12 *(**Order on Real Numbers**) If $x = (\boldsymbol{L} \wr \boldsymbol{R})$ and $y = (\boldsymbol{L'} \wr \boldsymbol{R'})$ are Real numbers , then $x < y$ if $\boldsymbol{L} \subset \boldsymbol{L'}$.*

Definition 13 *A non-empty set $S \subseteq \mathbb{R}$ is said to be **bounded above** if there is an $x \in \mathbb{R}$, such that for every $s \in S$, $s < x$. Such an x is called an **Upper Bound** of S.*

Theorem 11 *Every non empty set $S \subseteq \mathbb{R}$ which is bounded above has a* ***least upper bound (lub)****.i.e., $\exists y \in \mathbb{R}$ satsifying*

1. *$s \leq y$, $\forall\ s \in S$.*
2. *If y' satisfies $s \leq y'\ \forall\ s \in S$, then $y \leq y'$.*

Proof: Let $s = (L_s \wr R_s)$. Let $L = \cup_{s \in S}\ L_s$ and $R = \mathbb{Q} \setminus L$.

1. **CLAIM:** $y = (L \wr R)$ will do.
   First, $L \neq \emptyset$ and $R \neq \emptyset$, since every rational upper bound for S is not in L_s for any s. Further, for $a \in L$ and $b \in R$, we have $a \leq b$. Lastly, it is not hard to see that L has no greatest element since such an element must necessarily be in some of the L_s which is not possible by assumption. Now we have $L_s \subset L$ and therefore for every $s \in S$ we have $s \leq y$.
2. Suppose there exists another number y' such that $y' \geq s\ \forall s \in S$. Since $y' \geq s\ \forall s$, by definition we have $L_{y'} \supseteq L_s\ \forall s \in S$,

$$\Rightarrow L_{y'} \supseteq \cup_{s \in S} L_s$$

$$\Rightarrow y' \geq y.$$

♣

Example 2 *The real number $\sqrt{2}$ can be obtained as the following Dedekind cut $\boldsymbol{L} \wr \boldsymbol{R}$ defined as:*

$$\boldsymbol{L} = \{x \in \mathbb{Q}\,|x \leq 0 \text{ or } x^2 < 2\}$$

$$\boldsymbol{R} = \mathbb{Q} \setminus \boldsymbol{L}.$$

We will in fact prove something stronger and more general later.

1.5.1 Addition and Multiplication on $\mathbb{R}$

We define Addition on $\mathbb{R}$ in the following manner:

$$x = (L_x \wr R_x)$$

and

$$y = (L_y \wr R_y),$$

$$x + y = (L_{x+y} \wr R_{x+y}),$$

where

$$L_{x+y} := L_x + L_y := \{r + s \mid r \in L_x, s \in L_y\}.$$
$$R_{x+y} := \mathbb{Q} \backslash L_{x+y}.$$

We define Multiplication on $\mathbb{R}$ in the following manner:

First consider the case of positive cuts where $x, y \geq 0$.
Then

$$L_{x.y} := \{r \leq 0\} \cup \{a.b \mid a \in L_x, b \in L_y, a, b \geq 0\}$$

For the remaining cases, we need to define the notion of negative cut.

Definition 14 *If*

$$y = (L_y \wr R_y) \in \mathbb{R}$$

then

$$L_{-y} := \{r \mid r = -s \ for \ some \ s \in R_y, s \ is \ not \ min(R_y)\}$$

Now,

If $x \geq 0, y \leq 0$
$$x.y := -(x.(-y)).$$
If $x \leq 0, y \geq 0$
$$x.y := -((-x).y).$$
If $x \leq 0, y \leq 0$,
$$x.y := (-x).(-y).$$

One can prove (we omit the rather laborious proof) that

Theorem 12 $(\mathbb{R}, +, \cdot)$ *is a field.* ♣

The next proposition shows that the set of reals constructed this way does not have 'gaps' the way the set of rationals did.

Proposition 8 $\mathbb{R}$ *has no gaps.*

Proof: Suppose $A \wr B$ is a partition of $\mathbb{R}$ s.t

1. $\mathbf{A}$, $\mathbf{B} \neq \emptyset$, $\mathbf{A} \cup \mathbf{B} = \mathbb{R}$, $\mathbf{A} \cap \mathbf{B} = \emptyset$.
2. For $a \in \mathbf{A}$, $b \in \mathbf{B}$, we have $a < b$.

3. $\mathbf{A}$ has no largest element.

By (2) each $b \in \mathbf{B}$ is an upper bound for $\mathbf{A} \Rightarrow \mathbf{A}$ is bounded above $\Rightarrow$ $x = \text{lub}(\mathbf{A})$ exists.
Since $x = \text{lub}(\mathbf{A})$, we have

$$a \leq x,\ \textit{for all } a \in \mathbf{A} \text{ (} \textit{x is a least upper bound} \text{)},$$

$$x \leq b \textit{ for all } b \in \mathbf{B} \text{ (}\textit{each b is also an upper bound for } \mathbf{A}\text{)},$$

so it follows that

$$a \leq x \leq b \textit{ for all } a \in \mathbf{A}\ ,\ b \in \mathbf{B}.$$

Consider

$$\mathbf{A}' = \{r \in \mathbb{Q} \,\|\, r < x\}$$

$$\mathbf{B}' = \{r \in \mathbb{Q} \,\|\, r \geq x\}$$

$$x = \mathbf{A}' \wr \mathbf{B}'$$

Claim 7

$$\boldsymbol{A} = \{y \in \mathbb{R} \,\|\, y < x\}$$

$$\boldsymbol{B} = \{y \in \mathbb{R} \,\|\, y \geq x\}$$

Indeed,

$$\mathbf{A} \subseteq \{y \in \mathbb{R} \,\|\, y < x\}$$

clearly it is true as $x = \text{lub}(\mathbf{A})$ and moreover

$$\{y \in \mathbb{R} \,\|\, y < x\} \subseteq \mathbf{A}.$$

The last holds since otherwise there is $y < x,\ \ y \notin \mathbf{A} \Rightarrow y \in \mathbf{B}$. But each $b \in \mathbf{B}$ satisfies $x \leq b$ and this contradicts $y < x$.
The other parts are proven similarly. ♣

The following theorem (whose proof we skip as the details are laborious) consolidates our understanding of the reals.

Theorem 13 $(\mathbb{R}, (+), (.), <, >)$ *is a* ***Totally Ordered Field***. ♣

1.5.2 Another description for Real Numbers

Before we go there, here is a remarkable property of $\mathbb{R}$.

Definition 15 *A real number x is called IRRATIONAL if $x \in \mathbb{R} \smallsetminus \mathbb{Q}$*

Notation: Suppose $a < b, \quad a, b \in \mathbb{R}$.

$$(a, b) = \{x \in \mathbb{R} \parallel a < x < b\}.$$

$$[a, b] = \{x \in \mathbb{R} \parallel a \leq x \leq b\}.$$

$$(-\infty, a) = \{x \in \mathbb{R} \parallel x < a\}.$$

$$(-\infty, a] = \{x \in \mathbb{R} \parallel x \leq a\}.$$

$$(b, \infty) = \{x \in \mathbb{R} \parallel x > b\}.$$

$$[b, \infty) = \{x \in \mathbb{R} \parallel x \geq b\}.$$

Theorem 14 *For every $a < b$, the interval (a, b) contains both rational and irrational numbers.*

Proof: $\sqrt{2} \in \mathbb{R} \smallsetminus \mathbb{Q} \Rightarrow \frac{1}{\sqrt{2}} \in \mathbb{R} \smallsetminus \mathbb{Q}$.
Also we know that $\frac{1}{2} \in \mathbb{Q}$ and in fact, $\frac{1}{2} \in (0, 1)$ since if $\frac{1}{2} > 1$, then multiplying by 2 on both sides, we get $1 > 2$ which is a contradiction.
Thus for $a = 0, b = 1$ we are through.
Now let $a = A \wr B$, $b = A' \wr B'$ and $a < b$
$\Rightarrow A \subset A' \Rightarrow \exists r, s \in A' \smallsetminus A$.
WLOG, $r < s \Rightarrow a \leq r < s \leq b$
Now, consider the function, $f : (0, 1) \to (r, s)$ defined as $f(x) = r + x(s - r)$.
Clearly, f maps rationals to rationals and this property is also satisfied by its inverse, which exists as f is a bijection.
$\Rightarrow f(\frac{1}{2}) \in \mathbb{Q}$, $f(\frac{1}{\sqrt{2}}) \in \mathbb{R} \smallsetminus \mathbb{Q}$ and they lie in (a, b). ♣

Theorem 15 *For all $x > 0$, and $n \geq 2, n \in \mathbb{N}$, there exists a unique $y > 0$ such that $y^n = x$ or equivalently, we write $y = \sqrt[n]{x}$*

Proof: Suppose $x > 1$ (WLOG, this is the only case we have to deal with).
Let $R = \{r > 0 \| \; r^n \leq x\}$. Clearly, $R \neq \emptyset$ as $1 \in R$.

Claim 8 *R is bounded above.*

$x > r$ for all $r \in R$. A simple induction argument shows that $x^n \geq x$. Also, by definition, $x \notin R$ so R is bounded above. Hence $y = lub\ (R)$ exists.

Claim 9 $y^n = x$.

We will use the Trichotomy law for showing that equality holds, i.e., we shall show $y^n \geq x$ and $y^n \leq x$.
As $y = lub(R)$, we have $r \leq y$ for any $r \in R$. Now, for any $r \in R$ we have $r \leq y \Rightarrow r.r \leq y.r$ (as $r > 0$). Hence $r^2 \leq y^2$. Similarly, $r^n \leq y^n$. (by induction) $\Rightarrow y^n \geq x$.

For the other case consider $\mathbf{A} \ = \{(y-\epsilon)^n \| \ 0 < \epsilon < y\}$.

Claim 10 *x is an upper bound for $\boldsymbol{A}$.*

$$y = lub(\mathbb{R})$$

$\Rightarrow y - \epsilon$ *is not an upper bound* $for\, \mathbb{R}$.

Claim 11 $(y-\epsilon)^n \leq x$.

Since $y-\epsilon$ is not an upper bound for R, there exists $r \in R$ such that $(y-\epsilon)^n \leq r^n \leq x$ $\Rightarrow$ x is an upper bound for **A**.

Claim 12 $y^n = lub(\boldsymbol{A})$

Suppose not. Clearly, y^n is an upper bound.
Let $y^n - \delta$ be the lub for **A** for some fixed $\delta > 0$
So

$$\begin{aligned}
(y-\epsilon)^n &\leq y^n - \delta \text{ for all } 0 < \epsilon < y, \text{ which implies} \\
\delta &\leq y^n - (y-\epsilon)^n \\
&\leq \epsilon(y^{n-1} + y^{n-2}(y-\epsilon) + ...(y-\epsilon)^{n-1}) \\
&\leq \epsilon n y^{n-1}
\end{aligned}$$

for all $0 < \epsilon < y$. If $\epsilon = \frac{\delta}{2ny^{n-1}} < y$ then we have $\delta \leq \frac{\delta}{2}$ for $\delta > 0$ which is a contradiction.
So, $y^n = lub(\mathbf{A}) \leq x$, and this completes the proof that $y^n = x$. ♣

1.5.3 Archimedian Property of $\mathbb{R}$

Proposition 9 *For any real $x > 0$, there exists a unique $n \in \mathbb{N}$ s.t $n \leq x < n+1$.*

Proof: Let $x \in \mathbb{R}, x = \mathbf{A} \wr \mathbf{B}$. For $p, q \in \mathbb{N}$, pick

$$\frac{p}{q} \in \mathbf{B}, \frac{p}{q} > x > 0, \text{ and } p, q > 0.$$

Then

$$\frac{p}{q} < \frac{p}{q} + \frac{p}{q} + \frac{p}{q} + ... + \frac{p}{q} = p.$$

So, there exists an integer $p > x$. Consider

$$S = \{m \in \mathbb{N} \parallel m > x\}.$$

By the above, S$\neq \emptyset$. So, by the WOP, S has a least element $n+1$, say. So $n+1 > x$ since $n+1 \in \mathbf{S}$, and $n \in \mathbf{S}$ *so* $n <= x$. ♣

Corollary 1 *Every $x \in \mathbb{R}$ can be written as $x = n + x_0$ with $x_0 \in [0,1)$, and $n \in \mathbb{Z}$.*

Proof: Trivial. ♣

Definition 16 *A Sequence is a function whose domain is $\mathbb{N}$.*

Let $x \in (0,1)$. Define a Sequence in the following way:

$$\begin{aligned} a_1 &= 0 \ \ if \ 0 < x < \frac{1}{2} \\ &= 1 \ \ if \ x \geq \frac{1}{2} \\ a_2 &= 0 \ \ if \ 0 < x < \frac{1}{4} \ or \ \frac{1}{2} \leq x < \frac{3}{4} \\ &= 1 \ \ if \ \frac{1}{4} \leq x < \frac{1}{2} \ or \ x \geq \frac{3}{4} \end{aligned}$$

In general, at the k^{th} step, after having defined $(a_1, a_2, a_3, \ldots\ldots, a_{k-1})$, note that we have

$$x \geq \frac{a_1}{2} + \frac{a_2}{2^2} + \ldots\ldots + \frac{a_{k-1}}{2^{k-1}}.$$

If

$$x - (\frac{a_1}{2} + \frac{a_2}{2^2} + \ldots\ldots + \frac{a_{k-1}}{2^{k-1}}) < \frac{1}{2^k}$$

then define

$$a_k = 0$$

else

$$a_k = 1$$

This gives a function $f : (0,1) \mapsto \mathfrak{a}$ where $\mathfrak{a}$ is the set of all binary sequences.

Conversely, given a binary sequence $\mathbf{a} = (a_1, a_2, \ldots\ldots)$ define the set

$$S_{\mathbf{a}} = \left\{ \sum_{k=1}^{n} \frac{a_k}{2^k} \parallel n \in \mathbb{N} \right\}.$$

Since each element of S is strictly less than 1, lub(S) exists for each binary sequence.

To summarize: For every x , $0 < x < 1$ we can associate a binary sequence $\mathbf{a}_x$, and conversely, for each binary sequence $\mathbf{a}$ we can define a real number in (0,1) by $\mathbf{x} = lub(S_{\mathbf{a}})$.

Question 1: If $0 < x < y < 1$, then can we have $\mathbf{a}_x = \mathbf{a}_y$? In other words, if $x < y$ then is it true that $a_k(x) \neq a_k(y)$ for some k?

If $y - x = \delta > 0$, pick n such that $\delta > \frac{1}{2^n}$. Suppose $a_x = a_y$ i.e $a_x(i) = a_y(i)$ for all $i \in \mathbb{N}$. Then

$$x > \Sigma \frac{a_x(i)}{2^i} = \Sigma \frac{a_y(i)}{2^i} > y - \frac{1}{2^n},$$

so that we have

$$y \ - \ x \ < \ \frac{1}{2^n} \Rightarrow \delta < \frac{1}{2^n},$$

which is a contradiction. Hence, distinct elements of $(0, 1)$ give distinct binary sequences by this mapping. In other words, $x \to \mathbf{a}_x$ is an injection.

Question 2: If $a_1 \neq a_2$, is $lub(S_1) \neq lub(S_2)$?

This is not the case. Consider
$\mathbf{a}_1 = (1, 0, 0, 0...,)$ and $\mathbf{a}_2 = (0, 1, 1, 1...)$. Both 'represent' the same number.

Proposition 10 *The sequence* $\mathbf{a} = (1, 1, 1...)$ *has* $lub(S_a) = 1$.

Proof: Clearly 1 is an upper bound. Suppose $lub(S_a) = 1 - \delta \ \exists \ \delta > 0$. Then $\Sigma_{k \leq n} \frac{1}{2^k} \leq 1 - \delta$ for all $n \in \mathbb{N}$. But then $1 - \frac{1}{2^n} \leq 1 - \delta \Rightarrow \delta \leq \frac{1}{2^n}$ for all $n \in \mathbb{N}$. By the Archimedean property, there exists $N \in \mathbb{N}$ such that $N > \frac{1}{\delta}$. Now, one can prove by induction, that $2^N > N$ for all $N \in \mathbb{N}$, so this gives $2^N > \frac{1}{\delta}$. This is a contradiction. ♣

A consequence of the above proposition is the following. For any sequence $\mathbf{a}$ which *terminates* in ones, i.e. $a_i = 1$ for all $i \geq N_0$ for some N_0 there is another binary sequence $\mathbf{b}$ such that $b_i = 0$ for all $i \geq N_0$ and $lub(S_{\mathbf{a}}) = lub(S_{\mathbf{b}})$. Let us now denote by $\mathcal{A}$ the set of all binary sequences that do NOT terminate in ones.

We now can prove the following

Theorem 16 *The maps* F, G*,*

$$\begin{aligned} F : \mathcal{A} \ &\to \ (0,1), \ \textit{defined as } F(\mathbf{a}) = lub(S_{\mathbf{a}}), \\ G : (0,1) \ &\to \ \mathcal{A}, \ \textit{defined as } G(x) = \mathbf{a}_x \end{aligned}$$

as defined in class, are inverses of each other, i.e., $F(G(x)) = x, G(F(\mathbf{a})) = \mathbf{a}$.

Proof: The first statement essentially says that the lub of the binary sequence generated by x is x itself. Let

$$y = lub(S_{\mathbf{a}}),$$

and consider the sequence generated by y; call it $\mathbf{b}$. If $\mathbf{a} = \mathbf{b}$ then $y = x$, and we are done. So suppose they are not equal. Then the sequences $\mathbf{a}, \mathbf{b}$ must differ for the first time at some position, say the i^{th} position.

Case 1:

$$a_i = 1 \ and \ b_i = 0.$$

In this case

$$y \ - \ (\ \frac{b_1}{2} \ + \ \frac{b_2}{2^2} \ + \ ..... \ + \ \frac{b_{i-1}}{2^{i-1}} \ ) \ < \frac{1}{2^i} \text{ so that}$$
$$y \ < \ \frac{b_1}{2} \ + \ \frac{b_2}{2^2} \ + \ ..... \ + \frac{b_{i-1}}{2^{i-1}} + \ \frac{1}{2^i} \ .$$

But this term on the RHS is an element of $S_{\mathbf{a}}$, so y is not an upper bound for $S_{\mathbf{a}}$ contradicting that $y = lub(S_{\mathbf{a}})$.

Case 2:

$$a_i = 0 \ and \ b_i = 1.$$

Clearly x is an upper bound for $S_{\mathbf{a}}$. Moreover

$$y \ - \ (\ \frac{b_1}{2} \ + \ \frac{b_2}{2^2} \ + \ ..... \ + \ \frac{b_{i-1}}{2^{i-1}} \ ) \ \geq \frac{1}{2^i}$$

so

$$y \ \geq \ (\ \frac{b_1}{2} \ + \ \frac{b_2}{2^2} \ + \ ..... \ + \ \frac{1}{2^i} \ ).$$

But

$$x \ - \ (\ \frac{a_1}{2} \ + \ \frac{a_2}{2^2} \ + \ ..... \ + \ \frac{a_{i-1}}{2^{i-1}} \ ) \ < \frac{1}{2^i}$$

so that

$$x \ < \ \frac{a_1}{2} \ + \ \frac{a_2}{2^2} \ + \ ..... + \frac{a_{i-1}}{2^{i-1}} + \ \frac{1}{2^i} \ .$$

This implies that

$$x \ < \ \frac{b_1}{2} \ + \ \frac{b_2}{2^2} \ + \ ..... \ + \ \frac{1}{2^i},$$

because $\mathbf{a}$ and $\mathbf{b}$ differ for the first time at the i^{th} position. This proves $y > x$. But this is a contradiction because x is an upper bound for $S_{\mathbf{a}}$. Hence

$$F(G(x)) = x,$$

The details for the other part, namely,

$$G(F(\mathbf{a})) = \mathbf{a}$$

are similar. ♣

1.6 Cardinality

Until now, we have talked a great deal about natural numbers, integers, rationals and then the reals, but haven't had a look at the *sizes* of these sets. A measure of the number of elements in any set in general motivates the idea of cardinality.

Definition 17 *We say that two sets A and B have the same cardinality iff* $\exists$ *a bijection* $h : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$

Definition 18 *Set A is said to have cardinality n (for some* $n \in \mathbb{N}$ *if A has same cardinality as 0,1,2,...,n-1 (or equivalently 1,2,...,n). We write it as :* $|A| = n$

Remark 5 *We denote the set 1,2,...,m by [m].*

Proposition 11 *Given m,n*$\in \mathbb{N}$ *where* $m < n$*, there is no injection* $f : [n] \to [m]$.

Proof: Let S $= \{n \in \mathbb{N} \mid \exists m \in \mathbb{N}, m < n$ such that there is an injection $f_{nm} : [n] \to [m]\}$
We now need to prove that S $= \Phi$.
If not, then by the **W**ell-Ordering Principle, it must have a least element (say n_0) and a corresponding $m < n_0$.
Let h be the injection $h : \{1, 2, ..., n_0\} \to \{1, 2, ..., m\}$. Consider the value $h(n_0)$.

<u>Case I:</u> $h(n_0) = m$
Construct the function *h'* such that

$$h' : [n_0 - 1] \to [m - 1]$$
$$h'(i) = h(i) \,\forall\, i \in \{1, 2, ..., n_0 - 1\}$$

Then *h'* is also an injection. Hence $(n_0 - 1) \in$S which contradicts the fact that n_0 is the least element of S.

<u>Case II:</u> $h(n_0) \neq m$
Suppose $h(n_0) = k,\ k < m$
Let π be the map (km), i.e., the map that permutes k and m.
Then, we have an injection $\pi \circ h : [n] \rightarrowtail [m]$ with $(\pi \circ h)(n) = m$ which is a function satisfying Case I, hence again leading to a contradiction.
So, our assumption was wrong. S must be Φ. ♣

Corollary 2 *If* $A \rightarrowtail\!\!\!\!\twoheadrightarrow [n]$*, then we cannot have* $A \rightarrowtail\!\!\!\!\twoheadrightarrow [m]$ *for any* $m \neq n$.

Definition 19 *We say that a non-empty set A is INFINITE iff there is no* $n \in \mathbb{N}$ *such that* $|A| = n$.

Remark 6 $|\Phi| := 0$.

Proposition 12 $\mathbb{N}$ *is infinite.*

Proof: Suppose $f : \mathbb{N} \rightarrowtail\!\!\!\!\twoheadrightarrow [n]$ for some n. In particular, restrict f to $[n+1]$. This defines an injection $f' : [n+1] \rightarrowtail [n]$ which is a contradiction. ♣

Proposition 13 *Suppose* $A \subseteq \mathbb{N}$*. Then either* $|A| = n$ *for some* $n \in \mathbb{N}_0$ *or* A *has the same cardinality as* $\mathbb{N}$*.*

Proof: If $A \subseteq \mathbb{N}$ is finite, we say that $|A| = n$ for some $n \in \mathbb{N}_0$. Also, if $A = \mathbb{N}$, we are through. So, the only thing that remains to be proved is that if A is infinite and $A \neq \mathbb{N}$, there exists

$$f : A \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbb{N}$$

Since $A \neq \Phi$, it has a least element, say a_1. Let $A_0 = A$.
Consider $A_1 = A_0 \setminus \{a_1\}$. A_1 must be infinite. Because if it weren't, then A_0 would have just one more element than A_1 and hence also be finite.
So, there must exist a minimum element in A_1, say a_2. Construct $A_2 = A_1 \setminus \{a_2\}$.
Inductively, obtain $a_{n+1} = min(A_n)$; $A_{n+1} = A_n \setminus \{a_{n+1}\}$
The bijection is:

$$f : \mathbb{N} \to A$$
$$f(i) = a_i$$

f is clearly injective since $a_i < a_j \forall i < j$. Hence also observe that $a_i \geq i$ (can be proved using induction). We now need to prove surjectivity of f.
Suppose f is not surjective. Since $A \neq \mathbb{N}$, we must have an $x \in A$ such that $a_x > x$ and $x \neq a_i \forall i$. But in this case, $x \in A_{x-1}$ and $x < a_x$ which is a contradiction. So, f must be surjective and hence bijective too. ♣

Corollary 3 *There are as many primes as elements of* $\mathbb{N}$*.*

Remark 7 *We write* $|A| < \infty$ *if* $|A| = n$ *for some* $n \in \mathbb{N}_0$*.*
If $|A| < \infty, B \subset A$*, then* $|B| < |A|$*.*

Definition 20 *We say that a set is countable (countably infinite) if it has same cardinality as* $\mathbb{N}$*.*

Observations:

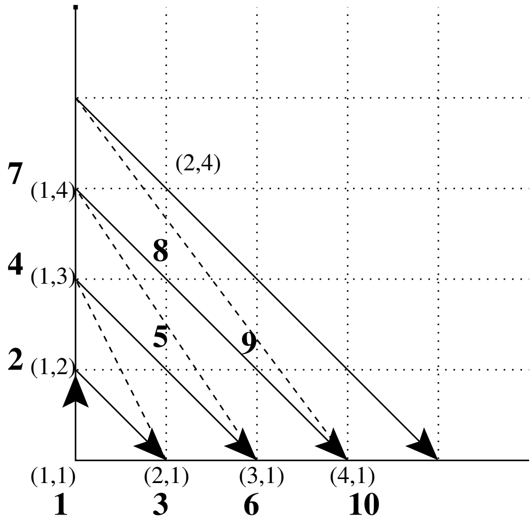
1. $|Even\ naturals| = |\,\mathbb{N}\,|$
   This is because we have the bijection

$$f : 2\,\mathbb{N} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbb{N}$$
$$f(2k) = k$$

2. $|\mathbb{Z}| = |\mathbb{N}|$
   We can build the bijection here as

$$f : \mathbb{N} \rightarrowtail\!\!\!\!\!\twoheadrightarrow \mathbb{Z}$$
$$f(2k) = k$$
$$f(2k+1) = -k$$

3. $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$ This can be shown using the bijection as depicted in the diagram.



Theorem 17 $\mathbb{Q}$ *is* ***C****OUNTABLE*

Proof: Every $x \in \mathbb{Q}$ is uniquely of the form $x = \dfrac{p}{q}$, $p \in \mathbb{Z}$, $q \in \mathbb{N}$, where p and q are relatively prime.

<u>Proof 1</u>: $\mathbb{Q} \subset \mathbb{Z} \times \mathbb{Z}$ and $|\mathbb{Z} \times \mathbb{Z}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$,
and a subset of $\mathbb{N}$ is either finite or countable.

<u>Proof 2</u>: We build a bijection between the sets $\mathbb{Q}^+$ and $\mathbb{N}$ where $\mathbb{Q}^+$ represents all positive rational numbers.
$x \in \mathbb{Q}^+ \Rightarrow x = \dfrac{m}{n}$, where m and n are relatively prime natural numbers.
Every $m, n \in \mathbb{N}$, $n > 1$ can be written (uniquely upto rearrangement) as a product of primes:

$$m = p_1^{\alpha_1} p_2^{\alpha_2} \ldots . p_k^{\alpha_k}$$
$$m = q_1^{\beta_1} q_2^{\beta_2} \ldots . q_l^{\beta_l}$$

So, we map

$$\frac{m}{n} \mapsto (p_1^{2\alpha_1} p_2^{2\alpha_2} ....p_k^{2\alpha_k} q_1^{2\beta_1 - 1} q_2^{2\beta_2 - 1} ....q_l^{2\beta_l - 1})$$

This is a bijection as an even exponent will map to a factor in the numerator while an odd one will give me one in the denominator. ♣

What about $\mathbb{R}$? Is it also countable?

In the following material, we attempt to answer this particular question.

Definition 21 *For a set A, the power set of A is defined as $\mathcal{P}(A) := \{B | B \subseteq A\}$.*

Theorem 18 $|\,\mathbb{N}\,| \neq |\mathcal{P}(\mathbb{N})|$

Proof: Here, we simply show that there can't be an injection from $\mathcal{P}(\mathbb{N})$ to $\mathbb{N}$.
Identify $\mathcal{P}(\mathbb{N})$ with binary sequences in the following way:
Given $A \subseteq \mathbb{N}$, for $i \in \mathbb{N}$, $a_A(i) = 1$ if $i \in A$ and 0 otherwise.
Suppose if possible that such an injection exists.

$$f : \mathbb{N} \mapsto BinarySequences$$

Construct the binary string b such that $b_i = 1$ if $f(i)$ has 0 at i^{th} position and o otherwise. Then, b can't be the image of any natural number and hence f cannot be surjective, leading to the fact that the injection we wanted can't exist. ♣

Observation: If $x \in \mathbb{R}$ is associated with binary string α where α ends in all zeroes, then $x \in \mathbb{Q}$.

Theorem 19 $\mathbb{R}$ *is not countable.*

Proof: We will show that there is no bijection from from the real numbers in (0,1) to $\mathbb{N}$. We once again use the idea of identifying elements (real numbers in this case) with strings. The only difference being that we use ternary strings instead of binary in order to avoid dealing with such strings that end in all 1s.

Each real number between 0 and 1 can be represented in ternary notation as a string of 0s, 1s and 2s, leading to a one-to-one mapping between (0,1) and the set T of all ternary strings on 0,1,2.
Suppose f is a bijection from $\mathbb{N}$ to T.

Using Cantor's diagonalization idea,
Construct a ternary string t wherein $t_i = 0$ if $f(i)$ has 1 or 2 in the i^{th} position and 1 otherwise. Then t cannot be the image of any natural number, ruling out the possibility of the existence of a bijection between (0,1) and $\mathbb{N}$. ♣

Proposition 14 $|(0,1)| = |\mathbb{R}|$

Proof: Construct the bijection:

$$f : (0,1) \to \mathbb{R}$$
$$x \mapsto \frac{2x-1}{2x(1-x)}$$

♣

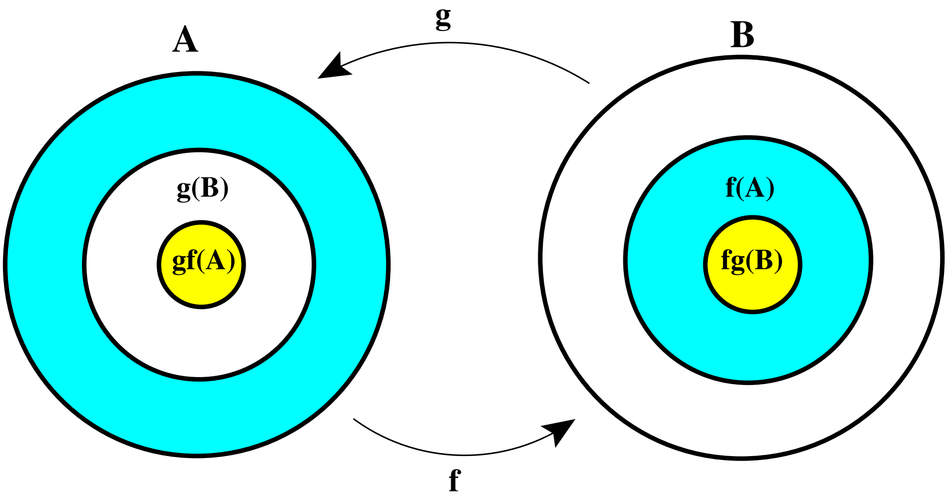
Definition 22 *We say that $|A| \geq |B|$ if there is an injection $f : B \rightarrowtail A$.*

Theorem 20 *(SCHRÖDER-BERNSTEIN)*
If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.

Proof: Consider two sets A and B such that there is an injection f from A to B and an injection g from B to A. Construct the sequences:

$A_0 = A$, $A_1 = g \circ f(A_0$, $A_2 = g \circ f(A_1$, ...
$B_0 = B$, $B_1 = f \circ g(B_0$, $B_2 = f \circ g(B_1$, ...

We can define our bijections between $A_{i-1} \setminus A_i$ and $B_{i-1} \setminus B_i$, $i \geq 1$.
In the figure that follows, the blue parts are in bijection and so are the ones in white.



Consider the blue parts. Under f, the range of A is $f(A)$ which is precisely the blue and yellow coloured parts in B. Also, the range of $g(B)$ is the yellow coloured part in B. So, since f is an injection, the range of $A \setminus g(B)$ must be the blue part in B, meaning that the two blue parts are in bijection.
By symmetry, the two white parts are also in bijection. Hence, extending this argument to the entire sequence, we will need to show that A_{res} and B_{res} are in bijection where

$$A_{res} = A \setminus (\bigcup_{i\geq 1} A_{i-1} \setminus A_i)$$
$$B_{res} = B \setminus (\bigcup_{i\geq 1} B_{i-1} \setminus B_i)$$

Claim: $f(A_{res}) = B_{res}$. Suppose there exists a $y \in B_{res}$ which does not have a pre-image in A_{res} under f. Say $f(x) = y$. Then $x \in (\bigcup_{i\geq 1} A_{i-1} \setminus A_i)$. But this means that $x \in A_i$ for some i. Then $y \in B_{res}$ is a contradiction.
Therefore, A_{res} and B_{res} must be in bijection and this completes our proof. ♣

Remark 8 *So, we say that A and B are in bijection if we have two injections, one from A to B and one from B to A.*

Examples: Consider the two sets (0,1) and [0,1]. The two injections are as follows:

$$f : (0, 1) \to [0, 1]$$
$$x \mapsto x \ g : [0, 1] \to (0, 1)$$
$$x \mapsto \frac{2x + 1}{4}$$

Theorem 21 *For any set A, $|A| < |\mathcal{P}(A)|$*

Proof: If A is finite, we have $|\mathcal{P}(A)| = 2^{|A|}$. Now, we need to prove the theorem for infinite sets. Suppose there is a bijection between A and $\mathcal{P}(A)|$ where A is infinite.

$$f : A \to \mathcal{P}(A) = \{B \subseteq A\} \ a \mapsto f(a)$$

Construct $B := a \in A| \notin f(a)$.
Suppose $f^{-1}(B) = b \Rightarrow B = f(b)$.
If $b \in B$, then $b \notin B$.
If $b \notin B$, then $b \in B$.
This is a contradiction. Hence such a bijection cannot exist. ♣

1.7 The Complex Numbers

A simple algebraic equation like $X^2 = -1$ may not have a real solution. Introducing complex numbers validates the so called fundamental theorem of algebra: every polynomial with a positive degree has a root (though we will not prove that here).

Definition 23 *A complex number is a pair $z = (a, b)$ where $a, b \in \mathbb{R}$ and we write $z = a + ib$.*

$\mathbb{R} \times \mathbb{R} \approx \mathbb{C}$ as sets.

Consider two complex numbers $z_1 = (a_1, b_1)$ and $z_2 = (a_2, b_2)$

1. $z_1 + z_2 = (a_1 + a_2, b_1 + b_2)$
2. $z_1 z_2 = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1)$
3. $\bar{z_1} = (a_1, -b_1)$. This is called the conjugate of z_1.
4. $|z_1| = \sqrt{a_1^2 + b_1^2}$. Note that $|z_1| = 0$ iff $a_1 = b_1 = 0$.
5. For $z_1 \neq 0$, $z_1^{-1} = \dfrac{\bar{z_1}}{|z|^2}$ and $z\, z^{-1} = 1$.

Proposition 15 *$\mathbb{C}$ along with the operations $+, \cdot$ as defined above make it a field with additive and multiplicative identities being $0, 1$ respectively. However, there is no total order on $\mathbb{C}$, satisfying the following order properties:*

- *$a \in \mathbb{C} \Rightarrow a > 0$, $a = 0$ or $a < 0$.*
- *$a, b > 0 \Rightarrow a + b > 0$.*
- *$a > 0$, $b > c \Rightarrow ab > ac$.*

Proof: We will only prove the latter statement here. Suppose if possible that there exists a total order on $\mathbb{C}$ satisfying the properties listed above.

Since $i \neq 0$, i must be either less than or greater than 0.
Suppose $i > 0$, then $i.i = i^2 = -1 > 0$ where $-1 \in \mathbb{C}$. Hence $-1 \times -1 = 1 > 0$ where $1 \in \mathbb{C}$. $-1 > 0$ and $1 > 0 \Rightarrow -1 + 1 = 0 > 0$ which is a contradiction. Hence, i must be less than 0 for which we can find a similar contradiction. ♣

2 Basic Topology

2.1 Metric Spaces

We start with some basic inequalities.

1. $x \in \mathbb{R} \implies x^2 \geq 0; x^2 = 0; \text{iff } x = 0.$
2. $x \in \mathbb{R} \implies$

$$|x| := \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

Clearly, $x \leq |x|$ for all $x \in \mathbb{R}$.

Theorem 22 ***The Triangle Inequality:***

$$x + y \in \mathbb{R} \implies |x + y| \leq |x| + |y|.$$

Proof: $x + y \leq |x| + |y|$, comes from adding $x \leq |x|, y \leq |y|$ for all $x \in \mathbb{R}$ and $y \in \mathbb{R}$. The other inequality, namely, $-(x + y) \leq |x| + |y|$, comes from adding $-x \leq |x|, -y \leq |y|$ for all $x \in \mathbb{R}$ and $y \in \mathbb{R}$. ♣

Definition 24 ***Metric Space:*** *A metric space is a set $\mathfrak{X}$, along with a function $d : \mathfrak{X} \times \mathfrak{X} \to \mathbb{R}^+$ ($\mathbb{R}^+ = \{x \in \mathbb{R} | x \geq 0\}$) satisfying:*

1. *$d(x, y) = d(y, x)$,for all $x, y \in \mathfrak{X}$.*
2. *$d(x, y) = 0$ if and only if $x = y$ for all $x, y \in \mathfrak{X}$.*
3. ***Triangle Inequality**: $d(x, y) \leq d(x, z) + d(y, z)$ for all $x, y, z \in \mathfrak{X}$.*

Here are some examples:

1. $\mathfrak{X} = \mathbb{R}, d(x, y) = |x - y|$.
2. $\mathfrak{X} = \mathbb{Q}, d(x, y) = |x - y|$.

3. $\mathfrak{X} = \mathbb{C}, d(z, w) = |z - w|$.

4. $\mathfrak{X} = \{0, 1\}^n$, $x, y \in \mathbb{X}$ i.e x, y are binary strings. $d(x, y) = \#\{1 \leq i \leq n | x_i \neq y_i\}$.

5. One of the most commonly used metrics is $\mathfrak{X} = \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, with the distance function defined as $d((x, y), (x', y')) = \sqrt{(x - x')^2 + (y - y')^2}$.

6. One also uses the other metric (called the Taxicab metric) on $\mathbb{R}^2$ with the distance function $d((x, y), (x', y')) = |x - x'| + |y - y'|$.

Definition 25 ***Limit point:*** *Suppose $A \subseteq \mathbb{R}$. We say that $x \in \mathbb{R}$ is a Limit point of A if for any $\epsilon > 0$ we have $(x - \epsilon, x + \epsilon) \cap (A - \{x\}) \neq \emptyset$.*

Definition 26 ***Cauchy sequence:*** *A sequence x_n in the reals is called a Cauchy sequence if given $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that for all $m, n \in N_\epsilon$ we have $|x_n - x_m| < \epsilon$.*

The following proposition is an immediate consequence.

Proposition 16 *If $\{x_n\}$ is Cauchy the the set $\{x_n\}$ is* ***bounded****. Equivalently, there exists $M > 0$ such that $x_n \in [-M, M]$ for all $n \in \mathbb{N}$.*

Proof: For $\epsilon = \frac{1}{2}$, there exists $N = N_{\frac{1}{2}}$ such that for all $n \geq N, x_n \in (x_N - \frac{1}{2}, x_N + \frac{1}{2})$. Let

$$M = |x_1| + |x_2| + \ldots |x_N| + \frac{1}{2}.$$

We claim that this value of M does the job as stated in the proposition.

Now for all $n < N, x_n < |x_n| + \frac{1}{2}$, so $x_n < M$. Also if $n \geq N$, then

$$\begin{aligned} x_n - x_N \in (-\frac{1}{2}, \frac{1}{2}) &\implies |x_n - x_N| < \frac{1}{2}, \\ &\implies |x_n| - |x_N| < \frac{1}{2}, \\ &\implies x_n < |x_n| < |x_N| + \frac{1}{2}, \\ &\implies x_n < M. \end{aligned}$$

♣

Example 3 $x_n = (-1)^n \frac{1}{n}$. *Is $\{x_n\}$ Cauchy?*

Suppose $n < m$. Then

$$\begin{aligned}|x_n - x_m| &= |(-1)^n\{\frac{1}{n} - \frac{(-1)^{m+n}}{m}\}|,\\ &= |\frac{1}{n} - \frac{(-1)^{m+n}}{m}\}|,\\ &\leq \frac{1}{n} + \frac{1}{m} < \frac{2}{n}.\end{aligned}$$

Consider an $\epsilon > 0$. By the archimedian property, $n_0 \leq \frac{2}{\epsilon} < n_0 + 1$ for some n_0. So $\frac{2}{n_0} \geq \epsilon \geq \frac{2}{n_0+1}$. Therefore $|x_n - x_m| < \epsilon$, for every $n > n_0$. Hence this sequence is indeed Cauchy.

Proposition 17 *Suppose* $\{x_n\}$*is Cauchy. Then the set* $\{x_n\}$ contains at most one limit point.

Proof: Suppose $x < y$ and x, y are both limit points of $\{x_n | n \in \mathbb{N}\}$. Since x is a limit point of $\{x_n | n \in \mathbb{N}\} = A$. For every small enough $\epsilon > 0$, the set $(x - \epsilon, x + \epsilon) \cap A$ is infinite. Similarly, $(y - \epsilon, y + \epsilon) \cap A$ is infinite. Since $\{x_n\}$ is Cauchy, there exists N_ϵ such that for all $m, n \geq N_\epsilon, |x_n - x_m| < \epsilon$. By (the first observation above there exists $n \geq N_\epsilon$, such that $x_n \in (x - \epsilon, x + \epsilon) \cap A$. In the same vein, there exists $m \geq N_\epsilon$, such that $x_m \in (y - \epsilon, y + \epsilon) \cap A$. This gives $|x_n - x_m| < \epsilon$. However, on the other hand, for any $\alpha \in (x - \epsilon, x + \epsilon)$,$\beta \in (y - \epsilon, y + \epsilon)$

$$\begin{aligned}|\alpha - \beta| &\geq (y - \epsilon) - (x + \epsilon),\\ &= y - x - 2\epsilon,\\ &\geq 98\frac{y - x}{100},\\ &= 98\epsilon,\end{aligned}$$

where $\epsilon = \frac{y-x}{100}$; this gives a contradiction and our proof is complete. ♣

Definition 27 *For any metric* $(\mathfrak{X}, d)$*, for any point* $x \in \mathfrak{X}$*,we define the open ball of radius* r *at* x *as*$B_r(x) := \{y \in \mathfrak{X} | d(x, y) < r\}$.

Remark 9 *One defines a limit point in an arbitrary metric space in the same manner as in the case of the real line. Indeed, for* $A \subseteq \mathfrak{X}$*, we say that* a *is a limit point of the set* A *if and only if* $B_a(\epsilon) \cap A \setminus \{a\} \neq \emptyset$ *for every* $r > 0$.

We also define a Cauchy sequence in a metric space in the same manner.

Definition 28 ***Cauchy Sequence*** *in* $\mathfrak{X}$*:* $\{x_n\}$*is Cauchy in* $\mathfrak{X}$ *iff for every* $\epsilon > 0$*, there exists* $N_\epsilon \in \mathbb{N}$ *such that for m,n*$\geq N_\epsilon$*, we have* $d(x_m, x_n) < \epsilon$.

The following proposition follows along the same lines as the real case.

Proposition 18 *A **Cauchy sequence** in $(\mathfrak{X}, d)$ has atmost one limit point.*

However, in the case of the reals, we are guaranteed a limit point as the following theorem states

Theorem 23 *Every Cauchy sequence in $\mathbb{R}$ has a limit point.*

Remark 10 *Not all metric spaces enjoy this privilege. Indeed, the metric $\mathfrak{X} = \mathbb{Q}$ has Cauchy sequences that do not have any limit in $\mathbb{Q}$.*

Proof: Suppose $\{x_n\}$ is Cauchy. As seen before, $\{x_n\}$ is bounded, i.e., there exists $M > 0$ such that $x_n \in [-M, M]$ for all $n \in \mathbb{N}$. Consider

$$A = \{a \in \mathbb{R} \mid a \leq x_n \text{ for infinitely many } n\}.$$

Clearly, $A \neq \emptyset$ since $-M \in A$. Also, $A \subseteq [-\infty, M]$. So, A is non-empty and bounded above. Let $(A) = x$.

Claim 13 *x is a limit point for the sequence $\{x_n\}$. In other words, for any $\epsilon > 0, (x - \epsilon, x + \epsilon)$ contains some x_n.*

Suppose not. Then there exists some $\epsilon > 0$ such that $(x - \epsilon, x + \epsilon)$ has no element of x_n. Since $x = lub\,(A)$ $x - \epsilon$ is not an upper bound for A, so there exist infinitely many $x_n > x - \epsilon$. But this implies that all elements of $\{x_n\}$ greater than $x - \epsilon$ must in fact be larger than $x + \epsilon$. This gives $x + \epsilon \in A$, a contradiction since $x = lub\,(A)$. This completes the proof of the claim, and consequently, every Cauchy sequence in $\mathbb{R}$ has a limit point. ♣

We say that a set $A \in \mathfrak{X}$ is bounded if $A \subset B_r(x)$ for some $x \in \mathfrak{X}$ and a suitable $r > 0$. The following proposition is also easy.

Proposition 19 *In a metric space, every Cauchy sequence is bounded.*

Definition 29 *Given $\{x_n\} \subseteq \mathfrak{X}$, we say that $\{x_n\}$ **converges to a limit** if it is Cauchy and has a limit point.*

In symbols, $\lim\limits_{n \to \infty} x_n = x$ if for any $\epsilon > 0$ there exists $N_\epsilon \in \mathbb{N}$ such that for all $n \geq N_\epsilon$ we have $d(x_n, x) < \epsilon$.

Example 4 $x_n = \frac{1}{n}$

Claim 14 $\lim_{n \to \infty} \frac{1}{n} = 0$
$|\frac{1}{n} - 0| = \frac{1}{n} < \epsilon$ *for large enough n (Archimedian Property)*

Proposition 20 $(\mathfrak{X} = \mathbb{R}, |.|)$ *Suppose* $\{x_n\}$*and* $\{y_n\}$*are sequences in* $\mathbb{R}$. *Let*

$$\lim_{n\to\infty} x_n = X, \lim_{n\to\infty} y_n = Y.$$

1. $\lim_{n\to\infty} x_n \pm y_n = X \pm Y$.
2. $\lim_{n\to\infty} \lambda.x_n = \lambda.X$.
3. $\lim_{n\to\infty} x_n.y_n = X.Y$.
4. *If* $y_n \neq 0$ *for all* $n > N$ *for some* N *and* $\lim_{n\to\infty} y_n \neq 0$. *then* $\lim_{n\to\infty} \frac{x_n}{y_n} = \frac{X}{Y}$.

Proof: We only write the proofs for statements (3), (4) in the proposition above.

$$\begin{aligned} \lim_{n\to\infty} x_n.y_n &= X.Y \\ |x_n.y_n - X.Y| &= |x_n.y_n - X.y_n + X.y_n - X.Y| \\ &\leq |y_n|.|x_n - X| + |X|.|y_n - Y| \end{aligned}$$

If $x = 0$, then the second term vanishes; otherwise choose N_{1_ϵ} such that $|y_n - Y| < \frac{\epsilon}{2|x|}$ for all $n \geq N_{1_\epsilon}$. Since $\lim_{n\to\infty} y_n = Y$, $|y_n| \leq M$ for some $M > 0$ for all $n \in \mathbb{N}$.

So, for the first term pick N_{2_ϵ} such that $|x_n - X| < \frac{\epsilon}{2M}$. Hence,

$$|x_n.y_n - X.Y| < M.\frac{\epsilon}{2M} + \frac{\epsilon}{2} = \epsilon.$$

We need to show that $\frac{1}{y_n}$ converges to $\frac{1}{y}$. Now,

$$|\frac{1}{y_n} - \frac{1}{Y}| = \frac{|Y - y_n|}{|Y||y_n|}.$$

Suppose $\lim_{n\to\infty} y_n = Y \neq 0$. Then there exists N_ϵ such that for all $n \geq N_\epsilon$ (where $\epsilon = \frac{|Y|}{2}$) ,we have $|y_n - Y| < \frac{|Y|}{2}$. This implies $y_n \geq \frac{|Y|}{2} \implies \frac{|Y-y_n|}{|Y||y_n|} \leq 2.\frac{|Y-y_n|}{|Y|^2}$ for all $n > N$. Pick $N^* \geq N$ such that $|Y - y_n| < \frac{|y|^2\epsilon}{2}$ for $n \geq N^*$. ♣

The following proposition is often useful in deducing the existence of a limit of a sequence.

Proposition 21 *Suppose* $\{x_n\}$ *is increasing and bounded above, then* $\lim_{n\to\infty} x_n$ *exists and equals* $lub(\{x_n\})$.

The same conclusion also holds if the sequence is decreasing and bounded below.

Example 5 *If $x > 0$ and $0 < x < 1$ let $x_n = x^n$. Note that $x^n > x^{n+1}$ (This can be proved by Induction.), so $\{x_n\}$ is decreasing. By the previous proposition, since $x > 0$ for all n, $\lim_{n\to\infty} x^n = L$ exists and $\lim_{n\to\infty} x^{n+1} = L$ as well. Hence,*

$$L = \lim_{n\to\infty} x^{n+1} = x. \lim_{n\to\infty} x^n = x.L,$$

so

$$\implies L.(1-x) = 0 \implies L = 0.$$

If $x > 1$ then $\{x_n\}$ is not bounded. Indeed, suppose not, i.e., suppose $x^n \leq M$ for all n. Then $x^n < x^{n+1}$ so x^n is increasing. But then by the proposition above the limit would exist, and then by the same argument the limit would be zero. But since $x > 1$ thesis leads to a contradiction.

Proposition 22 *$e := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$ exists*

Proof: Let $x_n = \left(1 + \frac{1}{n}\right)^n$. We show that x_n is increasing and bounded above. Indeed,

$$\begin{aligned}
x_n &= \left(1 + \frac{1}{n}\right)^n, \\
&= 1^n + n\frac{1}{n} + \frac{n(n-1)}{2!}\left(\frac{1}{n}\right)^2 + \ldots + \frac{n(n-1)\ldots 1}{n!}\left(\frac{1}{n}\right)^n, \\
&= 1 + \frac{n}{n} + \frac{\frac{n}{n}\left(\frac{n-1}{n}\right)}{2!} + \ldots + \frac{\frac{n}{n}\left(\frac{n-1}{n}\right)\ldots\frac{1}{1}}{n!},
\end{aligned}$$

Hence,

$$x_{n+1} = 1 + \frac{n+1}{n+1} + \frac{\left(\frac{n+1}{n+1}\right)\left(\frac{n}{n+1}\right)}{2!} + \ldots + \frac{\left(\frac{n+1}{n+1}\right)\left(\frac{n}{n+1}\right)\ldots\frac{1}{1}}{(n+1)!}.$$

We know that $1 - \frac{k}{n+1} > 1 - \frac{k}{n}$; also, x_{n+1} has more number of terms of than x_n, so it follows that $x_{n+1} > x_n$. Also,

$$x_n \leq 1 + \frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{n!} \leq 1 + \frac{1}{1} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots + \frac{1}{2^n} < 3.$$

Since the sequence is increasing and bounded above, by the above proposition the limit exists. ♣

Example 6 $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Write $\sqrt[n]{n} = 1 + x_n$. Then, we have

$$\begin{aligned}
n &= (1 + x_n)^n \\
&\geq x_n^2 \binom{n}{2} \\
&> \sqrt{\frac{2}{n-1}} \text{ for } n \geq 2.
\end{aligned}$$

Now, $\lim_{n\to\infty}\sqrt{\frac{2}{n-1}}$ *exists and is equal to* 0*. Also,* $x_n > 0$*. So, by the sandwich theorem,* $\lim_{n\to\infty} x_n = 0$ *and so* $\lim_{n\to\infty}\sqrt[n]{n} = 1 + 0 = 1$*.*

2.2 Subsequences

Definition 30 *Given a sequence* $\{x_n\}$ *in* $\mathbb{R}$ *any sequence* $\{x_{n_k}\}_{k\geq 1}$ *where* $\{n_k\}$ *is an increasing infinite subset of* $\mathbb{N}$ *is called a subsequence of* $\{x_n\}$*.*

Proposition 23 *Any subsequence of a convergent sequence converges to same value i.e., if* $x_n \to X$ *then for any subsequence* $x_{n_k} \to X$ *as* $k \to \infty$*.*

Proof: Given $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that $\mid x_n - x \mid < \epsilon$ for all $n \geq N_\epsilon$. Since $\{n_k\}$ is an increasing infinite subset of $\mathbb{N}$, there are atmost finitely many k such that $n_k < N_\epsilon$ i.e., for $k > k_\epsilon$(for some k_ϵ) , we must have $n_k \geq N_\epsilon$. Therefore for all $k > k_\epsilon$, we must have $\mid x_{n_k} - x \mid < \epsilon$. ♣

Remark 11 *The converse of this proposition is not true. For instance the sequence 0,1,0,1,0,1..... does not converge even though it has convergent subsequences.*

We now state one of the most important theorems in the context of bounded sequences.

Theorem 24 ***Bolzano-Weierstrass Theorem*** *Every bounded real sequence admits a convergent sub-sequence.*

Proof: Without loss of generality we may assume $0 \leq x_n \leq 1$ for all n. There are two cases to deal with.

Case 1: $\{x_n\}$ has only finitely many distinct values.

In this case one element x must repeat infinitely many times. Consider that subsequence $\{x, x, x, x \ldots\}$ which clearly converges to x.

Case 2: $\{x_n\}$ has infinitely many distinct values:

Let $A = \{x_n | n \in \mathbb{N}\}$. Since A is infinite either $A \cap [0, \frac{1}{2}]$ is infinite or $A \cap [\frac{1}{2}, 1]$ is infinite.

Let I_1 be one of the intervals $[0, \frac{1}{2}], [\frac{1}{2}, 1]$ which contains infinitely many elements of A.Inductively if $I_k = [a, b]$ then let I_{k+1} be one of $[a, \frac{a+b}{2}], [\frac{a+b}{2}, b]$ which contains infinitely many elements of A.

This gives a sequence of intervals $I_0 \supset I_1 \supset I_2 \supset I_3 \cdots$ such that each I_k contains infinitely many elements of A.Write

$$I_0 = [a_0, b_0], I_1 = [a_1, b_1], I_2 = [a_2, b_2] \ldots$$

Since the sequence of intervals is nested, it follows that

$$a_0 \leq a_1 \leq a_2 \leq \cdots < b_0 \text{ and } b_0 \geq b_1 \geq b_2 \geq \cdots > a_0.$$

Since $\{a_n\}$ is increasing and bounded, $\lim_{n\to\infty} a_n = a$ exists.

For each k, pick $x_{n_k} \in I_k$.

$\{b_n\}$ is decreasing and bounded below. So $b_n \to b$ for some b as $n \to \infty$.

$a_n < b_n$, we must have $a \leq b$. Note that $b_k - a_k = \frac{1}{2^k}$. Thus,

$$\lim_{k\to\infty}(b_k - a_k) = \lim_{k\to\infty}\frac{1}{2^k} = 0 = \lim(b_k - a_k) = \lim_{k\to\infty} b_k - \lim_{k\to\infty} a_k,$$

which gives $b = a$.

Since $a_k \leq x_{n_k} \leq b_k$, the sandwich theorem implies that $\{x_{n_k}\}$ is convergent. ♣

Proposition 24 *Suppose $\{a_n\}$ is a sequence in $[0,1]$ and $a_n \to a$ for some a, then $a \in [0,1]$.*

Proof: Suppose not. Without loss of generality suppose $a > 1$; then $a > 1 + \frac{1}{n}$ for some natural number n. Then we have $(a-\delta, a+\delta) \cap [0,1] \neq \emptyset$ for $\delta = a - (1 + \frac{1}{n})$. But since $a_n \in [0,1]$ $(a-\delta, a+\delta)$ cannot have any element of $\{a_n\}$, and this contradicts that a_n converges to a. ♣

Remark 12 *Suppose $I_1 \supset I_2 \supset \ldots \supset I_n \supset \ldots$ is a sequence of bounded closed intervals. Pick $x_i \in I_i$; by the Bolzano-Weierstrass theorem, there exists $x_{n_k} \to x$ for some $x \in \mathbb{R}$. By the preceding proposition $x \in I_1$. Similarly $\{x_{n_r}\}_{r\geq k} \to x$ and $\{x_{n_r}\}_{r\geq k} \subseteq I_k$. So, $x \in I_k$ by the preceding proposition, and therefore $x \in I_k$ for all k. This implies that $\bigcap_{i\geq 1} I_i \neq \emptyset$.*

Remark 13 *The above property is not true on all nested families of sets of reals. Consider $I_k = (0, \frac{1}{k})$. Here $I_1 \supset I_2 \supset \ldots \supset I_n \supset \ldots$. In this case, $\bigcap_{i\geq 1} I_i = \emptyset$.*

Definition 31

1. *A set $U \in \mathbb{R}$ is **open** if for each $x \in U$, there exists a $\delta = \delta_x > 0$ for which $(x-\delta, x+\delta) \subset U$.*
2. *A set $U \in \mathbb{R}$ is closed if $\overline{U}$ is open.*

Any open interval in $\mathbb{R}$ is an open set clearly. It is also easy to see that a closed interval in $\mathbb{R}$ is a closed set.

Some Generalities: The following properties are easily verified.

1. A is open if and only if $\overline{A}$ is closed.
2. $\emptyset$ is closed; it is also open.
3. A set can be neither open nor closed: $U = (0, 1]$ not open and $\overline{U} = (-\infty, 0] \cup (1, \infty)$ is not open.
4. If A_i are open then $\bigcup A_i$ is open.
5. If A_i are closed then $\bigcap A_i$ is closed.
6. Finite intersection of open sets is open.
7. Finite intersection of closed sets is closed.
8. $\{a\}$ is closed.

Remark 14 $\bigcap_{n=1}^{\infty}(-\frac{1}{2^n}, \frac{1}{2^n}) = \{0\}$. *This need not be true for infinite sets.*

2.3 Continuity

Definition 32 *Suppose* $U \in \mathbb{R}$ *is open.* $f : U \implies \mathbb{R}$ *is said to be continuous at a point* $a \in U$ *if given* $\epsilon > 0$, *there exists* $\delta > 0$ *such that* $\mid x - y \mid < \delta \implies \mid f(x) - f(y) \mid < \epsilon$.

Proposition 25 $f : U \to \mathbb{R}$ *is continuous and* $\{x_n\}$ *is a sequence in* U *converges to* $x \in U$. *Then* $f(x_n) \to f(x)$.

Proof: Given $\epsilon > 0$, we need to prove: There exists N_ϵ such that $\mid f(x_n) - f(x) \mid < \epsilon$ for all $n \geq N_\epsilon$.

Since f is continuous there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \delta$. Since U is open, there exists $\delta_1 > 0$ such that $(x - \delta_1, x + \delta_1) \subset U$. Since $x_n \to x$ there exists N_ϵ such that $\mid x_n - x \mid < \delta_1$ if $n \geq N_\epsilon$. Hence for $n \geq N_\epsilon$,we have $\mid x_n - x \mid < \delta \implies \mid f(x_n) - f(x) \mid < \epsilon$. ♣

Proposition 26 f *is continuous at* x *if and only if every sequence* $\{x_n\} \implies x$, *we also have* $f(x_n) \implies f(x)$.

Proof: It suffices to prove the sufficiency part of the proposition. Suppose it does not hold, i.e., suppose f is not continuous at some x i.e. given $\epsilon > 0$ there exists $\delta > 0$ such that $\mid x - y \mid < \delta \implies \mid f(x) - f(y) \mid < \epsilon$. This is the same as saying that there exists $\epsilon > 0$ such that for every $\delta > 0$, there is some y such that $\mid x - y_\delta \mid < \delta$, but $\mid f(x) - f(y_\delta) \mid \geq \epsilon$.

In particular, for $\delta = \frac{1}{n}$, (for every $n \in \mathbb{N}$), there exists y_n such that $\mid x - y_n \mid < \frac{1}{n}$ and $\mid f(x) - f(y_n) \mid \geq \epsilon$. Now, $\{y_n\} \to x$ and still $\{f(y_n)\}$ does not converge to $f(x)$ - this contradicts the hypothesis. ♣

Proposition 27 *A set $A \subseteq \mathbb{R}$ is closed if and only if it contains all its limit points.*

Proof: Recall that $x \in \mathbb{R}$ is a limit point of A if and only if for each $\delta > 0$ $(x-\delta, x+\delta) \cap (A - \{x\}) \neq \emptyset$.

Proof of Necessity:

Suppose A is closed. Let x be a limit point of A. If $x \notin A$ then $x \in \overline{A}$ which implies that there exists $\delta > 0$ such that $(x-\delta, x+\delta) \subset \overline{A}$. But then for this $\delta > 0$, $\{(A - \{x\}) \cap (x-\delta, x+\delta)\} = \emptyset$. but this contradicts the hypothesis that A is closed.

Proof of Sufficiency:

Suppose A contains all its limit points. It suffices to prove that $\overline{A}$ is open.
Pick $x \in \overline{A}$. We claim that there exists $\delta > 0$ such that $(x-\delta, x+\delta) \cap A = \emptyset$.

Suppose not,then for each $\delta = \frac{1}{n}$, there exists $x_n \in (x - \frac{1}{n}, x + \frac{1}{n}) \cap A$. Let $B = \{x_n | n \geq 1\}$. By definition, x is a limit point of B, but B is contained in A. In other words x is a limit point of $B \subseteq A$, and therefore $x \in A$, contradicting the hypothesis that $x \in \overline{A}$. ♣

2.4 Compactness

Recall: For any closed interval $[a, b]$ any sequence $\{x_n\} \subset [a, b]$ contains a convergent subsequence, whose limit is also in $[a, b]$. This comes from the fact that $[a, b]$ is a closed set and the Bolzano Weierstrass theorem.

Definition 33 ***Sequentially Compactness*** *Let $K \subset \mathbb{R}$. We say that K is sequentially compact if every sequence in K has a convergent subsequence whose limit is also in K .*

Example 7

1. *The interval $[a, b]$ is Sequentially Compact.*
2. *Consider $A = [0, 1] \cap \mathbb{Q}$. This is not sequentially compact since if we consider $x_n \to \frac{1}{\sqrt{2}}$, $\{x_n\} \subseteq A$ the limit of any of its subsequence does not lie in A.*
3. *$\mathbb{R}$ or $\mathbb{N}$. Neither of these is sequentially compact. Indeed, consider the sequence $x_n = n$. None of its subsequences converge.*

Lemma 4 *If a set $K \subseteq \mathbb{R}$ is sequentially compact, then it is closed.*

Proof: Suppose x is a limit point of K and it does not lie in K. Then for every $\delta > 0$ $(x-\delta, x+\delta) \cap K \neq \emptyset$, because x is a limit point. In particular for $\delta = \frac{1}{n}$, $x_n \in (x - \frac{1}{n}, x + \frac{1}{n}) \cap K$. So, x_n converges to x. So, any subsequence of that sequence would also converge to x. This in particular implies that the limit x should lie in K because K is sequentially compact. So, all limit points of K lie in K. That means K is closed. ♣

Lemma 5 *If a set $K \subseteq \mathbb{R}$ is sequentially compact, then it is bounded.*

Proof: Suppose K is not bounded, then there is no $M \in \mathbb{N}$ such that $K \subseteq [-M, M]$. So there exists a sequence in K such that for $x_m \in K \setminus [-M, M]$.

Consider $\{x_m\}_{m\geq 1} \subset K$. As K is sequentially compact, it has a convergent subsequence $\{x_{m_k}\}$. Let $\{x_{m_k}\} \to y$ as $k \to \infty$. Now $x_{m_k} \in (y-1, y+1)$ for all $k \geq k_0$. Hence, $x_{m_k} \in (-m_k, m_k)$ by the choice of the x_i', so

$$m_{k+1} \geq m_k,$$

$$(y-1, y+1) \in [-M, M], \text{ for some } M$$

and this yields $x_{m_k} \notin [-M, M]$, which is a contradiction. ♣

Another Characterization of Continuity: Let $U \subseteq \mathbb{R}$ be open in $\mathbb{R}$, $f : U \mapsto \mathbb{R}$ and suppose f is continuous i.e., for $x \in U$, given $\epsilon > 0$, there exists $\delta > 0$ such that $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$. Equivalently, for y$\in (x - \delta, x + \delta)$, we have $f(y) \in (f(x) - \epsilon, f(x) + \epsilon)$. For a set $V \in \mathbb{R}$,

$$f^{-1}(V) := \{u \in U | f(u) \in V\}$$

so that for $V = (f(x) - \epsilon, f(x) + \epsilon)$ we have

$$f^{-1}(f(x) - \epsilon, f(x) + \epsilon) \supset (x - \epsilon, x + \epsilon)$$

In other words if V is open in $\mathbb{R}$, $f^{-1}(V)$ is open in U. The converse is also valid, i.e if $f : U \to \mathbb{R}$ such that $f^{-1}$is open in U for every open set V in $\mathbb{R}$, then f is continuous.

For $v \in V$, there exists $\epsilon > 0$ such that $(v - \epsilon, v + \epsilon) \in V$ if $f^{-1}(v - \epsilon, v + \epsilon) = \emptyset$, then there is nothing to prove. Suppose $u \in U$ such that $f(u) = v$. Since $f^{-1}(v - \epsilon, v + \epsilon)$ is open in U, and contains u, there exists $\delta > 0$ such that $(u - \delta, u + \delta) \in f^{-1}(v - \epsilon, v + \epsilon)$.

For the next lemma, we need a definition.

Definition 34 *A collection of open sets $\mathfrak{U}$ in $\mathbb{R}$ is called an* ***open cover*** *for a set K if*

$$K \subset \bigcup_{U \in \mathfrak{U}} U.$$

Lemma 6 ***Lebesgue Number Lemma****: Given a set K that is sequentially compact and an open covering $\mathfrak{U}$ for K, there exists $\delta > 0$ such that for each $x \in K$, $(x - \delta, x + \delta) \subset U$ for some $U \in \mathfrak{U}$.*

Proof: Suppose not. For every $\delta > 0$, there exists $x \in K$ such that $(x - \delta, x + \delta) \nsubseteq U$ for any $U \in \mathfrak{U}$. In particular take $\delta = \frac{1}{n}$, get $x_n \in K$ such that $(x_n - \delta, x_n + \delta) \nsubseteq U$ for all $U \in \mathfrak{U}$ i.e $(x_n - \frac{1}{n}, x_n + \frac{1}{n}) \nsubseteq U$,for all $U \in \mathfrak{U}$.

Consider $\{x_n\}$ in K, this has a convergent subsequence $x_{n_k} \to x$ in K. Now if $x \in U_0$ for some $U_0 \in \mathfrak{U}$, U_0 is open , then there exists a $\delta_0 > 0$ such that $(x-\delta_0, x+\delta_0) \subset U_0$. Pick a k such that $\frac{1}{n_k} < \frac{\delta_0}{4}$ i.e $n > \frac{4}{\delta_0}$. For such n, clearly $(x_{n_k} - \frac{1}{n_k}, x_{n_k} + \frac{1}{n_k}) \subset (x-\delta_0, x+\delta_0)$. So $(x_{n_k} - \frac{1}{n_k}, x_{n_k} + \frac{1}{n_k}) \subset U_0$ which is a contradiction. ♣

Theorem 25 *If K is sequentially compact, then for every open cover $\mathfrak{U}$ for K, there exists a finite subcover. i.e., if $K \subseteq \mathfrak{U}$ there exists $U_1, U_2, \ldots, U_n$, for some $n \in \mathbb{N}$, such that $K \in \bigcup_{i=1}^{n} U_i$.*

Proof: Suppose not. Pick $x_1 \in K$, there exists $U_1 \in \mathfrak{U}$ such that $x_1 \in U_1$. Since there is no finite subcover, pick $x_2 \in K \smallsetminus U_1$such that $U_2 \in \mathfrak{U}$ and $(x_2-\delta, x_2+\delta) \subseteq U_2$. Proceeding inductively, pick $x_n \in K \smallsetminus \bigcup_{i=1}^{n} U_i$ and $U_n \in \mathfrak{U}$ such that $x_n \in U_n$ and $(x_n-\delta, x_n+\delta) \subseteq U_n$.

For the sequence $\{x_n\}_{n \geq 1}$, let x_{n_k} be a convergent subsequence, and let $x \in K$ be a limit of x_{n_k}. Let $U \in \mathfrak{U}$ such that $(x - \delta, x + \delta) \subset U$. Write $y_k = x_{n_k}$. We have $y_k \to x$ in K i.e given $\epsilon > 0$, there exists N_ϵ such that $|y_k - x| < \epsilon$ for every $k \geq N_\epsilon$. This implies $|y_k - y_{k+1}| < 2\epsilon$ for $k \geq N_\epsilon$. Pick $\epsilon = \frac{\delta}{2}$; then for $k \geq N_\epsilon$, we have

1. $|y_k - y_{k+1}| < 2\epsilon \implies y_{k+1} \in (y_k - \delta, y_k + \delta)$.
2. $(y_k - \delta, y_k + \delta) \subseteq U_k$.
3. $y_{k+1} \notin U_k$.

That is a contradiction. ♣

The Lebesgue Number Lemma has a very important consequence for continuous functions defined on compact sets, in particular, closed intervals.

Theorem 26 *Suppose $f : [0, 1] \to \mathbb{R}$ is continuous, then f is bounded, and attains maximum and minimum.*

Proof: Consider $f^{-1}(-n, n)$ as $n \in \mathbb{N}$. Let $U_n = f^{-1}(-n, n)$. As f is continuous, U_n is open in $[0, 1]$. Also $\bigcup_{i=1} \geq [0, 1]$. $[0, 1]$ is sequentially compact implies there is a finite subcover i.e., there exists $n_1, n_2, \ldots, n_r$ such that $[0, 1] \subseteq \bigcup_{i=1}^{r} f^{-1}(-n_i, n_i)$ implies f maps from $[0, 1]$ to some finite open sets. So, f is bounded.

Suppose $M = \sup(f(x))$. If M is never attained, then the continuous function $g(x); = M - f(x) > 0$ which implies $\frac{1}{M-f(x)} > 0$. Since a continuous function is bounded, it follows that $\frac{1}{M-f(x)} \leq k$ for all x which implies $f(x) \leq M - \frac{1}{k}$.

This is a contradiction that M is the supremum of f. So, M is attained. The proof of attainment of minimality is similar. ♣

Definition 35 *A set K is **Compact** if every open cover of K admits a finite subcover.*

In particular, for subsets of $\mathbb{R}$ sequentially compactness implies compactness.

Theorem 27 ***Heine-Borel Theorem:*** *If $K \subseteq \mathbb{R}$ is compact, then it is also sequentially compact. In particular, $K \subseteq \mathbb{R}$ is compact if and only if K is closed and bounded.*

Proof: Suppose $\{x_n\} \subseteq K$. There exists convergent subsequent i.e we want $\{x_{n_k}\}$ such that $x_{n_k} \to x$ for some $x \in K$.

We will instead prove the following: If $K \subseteq \mathbb{R}$ is compact then K is closed and bounded. By a previous theorem, this will prove that K is sequentially compact.

1. *K is compact $\implies$ K is bounded.*

   For each $x \in K$,$(x-1, x+1) = U_x$, $\mathfrak{U} = \bigcup U_x$. Clearly $\mathfrak{U}$ covers K. So, there exists a finite subcover. WLOG $x_1 < x_2 < \ldots < x_n$; $U_{x_1}, U_{x_2}, \ldots, U_{x_n}$ cover K. Hence

   $$[x_1 - 2, x_n + 2] \supseteq K$$

   Let $M = max\{|x_2 - 2|, |x_n + 2|\}$, so that $K \subseteq [-M, M]$, hence K is bounded.

2. *K is compact $\implies$ K is closed.*

   Suppose $\{x_n\} \subseteq K$ and $x_n \to x$ in $\mathbb{R}$ but $x \notin K$. WLOG we assume $x_n < x$. Since infinitely many x_n are less than x, that will do.

   Consider $U_n = (-\infty, x_n + \frac{|x-x_n|}{2})$ and $U = (x, \infty)$,$\bigcup_{n \geq 1} U_n = (-\infty, x)$. Let

   $$\mathfrak{U} = \{U\} \cup \{U_n | n \geq 1\}.$$

   $\mathfrak{U}$ is an open cover for K, so there is a finite subcover, say $U_{n_1}, U_{n_2}, \ldots, U_{n_k}$ with $n_1 < n_2 < \ldots < n_k$, and $(x, \infty) \cup \bigcup_{i=1}^{n} U_i$ covers K.

   But since $x_n \to x$ there exists k such that $|x - x_k| < \min_{i=1}^{n} \frac{|x-x_i|}{2}$ and that is a contradiction. So, $x \in K$, which implies K is closed.

♣

2.5 Induced Topology on Subsets of $\mathbb{R}$

We have thus far defined continuity of functions whose domain is open in $\mathbb{R}$ ($\epsilon - \delta$ definition). We also saw that this definition is equivalent to the following:

$f : (a, b) \to \mathbb{R}$ is continuous if and only if for any $U \subseteq \mathbb{R}$ open, we have that $f^{-1}(U)$ is open in (a, b) . We use this as a more general definition of continuity for functions whose domain is an arbitrary subset X of $\mathbb{R}$.

Definition 36 *Given $X \in \mathbb{R}$, we define the* ***Induced topology of open subsets in*** *X as follows. We say that $U \subseteq X$ is open in X if $U = X \cap V$, for some V open in $\mathbb{R}$. It is easy to verify the following.*

1. *$\emptyset, X$ are open in X.*
2. *$\{U_\alpha\} \subset X$ are open , so is $\bigcup_\alpha U_\alpha$*
3. *$U_1, U_2, \ldots, U_n$ are open in X, then $\bigcap_{i=1}^n U_i$ is also open in X.*

Definition 37 *For $f : X \to \mathbb{R}$, we say that X is continuous, if for any open set $U \subseteq \mathbb{R}$, $f^{-1}(U)$is open in X.*

Example 8
1. *$X = \mathbb{N}$. The topology induced by $\mathbb{R}$ gives $\{1\} = \mathbb{N} \cap (1 - \frac{1}{2}, 1 + \frac{1}{2})$ and $(1 - \frac{1}{2}, 1 + \frac{1}{2})$ is open in $\mathbb{R}$. So every $\{n\}, n \in \mathbb{N}$ is open in $\mathbb{N}$. So every function whose domain is $\mathbb{N}$ is continuous!*
2. *$X = [a, b]$. Here, intervals of the form $[0, \epsilon), (1 - \epsilon, 1]$ are also open in $[0, 1]$ other than the ones open in $\mathbb{R}$ and contained in $[0, 1]$.*

Remark 15 *We can define sequential compactness and compactness for arbitrary subsets of $\mathbb{R}$ in a similar fashion as we did for $\mathbb{R}$.*

Another feature of Compact Sets: We have already seen that continuous functions defined on a compact set attain maxima and minima. Here, we shall see that something else holds for continuous functions defined on a compact set.

Definition 38 *$f : [0, 1] \to \mathbb{R}$ is continuous, we say that f is* ***uniformly continuous*** *if for a given $\epsilon > 0$, there exists $\delta > 0$ such that $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon$.*

Example 9
1. *$f : [0, 1] \to \mathbb{R}$, $f(x) = x$,$f(x)$ is uniformly continuous. Indeed, take $\delta = \epsilon$, for a given $\epsilon > 0$.*
2. *$f : [0, 1] \to \mathbb{R}$, $f(x)$ is a constant function, then $f(x)$ is uniformly continuous. Take $\delta = k$ $k \in \mathbb{R}$,for a given $\epsilon > 0$.*
3. *$f : (0, 1) \to \mathbb{R}$, $f(x) = \frac{1}{x}$ is not uniformly continuous.*

Proof: The last statement needs a little clarification. Suppose for given , $\epsilon = 1$, there exists a $\delta > 0$ such that if $|x - y| < \delta$ implies $|f(x) - f(y)| < \epsilon$. Take $x = k = \frac{-\delta+\sqrt{\delta^2+8\delta}}{4}$ and $y = \frac{\delta}{2} + k$. Then

$$|f(x) - f(y)| = \frac{1}{k} - \frac{1}{k + \frac{\delta}{2}}$$
$$\implies |f(x) - f(y)| = \frac{\frac{\delta}{2}}{k^2 + k.\frac{\delta}{2}}$$
$$\implies |f(x) - f(y)| = 1$$

Since k is root of $k^2 +. \frac{k\delta}{2} = \frac{\delta}{2}$, $|f(x) - f(y)|$ is not less than ϵ. Hence $f(x)$ is not uniformly continuous. ♣

Theorem 28 *Suppose $f : [0, 1] \to \mathbb{R}$ is continuous, then f is uniformly continuous.*

Proof: It follows from continuity of f that for a given $\epsilon > 0$ and $x \in [0, 1]$ there exists $\delta_x > 0$ such that for all $y \in (x + \delta_x, x - \delta_x) \cap [0, 1]$, we have $|f(x) - f(y)| < \epsilon$.

Consider $\mathfrak{U} = \{(x - \delta_x, x + \delta_x)|x \in [0, 1]\}$; clearly, this is an open cover for $[0, 1]$ and since $[0, 1]$ is sequentially compact the Lebesgue Number lemma tells us that there is $\delta > 0$ such that for any $x \in [0, 1]$, $(x - \delta, x + \delta) \subset U_y$ for some $y \in [0, 1]$. For this δ suppose we have $|x - y| < \delta$ i.e $y \in (x - \delta, x + \delta) \subset U_z$ for some z. Then by continuity of f at z,

$$|f(x) - f(z)| < \epsilon$$

Also, $|f(z) - f(y)| < \epsilon$. Therefore

$$|f(x) - f(y)| = |f(x) - f(z) + f(z) - f(y)|$$
$$|f(x) - f(y)| \leq |f(x) - f(z)| + |f(z) - f(y)|$$
$$|f(x) - f(y)| \leq 2\epsilon.$$

♣

In any arbitrary metric space, consider

$$B_r(x) = \{y \in \mathfrak{X}|d(x, y) < r\}$$

Let K be compact in $\mathfrak{X}$. Suppose x is a limit point for K . If $x \notin K$, for each $y \in K$ pick $B_r(y)$ such that $x \notin B_r(y)$ (for some suitable r). Then $\{B_{r_y}(y)|y \in K\}$ is an open cover for K . Since K is sequentially compact, there is a finite subcover for K. Therefore, say

$$B_{r_1}(y_1) \cup B_{r_2}(y_2) \cup \ldots \cup B_{r_n}(y_n) \supseteq K$$

Let $min\{d(y_1,x),d(y_2,x),\ldots,d(y_n,x)\} = d_0$. Since x is limit point, $B_{\frac{d_0}{2}}(x)\cap K \neq \emptyset$. But d_0 is minimum.This is a contradiction.

Remark 16 *$f:[0,1]\to\mathbb{R}$ is continuous $\implies$ f is uniformly continuous.*
The same proof works for any K compact and $f:K\to\mathbb{R}$.

Proposition 28 *U is open in $\mathbb{R}$ then $U = \bigcup_{i=1}(a_i,b_i)$, where $(a_i,b_i)\cap(a_j,b_j) = \emptyset$ for $i\neq j$.*

Proof: Suppose $x\in U$, By definition there is some $(x-\delta,x+\delta)\subset U$ for some $\delta>0$. Let $a_x = \inf\{a|(a,x)\subset U\}$,$b_x = sup\{b|(x,b)\subset U\}$.

If $R_x = \{b|(x,b)\subset U\}$ is not bounded above, then $(x,\infty)\subset U$. Similarly, remark for $\{a|(a,x)\subset U\}$ is not bounded below, $(-\infty,x)\subset U$.

$x\in(a_x,b_x)$, then b_x cannot be in U because if it is in U, then $(b_x-\delta,b_x+\delta)$ will also be in U. So b_x will be not be supremum. So b_x does not lie in U. Similarly a_x also does not lie in U.

We claim that the set U is at most countably infinite. We can say this because there can be either finite sets which would be countable or infinite sets. In that case take all the infimums of those sets say $C = \{c_0,c_1,c_2,\ldots\}$. Then define a bijection from that set to $\mathbb{N}$ where each element of the set C represents one open set (suppose infimum of a set does not exist, the only case is that there is only one set of that sort in U i.e of the form $(-\infty,k), k\in\mathbb{R}$.So that does not effect the type of infinity). So, the number of such sets is at most countably infinite. ♣

2.6 Extending from the Reals to arbitrary Metric spaces

Many of the results we have considered in the case of Real numbers also extend with the 'same' proof (with some cosmetic changes, of course) an to arbitrary metric space. But some properties do not necessarily carry through. We tabulate them in this table for convenience.

Property	$\mathbb{R}$	$(\mathfrak{X},d)$
1. Limits of Sequences	✓	✓
2. Continuity of function	✓	✓
3. Sequential Compactness	✓	✓
4. Compactness	✓	✓
5. Sequential Compactness ⇔ Closed and Bounded	✓	×
6. Sequential Compactness ⇔ Compact	✓	✓

Example 10 *For Property 5, consider $\mathfrak{X} = \mathbb{Q}\bigcap[0,1]$ and $d(x,y) = |x-y|$. This is* ***closed and bounded*** *but there are Cauchy sequences with no limit point in $\mathfrak{X}$; for instance the rational sequences which converge to an irrational number in $\mathbb{R}$.*

Theorem 29 *For a metric space $(\mathfrak{X}, d)$ if $(\mathfrak{X}, d)$ is sequentially compact, then $(\mathfrak{X}, d)$ is compact.*

Proof: We have proved the Lebesgue number lemma for open covers which states that if K is sequentially compact and $\mathbf{U}$ is an open cover for K, then there exists $\delta > 0$ such that for all $x \in K, B_\delta(x) \subset u$ for some $u \in \mathbf{U}$. In fact, we have just replaced $(x-\delta, x+\delta)$ with $B_\delta(x)$. Both the proofs carry through with minimal changes. ♣

The question of the converse of this statement is quite a natural one. Unfortunately, our proof from the real case does not quite apply since the Heine-Borel theorem does not hold for an arbitrary metric space. We will see the validity of the converse in a slightly more general manner, in the next section. We do however, make one remark: If $\{x_n\}$ has a convergent subsequence, then it's limit must lie in $\mathfrak{X}$ because compact sets are necessarily closed in an arbitrary metric space.

Alternate way of looking at compactness: The following proposition is a consequence of interpreting the definition of compactness by taking the complements of the sets of any open cover. We omit the proof.

Proposition 29 *K is compact if and only if for any finite collection of closed sets $\mathcal{C}$ satisfying $C_1, C_2, \ldots C_n \in \mathcal{C}$ with $C_1 \cap C_2 \cdots \cap C_n \neq \emptyset$, we must have $\bigcap_{C\in\mathcal{C}} \neq \emptyset$.*

2.7 Connectedness

We have already seen that any open set in $\mathbb{R}$ is at most a countable union of pairwise disjoint open intervals. In particular, if $\emptyset \subsetneq \mathbf{U} \subsetneq \mathbb{R}$ is open in $\mathbb{R}$, then $\mathbf{U}$ is not closed. Equivalently, the only sets of $\mathbb{R}$ that are both open and closed are $\emptyset$ and $\mathbb{R}$. What about in general?

Example 11 *$\mathfrak{X} = \mathbb{Q}$, open sets are of type $\mathbb{Q} \cap \boldsymbol{U}$, where $\boldsymbol{U} \subseteq \mathbb{R}$ is open. The set $(-\infty, \sqrt{2}] \cap \mathbb{Q} = (-\infty, \sqrt{2}) \cap \mathbb{Q}$ is both closed and open in $\mathbb{Q}$. As witnessed over $\mathbb{Q}$, in the general case, this is not true.*

Definition 39 *We say that a metric space $(\mathfrak{X}, d)$ is* ***disconnected*** *if there exists U, V (open)$\subseteq \mathfrak{X}$, that satisfy*

1. *$\{U \cap V\} = \emptyset$.*

2. $\{U \cup V\} = \mathfrak{X}$.

3. $U, V \neq \emptyset$.

*And we say $\mathfrak{X}$ is **connected** if it is **not disconnected**. (If U, V witness a disconnection, we say that U, V separate $\mathfrak{X}$).*

Theorem 30 *$\mathbb{R}$ is connected and so is any interval of the type $(a, b], (a, b), [a, b], [a, b)$.*

Proof: We have already seen that $\mathbb{R}$ is connected. For the remaining ones, we shall only prove the case of $(a, b]$. The other cases follow similarly.

Consider $(a, b]$.Suppose this is not connected. U, V are open in $\mathbb{R}$ such that $U \cap V = \emptyset$, $U \cup V \supseteq (a, b]$ and $U, V \neq \emptyset$. Pick $a < c < d \leq b$ such that $c \in U, d \in V$.

Now consider $[c, d] \cap U, [c, d] \cap V$. This must give a separation of $[c, d]$. Pick $[c, d] \cap U$. This is bounded above and non-empty. Let $x \in \mathbb{R} = \sup([c, d] \cap U)$. ♣

Theorem 31 ***Intermediate Value Theorem***

Suppose $f : [a, b] \implies \mathbb{R}$ is continuous and $f(a) > 0, f(b) < 0$. Then there exists $c \in (a, b)$ such that $f(c) = 0$.

Suppose *not*. Then for all $x \in [a, b], f(x) > 0$ or $f(x) < 0$. Let $U = \{x \in [a, b] | f(x) > 0\}$ and $V = \{x \in [a, b] | f(x) < 0\}$.

$U, V \neq \emptyset$ since $a \in U, b \in V$.
$U \cup V = [a, b]$ *and* $U \cap V = \emptyset$.
$U = f^{-1}(0, \infty)$ *and* $V = f^{-1}(-\infty, 0) \implies U, V$ *are open in* $[a, b]$.

So, U, V separate $[a, b]$ which is a contradiction.

Note: Open in $\mathbb{R} \approx \bigcup_{i=1}^{\infty}(a_n, b_n)$.

This begs another question. Is it true that a closed set in $\mathbb{R}$ is a countable union of closed intervals? The answer is a resounding NO.

2.7.1 A Weird Closed Set in $\mathbb{R}$:

Consider

$$\begin{aligned} C_0 &= [0,1] \\ C_1 &= [0,\frac{1}{3}] \ \cap \ [\frac{2}{3},1] \\ C_2 &= [0.\frac{1}{9}] \ \cap \ [\frac{2}{9},\frac{3}{9}] \ \cap \ [\frac{6}{9},\frac{7}{9}] \ \cap \ [\frac{8}{9},1] \\ &\vdots \\ \mathcal{C} &= \bigcap_{i=0}^{\infty} C_i \end{aligned}$$

$\mathcal{C}$ is called the **Cantor Set**. We make a few observations.

1. $\mathcal{C}$ is **closed** since $\overline{\mathcal{C}}$ is a union of open intervals and hence it is open.
2. $0,1 \in \mathcal{C}$, since $0,1 \in C_i$,for all $i \in \mathbb{N}$.
3. No $(a,b) \subset \mathcal{C}$ where $a,b \in [0,1]$. Indeed, C_n is a finite union of closed intervals and the sum of lengths of these intervals is $(\frac{2}{3})^n$.
   If $(a,b) \subset \mathcal{C}$, then $(a,b) \subset C_n$ for all $n \in \mathbb{N}$. But $\lim_{n\to\infty} (\frac{2}{3})^n = 0$ which implies if $(\frac{2}{3})^n < \frac{b}{a}$, then we have a contradiction.
4. $\mathcal{C}$ is **uncountable!** More precisely $|\mathcal{C}| \sim |\mathbb{R}|$.
   Consider a ternary expansion for elements $x \in [0,1]$, where we pick only those expansions which don't terminate in 2.
   Note that $x \in C_1 \Leftrightarrow$ the first place after the decimal is either 0 or 2.
   $\vdots$
   Similarly $x \in C_n \Leftrightarrow$ the first n places after the decimal are all either 0 or 2.
   Replacing 2's in these ternary strings with 1's gives us the set of all binary strings whose cardinality is same as $|\mathbb{R}|$.

We return to a question we posed earlier, namely, the equivalence of the notions of Compactness and Sequential Compactness on Metric spaces. We have already seen that Sequential Compactness gives compactness in metric spaces, and we now wish to show that Compactness implies Sequential Compactness.

Over $\mathbb{R}$, Compactness of a set is equivalent to the set being Closed and Bounded which is in turn equivalent to Sequential Compactness. What is boundedness in the general metric space context? This is straightforward; we say that a metric space $(\mathfrak{X}, d)$ is **bounded** if there exists $x \in \mathfrak{X}$ and $r \in \mathbb{R}$ such that $B_r(x) \supseteq (\mathfrak{X}, d)$.

Unfortunately, in a general metric space, 'Closed and Bounded' does not guarantee Sequential Compactness. For instance, there is no guarantee that Cauchy sequences even converge.

Definition 40 *We say that a metric space $(\mathfrak{X}, d)$ is* ***Complete*** *if every Cauchy sequence actually converges to some member of $\mathfrak{X}$.*

2.7.2 Pathological Examples:

$\mathfrak{X} = \mathbb{Q} \cap [0,1], d(x,y) = |x-y|$. Here $\mathfrak{X}$ is not complete.

If $\mathfrak{X}$ is complete and bounded, is it also sequentially compact?

NO!

Suppose $\mathfrak{X} = \mathbb{N}$ and the discrete metric $d(x,y)$ where

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

This is complete as every Cauchy sequence on $\mathbb{N}$ will eventually have the same number coming up infinitely often.
This is bounded as $d(x,y)$ can be just 0 or 1. In particular, any ball of radius 1 contains every element in $\mathfrak{X}$. However the sequence $a_n := n$ has no convergent subsequence. So, it is not sequentially compact. And it is not compact either; $\{1\}, \{2\}, \{3\}, \ldots$ form an open cover with no finite subcover.

This begs a question:

Can we characterize compact sets in $(\mathfrak{X}, d)$?

Proposition 30 *If $\mathfrak{X}$ is compact then it is complete.*

Proof: Suppose $\{x_n\}$ is Cauchy in $\mathfrak{X}$. It suffices to find a limit point for this set $A = \{x_n | n \geq 1\}$. Suppose A has no limit point. Then note that A is closed. Consider

$$\begin{aligned} A_1 &= \{x_1, x_2, \ldots\} \\ A_2 &= \{x_2, x_3, \ldots\} \\ &\vdots \\ A_n &= \{x_n, x_{n+1}, \ldots\} \\ &\vdots \end{aligned}$$

For the same reason that A is closed, it follows that each A_i is also closed. Furthermore, $A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_n} \neq \emptyset$ for all $n \in \mathbb{N}$ and $i_1, i_2, \ldots, i_n$. But $\bigcap_{n \geq 1} A_n = \emptyset$.

But the above observation contradicts the Finite Intersection Property for compact sets. This completes the proof. ♣

Suppose $(\mathfrak{X}, d)$ is compact. The above theorem implies that it is complete.

For each $x \in \mathfrak{X}$, consider $B_1(x)$, and consider the open cover

$$\mathfrak{U} = \{B_1(x) | x \in \mathfrak{X}\}.$$

Since $\mathfrak{U}$ is an open cover $B_1(x_1) \cup B_1(x_2) \ldots \cup B_1(x_n) \supseteq \mathfrak{X}$ for some $x_1, x_2 \ldots, x_n$. In particular, for any sequence $\{y_n\}$ in $(\mathfrak{X}, d)$ it follows that there exists some subsequence $\{y_{n_i}\} \subseteq B_1(x_1)$.

Lemma 7 *Suppose $\mathfrak{X}$ is compact and $K \subseteq \mathfrak{X}$ is closed then K is also compact.*

Remark 17 *The set $B_r^{cl}(x) := \{y \in \mathfrak{X}, d(x, y) \leq r\}$ is closed. So, by this lemma $B_1^{cl}(x)$ is compact in $\mathfrak{X}$.*

Proof: (Of the lemma)

Suppose $\mathfrak{U}$ is an open cover for K. Then

$$\mathfrak{U}' = \mathfrak{U} \cup \{\mathfrak{X} \setminus K\},$$

is an open cover for $\mathfrak{X}$. This implies it has a finite subcover for $\mathfrak{X}$, which gives a finite subcover of $\mathfrak{U}$ for K. ♣

So far, we have $\{y_n^{(1)}\} \subseteq B_1 \subseteq B_1^{cl}$ which is compact. So, $\{B_{\frac{1}{2}}(x) | x \in B_1(x)\}$ is an open cover for $B_1(x)^{cl}$. Suppose (WLOG) that there exists an infinite $\{y_n^{(k)}\} \subseteq \{y_n^{(1)}\} \subset B_{\frac{1}{2}}(x_2) = B_2$.

Get $\{y_n^{(k)}\}$(infinite) such that $y_n^{(k)} \subseteq B_{(\frac{1}{2})_{k-1}}(x_k)$ for all $n \in \mathbb{N}$

Given $\epsilon > 0$

$$\begin{aligned} &\implies d(y_n^{(k)}, x_k) \leq \tfrac{1}{2^{k-1}} \\ &\implies d(y_n^{(k)}, y_m^{(k)}) \leq d(y_n^{(k)}, x_k) + d(x_k, y_m^{(k)}) \\ &\qquad\qquad\qquad\quad \leq \tfrac{1}{2^{k-2}} \end{aligned}$$

So,given $\epsilon > 0$, pick k such that $\frac{1}{2^{k-2}} < \epsilon$. So, the sequence $\{y_1^{(1)}, y_2^{(2)}, \ldots\}$ is Cauchy in $\mathfrak{X}$.

Remark 18 *Suppose $\mathfrak{X}$ has the property that for any $\epsilon \geq 0$, there is a finite open cover of $\mathfrak{X}$ by balls of radius ϵ, then the same proof essentially works.*

Definition 41 *We say $\mathfrak{X}$ is **totally bounded** if for any $\epsilon > 0, \mathfrak{X}$ is a union of finitely many balls of radius ϵ.*

Theorem 32 ***Generalized Heine-Borel Theorem:*** *A metric space $(\mathfrak{X}, d)$ is compact if and only if it is complete and totally bounded.*

2.8 Another Construction of $\mathbb{R}$ from $\mathbb{Q}$

Informally: Every Cauchy sequence of rationals represents a real number.

Definition 42 *Say $X = (x_1, x_2, x_3, \cdots)$ and $Y = (y_1, y_2, y_3, \cdots)$, where $x_i, y_i \in \mathbb{Q}$. We say $X \sim Y$ if for any rational $\epsilon > 0$, there exists N_ϵ such that $|x_n - y_n| < \epsilon$ for all $n > N_\epsilon$.*

Proposition 31 *$\sim$ is an equivalence relation.*

Proof: Reflexive: Trivial, because $|x_n - x_n|$ is zero.
Symmetric: Trivial, because $|x_n - y_n| = |y_n - x_n|$ for all x_n, y_n.
Transitive: X, Y, Z are such that if we are given an $\epsilon > 0$
there exists N_{ϵ_1} such that $|x_n - y_n| < \frac{\epsilon}{2}$ for all $n > N_{\epsilon_1}$, and
there exists N_{ϵ_2} such that $|y_n - z_n| < \frac{\epsilon}{2}$ for all $n > N_{\epsilon_2}$.
Let N_ϵ be the greater of N_{ϵ_1} and N_{ϵ_2}.
Using the triangle inequality, we have:
$|x_n - z_n| \leq |x_n - y_n| + |y_n - z_n| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$, for all $n > N_\epsilon$ .Hence, it is transitive. ♣

Let $\widehat{Q} = \{\ \widehat{X} \mid \widehat{X}$ is an equivalence class of the Cauchy sequence X, related by $\sim \}$
Let $\widehat{X} = \{x_n\}$ and $\widehat{Y} = \{y_n\}$. We can define the operations $+$ and $\cdot$ on $\widehat{Q}$:

- $\widehat{X} + \widehat{Y} = \widehat{X + Y} = \{x_n + y_n\}$
- $\widehat{X} \cdot \widehat{Y} = \widehat{X \cdot Y} = \{x_n \cdot y_n\}$

Similarly define:

- $-(\widehat{X}) = -\widehat{X} = \{-x_n\}$
- $\widehat{X}\ /\ \widehat{Y} = \widehat{X/Y} = \{\frac{x_n}{y_n}\}$, where $y_n \neq 0$ for all $n \geq 1$.

Since $+$ and . are well defined (proofs skipped), we have the following:

Theorem 33 *$(\widehat{Q}, +, \cdot, 0, 1)$ is a field, where $0 \equiv \{0,\ 0,\ 0, \cdots\}$ and $1 \equiv \{1,\ 1,\ 1, \cdots\}$.*

Definition 43 *We say $\widehat{A} > \widehat{B}$ if $a_n = b_n + \epsilon$ for some rational $\epsilon > 0$, for all $n > N_\epsilon$, where $a_n \in \widehat{A}$ and $b_n \in \widehat{B}$.*

Making use of this definition, we have:

Theorem 34 *$(\widehat{Q}, +, ., 0, 1, <, >)$ is an ordered field.*

Proof: $A \subseteq \widehat{Q}$ is bounded above, i.e., for all $\widehat{p} \subseteq A$, $\widehat{p} < \widehat{q}$ for some $\widehat{q} \in \widehat{Q}$.
WLOG, $q_n \geq p_n + \epsilon$ for some rational $\epsilon > 0$, for all $n > 0$. We want to produce an $\widehat{X}$ as an lub for A, i.e. a sequence $\{x_n\}$ where $x_n \in \mathbb{Q}$.
Pick $x_1 \in \mathbb{N}$ such that x_1 is the least element which is an upper bound for A.
Pick x_2 as the least element of the form $\frac{n}{2}$, such that it is an upper bound for A, n $\in \mathbb{N}$.
Pick x_2 as the least element of the form $\frac{n}{2^2}$, such that it is an upper bound for A, n $\in \mathbb{N}$.

$\vdots$

Continue this process for all n

Claim: $\{x_n\}$ is Cauchy.

$x_1 \geq x_2 \geq x_3 \geq \cdots$

Also, $x_n - x_{n-1} \leq \frac{1}{2^n}$
Suppose m $<$ n:

$$\begin{aligned}
|x_m - x_n| =&|(x_m - x_{m+1}) + (x_{m+1} - x_{m+2}) + \cdots + (x_{n-1} - x_n)| \\
\leq&|x_m - x_{m+1}| + |x_{m+1} - x_{m+2}| + \cdots + |x_{n-1} - x_n| \\
\leq&\frac{1}{2^m} + \frac{1}{2^{m+1}} + \cdots + \frac{1}{2^{n-1}} \\
\leq&\frac{1}{2^{m-1}}
\end{aligned}$$

So, $X = x_n$ is Cauchy.

Claim: $\widehat{X}$ is an lub for A.

We prove it in two steps:

(i)$\widehat{X}$ is an upper bound for A.

(ii)$\widehat{X}$ is the least upper bound.

Suppose $\widehat{X}$ is NOT an upper bound for A, i.e. $\widehat{p} > \widehat{X}$ for some $\widehat{p} \in A$.

So, there exists $\epsilon > 0$ such that $p_n \geq x_n + \epsilon$, for all $n \geq n_0$. ———— $(*)$
For $m \geq n$, this means $p_m \geq x_m + \epsilon$.
Pick $n \geq n_0$ such that $\frac{1}{2^{n-1}} < \epsilon$
But x_n was chosen such that $x_n \geq p_m$ for all large m.

Then,

$x_n \geq p_m \geq x_m + \epsilon$

or, $\epsilon < x_m - x_n < \frac{1}{2^{n-1}} < \epsilon$

which is a contradiction.

So, $\widehat{X}$ is an upper bound for A.
2. Suppose $\widehat{q}$ is a SMALLER upper bound for A than $\widehat{X}$,
i.e. $\widehat{q} \geq \widehat{p}$ for all $\widehat{p} \in A$, and $\widehat{X} \geq \widehat{q} + \epsilon$ for some $\epsilon > 0$.
Pick n such that $\frac{1}{2^{n-1}} < \epsilon$.
We claim that at the n^{th} step in the construction of X, we could then, in light of $(*)$, have done strictly better and that would contradict the choice of x_n.
For each $\widehat{p} \in A$,

$$q_n \geq p_n forall n \geq n(p)$$

$$x_n \geq x_n - \epsilon \geq q_n \geq p_n$$

But x_n was chosen to be the least rational of the form $\frac{a}{2^{n-1}}$ that is an upper bound for A. Since $\epsilon > \frac{1}{2^{n-1}}$, the element $\frac{a-1}{2^{n-1}}$ must also satisfy:

$$\frac{a-1}{2^{n-1}} \geq q_n \geq p_n$$

This contradicts the minimality of x_n.
So, X is the least upper bound for A.
Hence, if $A \subseteq \widehat{Q}$ is bounded above, then it has an lub in $\widehat{Q}$. ♣

We have seen that $\mathbb{R}$ can be constructed from $\mathbb{Q}$ by taking all Cauchy sequences.
We have also seen before (from the Dedekind cuts definition of $\mathbb{R}$), that Cauchy sequences in $\mathbb{R}$ are convergent.
Question: Can we do the same for an arbitrary metric space $(\mathfrak{X}, d)$?
Answer: Yes. Suppose $(x_1, x_2, \cdots) = X$
$(y_1, y_2, \cdots) = Y$ are Cauchy sequences in $(\mathfrak{X}, d)$.
Define: $X \sim Y$ iff
$lim_{n \Longrightarrow \infty} d(x_n, y_n) = 0$. We can check that $\sim$ is an equivalence relation.
Let $\widehat{\mathfrak{X}} = \{ \ \widehat{X} \mid$ Equivalence classes of Cauchy sequences in $\mathfrak{X}$ $\}$
Define: $\widehat{d}\ (\widehat{X}, \widehat{Y}) := lim_{n \Longrightarrow \infty} d(x_n, y_n)$.
where,
$\widehat{X} = \{x_1, x_2, \cdots\}$,
$\widehat{Y} = \{y_1, y_2, \cdots\}$. We can check that:

1. $\widehat{d}$ is well defined.

2. $\widehat{d}$ gives a metric space, i.e.

- $\widehat{d}\ (\widehat{X}, \widehat{X}) = 0$
- $\widehat{d}\ (\widehat{X}, \widehat{Y}) = \widehat{d}\ (\widehat{Y}, \widehat{X})$
- $\widehat{d}\ (\widehat{X}, \widehat{Y}) \leq \widehat{d}\ (\widehat{X}, \widehat{Z}) + \widehat{d}\ (\widehat{Z}, \widehat{Y})$

Lemma 8 *If $\{x_n\}$, $\{y_n\}$ are Cauchy in $\mathfrak{X}$, then $lim_{n \Longrightarrow \infty}\ d(x_n, y_n) = \alpha_n$ exists, $\{\alpha_n\} \subseteq \mathbb{R}$.*

Proof: It is enough to check that $d(x_n, y_n)$ is Cauchy in $\mathbb{R}$.
WTS: Given $\epsilon > 0$ there exists N_ϵ such that if m, $n \geq N_\epsilon$,
$|\alpha_n - \alpha_m| < \epsilon$.
The triangle inequality gives:

$$d(x_m, y_m) \leq d(x_m, x_n) + d(x_n, y_n) + d(y_n, y_m)$$

$\{x_n\}$ and $\{y_n\}$ are Cauchy. Pick N_ϵ large so that $d(x_m, x_n) < \frac{\epsilon}{2}$ and $d(y_m, y_n) < \frac{\epsilon}{2}$. So, $d(x_m, y_m) - d(x_n, y_n) < \epsilon$, and
$d(x_n, y_n) - d(x_m, y_m) < \epsilon \implies \{\alpha_n\}$ is Cauchy in $\mathbb{R}$. ♣

If $X \sim X'$ and $Y \sim Y'$, then
$\lim_{n \implies \infty} d(x_n, y_n) = \lim_{n \implies \infty} d(x'_n, y'_n)$
And since $X \sim X'$, we have
$\lim_{n \implies \infty} d(x_n, x'_n) = \lim_{n \implies \infty} d(y_n, y'_n) = 0$
i.e., given $\epsilon > 0$, there exists N_ϵ such that if $n > N_\epsilon, d(x_n, x'_n) < \frac{\epsilon}{2}$ and $d(y_n, y'_n) < \frac{\epsilon}{2}$. As before,

$$\begin{aligned} d(x_n, y_n) &\leq d(x_n, x'_n) + d(x'_n, y'_n) + d(y'_n, y_n) \\ &\leq d(x'_n, y'_n) + \epsilon \end{aligned}$$

Taking limits:
$lim_{n \implies \infty}\ d(x_n, y_n) \leq lim_{n \implies \infty}\ d(x'_n, y'_n) + \epsilon$ for any $\epsilon > 0$
$\implies lim_{n \implies \infty}\ d(x_n, y_n) \leq lim_{n \implies \infty}\ d(x'_n, y'_n)$.

Supposing (x_n, x'_n) and (y_n, y'_n) in the above, we get the reverse inequality.
$\implies \widehat{d}$ is well defined.
Clearly,

$$\begin{aligned} \mathfrak{X} &\hookrightarrow \widehat{\mathfrak{X}} \ by \\ x &\mapsto (\ (\widehat{x, x, \cdots}), (\widehat{y, y, \cdots})\ ) \\ d(x, y) &\mapsto \widehat{d}\ (\ (\widehat{x, x, \cdots}), (\widehat{y, y, \cdots})\ ) \\ lim_{n \implies \infty} d(x_n, y_n) &= lim_{n \implies \infty} d(x, y) \end{aligned}$$

Theorem 35 $(\widehat{\mathfrak{X}}, \widehat{d})$ *contains an* ***isometric copy*** *of* $(\mathfrak{X}, d)$ *and is* ***complete***.

ISOMETRIC: For any x, $y \in \mathfrak{X}$, $\widehat{d}\ (\widehat{x}, \widehat{y}) =$

$$\begin{aligned} \textbf{where } \widehat{x} &= (x, x, \cdots) \\ \widehat{y} &= (y, y, \cdots) \end{aligned}$$

UPSHOT: Every metric space can be completed. **Proof:** We need to show that every Cauchy sequence in $\widehat{\mathfrak{X}}$ has a limit in $\widehat{\mathfrak{X}}$,
i.e. given $\epsilon > 0$, there exists N_ϵ such that $\widehat{d}\ (\widehat{X_m}, \widehat{X_n}) < \epsilon$ for all $m, n > N_\epsilon$.
$\widehat{X_1} = x_1^{(1)}, x_2^{(1)}, \cdots, x_{m'}^{(1)}, \cdots, x_{n'}^{(1)}, \cdots$
$\widehat{X_2} = x_1^{(2)}, x_2^{(2)}, \cdots, x_{m'}^{(2)}, \cdots, x_{n'}^{(2)}, \cdots$
$\vdots$
$\widehat{X_m} = x_1^{(m)}, x_2^{(m)}, \cdots, x_{m'}^{(m)}, \cdots, x_{n'}^{(m)}, \cdots$
$\vdots$

where $\widehat{X_1}$, $\widehat{X_2}$, $\cdots$, $\widehat{X_m}$, $\cdots$ are all Cauchy sequences,
and $(\widehat{X_1}, \widehat{X_2}, \cdots, \widehat{X_m}, \cdots)$ is a Cauchy sequence.

Take $\widehat{X} = (x_1^{(1)}, x_2^{(2)}, \cdots)$
Is $\widehat{X}$ Cauchy?
Pick N_ϵ such that:

$$\begin{aligned} d(x_m^{(n)}, x_n^{(m)}) &< \frac{\epsilon}{2} \\ d(x_n^{(m)}, x_m^{(n)}) &< \frac{\epsilon}{2} \\ \implies d(x_m^{(m)}, x_m^{(n)}) &< \epsilon \qquad\qquad \textbf{(triangle inequality)} \end{aligned}$$

We need to show that for $\epsilon > 0$, $d(X_n, X) < \epsilon$ for all $n > N_\epsilon$, for some N_ϵ, i.e., $\lim_{k \implies \infty} d(x_k^{(n)}, x_k^{(k)}) < \epsilon$, or
$d(x_k^{(n)}, x_k^{(k)}) < \epsilon$ for all $k \geq k_\epsilon$, $n \geq N_\epsilon$. ♣

2.9 Returning to the Cantor Set

$$\begin{aligned} C_0 &= [0,1] \\ C_1 &= [0, \frac{1}{3}] \ \cup \ [\frac{2}{3}, 1] \\ C_2 &= [0, \frac{1}{9}] \ \cup \ [\frac{2}{9}, \frac{1}{3}] \ \cup \ [\frac{2}{3}, \frac{7}{9}] \ \cup \ [\frac{8}{9}, 1] \\ &\vdots \end{aligned}$$

Let $\mathcal{C}^* = \cap_{n \geq 0} C_n$. We have the following:

1. $\mathcal{C}^*$ is closed, uncountable and bounded.

2. Elements of $\mathcal{C}^*$ are those with a ternary expansion involving no ones.

3. Consider $f\colon \mathcal{C}^* \implies [0,1]$, where
   $x = (0.x_1 x_2 \cdots)_3 \mapsto (0.\frac{x_1}{2}\frac{x_2}{2}\cdots)_2$
   It is a continous, surjective map.

4. $\mathcal{C}^*$ has measure zero.

Theorem 36 ***(Cantor Surjection Theorem)***
If $\mathfrak{X}$ is a COMPACT metric space, then there exists a continuous surjection $f\colon \mathcal{C}^* \implies \mathfrak{X}$.

3 Differentiation

3.1 Differentiation of Real Valued Functions

Definition 44 *Suppose there is a function $f:\ (a,b) \to \mathbb{R}$, we say f is differentiable at a_0 if*

$$\lim_{h\to 0} \frac{f(a_0+h)-f(a_0)}{h}$$

exists; that is, for a given $\epsilon > 0$ there exists $\delta > 0$(δ may depend on a_0) and a real L such that

$$0 < |a_0 - x| < \delta \Rightarrow |\frac{f(x)-f(a_0)}{x-a_0} - L| < \epsilon.$$

Remark 19 *The derivative of f at a_0 is what we will call the number L in this definition. It heuristically captures the notion of slope of a tangent line to the curve y = f(x) at the point $(a_0, f(a_0))$.*

Remark 20 *The differentiability of f is a local property, that is f is differentiable on (a,b) if and only if it is differentiable at each $a_0 \in (a,b)$.*

Proposition 32 *If a function $f\colon\ (a,b) \to \mathbb{R}$ is differentiable at a_0, then f is continuous at a_0.*

Proof: Let $\epsilon > 0$, we need to find $\delta > 0$ such that

$$|x-a_0| < \delta \Rightarrow |f(x)-f(a_0)| < \epsilon.$$

Since f is differentiable at a_0 for this given $\epsilon > 0$, there exists $\delta > 0$ and L$\in \mathbb{R}$ such that

$$0 < |a_0 - x| < \delta \Rightarrow |\frac{f(x)-f(a_0)}{x-a_0} - L| < \epsilon$$

that is

$$L - \epsilon < \frac{f(x)-f(a_0)}{x-a_0} < L + \epsilon$$

which implies

$$(L-\epsilon)|x-a_0| < \frac{(f(x)-f(a_0))|x-a_0|}{x-a_0} < (L+\epsilon)|x-a_0|$$

We shall assume $\epsilon < |L|$. Therefore $0 \leq |f(x) - f(a_0)| \leq (L+\epsilon)|x - a_0|$ because $|x - a_0| < \delta_1$. therefore $|f(x) - f(a_0)| < (|L| + \epsilon)\delta$. This completes the proof. ♣

Proposition 33 *Suppose $f, g\colon (a,b) \to \mathbb{R}$ are differentiable at a_0.*

1. $(f+g)'(a_0) = f'(a_0) + g'(a_0)$.
2. $(fg)'(a_0) = f'(a_0)\, g(a_0) + g'(a_0)\, f(a_0)$.
3. $(\lambda f)'(a_0) = \lambda f'(a_0)$.
4. $\left(\frac{f}{g}\right)'(a_0) = \frac{f'(a_0)g(a_0) - g'(a_0)f(a_0)}{(g(a_0))^2}$. *if $g(x)$ is never zero in some small neighborhood of a_0.*
5. $(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$.

Proof:

1. Consider $(f+g)(x) = f(x) + g(x)$.

   We want to show:

   $\lim_{x\to a_0} \frac{(f+g)(x) - (f+g)(a)}{x - a_0}$ exists and is equal to $f'(a_0) + g'(a_0)$. Now,

   $$|\frac{(f(x) - f(a_0)) + (g(x) - g(a_0))}{x - a_0} - f'(a_0) - g'(a_0)| \leq |\frac{f(x) - f(a_0)}{x - a_0} - f'(a_0)| + |\frac{g(x) - g(a_0)}{x - a_0} - g'$$

   and these two terms are bounded above by ϵ_1, ϵ_2 respectively.

2. $|\frac{f(x)g(x) - f(a_0)g(a_0)}{x - a_0} - f(a_0)\, g'(a_0) - f'(a_0)\, g(a_0)|$

   $= |\frac{f(x)g(x) - f(x)g(a_0) + f(x)g(a_0) - f(a_0)g(a_0)}{x - a_0} - f(a_0)\, g'(a_0) - f'(a_0)\, g(a_0)|$

   $\leq |f(x)\left(\frac{g(x) - g(a_0)}{x - a_0}\right) - f(a_0)\, g'(a_0)| + |g(a_0)|\,|\frac{f(x) - f(a_0)}{x - a_0} - f'(a_0)|$

   Since g(x) is differentiable in the interval so g(x) is bounded (by a positive number M) and f(x) is differentiable

   $\therefore |g(a_0)|\,|\frac{f(x) - f(a_0)}{x - a_0} - f'(a_0)| < M\epsilon_1 \forall \epsilon_1 \in \mathbb{R}_+$

   Hence it suffices to show that the other term is small enough.

   Now, $|f(x)\left(\frac{g(x) - g(a_0)}{x - a_0}\right) - f(a_0)\, g'(a_0) + f(x)\, g(a_0) - f(x)\, g(a_0)|$

   $\leq |f(x)|\,|\left(\frac{g(x) - g(a_0)}{x - a_0}\right) - g'(a_0)| + |g'(a_0)|\,|f(x) - f(a_0)|$

Since $f'(a_0)$ exists and f is also continuous which implies $|f(x)| \leq f(\alpha)$ for some constant α in $|x - a_0| < \delta$ for some small $\delta > 0$.

3. $|\frac{\lambda f(x) - \lambda f(x)}{x - a_0} - \lambda f'(a_0)|$

   $= |\lambda||\frac{f(x) - f(x)}{x - a_0} - f'(a_0)|$

   $< |\lambda|\epsilon_1$

   Choose $\epsilon_1 \leq \frac{\epsilon}{2\lambda}$; this gives

$$|\lambda||\frac{f(x) - f(x)}{x - a_0} - f'(a_0)| < \frac{\epsilon}{2}.$$

4. **QUOTIENT RULE :**

$$\begin{aligned}
|\frac{\frac{1}{g(x)} - \frac{1}{g(a_0)}}{x - a_0} + \frac{g'(a_0)}{g^2(a_0)}| &= |\frac{g(a_0) - g(x)}{(x - a_0)\, g(x)\, g(a_0)} + \frac{g'(a_0)}{g^2(a_0)}| \\
&= \frac{1}{|g(a_0)|}|\frac{g(a_0) - g(x)}{(x - a_0)\, g(x)} + \frac{g'(a_0)}{g(a_0)}| \\
&= \frac{1}{|g(a_0)|}|\frac{g(x) - g(a_0)}{(x - a_0)\, g(x)} - \frac{g'(a_0)}{g(a_0)}| \\
&= \frac{1}{|g(a_0)|}|\frac{g'(a_0)}{g(a_0)} - \frac{\frac{g(x) - g(a_0)}{(x - a_0)}}{g(x)}|
\end{aligned}$$

5. Define variable $v = \frac{g(x+h) - g(x)}{h} - g'(x)$ and $w = \frac{f(y+k) - f(y)}{k} - f'(y)$

   We see that $v \to 0$ as $h \to 0, w \to 0$ as $k \to 0$.

   We can rewrite the above equations as

$$\begin{aligned}
g(x + h) &= g(x) + [g'(x) + v]h, \\
f(y + k) &= f(y) + [f'(y) + w]k
\end{aligned}$$

   Taking f of the first equation, we get

   $f(g(x + h)) = f(g(x) + [g'(x) + v]h)$

   Now, put $y = g(x)$ and $k = [g'(x) + v]h$ in the second equation

   $f(y + k) = f(g(x) + [g'(x) + v]h) = f(g(x)) + [f'(g(x)) + w].[g'(x) + v]h$

   So, $\frac{f(g(x+h)) - f(g(x))}{h} = \frac{f(g(x)) + [f'(g(x)) + w].[g'(x) + v].h - f(g(x))}{h}$

$= \frac{[f'(g(x))+w].[g'(x)+v].h}{h}$

$= [f'(g(x)) + w].[g'(x) + v].h$

$\lim_{h\to 0} \frac{f(g(x+h))-f(g(x))}{h} = \lim_{h\to 0}[f'(g(x)) + w].[g'(x) + v]$

$= (\lim_{h\to 0} f'(g(x)) + \lim_{h\to 0} w).(\lim_{h\to 0} g'(x) + \lim_{h\to 0} v)$

$= f'(g(x)).g'(x)$

since $v \to 0$ as $h \to 0$ and $w \to 0$ as $h \to 0$.

♣

Remark 21 *Suppose* $f\colon (0,1) \to \mathbb{R}$ *is differentiable at* $a \in (0,1)$. *Then by definition, given* $\epsilon > 0$ *there exists* $\delta > 0$ *such that*

$0 < |x - a_0| < \delta \Rightarrow |\frac{f(x)-f(a_0)}{x-a_0} - f'(a_0)| < \epsilon$

$\Rightarrow f(a_0) - \epsilon < \frac{f(x)-f(a_0)}{x-a_0} < f(a_0) + \epsilon$ *for all* $x \in (a_0 - \delta, a_0 + \delta) - \{a_0\}$

In particular, for $x \in [a_0, a_0 + \delta)$,

$(f'(a_0) - \epsilon)(x - a_0) \leq f(x) - f(a_0) \leq (f'(a_0) + \epsilon)(x - a_0)$

$\Rightarrow (f'(a_0) - \epsilon)(x - a_0) + f(a_0) \leq f(x) - f(a_0) + f(a_0) \leq (f'(a_0) + \epsilon)(x - a_0) + f(a_0)$

that is $f(x) = f'(a_0)(x - a_0) + f(a_0)$
Similar analysis can be done for $x \in (a_0 - \delta, a_0]$.

This is called ***(NEWTON'S APPROXIMATION)*** *or* ***(THE FIRST ORDER EXPANSION)***

Theorem 37 ***(Rolle's theorem)*** *Suppose* $f:\ (0,1) \to \mathbb{R}$ *is continuous and suppose f is differentiable at each* $x \in (0,1)$. *Further suppose f(0) = f(1), then there exists* $\xi \in (0,1)$ *such that* $f'(\xi) = 0$

Proof: Without loss of generality, we assume $f(0) = f(1) = 0$.

1. Suppose $f(x) = 0$ for all $x \in (0,1)$, then $f'(x) = 0$ everywhere in the interval. Hence this part is trivial.

2. Suppose $f(x)$ is not identically 0.

   Without loss of generality, let $f(\alpha) > 0$ for some $\alpha \epsilon (0,1)$. Since $f\colon (0,1) \to \mathbb{R}$ is continuous, it attains maximum. Since $f(\alpha) > 0$, if f attains maximum at some $\xi \in [0,1]$, then $f(\xi) > 0$, in particular for $\xi \in (0,1)$.

Claim 15 $f'(\xi) = 0$.

Suppose not, that is $f'(\xi) > 0$ (this can be assumed without loss of generality). Then given $\epsilon > 0$ there exists $\delta > 0$ such that for all $x \in (\xi - \delta, \xi + \delta)$

$f'(\xi) - \epsilon < \frac{f(x) - f'(\xi)}{x - \xi} < f'(\xi) + \epsilon$

Pick $\epsilon = \frac{f'(\xi)}{2} - \epsilon$ and pick a corresponding $\delta > 0$, then

$\frac{f'(\xi)}{2} < \frac{f(x) - f(\xi)}{x - \xi} < \frac{3f'(\xi)}{2}$ holds for all $x \in (\xi - \delta, \xi - \delta)$.

In particular, at $x = \xi + \frac{\delta}{2}$,

$\frac{f'(\xi)}{2}\frac{\delta}{2} < f\left(\xi + \frac{\delta}{2}\right) - f(\xi)$, contradicting that f attains maximum at ξ

The case where $f'(\xi) < 0$ is similarly dealt with.

♣

Remark 22 *The above proof actually shows that if $f'(a_0) > 0$, then for element to the right of a_0 in its small neighbourhood, the function value is higher than that at a_0, the function value is higher than that at a_0; likewise for all x to the left of a_0 and sufficiently close to a_0, $f(x) < f(a_0)$.*

3.2 The Mean Value Theorems and Consequences

Theorem 38 ***(Mean Value Theorem):*** *Given f: $[a, b] \to \mathbb{R}$, continuous on $[a, b]$ and differentiable on (a, b), there exists $\xi \in (a, b)$ such that $f'(\xi) = \frac{f(b) - f(a)}{b - a}$.*

Proof: Consider $g(x) = f(x) - \frac{f(b) - f(a)}{b - a}$ define on $[a, b]$ Then g is continuous on $[a, b]$,differentiable on (a, b) and $g(a) = g(b)$. By applying Rolle's theorem on $g(x)$, we get $g'(\xi) = 0$ for some $\xi \ \epsilon (a, b)$. So,

$$f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0$$

$\Rightarrow$

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

♣

Proposition 34 ***Ratio Mean Value Theorem*** *Suppose f,g: $[a, b] \to \mathbb{R}$ are continuous on $[a, b]$, differentiable on (a, b), then there exists $\xi \in (a, b)$ such that*

$$(f(b) - f(a))g'(\xi) = (g(b) - g(a))f'(\xi).$$

Proof: Consider the function, $h(x) = \frac{f(x)-f(a)}{f(b)-f(a)} + \frac{g(b)-g(x)}{g(b)-g(a)} - 1$

Since $h(a) = 0$ and $h(b) = 0$, applying mean value theorem to h(x), we get

$$h'(\xi) = 0$$

$\Rightarrow$

$$\frac{f'(\xi)}{f(b) - f(a)} = \frac{g'(\xi)}{g(b) - g(a)}$$

$\Rightarrow$

$$(f(b) - f(a))g'(\xi) = (g(b) - g(a))f'(\xi)$$

♣

Proposition 35 *Suppose f: $[0, 1] \to \mathbb{R}$ is continuous on $[0, 1]$, differentiable on $(0, 1)$ and suppose that $f'(x) > 0$ for all $x\epsilon(0, 1)$. Prove that f is strictly increasing.*

Proof: Let us suppose that $f(a) \geq f(b)$ and $a < b$ for some a,b$\epsilon(0, 1)$

Applying Mean value theorem on the interval (a, b), for some $\xi \in (a, b)$

$$\frac{f(b) - f(a)}{b - a} = f'(\xi)$$

Since $f'(x) > 0$ for all x in the interval, we have

$$\frac{f(b) - f(a)}{b - a} > 0$$

which implies

$$f(b) > f(a)$$

which is a contradiction. ♣

Remark 23 *The converse is false. Take for example the function $f(x) = x^3$.*

3.2.1 A Theorem of Darboux

The following theorem, due to Darboux tells us a rather interesting feature of the first derivative. In particular, it tells that arbitrary functions cannot be candidates for the derivative function.

Theorem 39 *If f: $[a, b] \to \mathbb{R}$ is differentiable ,then function f' satisfies the Intermediate Value Theorem.*

Corollary 4 *f' cannot have JUMP DISCONTINUITIES.*

Proof: Pick $\epsilon > 0$ such that $f'(x_1) + \frac{\epsilon}{2} < \gamma < (x_2) - \frac{\epsilon}{2}$ Let $h_1 > 0$ such that $|\frac{f(x_1+h)-f(x_1)}{h} - f(x_1)| < \frac{\epsilon}{2}$ and $|\frac{f(x_2+h)-f(x_2)}{h} - f(x_2)| < \frac{\epsilon}{2}$

In particular $\frac{f(x_1+h)-f(x_1)}{h} < f'(x_1) + \frac{\epsilon}{2} < \gamma$ and $\frac{f(x_2+h)-f(x_2)}{h} > f'(x_2) - \frac{\epsilon}{2} > \gamma$

Let $g(x) = \frac{f(x+h)-f(x)}{h}$ for $x \in [x_1, x_2]$. By the continuity of f on (a, b) it follows that g is also continuous.

Hence ,by I.V.P for g ,it follows that there exists $y \in [x_1, x_2]$ such that

$$\left|\frac{f(y+h) - f(y)}{h}\right| = \gamma.$$

Now restrict f to $[y, y+h]$ and apply the mean value theorem. ♣

3.2.2 The L'Hôpital Rule

The L'Hôpital Rule is one of the most convenient-to-use results in Differential Calculus. But surprisingly, history seems to suggest that the result was most probably discovered by Bernoulli, and yet does not bear his name.

Theorem 40 ***(The L'Hôpital Rule):*** *Suppose f,g* $(a, b) \to \mathbb{R}$ *such that* $\lim_{x\to 0} f(x) \to 0, \lim_{x\to 0} g(x) \to 0$ *and if* $\lim_{x\to 0} \frac{f'(x)}{g'(x)} = L$ *for some real L,then* $\lim_{x\to 0} \frac{f(x)}{g(x)} = L$.

Proof: Given $\epsilon > 0$ we need to show that

$$|\frac{f(x)}{g(x)} - L| < \epsilon$$

We note that by Ratio Mean Value Theorem for any interval (y, x), we have

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(\xi)}{g'(\xi)}$$

for some $\epsilon(y, x)$ and ξ and $g(x) - g(y) \neq 0$. We can see by the triangle inequality that

$$|\frac{f(x)}{g(x)} - L| \leq |\frac{f(x)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)}| + |\frac{f(x) - f(y)}{g(x) - g(y)} - L|.$$

Since $\frac{f'(x)}{g'(x)}$ converges as x goes to 0, we can pick $\delta_0 > 0$ small enough so that $|\frac{f'(x)}{g'(x)} - L| < \frac{\epsilon}{2}$. The second term in the inequality above can be written

$$|\frac{f(x)}{g(x)} - L| = |\frac{f'(\xi_{x,y})}{g'(\xi_{x,y})} - L|,$$

where $y < \xi_{x,y} < x$, so by choosing $x < \delta_0$, we can make $|\frac{f(x)}{g(x)} - L| < \frac{\epsilon}{2}$. In order to bound the first term, note that

$$|\frac{f(x)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)}| = |\frac{f(y)g(x) - f(x)g(y)}{g(x)(g(x) - g(y))}|$$

Now since $g(x) \to 0$ as $x \to 0$ for a fixed $x < \delta_0$ we can choose $y < y_0 < x$ such that $|g(x) - g(y)| > \frac{|g(x)|}{2}$. Consequently,

$$|\frac{f(y)g(x) - f(x)g(y)}{g(x)(g(x) - g(y))}| \leq |\frac{2(|f(x)| + |g(x)|)}{g^2(x)}|(|f(y)| + |g(y)|)$$

Again, since both f, g go to 0 as y approaches 0, pick $y < y_1$ so that

$$|f(y)| + |g(y)| < (\frac{g^2(x)}{2(|f(x)| + |g(x)|)})(\frac{\epsilon}{2})$$

which makes the first term less than $\frac{\epsilon}{2}$.This completes the proof. ♣

Corollary 5 *Suppose $f, g : [a, b] \to \mathbb{R}$ are differentiable on (a, b) and suppose $\frac{f'}{g'}$ converges to a finite limit L, as $x \to a$.Suppose that $\frac{f(x)}{g(x)}$ is of the form ∞/∞ as $x \to a$, then,*

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} = L$$

Proof:
Consider $u(x) = \frac{1}{f(x)}$ and $v(x) = \frac{1}{g(x)}$, and let $\lim_{x\to a} \frac{f(x)}{g(x)} = L_0$.
Then $\frac{v(x)}{u(x)}$ is of the form 0/0 and $\frac{v'}{u'}$ will also converge to the finite limit.
Using L'Hôpital Rule, we will have,

$$\lim_{x\to a} \frac{v(x)}{u(x)} = \lim_{x\to a} \frac{v'(x)}{u'(x)}$$

$$\Rightarrow \lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{g'(x)}{f'(x)} \frac{(f(x))^2}{(g(x))^2}$$

Since, $\frac{f'(x)}{g'(x)}$ converges to a finite limit L and $\frac{f(x)}{g(x)}$ to L_0, as $x \to a$, we can write this as,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{g'(x)}{f'(x)} \lim_{x\to a} \frac{(f(x))^2}{(g(x))^2}$$

which gives $L_0 = \frac{1}{L} L_0^2$, so we have $L = L_0$. ♣

3.3 Series

Definition 45 *Given a real sequence* $\{a_n\}$, *if the sequence of partial sums* $\sum_{k=1}^{n} a_k$ *converges,we say that the series* $\sum_{k=1}^{n} a_k$ *converges .*

Example 12 $a_n = x^n$,$0 < x < 1$
The partial sums are

$$\sum_{k=1}^{n} x_k = x + x_2 + x_3 + ... + x_n = \frac{x\,(1 - x^n)}{(1 - x)}.$$

As $n \to \infty$ *we have* $x_n \to 0$, *so* $\sum_{k=1}^{n} x_k = \frac{x}{(1 - x)}$.

Definition 46 *We say* $\sum_{k=1}^{n} x_k$ *diverges to infinity,* $\{S_n\}$ *of partial sums satisfies :*
Given $M > 0$ *,there exists* $n_0 \in \mathbb{N}$ *such that for all* $n \geq n_0$, *we have* $S_n > M$.

Example 13 $a_n = 1$ *for all* $n_0 \in \mathbb{N}$.

Proposition 36 *Suppose* $\sum_{k=1}^{\infty} a_k$ *converges, then* $a_n \to 0$ *as* $n \to \infty$.

Proof: Given $\epsilon > 0$,let n_0 be such that for all $n \geq n_0$ $|S_n - L| < \frac{\epsilon}{2}$ for some real L.
In particular for $n \geq n_0$, we have $|S_n + 1 - L| < \frac{\epsilon}{2}$ and $|S_{n+1} - L| < \frac{\epsilon}{2}$, so
$\Rightarrow |S_{n+1} - S_n| < \epsilon$,
$\Rightarrow |a_{n+1}| < \epsilon$. ♣

Remark 24 *THE CONVERSE IS NOT TRUE;*

Example 14 : $a_n = \frac{1}{n}$, *The Harmonic series. It is easy to see that this series diverges as*

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} > \frac{1}{2}$$

holds for all n.

Theorem 41 *Suppose* $\{a_n\}$,$\{b_n\}$ *are non-negative sequences. Suppose* $a_n \leq b_n$ *for all* n *large enough, i.e.,* $(n \geq n_0)$.

1. *If* $\sum_{k=1}^{n} b_k$ *converges ,then so does* $\displaystyle\sum_{k=1}^{\infty} a_n$.

2. *If* $\sum_{k=1}^{n} a_k$ *diverges to infinity, so does* $\displaystyle\sum_{k=1}^{\infty} b_n$.

Proof:

1. $\sum_{k=1}^{n} b_k$ converges ,so in particular there exists $n_1 \in \mathbb{N}$ such that $|\sum_{k=1}^{n} b_k - L| < 1 \ \forall n \geq n_1$,for some real L.
   In particular, $\sum_{k=1}^{n} b_k < L + 1$ for all $n \geq n_1$,
   $\Rightarrow \sum_{k=1}^{n} a_k < L + 1$ for all $n \geq n_1$,
   $\Rightarrow \{\sum_{k=1}^{n} a_k\}$ is bounded above.
   Since $\sum_{k=1}^{n} a_k$ is increasing and bounded above it follows that $\{\sum_{k=1}^{n} a_k\}$ converges.

2. $\sum_{k=1}^{n} a_k$ diverges ,so in particular there exists $n_1 \in \mathbb{N}$ such that $\sum_{k=1}^{n} a_k > M$ for all $n \geq n_1$,for some M.
   $\Rightarrow \sum_{k=1}^{n} b_k > M$ for all $n \geq n_1$.
   Hence $\sum_{k=1}^{n} b_k$ diverges.

♣

Theorem 42 *Suppose* $a_n \geq 0$ *.If* $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L$*,then*

1. *If* $0 \leq L < 1$ *,then* $\sum_{k=1}^{n} a_k$ *converges.*

2. *If* $L > 1$ *,then* $\sum_{k=1}^{n} a_k$ *diverges.*

3. *If* $L = 1$*, then the test is inconclusive.*

Proof:

1. Suppose $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L < 1$
   $\Rightarrow$ there exists n_0 such that if $n \geq n_0$,then
   $(L - \epsilon)a_n < a_{n+1} < (L + \epsilon)a_n$ where $L + \epsilon < 1$
   $a_{n_0+1} < (L + \epsilon)a_{n_0}$
   $a_{n_0+2} < (L + \epsilon)a_{n_0+1}$
   $a_{n_0+3} < (L + \epsilon)a_{n_0+2}$
   In general,
   $a_{n_0+k} < (L + \epsilon)^k a_{n_0}$
   $\Rightarrow \sum_{k\geq 1} a_{n_0+k} < \sum_{k\geq 1} (L + \epsilon)^k a_{n_0}$
   But since $\sum_{k\geq 1} (L + \epsilon)^k$ is convergent $(L + \epsilon < 1)$

we must have covergence of $\sum_{k\geq 1} a_n$ as well.

2. Suppose $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L > 1$
   $\Rightarrow$ there exists n_0 such that if $n \geq n_0$,then
   $(L-\epsilon)a_n < a_{n+1}$ where $L-\epsilon > 1$
   $a_{n_0+1} > (L-\epsilon)a_{n_0}$
   In general
   $a_{n_0+k} > (L-\epsilon)^k a_{n_0}$
   Since $\sum_{k\geq 1}(L+\epsilon)^k$ is divergent $(L-\epsilon > 1)$ it follows that $\sum_{k\geq 1} a_n$ diverges.

♣

Remark 25 *For L=1, RATIO TEST is inconclusive, that is there are instances of series where* $\frac{a_{n+1}}{a_n} \to 1$ *and* $\sum_{n\geq 1} a_n$ *converges and also instances where* $\frac{a_{n+1}}{a_n} \to 1$ *but* $\sum_{n\geq 1} a_n$ *diverges.*

Theorem 43 *Suppose* $a_n \geq 0$ *and* $\lim_{n\to\infty} a_n^{\frac{1}{n}} = L$ *.Then*

1. *If* $0 \leq L < 1$ *,then* $\sum_{n\geq 1} a_n$ *converges*

2. *If* $L > 1$*,then* $\sum_{n\geq 1} a_n$ *diverges*

Proof:

1. Fix $\epsilon > 0$,such that $L+\epsilon < 1$
   Get $n_0 \in \mathbb{N}$ such that for $n \geq n_0$,we have
   $L-\epsilon < a_n^{\frac{1}{n}} < L+\epsilon$
   $(L-\epsilon)^n < a_n < (L+\epsilon)^n$
   Now,we get
   $\sum_{k=1}^{N} a_{n_0+k} < (L+\epsilon)^{n_0}\left\{\sum_{k=1}^{N}(L+\epsilon)^K\right\}$
   But since $\sum_{k\geq 1}(L+\epsilon)^k$ is convergent $(L+\epsilon < 1)$
   $\sum_{k\geq 1} a_n$ converges as well.

2. The proof of this part is similar, and we omit the details.

♣

Remark 26 *1. Consider $x_n = a_n x^n$ for some $x > 0$ a real number for some positive sequence a_n. If $x \lim_{n\to\infty} a_n{}^{\frac{1}{n}} < 1$, then $\sum x_n$ converges that is,*

if

$$0 < x < \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{a_n}},$$

then $\sum x_n = \sum a_n x^n$ converges.

2. A general technique: Calculate $L = \lim_{n\to\infty} a_n{}^{\frac{1}{n}}$. If this limit is finite for all $0 \leq x < \frac{1}{L}$ then series $\sum x_n$ converges. We may define a function f: $\left[0, \frac{1}{L}\right) \to \mathbb{R}$.

$f(x) = \sum a_n x^n$

Theorem 44 *Suppose $a_n \in \mathbb{R}$. We can still define convergence of $\sum a_n$ .Suppose $\sum |a_n|$converges. then $\sum a_n$ also converges.*

Proof: We can check the Cauchyness of $\{\sum S_n\}$, that is for given $\epsilon > 0$,we must find $n_0 \in \mathbb{N}$ such that $|S_m - S_n| < \epsilon$ for all $m, n \geq n_0$ Without Loss of Genrality, $m > n$.
$|S_m - S_n| < |a_m + a_{m-1} + ... + a_{n+1}| \leq |a_m| + |a_{m-1}| + ... + |a_{n+1}|$ $(*)$
since $\sum |a_n|$ converges, there exists $n_0 \in \mathbb{N}$ such that $\sum_{n>n_0}^{n_0+N} |a_n| < \epsilon$ for any $N \geq 1$
Clearly, this same n_0 establishes that R.H.S of $(*) < \epsilon$

♣

Definition 47 *We say that a series $\sum a_n$ is absolutely convergent if $\sum |a_n|$ converges.*

Remark 27 *Consider $\sum a_n x^n$. The Root test applies to $|a_n x^n| = b_n$ In general if $a_n \downarrow 0$,then $\sum (-1)^n a_n$ converges.*

Proposition 37 *Rearrangements of absolutely convergent sums are absolutely convergent. i.e. given an absolutely convergent series $\sum a_n$ and a bijection $\sigma : \mathbb{N} \to \mathbb{N}$, $\sum a_{\sigma(n)}$ is absolutely convergent.*

Proof:
Pick any $\epsilon > 0$. Since $\sum a_n$ is absolutely convergent, there exists an $n_0 \in \mathbb{N}$ such that

$$\sum_{k=n}^{m} |a_k| < \epsilon \;\; \forall n, m \geq n_0$$

Now, consider the set $S = \{t | \sigma(t) < n_0, t \in \mathbb{N}\}$. Since S is a finite set (it has exactly n_0 elements, σ being just a rearrangement), so S has a maximum element, say M_S. Let $n_0' = M_S + 1$.

So, for any $u \geq n_0'$, we have $\sigma(u) \geq n_0$. (Since u clearly does not belong to S). Thus, for any $m \geq n \geq n_0'$,

$$\sum_{k=n}^{m} |a_{\sigma(k)}| < \sum_{t \geq n_0} |a_t| < \epsilon$$

♣

Proposition 38 *Rearrangements of a series do not necessarily have the same sum as the series itself. i.e. if $\sum a_n$ is a convergent series, and $\sigma : \mathbb{N} \to \mathbb{N}$ is a bijection such that $\sum a_{\sigma(n)}$ is also convergent, $\sum a_{\sigma(n)}$ may not be equal to $\sum a_n$.*

Proof: Consider the series $a_n = \frac{(-1)^n}{n}$, which sums to S. We propose a rearrangement σ such that $a_{\sigma(n)}$ will sum to $2S$.

$$S = \sum_{n \geq 1} \frac{(-1)^n}{n}$$

$$\frac{S}{2} = \sum_{n \geq 1} \frac{(-1)^n}{2n}$$

$$\frac{S}{2} = \sum_{n \geq 1} b_n$$

where b_n is 0 if n is odd, and $b_n = \frac{(-1)^{\frac{n}{2}}}{n}$ if n is even. (Basically, we've just added a '+0' after every term in the series). Also,

$$\frac{S}{2} = \sum_{n \geq 1} c_n$$

where c_n is 0 if n is not divisible by 4, and $c_n = \frac{(-1)^{\frac{n}{4}}}{n}$ if n is divisible by 4. (Here, we've added a '+0' after every term in b_n). So,

$$2S = S + \frac{S}{2} + \frac{S}{2}$$

$$= \sum_{n \geq 1} (a_n + b_n + c_n)$$

Let $d_n = a_n + b_n + c_n$ for all $n \in \mathbb{N}$. As we have just shown, $\sum d_n = 2S$. However, it can be seen that d_n is just a rearrangement of a_n, where every even term occurs at every fifth place, and odd terms occur at the rest of the places. ♣

3.3.1 Power Series

Definition 48 ***(Radius of Convergence)*** *Given a series* $\sum_{n\geq 0} a_n x^n$ *, its* ***radius of convergence*** *is defined to be*

$$R = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}}$$

Theorem 45 *Suppose* $f(x) = \sum\limits_{n\geq 0} a_n x^n$ *has radius of convergence* $R > 0$*, then*

1. $g(x) = \sum_{n\geq 1} n a_n x^{n-1}$ *also has radius of convergence* R.
2. f *is differentiable on* (R, R) *and* $f'(x) = g(x)$ *for all* $x \in (-R, R)$.

Proof:
Part 1:
The radius of convergence of this series is

$$\frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_{n+1}|(n+1)}} = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_{n+1}|}} \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{n+1}}$$

$$= R \cdot 1$$

$$= R.$$

Part 2:
Pick a certain $x \in (-R, R)$
We need to show that for every $\epsilon > 0$ there exists a $\delta > 0$ such that $|\frac{f(x)-f(t)}{x-t} - g(x)| < \epsilon$ provided $|x - t| < \delta$; that is, as $x \to t$ we have $|\frac{f(x)-f(t)}{x-t} - g(x)| \to 0$
Without loss of generality, we can assume $x < t < \beta < R$

$$|\frac{f(x) - f(t)}{x - t} - g(x)| = |\sum_{n\geq 1} a_n (\frac{x^n - t^n}{x - t} - nx^{n-1})|$$

$$= |\sum_{n\geq 1} a_n [(x^{n-1} + tx^{n-2} + t^2 x^{n-3} + ... + t^{n-1}) - nx^{n-1}]|$$

$$= \sum_{n\geq 1} |a_n||(x^{n-1} - x^{n-1}) + (tx^{n-2} - x^{n-1}) + ... + (t^{n-1} - x^{n-1})|$$

$$= \sum_{n\geq 1} |a_n||x^{n-2}(t - x) + ... + x^{n-r-1}(t^r - x^r) + ... + (t^{n-1} - x^{n-1})|$$

$$= \sum_{n\geq 1} |a_n||(t - x)[x^{n-2} + x^{n-3}(t + x) + ... + (t^{n-2} + t^{n-3}x + ... + x^{n-2})]|$$

$$\leq \sum_{n\geq 1} |a_n||t-x|\frac{n(n-1)}{2}\beta^{n-2}$$

$$= \frac{|t-x|}{\beta^2}(\sum_{n\geq 1} |a_n|\frac{n(n-1)}{2}\beta^n)$$

Since $\beta < R$, the *root test*, tells us that

$$\limsup \sqrt[n]{|a_n|\frac{n(n-1)}{2}} = \frac{1}{R}\cdot 1 = \frac{1}{R},$$

and we have

$$|\frac{f(x)-f(t)}{x-t} - g(x)| \leq \frac{|x-t|}{\beta^2}(\sum_{n\geq 1} |a_n|\frac{n(n-1)}{2}\beta^n).$$

Thus as $x \to t$ or $|x-t| \to 0$,

$$|\frac{f(x)-f(t)}{x-t} - g(x)| \to 0.$$

♣

Corollary 6 *A function defined by a power series $f(x) = \sum_{n\geq 0} a_n x^n$ is infinitely often differentiable, and the derivatives can be computed by differentiating the series term-by-term.*

In particular, we recall some 'well-known' functions' from elementary calculus; here are their formal definitions!

Definition 49

- $$e^x := \sum_{n\geq 0} \frac{x^n}{n!}.$$
- $$\cos(x) := \sum_{n\geq 0} \frac{(-1)^n}{(2n)!}x^{2n}.$$
- $$\sin(x) := \sum_{n\geq 0} \frac{(-1)^n}{(2n+1)!}x^{2n+1}.$$

By the theorem above,

- $(e^x)' = e^x$.
- $(\sin(x))' = \cos(x)$.
- $(\cos(x))' = \sin(x)$.

Consider $f(x) = \sum_{n\geq 0} a_n x^n$, then on differentiating it repeatedly we obtain

$$f(0) = a_0,\ f'(0) = a_1,\ \frac{f''(0)}{2!} = a_2,\ \frac{f'''(0)}{3!} = a_3$$

and so on. In general,

$$a_n = \frac{f^{(n)}(0)}{n!}$$

In short, if f has a power series in some neighborhood of 0, then power series expansion for f can be uniquely determined.

If we are given a function f, can we determine if it has a *power series expansion*? Suppose f has derivatives of all order, the *candidate* for the power series of f is

$$f(x) = \sum_{n\geq 0} \frac{f^{(n)}(0)}{n!} x^n$$

Definition 50 *A function is said to be* ***analytic*** *on an open interval I, if*

$$f(x) = \sum_{n\geq 0} a_n (x-\alpha)^n$$

which holds for all $x \in I$, for some $\alpha \in I$.

NOTE: The expansion $\sum_{n\geq 0} \frac{f^{(n)}(0)}{n!} x^n$ is called the **TAYLOR SERIES** for f.

Definition 51 *A function is said to be* ***smooth*** *if all its derivatives exist.*

FACT: There exist **smooth** and **non-analytic** functions defined on $\mathbb{R}$.

3.3.2 Taylor Series and Taylor Approximation

Theorem 46 ***(Taylor Approximation)***
Suppose $f^{(i)}(x)$ exists for all $x \in (a,b)$ and $0 \leq i \leq r$ for some $r \in \mathbb{N}$.
Define $R(h) = f(x+h) - P_{Taylor}(h)$ where,

$$P_{Taylor}(h) = \sum_{k=0}^{r} \frac{f^{(k)}(x)}{k!} h^k$$

Then,

$$\lim_{h\to 0} \frac{R(h)}{h^r} = 0.$$

Proof:
Suppose $h > 0$
Now we have $R(0) = R'(0) = R''(0) = ... = R^{(r)}(0) = 0$.
Thus by the *mean value theorem*, we have a θ_1 where $0 \leq \theta_1 \leq h$ such that,

$$R(h) = R(h) - R(0) = R'(\theta_1)h$$

Similarly we will have,

$$R'(\theta_1) = R'(\theta_1) - R'(0) = R''(\theta_2)\theta_1$$

$$R''(\theta_2) = R''(\theta_2) - R''(0) = R'''(\theta_3)\theta_2$$

$$\vdots$$

$$R^{(r-2)}(\theta_{r-2}) = R^{(r-2)}(\theta_{r-2}) - R^{(r-2)}(0) = R^{(r-1)}(\theta_{r-1})\theta_{r-2}$$

where, $h \geq \theta_1 \geq \theta_2 \geq ... \geq \theta_{r-1} \geq 0$.

$$|\frac{R(h)}{h^r}| = |\frac{R'(\theta_1)h}{h^r}| = ... = |\frac{\theta_1\theta_2...\theta_{r-2}hR^{(r-1)}(\theta_{r-1})}{h^r}| = |\frac{\theta_1}{h}\frac{\theta_2}{h}...\frac{\theta_{r-2}}{h}\frac{R^{(r-1)}(\theta_{r-1})}{h}|$$

Since we have $h \geq \theta_1 \geq \theta_2 \geq ... \geq \theta_{r-1} \geq 0$, so each $\frac{\theta_i}{h}$ will be less than 1. Thus,

$$|\frac{R(h)}{h^r}| \leq |\frac{R^{(r-1)}\theta_r}{h}| \leq |\frac{R^{(r-1)}(\theta_{r-1})}{\theta_{r-1}}| = |\frac{R^{(r-1)}(\theta_{r-1}) - R^{(r-1)}(0)}{\theta_{r-1}}|$$

Now, as $h \to 0$, we also have $\theta_{r-1} \to 0$. So,

$$\lim_{h\to 0}|\frac{R(h)}{h^r}| = \lim_{h\to 0}|\frac{R^{(r-1)}(\theta_{r-1}) - R^{(r-1)}(0)}{\theta_{r-1}}| = \lim_{\theta_{r-1}\to 0}|\frac{R^{(r-1)}(\theta_{r-1}) - R^{(r-1)}(0)}{\theta_{r-1}}| = |R^{(r)}(0)| = 0$$

Therefore,

$$\lim_{h\to 0}\frac{R(h)}{h^r} = 0.$$

♣

Theorem 47 *Suppose in addition to the hypothesis of the above theorem, we also have that $f^{(r+1)}(x)$ exists, that is $f^{(i)}(x)$ exists for all $x \in (a,b)$ and $0 \leq i \leq r+1$, then*

$$R(h) = \frac{f^{(r+1)}(\theta)}{(r+1)!}h^{r+1},$$

for some $\theta \in (x, x+h)$.

Proof:
Suppose $h > 0$. Consider for $0 \leq t \leq h$

$$g(t) = f(x+t) - \sum_{k=0}^{r} (\frac{f^{(k)}(x)}{k!} t^k) - \frac{R(h)}{h^{r+1}} t^{r+1}$$

Now we have $g(0) = g'(0) = ... = g^{(r)}(0) = 0$. So using the *mean value theorem*, we get $g(0) = 0$ and $g(h) = 0 \Rightarrow$ there exists θ_1, $0 \leq \theta_1 \leq h$ such that $g'(\theta_1) = 0$
Hence there exists θ_2, $0 \leq \theta_1 \leq \theta_1$ such that $g''(\theta_2) = 0$
Continuing in this vein, we see that there exists θ_r, $0 \leq \theta_r \leq \theta_{r-1}$ such that $g^{(r)}(\theta_r) = 0$.

Also $g^{(r)}$ is differentiable, so,

$$g^{(r+1)}(t) = f^{(r+1)}(x+t) - \frac{R(h)}{h^{r+1}}(r+1)!$$

Since $g^{(r)}(0) = g^{(r)}(\theta_r) = 0$, by *mean value theorem*, there exists $\theta_{r+1} \in (0, \theta_r)$ such that

$$g^{(r+1)}(\theta_{r+1}) = 0$$

$$\Rightarrow f^{(r+1)}(x + \theta_{r+1}) - \frac{R(h)}{h^{r+1}}(r+1)! = 0$$

Thus for $\theta = x + \theta_{r+1}$, we have $\theta \in (x, x+\theta_r) \subseteq (x, x+h)$ such that

$$f^{(r+1)}(\theta) - \frac{R(h)}{h^{r+1}}(r+1)! = 0$$

$$\Rightarrow R(h) = \frac{f^{(r+1)}(\theta)}{(r+1)!} h^{r+1}$$

♣

3.4 Uniform Convergence

Definition 52 *A sequence $\{f_n\}$ is said to* ***converge uniformly*** *to f on $[0,1]$ if given $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \epsilon$ whenever $n \geq n_0$ for all $x \in [0,1]$. This is denoted by writing $f_n \rightrightarrows f$.*

Theorem 48 *Suppose $f_n : [0,1] \to \mathbb{R}$ and $f_n \rightrightarrows f$ on $[0,1]$. If f_n are continuous at x_0 for some $x_0 \in [0,1]$, then f is also continuous at x_0.*

Proof:
Given an $\epsilon > 0$, we have to show that there exists $\delta > 0$ such that $|x_0 - x| < \delta \Rightarrow |f(x_0) - f(x)| < \epsilon$.

By the *triangle inequality*,

$$|f(x_0) - f(x)| \leq |f(x_0) - f_n(x_0)| + |f_n(x_0) - f_n(x)| + |f_n(x_0) - f(x)|$$

By *uniform convergence*, there exists $n_0 \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \frac{\epsilon}{3}$ for all $n \geq n_0$ and $x \in [0, 1]$. Now since f_n is continuous at x_0 there exists $\delta > 0$ such that $|f_n(x_0) - f_n(x)| < \frac{\epsilon}{3}$ for all $|x_0 - x| < \delta$. Therefore

$$|f(x_0) - f(x)| \leq \epsilon.$$

♣

Remark 28 *The converse of the above theorem is not true.*

Proposition 39 *Let $f(x) = \sum_{n\geq 1} f_n(x)$. Suppose there exists a sequence $\{M_n\}$ of non-negative reals such that $|f_n(x)| \leq M_n$ for all x and $\sum_{i=1}^{n} M_i$ converges, then $\sum_{i=1}^{n} f_i$ converges uniformly to f.*

Proof: In order to prove the result, it is sufficient to show that for every $\epsilon > 0$ there exists $n_0 \notin \mathbb{N}$ such that $|\sum_{n=n_0}^{n_0+m} f_n(x)| < \epsilon$ for all x.
Clearly,

$$|\sum_{n=n_0}^{n_0+m} f_n(x)| \leq \sum_{n=n_0}^{n_0+m} |f_n(x)| \leq \sum_{n=n_0}^{n_0+m} M_n \leq \sum_{n\geq n_0} M_n$$

As M_n converges, given $\epsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that $\sum_{n\geq n_0} M_n < \epsilon$.
This establishes uniform convergence of $\sum f_n$ to f. ♣

Corollary 7 ***(Weierstrass' M-test)***
For $f(x) = \sum_{n\geq 0} a_n x^n$, consider any interval $I \subset (-R, R)$ where R is the radius of convergence. Then $\sum_{n=0}^{N} a_n x^n$ converge uniformly to f on I.

Proof: Without loss of generality, consider I symmetric about 0. Fix $0 < \beta < R$ lying outside I. Let $f_n(x) = a_n x^n$, so $|f_n(x)| \leq |a_n|\beta^n$. Also $\sum |a_n|\beta^n$ converges. So by the previous theorem, we're done. ♣

3.4.1 The Metric Space $(C[0, 1], d)$

$C[0, 1]$ is the set of all continuous functions defined on $C[0, 1]$.
Define

$$d(f, g) := \sup_{x\in[0,1]} |f(x) - g(x)|$$

where $f, g \in C[0, 1]$. Here d is called the *sup-norm metric*.

Proposition 40 *$(C[0,1], d)$ is a metric space.*

Proof:

$$d(f,g) := \sup_{x\in[0,1]} |f(x) - g(x)|$$

where $f, g \in C[0,1]$. Clearly,

$$d(f,f) = 0$$

and,

$$d(f,g) = 0 \Leftrightarrow f(x) = g(x) \text{ for all } \ x \in [0,1]$$

Now,

$$d(f,h) = \sup_{x\in[0,1]} |f(x) - h(x)| = \sup_{x\in[0,1]} |f(x) - g(x) + g(x) - h(x)|.$$

Hence,

$$\begin{aligned} d(f,h) &\leq \sup_{x\in[0,1]} (|f(x) - g(x)| + |g(x) - h(x)|) \\ &\leq \sup_{x\in[0,1]} |f(x) - g(x)| + \sup_{x\in[0,1]} |g(x) - h(x)| \\ &= d(f,g) + d(g,h). \end{aligned}$$

This proves the result. ♣

Remark 29 *Define*

$$\| f \|_{\sup} := \sup_{x\in[0,1]} |f(x)|.$$

So the corresponding distance $d(f,g) = \| f - g \|_{\sup}$.

Remark 30 *Since $C[0,1]$ is a metric space, $f_n \rightrightarrows f \Leftrightarrow d(f_n, f) \to 0$ as $n \to \infty$.*

Theorem 49 *$C[0,1]$ is* ***complete*** *with respect to the sup-norm metric.*

Proof:
Suppose $\{f_n\}$ is Cauchy in $C[0,1]$ with respect to the sup-norm metric.
Given an $\epsilon > 0$, we have to find an n_0 such that $|f_n(x) - f(x)| \leq \epsilon$, for all $x \in [0,1]$, whenever $n \geq n_0$.
Since $\{f_n\}$ is Cauchy, we can find an n_0 such that

$$\sup_{x\in[0,1]} |f_n(x) - f_m(x)| < \frac{\epsilon}{2}$$

whenever $m, n \geq n_0$.

So, for each fixed $x \in [0,1]$, $\{f_n(x)\}$ is a real Cauchy sequence. Define

$$f(x) := \lim_{n\to\infty} f_n(x)$$

Now, for any $x \in [0,1]$, take M_x such that

$$|f_m(x) - f(x)| \leq \frac{\epsilon}{2}$$

for all $m \geq M_x$.
Now, for any $n \geq n_0$,

$$|f_n(x) - f(x)| \leq |f_n(x) - f_{M_x}(x)| + |f_{M_x}(x) - f(x)| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

♣

3.4.2 Theorems of Weierstrass

Theorem 50 ***(Weierstrass' Approximation Theorem)***
Suppose $f \in C[0,1]$ and $\epsilon > 0$. There is a polynomial $P \in C[0,1]$ such that $\| f - P \|_{\sup} < \epsilon$. In other words, if $\mathcal{P}$ denotes the set of polynomials in $C[0,1]$, then $\mathcal{P}$ is dense in $C[0,1]$.

Proof: *(Bernstein)*
$f \in C[0,1] \Rightarrow$ given $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $|x - y| < \delta$.
Let $\| f \|_{\sup} = M$. Consider

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

$$\begin{aligned}|f(x) - P(x)| &= |(x + 1 - x)^n f(x) - P(x)| \\ &= |\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} (f(x) - f(\frac{k}{n}))|\end{aligned}$$

Suppose X is a random variable $X \sim Bin(n, x)$. Then

$$\binom{n}{k} x^k (1-x)^{n-k} = P(X = k)$$

By *Chebychev's inequality,*

$$P(|X - nx| > t) \leq \frac{nx(1-x)}{t^2} \leq \frac{n}{4t^2}$$

Taking $t = n^{\frac{2}{3}}$,

$$P(|X - nx| > t) \leq \frac{1}{4n^{\frac{1}{3}}}$$

Call k as *deviant* if $|k - nx| > n^{\frac{2}{3}}$. So,

$$\sum_{deviant\ k} \binom{n}{k} x^k (1-x)^{n-k} < \frac{1}{4n^{\frac{1}{3}}}$$

$$\Rightarrow |\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} (f(x) - f(\frac{k}{n}))|$$

$$\leq \sum_{deviant\ k} \binom{n}{k} x^k (1-x)^{n-k} |(f(x) - f(\frac{k}{n}))| + \sum_{non-deviant\ k} \binom{n}{k} x^k (1-x)^{n-k} |(f(x) - f(\frac{k}{n}))|$$

Now,

$$\sum_{deviant\ k} \binom{n}{k} x^k (1-x)^{n-k} |(f(x) - f(\frac{k}{n}))| \leq \frac{2M}{4n^{\frac{1}{3}}}$$

and,

$$\sum_{non-deviant\ k} \binom{n}{k} x^k (1-x)^{n-k} |(f(x) - f(\frac{k}{n}))|$$

$$\leq \left(\sum_{non-deviant\ k} \binom{n}{k} x^k (1-x)^{n-k} \right) \left(\sup_{non-deviant\ k} |f(x) - f(\frac{k}{n})| \right)$$

$$< \sup_{non-deviant\ k} |f(x) - f(\frac{k}{n})|$$

Pick $\delta > 0$ such that $|f(x) - f(y)| < \frac{\epsilon}{2}$ whenever $|x - y| < \delta$ and pick n such that $n^{\frac{1}{3}} > \frac{2}{\delta}$ and $\frac{M}{2n^{\frac{1}{3}}} < \frac{\epsilon}{2}$.
This will guarantee that P is the required polynomial. ♣

Theorem 51 ***(Weirstrass)***
There exists $f : \mathbb{R} \to \mathbb{R}$ such that f is continuous but ***nowhere differentiable****.*

Proof: Let $f_0(x) : \mathbb{R} \to \mathbb{R}$

$$f_0(x) = \begin{cases} \{(x)\} & \text{on } [2n, 2n+1] \\ \{(1-x)\} & \text{on } [2n+1, 2n+2] \end{cases} \forall n \in \mathbb{Z}$$

and $f_k(x) = \alpha^k f_0(4^k x)$, for some $0 < \alpha < 1$.

Let $f(x) = \sum_{k \geq 0} f_k(x)$

Since $||f_k|| \leq \alpha^k$ and $\sum_{k \geq 0} \alpha^k < \infty$, by the M-test it follows that f is continuous.

Claim 16 *f is any NOT differentiable at any $x \in \mathbb{R}$.*

Observe : 'Slopes' appearing on $f_0 \in \{1, -1\}$.'Slopes' appearing on $f_k \in \{(4\alpha)^k, -(4\alpha)^k\}$. So we want $\alpha > \frac{1}{4}$ but $\alpha < 1$ (if f can be guaranteed to be non-differentiable).

It suffices to show that there is a sequence $\delta_n \to 0 (\delta_n \geq 0)$ such that

$$|\frac{f(x \pm \delta_n) - f(x)}{\delta_n}| \geq n$$

Also notice that $f_0(x+2) = f_0(x)$ for all x.

$$f_k(x + \frac{2}{4}k) = \alpha^k f_0(4^k(x + \frac{2}{4}k))$$

$$= \alpha^k f_0(4^k x + 2)$$

$$= \alpha^k f_0(4^k x)$$

$$= f_k(x)$$

Choose $\delta_n = \frac{1}{2(4^n)}$

$$\frac{f(x \pm \delta_n) - f(x)}{\delta_n} = \sum_{k \geq 0} \alpha^k \{\frac{f_0(4^k(x \pm \delta_n)) - f_0(4^k x)}{\delta_n}\}.$$

For $k > n$, coefficients are zero, therefore

$$\frac{f(x \pm \delta_n) - f(x)}{\delta_n} = \sum_{k \geq 0}^{n} \alpha^k \{\frac{f_0(4^k(x \pm \delta_n)) - f_0(4^k x)}{\delta_n}\}.$$

At $k = n$, $\alpha^k \{\frac{f_0(4^k x \pm \frac{1}{2}) - f_0(4^k x)}{\delta_n}\} = (4\alpha)^k$.

For $k < n$,

$$\alpha^k \{\frac{f_0(4^k(x \pm \frac{1}{2.4^n})) - f_0(4^k x)}{\frac{1}{2.4^n}}\} \leq (4\alpha)^k,$$

$$|\frac{f(x \pm \delta_n) - f(x)}{\delta_n}| = |\sum_{k<n} \frac{f_k(x \pm \delta_n) - f(x)}{\delta n} \pm (4\alpha)^n| \geq (4\alpha)^n - \sum_{0 \leq k < n} (4\alpha)^k.$$

Now,

$$\sum_{0}^{n-1} (4\alpha)^k = \frac{(4\alpha)^n - 1}{4\alpha - 1},$$

so

$$(4\alpha)^n - (\frac{(4\alpha)^n - 1}{4\alpha - 1}) > (\frac{4\alpha - 2}{4\alpha - 1}).(4\alpha)^n.$$

This shows that $f'(x) \to \infty$ as $n \to \infty$. As x was arbitrary, this shows that $f'(x)$ **does not exist** for any $x \in \mathbb{R}$. ♣

We denote $C[a, b]$ as the set of continuous functions in the interval $[a, b]$. The previous result in fact shows that there are members of $C[0, 1]$ that are nowhere differentiable in $[a, b]$.

One might wonder if the example of Weierstrass is a pathological extreme, and is not the norm. The next result however shows that far from being the norm, most members of $C[a, b]$ are in fact of this type, i.e., a 'generic' element of $C[a, b]$ is nowhere differentiable.

Theorem 52 *Consider $f \in C[0, 1]$. A 'generic' element of $C[0, 1]$ is nowhere differentiable.*

Consider for some $h_{n,x} > 0$,

$$R_n = \left\{ f \in C[0,1] \middle| \left|\frac{f(x \pm h_{n,x}) - f(x)}{h_{n,x}}\right| \geq n \quad \forall x \in [0, 1 - \frac{1}{n}] \right\}$$

Note that if f is differentiable at $x \in (0, 1)$, then $f \notin R_n$ for all $n \geq n_0$ for some $n_0 \in \mathbb{N}$. If $f \in \bigcap_{n \geq 1} R_n$, then f is nowhere differentiable (not differentiable at every $x \in [0, 1]$). We shall first prove a lemma.

Lemma 9 *Define R_n as,*

$$R_n := \{f \in C[0,1] \Big| \left|\frac{f(x \pm h_{n,x}) - f(x)}{h_{n,x}}\right| \geq n \ \forall x \in [0, 1 - \frac{1}{n}].$$

Then R_n is open and dense in $C[0, 1]$.

Note 3 *In general, in a metric space $(\mathcal{X}, d)$ if $U_n \subset \mathcal{X}$, are open in $\mathcal{X}$, then one could have $\bigcap_{n \geq 1} U_n = \emptyset$.*

The theorem follows as a consequence of the Baire Category theorem, which we shall do in the next section, and the lemma above. This incidentally also defines what we mean by 'generic'.

Proof: There are two things we need to prove:

1. R_n is open in $C[0, 1]$.
2. R_n is dense in $C[0, 1]$.

We will address both of these below.

1. We have to show that, given $f \in R_n$ we want $\epsilon > 0$ such that for all $g \in C[0,1]$ satisfying $\|f - g\|_{\sup} < \epsilon$ we also have $g \in R_n$.

   For each x, there is a $h_x = h$ such that

   $$|\frac{f(x \pm h) - f(x)}{h}| > n.$$

   The functions $\frac{f(x \pm h) - f(x)}{h}$ are continuous and have absolute value greater than n at x.

   By continuity, there is some small interval I_x containing x and a $\delta_x > 0$ such that $|\frac{f(y \pm h_x) - f(y)}{h_x}| \geq n + \delta_x$ for all $y \in I_x$ and corresponding h_x for an appropriate choice of $+$ or $-$.

   Consider

   $$\mathcal{U} = \left\{ I_x \mid x \in \left[0, 1 - \frac{1}{n}\right] \right\}.$$

   This clearly covers $\left[0, 1 - \frac{1}{n}\right]$. So by compactness of $\left[0, 1 - \frac{1}{n}\right]$, there is a finite sub-cover, i.e., there exist $x_1, x_2 \ldots\ldots x_r$ such that $\bigcup_{i=1}^r I_{x_i}$ covers $\left[0, 1 - \frac{1}{n}\right]$. In other words, for all $x \in \left[0, 1 - \frac{1}{n}\right]$, $|\frac{f(x \pm h) - f(x)}{h}| > n + \min\{\delta_1, \delta_2, \ldots \delta_r\}$

   and a suitable h for each x.

   In particular, if $\|g - f\| < \epsilon$, then

   $$|\frac{g(x+h) - g(x)}{h} - \frac{f(x+h) - f(x)}{h}| = |\frac{(g(x+h) - f(x+h)) - (g(x) - f(x))}{h}| \geq n.$$

   This completes the proof.

2. We have to show that, given $f \in C[0,1]$ and $\epsilon > 0$, we want a $g \in R_n$ and $\|g - f\| < \epsilon$. We shall think of $g = f + a$ where h satisfies $\|a\| < \epsilon$, where a is a suitable sawtooth function.

   $$\begin{aligned} |\frac{f(x+h) + a(x+h) - f(x) - a(x)}{h}| &> n \\ = |\frac{f(x+h) - f(x)}{h} &- \frac{a(x+h) - a(x)}{h}| \\ a(x) = \left(\frac{3}{4}\right)^k & f_o\left(4^k x\right), \end{aligned}$$

   for a large enough k.

   So slopes occurring in a are $\pm 3^k$.

   If f is piecewise linear, then f consists of finite number of straight line segments.

Let $M = \max(\text{slopes of line segments occurring in } f)$. Pick k large so that $3^k - M > n$ and we are through.

If $P \in C[0,1]$ is the set of all piecewise linear functions in $C[0,1]$, then $(R_n \bigcap P)$ is dense in P.

Hence, P is dense in $C[0,1]$ and therefore R_n is dense in $C[0,1]$. This completes the proof of the lemma.

♣

3.4.3 Baire Category Theorem

In this section, we state and prove the Baire Category theorem (in the context of a Complete Metric space) which completes the proof of the previous theorem. It states that in a complete metric space, dense open sets are in some sense, 'large'. This sense of largeness is not set-theoretic (i.e., in cardinality terms) but is something topological.

Theorem 53 ***(Baire Category Theorem)***
In a complete metric space $(\mathcal{X}, d)$, if $U_n \subset \mathcal{X}$ are open and dense, then $\bigcap_n U_n$ is dense in $\mathcal{X}$.

Proof: Let $x \in \mathcal{X}$ and let $\epsilon > 0$. U_1 is dense and open, so there exists $P_1 \in B_\epsilon(x)$ and $r_1 > 0$ such that

$$B_{r_1}(P_1) \subset B_\epsilon(x) \bigcap U_1.$$

U_2 is open and dense. So there exists P_2 and r_2 such that $B_{r_2}(p_2) \subset B_{r_1}(p_1) \bigcap U_2 \subseteq U_1 \bigcap U_2$.

Get $P_n \in \mathcal{X}, r_n \in \mathbb{R}$ such that $B_{r_n}(P_n) \subset B_{r_{n-1}}(P_{n-1}) \bigcap U_n$ for each $n \geq 2$.

Consider $\{P_n\}$. Note that $B_{r_n}(P_n)$ contains P_m for all $m \geq n$. r_n can be chosen as $r_n < \frac{r_{n-1}}{2}$. So $r_n < \frac{\epsilon}{2^{n-1}}$ if $m, k \geq n$, so $\{P_n\}$ is Cauchy. Since $\mathcal{X}$ is complete, $P_n \to p$ for some p.

Claim 17 *$p \in U_n$ for all $n \in \mathbb{N}$.*

Note that: $d(p,x) < 2\epsilon$. So, $d(p,x) \leq d(x,p_x) + d(p_x,p)$. Pick ℓ such that $d(p_\ell, p) < \frac{\epsilon}{4}$. $d(x, p_\ell) < \epsilon$ so $d(p,x) < 2\epsilon$.
We want to show that $p \in U_n$ for all n; that will complete the proof. To see for instance that $p \in U_1$, note that

$$P_n \in B_{r_2}(P_2) \subset U_1$$

for all $n \geq 2$. Hence, $p \in U_1$ since p is a limit point of P_n. Similarly we can show that $p \in U_n$ for all $n \in \mathbb{N}$. ♣

Remark 31 *As commented before, the last lemma of the previous section and the Baire category theorem tell us that functions that are nowhere differentiable are in fact* ***dense*** *in $C[0,1]$.*

Definition 53 *A set A is of* ***CATEGORY 1*** *if A can be written as countable union of "nowhere dense" sets.*

(A set A is nowhere dense if A does not contain any non-trivial open ball.)

Examples :

1. $\mathcal{X} = \mathbb{R}$ and A is any linear set.
2. $\mathcal{X} = \mathbb{R}$ $A = \mathbb{Z}$.
3. $\mathcal{X} = \mathbb{R}$ $A = \mathbb{Q}$.

Definition 54 *A set is said to be of* ***CATEGORY 2*** *if it is not CATEGORY 1.*

Proposition 41 *Suppose $\mathcal{X}$ is a metric space. The following are equivalent.*

- *$\mathcal{X}$ is CATEGORY 2.*
- *$U_n \subset \mathcal{X}$ are dense open $\Rightarrow \bigcap_n U_n$ is dense in $\mathcal{X}$.*

Examples:

1. Cantor set is of category 1.
2. Real line($\mathbb{R}$) is of category 2.

4 Riemann Integration

4.1 Integrals according to Riemann and Darboux

Definition 55 *A **Partition** P of $[a,b]$ is a set*

$$\{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$$

Definition 56 *For a bounded function f, let*

$$m_i = \inf_{x \in [x_{i-1}, x_i]} f(x)$$

*where $i = 1, 2, ..., n$. Then the **Lower Sum** denoted by $L(P,f)$ is given by*

$$L(P,f) := \sum_{i=1}^{n} m_i (x_i - x_{i-1}).$$

Definition 57 *For a bounded function f, let*

$$M_i = \sup_{x \in [x_{i-1}, x_i]} f(x)$$

*where $i = 1, 2, ..., n$. Then the **Upper Sum** denoted by $U(P,f)$ is given by*

$$U(P,f) := \sum_{i=1}^{n} M_i (x_i - x_{i-1}).$$

Definition 58 *A partition Q is a **refinement** of P if $Q \supseteq P$ and this is denoted by $Q \succeq P$.*

Observations:

1. If $Q \succeq P$,then $U(Q,f) \leq U(P,f)$.

   Proof: The case where $Q = P$ is trivial. Otherwise, Q contains some additional points to that of P. As they must belong to some interval of P,let us denote them as $y_i \in [x_{i-1}, x_i]$. Now

$$M_i = \sup_{[x_{i-1}, x_i]} f(x) \ \geq \ M_i^{(1)} = \sup_{[x_{i-1}, y_i]} f(x) \tag{4.1}$$

$$M_i = \sup_{[x_{i-1}, x_i]} f(x) \ \geq \ M_i^{(2)} = \sup_{[y_i, x_i]} f(x). \tag{4.2}$$

Hence,

$$M_i(x_i - x_{i-1}) = M_i(x_i - y_i) + M_i(y_i - x_{i-1}) \geq {M_i}^{(2)}(x_i - y_i) + {M_i}^{(1)}(y_i - x_{i-1}).$$

Therefore

$$U(Q, f) \leq U(P, f).$$

2. If $Q \succeq P$,then $L(Q, f) \geq L(P, f)$.

   Proof of this is similar to the above case except for M_i being replaced with m_i and the inequalities reversed.

Definition 59 *For a partition P,the **Norm** of P(written as $\mathcal{N}(P)$) equals*

$$\max\{x_i - x_{i-1} | 1 \leq i \leq n\}.$$

Definition 60 ***Darboux Integrability:*** *We say $f\colon\ [a, b] \to \mathbb{R}$ is (Darboux) integrable if there exists a real I such that the following holds: Given $\epsilon > 0$, there exists $\delta > 0$ such that for any partition P with $\mathcal{N}(P) < \delta$,we have*

$$|U(P, f) - I| < \epsilon \text{ and } |I - L(P, f)| < \epsilon.$$

Observe that f is Darboux integrable iff given $\epsilon > 0$, there exists δ >0 such that for every P with $\mathcal{N}(P) < \delta$, we have

$$U(P, f) < L(P, f) + \epsilon.$$

This is because both $U(P, f)$ and $L(P, f)$ tend to the same I as $\mathcal{N}(P)$ tends to zero.

Definition 61 ***Riemann Integrability:*** *Given P, let T be a subset of $[a, b]$ such that*

$$T = \{t_1 \leq t_2 \leq ... \leq t_n\} \text{ and } t_i \in [x_{i-1}, x_i]$$

Define

$$R(P, T, f) = \sum_{i=1}^{n} (x_i - x_{i-1}).f(t_i).$$

f is (Riemann) integrable if there exists I real such that given $\epsilon > 0$, there exists $\delta > 0$ such that for any P with $\mathcal{N}(P) < \delta$ and any T of P,

$$|R(P, T, f) - I| < \epsilon.$$

The above sum $R(P, T, f)$ is called Riemann sum of f w.r.t partition P.

Theorem 54 *Darboux integrability is equivalent to Riemann integrability.*

Proof: Darboux integrability $\Rightarrow$ Riemann integrability:

$m_i \leq f(t_i) \leq M_i$ for every $t_i \in [x_{i-1}, x_i]$.Therefore

$$\begin{aligned} m_i(x_i - x_{i-1}) &\leq f(t_i)(x_i - x_{i-1}) \leq M_i(x_i - x_{i-1}) && (4.3) \\ L(P, f) &\leq R(P, T, f) \leq U(P, f) && (4.4) \end{aligned}$$

By Darboux integrability, we mean that both $L(P, f), U(P, f)$ tend to I as $\mathcal{N}(P)$ tends to zero. From (3),(4), by the Sandwich Theorem, $R(P, T, f)$ tends to I which is equivalent to

$$|R(P, T, f) - I| < \epsilon.$$

Riemann integrability $\Rightarrow$ Darboux integrability:

Riemann integrability provides us a P such that $|R(P, T, f) - I| < \frac{\epsilon}{2}$ for every T of P. Choose $T_1 = \{t_{1_1} \leq t_{2_1} \leq ... \leq t_{n_1}\}$ such that for every $t_{i_1} \in [x_{i-1}, x_i]$,

$$f(t_{i_1}) + \frac{\epsilon}{2(b-a)} > \sup_{[x_{i-1}, x_i]} f(x).$$

Then

$$U(P, f) - R(P, T_1, f) < \frac{\epsilon}{2(b-a)} \sum_{i=1}^{n} (x_i - x_{i-1}) = \frac{\epsilon}{2}$$

Similarly,choose $T_2 = \{t_{1_2} \leq t_{2_2} \leq ... \leq t_{n_2}\}$ such that for every $t_{i_2} \in [x_{i-1}, x_i]$,

$$f(t_{i_2}) - \frac{\epsilon}{2(b-a)} < \inf_{[x_{i-1}, x_i]} f(x)$$

Then

$$R(P, T_2, f) - L(P, f) > \frac{\epsilon}{2}$$

Hence,by Triangle inequality, we have $|U(P, f) - I| < \epsilon$ and $|I - L(P, f)| < \epsilon$. ♣

Notation 1: We define the set $R[a, b]$ to denote the set of all Riemann integrable functions on $[a, b]$.

Proposition 42 *If f is continuous on $[a, b]$, then f is integrable on $[a, b]$, i.e., $f \in C[a, b] \Rightarrow f \in R[a, b]$.*

Proof: Given $\epsilon > 0$, we need to show that $U(P, f) - L(P, f) < \epsilon$ for $\mathcal{N}(P) < \delta$. Given that f is continuous on $[a, b]$, it follows that f is uniformly continuous on $[a, b]$, i.e., given $\epsilon > 0$, there exists $\delta > 0$ such that

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

Here, $\dfrac{\epsilon}{b-a}$ is also another positive number. Hence ϵ can be replaced by $\dfrac{\epsilon}{b-a}$. For any partition P with $\mathcal{N}(P) < \delta$, we have

$$\max_{x \in [x_{i-1}, x_i]} f(x) - \min_{x \in [x_{i-1}, x_i]} f(x) < \frac{\epsilon}{b-a}.$$

$$U(P,f) - L(P,f) = \sum_{i=1}^{n} (M_i - m_i)(x_i - x_{i-1}) < \frac{\epsilon}{b-a} \sum_{i=1}^{n} (x_i - x_{i-1}) = \frac{\epsilon}{b-a}(b-a) = \epsilon.$$

This completes the proof. ♣

Example 15

$$f(x) = \begin{cases} 1 & : x \in [0,1] \\ 0 & : x \in (1,2] \end{cases}$$

This is also Riemann integrable. To see this, note that for a partition P, any interval of P does not contain $x = 1$,we are through. Else, as the norm of P gets lower and lower, it essentially adds to zero and hence is Riemann integrable.

Example 16

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

M_i on any subinterval is 1 and m_i on any subinterval is 0. Hence $U(P,f)$ does not tend to

$L(P,f)$ for any partition. This is therefore not Riemann integrable.

What kind of discontinuities admit Riemann Integrability? Before answering this question,we'll look into the concept of **measure zero**.

4.2 Measure Zero sets and Riemann Integrability

Concept of Measure Zero: A set A is said to be of Measure Zero if given $\epsilon > 0$, there exists a sequence of open intervals $\{(a_i, b_i)\}_{i \geq 1}$ such that $\bigcup_{i \geq 1} (a_i, b_i)$ covers A and $\sum_i (b_i - a_i) < \epsilon$.

Examples:

1. Finite sets.
2. Countable sets.
3. Cantor set.

Some Useful Facts:

1. $[0, 1]$ is **not** a set of measure zero. In fact, any proper closed interval $[a, b]$ is **not** of measure zero.

2. Subset of a set of measure zero also has measure zero.

3. If $D_n (n \geq 1)$ are of measure zero, then $\bigcup_{n=1}^{\infty} D_n$ also has measure zero, i.e., countable union of measure zero sets is also a measure zero set.
   Proof: We give a proof of the last statement here. Let $\epsilon > 0$ be given. D_1 is of measure zero, so we have an open cover $\{(a_{1_i}, b_{1_i})\}$ such that

$$\sum_{i \geq 1} (b_{1_i} - a_{1_i}) < \frac{\epsilon}{2}.$$

   D_n is of measure zero, so we similarly have an open cover $\{(a_{n_i}, b_{n_i})\}$ such that

$$\sum_{i \geq 1} (b_{n_i} - a_{n_i}) < \frac{\epsilon}{2^n}.$$

   Then $\bigcup_{n=1}^{\infty} D_n$ has the open cover

$$\bigcup_{n=1}^{\infty} \bigcup_{i=1}^{\infty} (a_{n_i}, b_{n_i})$$

   and

$$\sum_{i \geq 1} (b_i - a_i) \leq \sum_{i \geq 1} \frac{\epsilon}{2^i} = \epsilon.$$

♣

Notation 2: If $f\colon [a, b] \to \mathbb{R}$ is Riemann integrable, then we denote corresponding real number I by $\displaystyle\int_a^b f(x)\,\mathrm{d}x$.

Theorem 55 ***(Riemann-Lebesgue Theorem:)*** *If f is bounded in $[a, b]$, $f \in R[a, b]$ iff the set of discontinuities of f is a set of measure zero.*

Proof:
Riemann integrability$\Rightarrow$ set of discontinuities is of measure zero

Suppose $f \in R[a, b]$. First observe that if f is a real valued function discontinuous at

x_0, then there exists $\epsilon > 0$ such that

$$\sup_{y\in(x-r,x+r)} f(y) - \inf_{y\in(x-r,x+r)} f(y) \geq \epsilon$$

for every $r \geq 0$, i.e.,

$$osc_f(x) := \lim_{r\to 0} \operatorname{diam}(f(x-r,x+r)) \geq \epsilon.$$

Fact: f is discontinuous at x if and only if $osc_f(x) > 0$.
Let $D(f)$ be the set of discontinuities of f. Then $D(f) = \bigcup\limits_{n\geq 1} D_f(n)$ where

$$D_f(n) = \{x : osc_f(x) \geq \frac{1}{n}\}.$$

So, if $f \in R[a,b]$, it suffices to show that for each $n \geq 1$, $D_f(n)$ has measure zero. As $f \in R[a,b]$, given $\epsilon > 0$, there exists $\delta > 0$ such that for P satisfying $\mathcal{N}(P) < \delta$, we have

$$U(P,f) - L(P,f) < \frac{\epsilon}{n}.$$

In particular, take the open intervals (x_i, x_{i+1}) determined by the partition P such that $D_f(n) \subset \bigcup\limits_i (x_i, x_{i+1})$; call i *BAD* if $(x_i, x_{i+1}) \cap D_f(n) \neq \emptyset$. Therefore,

$$\sum_i \frac{(x_{i+1}-x_i)}{n} \leq \sum_{BADi} (M_i - m_i)(x_{i+1}-x_i) \leq U(P,f) - L(P,f) < \frac{\epsilon}{n}$$

$$\sum_i (x_{i+1}-x_i) < \epsilon.$$

Hence, $D_f(n)$ is of measure zero. Hence $\bigcup\limits_{n\geq 1} D_f(n)$, a countable union of measure zero sets, is also of measure zero.

<u>Proof of Converse:</u>
Given D_f has measure zero. Note that if $x \in D_f \setminus D_f(n)$, then $osc_f(x) < \frac{1}{n}$. Consider $D_1 = \bigcup\limits_{i=1}^{n} D_f(i)$ for some large n (which will be determined later) and $D_2 = D_f \setminus D_1$.
Now $D_1 = \bigcup\limits_{i=1}^{n} D_f(i)$ is a set of measure zero. So, there's a collection of open intervals $\{(a_i,b_i)\}$ such that $D_1 \subseteq \bigcup\limits_i (a_i,b_i)$ and $\sum\limits_i (b_i - a_i) < \epsilon$ for any $\epsilon > 0$.
We want a $\delta > 0$ such that any P with $\mathcal{N}(P) < \delta$ satisfies $U(P,f) - L(P,f) < \epsilon$. Note

that for any x, and any interval $[\alpha, \beta]$ containing x,

$$(M_{[\alpha,\beta]} - m_{[\alpha,\beta]})(\beta - \alpha) \leq 2M(\beta - \alpha)$$

$M_{[\alpha,\beta]} = \sup_{x \in [\alpha,\beta]} f(x)$ and $m_{[\alpha,\beta]} = \inf_{x \in [\alpha,\beta]} f(x)$. Consider the set $D_f(n)$.It is it closed?.

Note 4 $D_f(1) \subseteq D_f(2) \subseteq \ldots$.

Since $D_f(n)$ is closed and bounded, it is compact(proof given below). As $D_f(n)$ has measure
zero, there exists $\{(a_i, b_i)\}$ such that $D_f(n) \subseteq \bigcup_i (a_i, b_i)$ and $\sum_i (b_i - a_i) < \epsilon$ for any $\epsilon > 0$.
Since $D_f(n)$ is compact, the open cover $\mathcal{U} = \{(a_i, b_i)\}$ admits a Lebesgue number $\delta > 0$, i.e., for any $x \in D_f(n)$, $(x - \delta, x + \delta) \subset (a_i, b_i)$ for some i.
We claim that as long as $\mathcal{N}(P) < \dfrac{\delta}{2}$, we are through. Let P be a partition with $\mathcal{N}(P) < \dfrac{\delta}{2}$.
Want $S = \sum_{i=1}^{N} (M_i - m_i)(x_{i+1} - x_i)$ to be small. We say i is BAD if $[x_i, x_{i+1}]$ contains an element of $D_f(n)$; else say i is $GOOD$.

$$S = \sum_{GOODi} (M_i - m_i)(x_{i+1} - x_i) + \sum_{BADi} (M_i - m_i)(x_{i+1} - x_i)$$

i is $GOOD$ $\Rightarrow$ for any $x \in (x_i, x_{i+1})$, $osc_f(x) < \dfrac{1}{n} \Rightarrow \sum_{GOODi} (M_i - m_i)(x_{i+1} - x_i) <$ $\sum_{i=1}^{n} \dfrac{(x_{i+1} - x_i)}{n} = \dfrac{(b - a)}{n}$.

$$\sum_{BADi} (M_i - m_i)(x_{i+1} - x_i) \leq 2M \sum_{BADi} (x_{i+1} - x_i)$$

The subcovers $\bigcup_{BADi} (x_i, x_{i+1})$ contains $D_f(n)$ and each of these is contained in one of the $(a_i, b_i)s$ of the cover $\mathcal{U}$. Therefore,

$$\sum_{BADi} (x_{i+1} - x_i) \leq \sum (b_i - a_i) < \epsilon$$

Hence $U(P, f) - L(P, f) \leq (2M + 1)\epsilon$, where n is chosen large enough such that $\dfrac{b - a}{n} < \epsilon$.
If we choose $\dfrac{\epsilon}{2M + 1}$ instead of ϵ since the start of the proof (this is allowed as $\dfrac{\epsilon}{2M + 1}$ is

also another positive number), we end up getting

$$U(P,f) - L(P,f) \leq \epsilon$$

♣

Proposition 43 *Let f have the domain $[a,b]$. Then for any n, $D_n(f) = \{x \in [a,b]$ such that $osc_f(x) \geq \frac{1}{n}\}$ is compact.*

Proof: Our claim is that ${D_n}^c$ is open in $[a,b]$. For every $x_0 \in {D_n}^c$,$osc_f(x_0) = t(\text{say}) < \frac{1}{n}$. So,

$$t = \lim_{h\to 0} \sup |f(x_1) - f(x_2)|$$

for $x_1, x_2 \in (x_0 - h, x_0 + h) \cap [a,b]$.Therefore, for $\epsilon = \dfrac{1-nt}{2n} > 0$ there exists $\delta > 0$ such that if $0 < h < \delta$, then,

$$|\sup |f(x_1) - f(x_2)| - t| < \epsilon$$

for all $x_1, x_2 \in (x_0 - h, x_0 + h) \cap [a,b]$.Therefore,

$$\sup |f(x_1) - f(x_2)| < t + \epsilon = \frac{t}{2} + \frac{1}{2n} < \frac{1}{2n} + \frac{1}{2n} < \frac{1}{n}$$

That is, $(x_0 - \delta, x_0 + \delta) \cap [a,b] \subset D_f(n)^c$. Hence, $D_f(n)^c$ is open in $[a,b]$ and hence $D_f(n)$ is closed.Also $D_f(n) \subset [a,b]$ is bounded. Hence, it is compact. ♣

4.3 Consequences of the Riemann-Lebesgue Theorem

1. If $f \in R[a,b] \bigcap R[b,c]$, then $f \in R[a,c]$ $(a < b < c)$.
2. $f, g \in R[a,b] \Rightarrow f \pm g \in R[a,b]$ and $fg \in R[a,b]$.
3. If $g > 0$ in $[a,b]$ and $g \in R[a,b]$, then $\frac{1}{g} \in R[a,b]$.

Proposition 44 *If $f, g \in R[a,b]$,then*

1. $\displaystyle\int_a^b (f+g)(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x$.
2. $\displaystyle\int_a^b \lambda f(x)\,\mathrm{d}x = \lambda \int_a^b f(x)\,\mathrm{d}x$.

3. $\int_a^b f(x)\,\mathrm{d}x \le (sup_{[a,b]}|f|)(b-a)$.

Proof: (1)$f \in R[a,b] \Rightarrow$ we can find a partition P_1 such that

$$U(P_1,f) - L(P_1,f) < \frac{\epsilon}{2}$$

Similarly, we can find another partition P_2 such that

$$U(P_2,g) - L(P_2,g) < \frac{\epsilon}{2}$$

Now on $P = P_1 \cup P_2 = \{a = x_0 < x_1 < ... < x_n = b\}$. $\sup_{x\in[x_{i-1},x_i]}(f+g) \le \sup_{x\in[x_{i-1},x_i]} f +$
$\sup_{x\in[x_{i-1},x_i]} g$ and $\inf_{x\in[x_{i-1},x_i]}(f+g) \ge \inf_{x\in[x_{i-1},x_i]} f + \inf_{x\in[x_{i-1},x_i]} g$. Therefore, $U(P,f+g) \le U(P,f) + U(P,g)$ and $-L(P,f+g) \le -L(P,f) - L(P,g)$.Now, $P_1, P_2 \subseteq P$.Hence, $U(P_1,f) \ge U(P,f)$ and $-L(P_1,f) \ge -L(P,f)$. Similar thing holds for P_2, g.

Now,

$$\begin{aligned}
U(P,f+g) - L(P,f+g) &\le U(P,f) + U(P,g) - L(P,f) - L(P,g)\\
&= U(P,f) - L(P,f) + U(P,g) - L(P,g)\\
&\le U(P_1,f) - L(P_1,f) + U(P_2,g) - L(P_2,g)\\
&< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

(2),(3) can be similarly proved. ♣

If $f \in R[a,b]$, then by Riemann-Lebesgue theorem, $[a,b] \bigcap D(f)$ is of measure zero $\Rightarrow$ for any $x \in (a,b)$, $[a,x] \bigcap D(f)$ has measure zero$\Rightarrow$ $f \in R[a,x]$ for every $x \in (a,b]$.
Consider $F(x) = \int_a^x f(x)\,\mathrm{d}x$; $F(a) = 0$ (denoted by $\int_a^a f(x)\,\mathrm{d}x = 0$ for any a.)

Theorem 56 *(a) $F\colon [a,b] \to \mathbb{R}$ is continuous.*
*(b) If f is continuous at x_0, then F is differentiable at x_0 and $F'(x_0) = f(x_0)$(**Fundamental Theorem of Calculus**).*

Proof: (a) We need to show that given $\epsilon > 0$, there exists a $\delta > 0$ such that $|x-y| < \delta \Rightarrow |F(x) - F(y)| < \epsilon$. Without loss of generality, let $a < x < y < b$.

Note 5 *If $f \in R[a,b], f \in R[b,c]$, then $f \in R[a,c]$ and*

$$\int_a^c f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^c f(x)\,\mathrm{d}x$$

From the above result,

$$F(x) - F(y) = \int_y^x f(t)\,\mathrm{d}t$$

$$\left|\int_y^x f(t)\,\mathrm{d}t\right| \le \sup_{t\in[x,y]} |f(t)|.(y-x) \le \sup_{t\in[a,b]} |f(t)|.(y-x) = M.(y-x)$$

for some $M > 0$. Therefore,

$$|F(y) - F(x)| \le M|y-x|$$

We can therefore choose $\delta = \dfrac{\epsilon}{M}$ and this completes the proof of (a).
Now, let $h > 0$.

$$\left|\frac{F(x_0+h)-F(x_0)}{h} - f(x_0)\right| = \left|\frac{1}{h}\int_{x_0}^{x_0+h} f(t)\,\mathrm{d}t - f(x_0)\right| = \left|\frac{1}{h}\int_{x_0}^{x_0+h} (f(t)-f(x_0))\,\mathrm{d}t\right|$$

Since f is continuous at x_0, given $\epsilon > 0$, there exists $\delta > 0$ such that for every $t \in (x_0, x_0+\delta)$, $|f(t) - f(x_0)| < \epsilon$.

$$\left|\int_a^b f(t)\,\mathrm{d}t\right| \le \int_a^b |f(t)|\,\mathrm{d}t$$

. Hence, for $h < \delta$

$$\left|\frac{F(x_0+h)-F(x_0)}{h} - f(x_0)\right| \le \frac{1}{h}\int_{x_0}^{x_0+h} |f(t)-f(x_0)|\,\mathrm{d}t < \frac{\epsilon.h}{h} = \epsilon$$

This proves (b). ♣

Note 6 $f'(x) = 0$ *for every* $x \Rightarrow f(x) = constant$*. So, in particular,*

$$f(x) = sin^2x + cos^2x \Rightarrow f'(x) = 2\sin x\cos x + 2\cos x(-\sin x) = 0$$

Therefore $f(x)$*=constant. Plugging* $x=0$*,we get* $f(x) = 1$*.*

4.4 Antiderivatives and some 'well known' Integral Calculus techniques

Definition 62 *Given a real valued function f, a real valued function F is an* ***Antiderivative*** *for f if F is differentiable and $F' \equiv f$.*
If F is antiderivative for f, then so is $G = F + c$ where c is any constant.

Proposition 45 *Any two antiderivatives differ by a constant.*

Proof: If $F' = G' = f$, then $(F - G)' = 0$. Hence by note 3, $F - G = constant$. ♣

Theorem 57 ***(Integration by Parts)*** *Suppose f, g are differentiable on (a, b) and continuous on $[a, b]$ and suppose $f', g' \in R[a, b]$, then*

$$\int_a^b f'g(t)\,\mathrm{d}t = fg|_a^b - \int_a^b fg'(t)\,\mathrm{d}t.$$

Proof: By consequences of Riemann-Lebesgue theorem, $f'g$, $fg' \in R[a, b]$. By the Product rule of differenciation,

$$(fg)' = f'g + fg' \tag{4.5}$$

Hence, $(fg)' \in R[a, b]$. Pick antiderivatives $f(x)g(x) - (constant)$, $\int_a^x f'g(t)\,\mathrm{d}t$, $\int_a^x fg'(t)\,\mathrm{d}t$ respectively of the three functions in (5) and substitute b for x. Constant can be evaluated

by substituting a in x and this yields $c = f(a)g(a)$. This completes the proof. ♣

Theorem 58 ***(Integration by Substitution)*** *If $f \in R[\alpha, \beta]$, $g' > 0$ and continuous for*

every $x \in [\alpha, \beta]$ and $g : [a, b] \longrightarrow [\alpha, \beta]$ bijectively, then

$$\int_a^b f(g(t))g'(t)\,\mathrm{d}t = \int_\alpha^\beta f(t)\,\mathrm{d}t$$

Proof: By M.V.T, there exist $t_i s \in [x_{i-1}, x_i]$ such that

$$g(x_i) - g(x_{i-1}) = g'(t_i)(x_i - x_{i-1})$$

Pick these t_is and form the corresponding Riemann sum

$$R_1(P, T, f) = \sum_{i=1}^{n} f(g(t_i))g'(t_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} f(g(t_i))(g(x_i) - g(x_{i-1}))$$

on $[a, b]$. Let $y_i = g(x_i)$ and $s_i = g(t_i)$. Note that as $g'(x) > 0$, $\alpha = y_0 < y_1 < .... < y_n = \beta$ and $y_{i-1} < s_i < y_i$. Now, the sum will simplify to $R_1 = \sum_{i=1}^{n} f(s_i)(y_i - y_{i-1})$ which is the same as the one on the other hand, i.e;

$$\int_\alpha^\beta f(t)\,\mathrm{d}t = \sum_{i=1}^{n} f(s_i)(y_i - y_{i-1}) = R_1 = \int_a^b f(g(t))g'(t)\,\mathrm{d}t.$$

♣

Some facts about $\sin x$ **and** $\cos x$**:** We know some of the properties of the trigonometric functions (such as the sin of a sum and so forth); here, we look at some of them, from their formal definitions, and see how we might go about proving them. Recall that

1. $\cos x = \sum_{n\geq 0} \frac{(-1)^n x^{2n}}{(2n)!}$.
   $\sin x = \sum_{n\geq 0} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$.

Here we consider some well-known properties.

1. Consider $f(x) = \sin^2 x + \cos^2 x$. By differentiating a peer series term-by-term, we know that $(\sin x)' = \cos x, (\cos x)' = -\sin x$. Hence it follows that $f'(x) = 0$. By the fundamental theorem of calculus, it follows that $f(x) = Const$. At $x = 0, \sin x = 0$ and $\cos x = 1$, hence this constant is 1. In particular, $\sin^2 x + \cos^2 x = 1$ for all x, and $\cos x$ attains its maximum value 1 at $x = 0$.

2. $\sin(-x) = -\sin x$ (substituting $-x$ in power series).

3. $\cos 1 < \sin 1$.
   To see this,
   $$\sin 1 = 1 - \frac{1}{3!} + \frac{1}{5!} - \frac{1}{7!} + ...$$
   $$\cos 1 = 1 - \frac{1}{2!} + \frac{1}{4!} - \frac{1}{6!} + ...$$
   $$\sin 1 - \cos 1 = \frac{1}{2!} - \frac{1}{3!} - \frac{1}{4!} + \frac{1}{5!} + \frac{1}{6!} - \frac{1}{7!} - \frac{1}{8!} + ... = \left(\frac{1}{2!} - \frac{1}{3!}\right) - \left(\frac{1}{4!} - \frac{1}{5!}\right) + ...$$
   $$\sin 1 - \cos 1 = \left(\frac{2}{3!} - \frac{4}{5!}\right) + \left(\frac{6}{7!} - \frac{8}{9!}\right) + ...$$
   $$\frac{2n}{(2n+1)!} - \frac{2n+2}{(2n+3)!} = \frac{(2n+2)(4n^2+6n-1)}{(2n+3)!} > 0.$$
   Hence,
   $$\sin 1 > \cos 1.$$
   Define $f(x) = \cos x - \sin x . f(0) = 1 > 0$ and $f(1) < 0$. Therefore by Intermediate Value Property, $f(x) = 0$ for some $x \in (0, 1)$ Hence, there is some x (denoted by $\frac{\pi}{4}$) where $\sin x = \cos x$ for the first time in $(0, \infty)$.

4. Consider the following equations which will be proved later.
   $$\sin(x \pm a) = \sin x \cos a \pm \cos x \sin a \tag{4.6}$$
   $$\cos(x \pm a) = \cos x \cos a \mp \sin x \sin a. \tag{4.7}$$

a) If $x = a$, then first equation gives us $\sin 2x = 2\sin x \cos x$. Plugging $x = \dfrac{\pi}{4}$ gives
$$\sin\left(\frac{\pi}{2}\right) = 2\sin\left(\frac{\pi}{4}\right)\cos\left(\frac{\pi}{4}\right) = 2\sin^2\left(\frac{\pi}{4}\right).$$
But
$$\sin^2\left(\frac{\pi}{4}\right) + \cos^2\left(\frac{\pi}{4}\right) = 2\sin^2\left(\frac{\pi}{4}\right) = 1.$$
hence $\sin^2\left(\frac{\pi}{4}\right) = \frac{1}{2}$. Therefore, $\sin\left(\frac{\pi}{2}\right) = 1$.

b) $\cos^2\left(\frac{\pi}{2}\right) = 1 - \sin^2\left(\frac{\pi}{2}\right) = 0$, so $\cos\left(\frac{\pi}{2}\right) = 0$.

c) $\sin\left(x + \frac{\pi}{2}\right) = \cos x$ (from eq.n(6))
$\cos\left(x + \frac{\pi}{2}\right) = -\sin x$.(from eq.n(7))

d) $\sin(x + \pi) = \sin\left(x + \frac{\pi}{2} + \frac{\pi}{2}\right) = -\sin x$

e) $\sin(x+2\pi) = \sin(x+\pi+\pi) = (-1)(-1)\sin x = \sin x$. We can similarly obtain $\cos(x + 2\pi) = \cos x$. Hence we conclude that $\sin x$ and $\cos x$ are periodic.

5. To prove that $\sin(x + a) = \sin x \cos a + \cos x \sin a$, for instance, consider $f(x) = \sin(x+a) - \sin x \cos a - \cos x \sin a$. Then $f'(x) = \cos(x+a) - \cos x \cos a + \sin x \sin a$ and $f''(x) = -\sin(x + a) + \sin x \cos a + \cos x \sin a$. Hence $f''(x) = -f(x)$.

   Note that plugging $x = 0$ gives $f'(0) = f(0) = 0$. Also, $f(x)$ is analytic on $\mathbb{R}$ as $\sin x, \cos x$ are analytic in $\mathbb{R}$ as discussed in the previous chapter. Hence, we can differentiate it further and obtain the relation $f^{(2n)}(x) = (-1)^n f(x)$.Therefore, $f^{(2n)}(0) = (-1)^n f(0) = 0$. Also, $f^{(2n+1)}(x) = (-1)^n f'(x)$. Hence, $f^{(2n+1)}(0) = (-1)^n f'(0) = 0$. Hence $f(x) = \sum_{n\geq 1} \frac{f^{(n)}(0)x^n}{n!} = 0$. Therefore,
$$\sin(x + a) = \sin x. \cos a + \cos x. \sin a.$$
   By substituting $-a$ for a in the above result we get $\sin(x-a) = \sin x \cos a - \cos x \sin a$ and differentiating these two results gives $\cos(x \pm a) = \cos x. \cos a \mp \sin x \sin a$.

Some facts about e^x and the $\log$ function:

1. We have seen that $e^x = \sum_{n\geq 0} \frac{x^n}{n!}$ and $(e^x)' = e^x$.Define $f(x) = e^x.e^{-x}$. Hence
$$f'(x) = (e^x.e^{-x})' = e^x.e^{-x} + e^x.(e^{-x}(-1)) = 0,$$
   so by the fundamental theorem of calculus,
$$f(x) = c.$$

But $f(0)$ =1. Hence $f(x) \equiv 1$ for all x. Hence

$$e^{-x} = \frac{1}{e^x}$$

For $x > 0$, it is easy to see that $e^x > 0$ (from expansion). Hence $e^{-x} > 0$. This proves that e^x is positive throughout and is increasing (Derivative is itself which is positive$\Rightarrow$ increasing throughout) and hence f has an inverse. We'll denote it by $y = \log x$. In particular, we have $e^y = x$. Since the inverse of a differentiable function is also differentiable, we have (by results of the preceding chapter), the following: If $y_0 = f(x_0)$,then

$$(f^{-1}(y_0))' = \frac{1}{f'(x_0)}.$$

In particular, $\log x$ is differentiable and

$$(\log x)' = \frac{1}{e^{\log x}} = \frac{1}{x}.$$

2. $e^{x+y} = e^x e^y$.
   This will be proved after the following proposition.

Proposition 46 *Suppose g is analytic on $\mathbb{R}$ and $g'(x) = g(x)$ for every x, then $g(x) = ce^x$ for some constant c.*

Proof: For every x, we have the following two equations.

$$g(x) = \sum_{n\geq 0} \frac{g^{(n)}(0)x^n}{n!} \tag{4.8}$$

$$g'(x) = g(x) \tag{4.9}$$

$g'(0) = g(0) = c$ (say). $g''(x) = g'(x) = g(x)$. Hence $g''(0) = c$ as well. Proceeding further by differentiating the equation $g'(x) = g(x)$, by induction, we obtain that $g^{(n)}(0) = c$ for every n. Hence, substituting $g^{(n)}(0)$ back in eq.n(8),we get the desired form, i.e.,

$$g(x) = ce^x.$$

♣

Now, consider $f(x) = e^{x+a} - e^x.e^a$. $f'(x) = e^{x+a} - e^x.e^a = f(x)$ for every $x \in \mathbb{R}$ and $f(x)$ is analytic throughout $\mathbb{R}$ as discussed in topic 3. Hence by above proposition, $f(x) = ce^x$. Substituting $x = 0$ in this equation, $f(0) = c$. But $f(0) = e^a - e^a(1) = 0$. Hence $c = 0$ and so $f(x) \equiv 0$.

4.5 The Integral Test

Notation: By $\displaystyle\int_a^\infty f(x)\,\mathrm{d}x$, we mean $\displaystyle\lim_{N\to\infty}\int_a^N f(x)\,\mathrm{d}x$ provided the limit exists and is finite.

Theorem 59 ***(Integral Test:)*** *Suppose $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ is monotonically decreasing and continuous, then $\displaystyle\sum_{n\geq k} f(n)$ converges or diverges accordingly as $\displaystyle\int_k^\infty f(x)\,\mathrm{d}x$ exists or not respectively.*

Proof: Note that since $F(t) = \displaystyle\int_{t_0}^t f(t)\,\mathrm{d}t$ is differentiable, using M.V.T,

$$\begin{aligned} F(t) &= F(t) - F(t_0) \\ &= F'(\xi)(t - t_0) \\ &= f(\xi)(t - t_0) \end{aligned}$$

for some $\xi \in (t_0, t)$. Now,

$$\begin{aligned} \int_N^{N+1} f(x)\,\mathrm{d}x &= f(\xi)(N + 1 - N) \\ &= f(\xi) \end{aligned}$$

for some $\xi \in (N, N + 1)$. But as f is monotonically decreasing,

$$f(N) > f(\xi) > f(N + 1) \text{ for every } \xi \in (N, N + 1).$$

Therefore,

$$f(N) \geq \int_N^{N+1} f(x)\,\mathrm{d}x \geq f(N + 1).$$

Hence we have

$$\begin{aligned} \sum_{N=1}^K f(N) &\geq \sum_{N=1}^K \int_N^{N+1} f(x)\,\mathrm{d}x \geq \sum_{N=1}^K f(N + 1) \\ \sum_{N=1}^K f(N) &\geq \int_1^{K+1} f(x)\,\mathrm{d}x \geq \sum_{N=1}^K f(N + 1) \\ \sum_{N=1}^K f(N) &\geq \int_1^{K+1} f(x)\,\mathrm{d}x \geq \sum_{N=2}^{K+1} f(N) \end{aligned}$$

for every $K \in \mathbb{N}$. As $K \to \infty$, if the integral diverges, then so does the series, as $\sum_{N=1}^{K} f(N) \geq \int_1^{K+1} f(x)\,\mathrm{d}x$. Similarly, if the integral converges, then so does the series as $\int_1^{K+1} f(x)\,\mathrm{d}x \geq \sum_{N=2}^{K+1} f(N)$. ♣

Note 7 *In the convergent case,* $\sum_{N=1}^{\infty} f(N) = f(1) + \sum_{N=2}^{\infty} f(N)$. *Addition of a constant doesn't disturb the convergence of the sequence.*

Example 17 *Let f(x)=*$\frac{1}{x^\alpha}$. *When does* $\sum_{n\geq 1} \frac{1}{n^\alpha}$ *converge?*

We have seen in the previous chapter that it diverges for $\alpha = 1$ *and converges for* $\alpha = 2$. *Now let* $\alpha > 1$. *The function*

$$f(x) = \frac{1}{x^\alpha}$$

is continuous and monotonically decreasing $(x > 0)$. *Hence the Integral test is applicable:*

$$\int_1^\infty \frac{1}{x^\alpha}\,\mathrm{d}x = \int_1^\infty x^{-\alpha}\,\mathrm{d}x = \left.\frac{x^{-\alpha+1}}{-\alpha+1}\right|_1^\infty.$$

The integral above converges when $-\alpha + 1 < 0$, *i.e., when* $\alpha > 1$.

4.6 Weak version of Stirling's formula for $n!$

Theorem 60 ***Stirling's Approximation:*** $n! \approx \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}}$.

We will prove that a weaker version which states $n! \approx Ce^{-n}n^{n+\frac{1}{2}}$. The approximation is more precisely stated as follows:

Theorem 61

$$\lim_{n\to\infty} \frac{n!}{e^{-n}n^{n+\frac{1}{2}}} = C.$$

for some constant C.

Proof: Consider the following equation(for $a, b > 0$) whose proof is given below:

$$\log ab = \log a + \log b.$$

Therefore,

$$\log(n!) = \sum_{k=1}^{n} \log k.$$

Since $\log x$ is monotonically increasing function, from eq.n(10) of the proof for the Integral test,

$$\log n \leq \int_n^{n+1} \log x \, \mathrm{d}x \leq \log n + 1.$$

Hence we obtain

$$\sum_{k=1}^{n} \log k \leq \int_1^{n+1} \log x \, \mathrm{d}x \leq \sum_{k=2}^{n+1} \log k.$$

Now,

$$\int_1^{n+1} \log x \, \mathrm{d}x = (x \log x - x)|_a^b = (n+1)\log(n+1) - n,$$

which gives us

$$\log n! \leq (n+1)\log(n+1) - n \leq \log(n+1)!,$$
$$(n+1)\log(n+1) - n \leq \log n! + \log(n+1)$$

which implies

$$n \log(n+1) - n \leq \log n!$$
$$n \log n - n < n \log(n+1) - n \leq \log n! \leq (n+1)\log(n+1) - n$$

Therefore

$$n \log n - n < \log n! \leq (n+1)\log(n+1) - n.$$

Vaguely speaking,

$$\log n! \approx (n + \xi) \log n - n$$

for some $\xi \in (0,1)$. Let us examine how $\xi = \dfrac{1}{2}$ works.

Consider $y_n = \log n! - \left(n + \dfrac{1}{2}\right) \log n + n$

$$y_n - y_{n+1} = \left(n + \frac{1}{2}\right) \log\left(1 + \frac{1}{n}\right) - 1 \tag{4.10}$$

Using (the following eqns will be analyzed later)

$$\begin{aligned} \log(1+x) &= \frac{x}{1} - \frac{x^2}{2} + \frac{x^3}{3} - .... \\ \log(1-x) &= -\left(\frac{x}{1} + \frac{x^2}{2} + \frac{x^3}{3} + ...\right) \end{aligned}$$

simplifies eqn.(11) to

$$y_n - y_{n+1} = \frac{1}{3(2n+1)^2} + \frac{1}{5(2n+1)^4} + ... > 0 \tag{4.11}$$

$$0 < y_n - y_{n+1} < \frac{1}{3(2n+1)^2}\left\{1 + \frac{1}{(2n+1)^2} + ...\right\} = \frac{1}{3\{(2n+1)^2 - 1\}} = \frac{1}{12}\left(\frac{1}{n} - \frac{1}{n+1}\right).$$

Consequently,

$$y_n - \frac{1}{12n} < y_{n+1} - \frac{1}{12(n+1)} \tag{4.12}$$

y_n is decreasing(from eqn.(12)) and $y_n - \frac{1}{12n}$ is increasing(from eqn.(13)). Now assume that $y_n - \frac{1}{12n}$ is not bounded. Then, for every $M > 0$, there exists an n_M such that $y_n - \frac{1}{12n} > M$ for all $n \geq n_M$. Hence, $y_n > M + \frac{1}{12n} > M$ for all $n \geq n_M$. Hence, for all $M > 0$, we have $y_n > M$ for all $n \geq n_M$.This shows that y_n has no upper bound. But y_n is a decreasing sequence. Hence $y_n \leq y_1$ for all n. This is a contradiction. Therefore $\lim_{n\to\infty} y_n$ exists. Let the limit be c.

$$\log n! - \left(n + \frac{1}{2}\right)\log n + n \approx c,$$

$$\frac{n!e^n}{n^{n+\frac{1}{2}}} \approx e^c.$$

As e^c is also a constant, we'll replace it with C. Hence

$$n! \approx Cn^{n+\frac{1}{2}}e^{-n}$$

♣

Proposition 47 $\log a + \log b = \log ab$.

Proof: Let $\log a = x$ and $\log b = y$. Then by definition, $a = e^x$ and $b = e^y$. Hence,

$$a.b = e^x e^y = e^{x+y}$$

as seen earlier. Therefore,

$$\log ab = \log(e^{x+y}) = x + y = \log a + \log b.$$

♣

4.7 Convergence of sequences of functions and Integrals

The following questions have been implicitly made during some of the calculations of the preceding section: **Questions:**

1. Why and for what values of x is $\log(1+x) = \frac{x}{1} - \frac{x^2}{2} + \frac{x^3}{3} - ....?$
   More generally, if
$$f(x) = \sum_{n=01}^{\infty} a_n x^n$$
   is an analytic function defined on a closed interval $[a, b]$, then can we integrate term-by-term? The previous questions (with the log function) can be easily answered if such were the case. We can take the expansion of $\frac{1}{1+x}$ and integrate. Thus the relevant question really is:

   If we have a power series $f(x) = \sum_{n\geq 0} a_n x^n$ in some interval (-R,R), is it true that for $-R < a < b < R$, $\int_a^b f(x)\,\mathrm{d}x = \sum_{n\geq 0} a_n \int_a^b x^n\,\mathrm{d}x$?

2. Above question can more generally be posed. Suppose $\{f_n\}$ is a sequence where $f_n \in R[a, b]$ and suppose $f(x) = \lim_{n\to\infty} f_n(x)$ for every $x \in [a, b]$. Then when is
$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x?$$

3. Further more basically, if $f_n \in R[a, b]$ and $f = \lim_{n\to\infty} f_n$,does $f \in R[a, b]$?

If f_n has discontinuity set D_n, then f_n is continuous at all $x \in [a, b] \setminus D_n$. Consider $[a, b] \setminus \bigcup_{n\geq 1} D_n$. By the Riemann-Lebesgue theorem, the sets D_n all have measure zero, so $\bigcup_{n\geq 1} D_n$ has measure zero. So if f is continuous at all $x \in [a, b] \setminus \bigcup_{n\geq 1} D_n$, then we could conclude that $f \in R[a, b]$. Now, each f_n is continuous at every element of $[a, b] \setminus \bigcup_{n\geq 1} D_n$. If f_ns uniformly converge to f, then we know that f is also continuous at $x \in [a, b] \setminus \bigcup_{n\geq 1} D_n$.

Indeed,
$$\left|\int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f_n(x) - f(x)|\,\mathrm{d}x \leq ||f_n - f||_{\text{sup}}(b-a) \longrightarrow 0,$$
as f_n uniformly converges to f. This (partially) answers the above question and proves the following theorem.

Theorem 62 *If f_n uniformly converge to f and $f_n \in R[a, b]$,then $f \in R[a, b]$ and*
$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x.$$

Remark 32 *Without Uniform convergence, we may have a problem. We can have* f_n *that tend to zero(thereby* $\int_a^b f(x)\,\mathrm{d}x = 0$*) but* $\int_a^b f_n(x)\,\mathrm{d}x = 1$ *for every n. Indeed the following example shows that such a possibility is eminent.*

Example 18 *Define a sequence as follows.*

$$f_n(x) = \begin{cases} 4n^2x & : x \in \left[0, \dfrac{1}{2n}\right] \\ 4n - 4n^2x & : x \in \left(\dfrac{1}{2n}, \dfrac{1}{n}\right] \\ 0 & : x > \dfrac{1}{n} \end{cases}$$

For any $x > 0$*, there exists* $N \in \mathbb{N}$ *such that* $\frac{1}{n} < x$ *for every* $n \geq N$*. Further* $f_n(0) = 0$ *for every* n*. Hence* $\lim_{n\to\infty} f_n(x) \equiv 0$ *for every* $x > 0$*. But each* f_n *encloses a region of area* 1 *with* x*-axis.*

5 Measures of sets and a peek into Lebesgue Integration

The idea of Lebesgue Integration as an alternate viewpoint to the theory of Riemann Integration is motivated by the following analogy. Suppose we have a pile of coins and we wish to count the total money. One way we can do it is go over all the coins one by one, and add the denomination of the coin in consideration, to the running total. Another way would be to count the number of coins of each denomination, and sum these over all the denominations. The first way of doing it is analogous to the idea of Riemann integrals; we shall now pursue the other way of calculating 'integrals'. This is the spirit of Lebesgue integration.

5.1 Measure for subsets of $\mathbb{R}$

In order to make this point more meaningful, suppose we have a function f defined on a closed interval $[a, b]$ and we wish to give an alternate perspective on the idea of calculating the area under the curve f. To make the earlier analogy more relevant, we pick each $c \in \text{Image}(f)$ and consider the sets $f^{-1}(c)$. The idea now is to be able to 'measure' the content of this set $f^{-1}(c)$ in order to crystallize a meaningful version of the earlier analogy. So that is our primary question: How do we define a measure of an arbitrary subset of $\mathbb{R}$?

Firstly, we make a list of some of the properties, this measure ought to possess.

Desirable properties of MEASURE: We would like to determine a function $\mu : \mathcal{P}(\mathbb{R}) \to [0, \infty]$ $(:= [0, \infty) \cup \{\infty\})$ which satisfies the following properties.

1. $\mu(\phi) = 0$.

2. $\mu([a, b]) = b - a$. This comes from the intent to have compatibility with the corresponding Riemann Integral. Indeed, the function $f(x) = 1$ on the interval $[a, b]$ has Riemann integral equal $b - a$, and this new notion of the integral needs to agree with the Riemann integral.

3. $A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$.

4. $\mu\left(\overline{A} \cap [a,b]\right) = b - a - \mu(A \cap [a,b])$.
   This can be generalized as: If $A_1, A_2, ...$ are pairwise disjoint, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Remark 33 *If $\{A_n\}$ is increasing sequence of sets, then*

$$\begin{aligned}
\mu\left(\bigcup_{n\geq 1} A_n\right) &= \mu(A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup .....(A_n \setminus A_{n-1})....) \\
&= \mu(A_1) + \mu(A_2 \setminus A_1) + \mu(A_3 \setminus A_2) + ... \\
&= \mu(A_1) + (\mu(A_2) - \mu(A_1)) + (\mu(A_3) - \mu(A_2)) \\
&= \lim_{n\to\infty} \mu(A_n)
\end{aligned}$$

In particular, take $A_n = \left[a + \frac{1}{n}, b - \frac{1}{n}\right]$. By the above remark, we get $\mu((a,b)) = b - a$.

Now we will examine a candidate for μ.

Definition 63 *The **outer measure** of a set $A \subseteq \mathbb{R}$ as*

$$\mu^*(A) = \inf\left(\sum_{n\geq 1} |I_n|\right) \text{ where } I_n s \text{ are open intervals covering } A, \text{ i.e., } A \subseteq \bigcup_{n\geq 1} I_n$$

where $|I|$=length of the open interval I.

Proposition 48 *1.* $\mu^*(\phi) = 0$.

2. $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$.

3. $\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$.

Note 8 *The last desirable property for a measure was for pairwise disjoint sets. But, in the above proposition, the last point corresponds to any sets and this property is called **sub-additivity**.*

Proof: Part 1 is trivial. For part 2, note the fact that any open cover for B will also cover A. Hence the infimum of sum of lengths of intervals for A will be less than that of B. For part 3, let $\epsilon > 0$. For each A_n, consider the cover $\{I_{n_k}\}_{k\geq 1}$ such that

$$\sum_{k\geq 1} |I_{n_k}| \leq \frac{\epsilon}{2^{n+1}} + \mu^*(A_n).$$

Now $\bigcup_n \bigcup_k I_{n_k}$ covers $\bigcup_n A_n$, hence

$$\sum_n \sum_k |I_{n_k}| \leq \sum_n \mu^*(A_n) + \epsilon.$$

Now, as ϵ is an arbitrary positive number,

$$\inf_{n,k} \left(\sum_n \sum_k |I_{n_k}| \right) \leq \sum_n \mu^*(A_n).$$

Hence, μ^* is sub-additive. ♣

When does the equality hold in the last step of previous proof? As per the desired property, it must hold for pairwise disjoint sets. But does μ^* accomplish this? Unfortunately, the answer is no! It turns out, that an arbitrary subset of $\mathbb{R}$ can be a lot more 'funky' than we imagine. This is explained in the next proposition formally.

Proposition 49 *μ^* is **not** countably additive.*

Before looking into the proof of this proposition, we will need the Axiom of Choice.

Axiom of Choice: Given $\{A_\alpha\}_{\alpha\in\Lambda}$, where Λ is an arbitrary set,

$$\prod_{\alpha\in\Lambda} A_\alpha \neq \phi$$

More explicitly, it states that for every indexed family $(S_i)_{i\in I}$ of non-empty sets there exists an indexed family $(x_i)_{i\in I}$ of elements such that $x_i \in S_i$ for every $i \in I$.

Proof: Let us first define an equivalence relation on $\mathbb{R}$ as follows

$$x \sim y \text{ iff } x - y \in \mathbb{Q}\,.$$

Consider the equivalence classes of this relation. Let E be the set that contains exactly one element from each equivalence class. The validity of such a set is given by the Axiom of Choice. Further, as all real numbers have rational translates in $[0, 1]$ we'll pick an element from each equivalence class that lies in $[0, 1]$.

Define $E + r := \{e + r | e \in E\}$ where $r \in \mathbb{Q}$. Also note that if $r \neq s$, then $(E + r) \cap (E + s) = \phi$. This is because if there is a common element e, then $e - r \in E$ and $e - s \in E$. But $e - r$, $e - s$ are rational translates Therefore, they must belong to same equivalence class which contradicts our construction of E that it must contain only one element from each class.

Now, define $E^* = \bigcup\limits_{r\in\mathbb{Q}\cap[-1,1]} E + r$. Now, $E \subseteq [0, 1]$. E^* is a translation of E to right by a maximum of 1 and to left by again a maximum of 1 unit. Hence, $E^* \subseteq [-1, 2]$. Now, consider an element of E. any other element of its equivalence class in $[0, 1]$ is a rational translate by at most 1 on either side. As E^* includes all such translates, $E^* \supseteq [0, 1]$.

We can consolidate the above discussion as

$$[0, 1] \subseteq E^* \subseteq [-1, 2].$$

As E^* is a countable union of pairwise disjoint sets, if μ^* is countably additive,

$$\mu^*(E^*) = \sum_{r\in\mathbb{Q}\cap[-1,1]} \mu^*(E + r).$$

But, if $\{I_n\}$ covers $E + r$, then $\{I_n + s - r\}$ covers $E + s$. This happens for all such open covers and the lengths are not altered. Hence, $\mu^*(E + r) = \mu^*(E + s)$.

If each of $\mu^*(E + r) = 0$, then $\mu^*(E^*) = 0$ (countable summation of zeroes) and we run into a contradiction as $E^* \supseteq [0, 1]$. If each of $\mu^*(E + r)$ is finite nonzero value, then

$\mu^*(E^*)$ is a countably infinite sum of terms each of which is a nonzero constant, and this gives a contradiction since we must have $\mu^*(E^*) \le 3$ as $E^* \subseteq [-1, 2]$.

Hence μ^* isn't countably additive. ♣

Remark 34 *μ^* is not even finitely additive. To see this, in the above example, take E^* to be the union of n translations where n is such that $n > 3/\mu^*(E)$. Here, if each of the n measures is zero, we have $\mu^*(E^*) = 0$ which is a contradiction. Else, although $\mu^*(E^*)$ doesn't shoot up to ∞, it exceeds 3 which is again a contradiction.*

Remark 35 *In the above construction of a pathological example, we only used one of the important properties of μ^* that it is translation invariant. But that is a desired property for any μ because of the intent to make it compatible with the corresponding Riemann Integral. Hence, there's no μ satisfying all desired properties of a measure! This looks like our exercise of defining a suitable 'measure' for al subsets of $\mathbb{R}$ has been one in vain.*

As a remedy, we seek to define μ only on a (suitably large) subset of $\mathcal{P}(\mathbb{R})$ and this collection should not include any sets like the previous E^*. Recall that our choice of E didn't satisfy the equation

$$\mu^*(E) + \mu^*([0, 1] \setminus E) = \mu^*([0, 1]) = 1$$

This is the motivation to define measurable sets and to distinguish them from sets like E^*.

Definition 64 *A set $A \subseteq \mathbb{R}$ is **good/splitter/measurable** if for any $X \subseteq \mathbb{R}$*

$$\mu^*(X) = \mu^*(X \cap A) + \mu^*(X \cap \overline{A})$$

μ^* is referred to as the Lebesgue outer measure.

Observations:

1. $\emptyset$ is measurable.
2. A is good, then $\overline{A}$ is good.

Proposition 50 *If A_1, A_2 are good, so is $A_1 \cup A_2$.*

Proof: We want to show: For every set X, $\mu^*(X) \ge \mu^*(X \cap A_1 \cap A_2) + \mu^*(X \cap (\overline{A_1 \cap A_2}))$, as we already have other inequality by sub-additivity property of μ^*
Since, A_1 and A_2 are measurable,

$$\mu^*(X) = \mu^*(X \cap A_1) + \mu^*(X \cap (\overline{A_1})) \tag{5.1}$$

$$\mu^*(X) = \mu^*(X \cap A_2) + \mu^*(X \cap (\overline{A_2})) \tag{5.2}$$

$$X \cap (A_1 \cup A_2) = (X \cap A_1) \cup (X \cap \overline{A_1} \cap A_2) \tag{5.3}$$

Using $X \cap \overline{A}_1$ in place of X in (2) gives

$$\mu^*(X \cap \overline{A}_1) = \mu^*(X \cap \overline{A}_1 \cap A_2) + \mu^*(X \cap \overline{A}_1 \cap \overline{A}_2) = \mu^*(X \cap \overline{A}_1 \cap A_2) + \mu^*(X \cap (\overline{A_1 \cup A_2})).$$

From (3) we have :
$\mu^*(X \cap (A_1 \cup A_2)) = \mu^*((X \cap A_1)) \cup \mu^*((X \cap \overline{A_1} \cap A_2))$
Adding $\mu^*(X \cap (\overline{A_1 \cup A_2}))$ on both sides gives:

$$\mu^*(X \cap (A_1 \cup A_2)) + \mu^*(X \cap (\overline{A_1 \cup A_2})) \leq \mu^*((X \cap A_1)) + \mu^*((X \cap \overline{A_1} \cap A_2)) + \mu^*(X \cap \overline{A_1 \cup A_2})$$

Due to sub-additivity of μ^*, we have

$$\mu^*((X \cap \overline{A_1} \cap A_2)) + \mu^*(X \cap (\overline{A_1 \cup A_2})) \leq \mu^*(X \cap \overline{A_1})$$

Summarily we have,

$$\begin{aligned} \mu^*(X \cap (A_1 \cup A_2)) + \mu^*(X \cap \overline{A_1 \cup A_2}) &\leq \mu^*((X \cap A_1)) + \mu^*((X \cap \overline{A_1} \cap A_2)) + \mu^*(X \cap (\overline{A_1 \cup A_2})) \\ &\leq \mu^*((X \cap A_1)) + \mu^*(X \cap \overline{A_1}) \\ &= \mu^*(X), \end{aligned}$$

(from (1)) proving the required inequality. ♣

Corollary 8 *If A_1, A_2 are good, then so are $A_1 \setminus A_2$, $A_1 \cap A_2$.*

Proof: A_1, A_2 are good, therefore, $\overline{A_1}, \overline{A_2}$ are good. Therefore, by the previous proposition, $\overline{A_1} \cup \overline{A_2}$ is good. But, this implies

$$A_1 \cap A_2 = \overline{(\overline{A_1} \cup \overline{A_2})}$$

is good. Now, $A_1 \backslash A_2 = A_1 \cap \overline{A_2}$. So, we are done since if A_1, A_2 are good, then A_1, $\overline{A_2}$ are good and therefore, $A_1 \cap \overline{A_2}$ is good. ♣

Corollary 9 *If $A_1, A_2, ..., A_n$ are good, then so are $\bigcup_{i=1}^{n} A_i$, $\bigcap_{i=1}^{n} A_i$.*

Proof: Follows by induction and the above corollary. ♣

Proposition 51 *If A_1, A_2 are good, and $A_1 \cap A_2 = \emptyset$, then*

$$\mu^*(A_1 \cup A_2) = \mu^*(A_1) + \mu^*(A_2)$$

Proof: Let $X = A_1 \cup A_2$

Using the fact that A_1 is good, we get

$$\begin{aligned}\mu^*(X) &= \mu^*(X \cap A_1) + \mu^*(X \cap \overline{A_1}) \\ &= \mu^*(A_1) + \mu^*(A_2).\end{aligned}$$

♣

Corollary 10 *If $A_1, A_2, ..., A_n$ are good and pairwise disjoint, then*

$$\mu^*\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu^*(A_i)$$

Proof: This follows by induction again. ♣

Proposition 52 *$\{A_i\}_{i=1}^{\infty}$ are measurable, then $\bigcup_{i=1}^{\infty} A_i$ is also measurable.*

Proof: Without loss of generality, A_i are pairwise disjoint; else, consider

$$\begin{aligned}A_1' &= A_1 \\ A_2' &= A_2 \backslash A_1 \\ &\vdots\end{aligned}$$

Let $A = \bigcup_{i=1}^{\infty} A_i$.

By the sub-additivity of μ^*, we already have,

$$\mu^*(X) \leq \mu^*(X \cap A) + \mu^*(X \cap \overline{A})$$

It suffices to show: For any X,

$$\mu^*(X) \geq \mu^*(X \cap A) + \mu^*(X \cap \overline{A})$$

For $A_1, A_2, ..., A_n$, $\bigcup_{i=1}^{n} A_i$ is good. Therefore,

$$\begin{aligned}\mu^*(X) &= \mu^*\left(X \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) + \mu^*\left(X \cap \left(\overline{\bigcup_{i=1}^{n} A_i}\right)\right) \\ &\geq \mu^*\left(X \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) + \mu^*\left(X \cap \overline{A}\right)\end{aligned}$$

Hence, from the previous proposition, we get,

$$\mu^*(X) \geq \sum_{i=1}^{n} \mu^*(X \cap A_i) + \mu^*\left(X \cap \overline{A}\right) \quad \text{for all } n \in \mathbb{N}.$$

Taking the limit as n tends to infinity, we get

$$\mu^*(X) \geq \mu^*(X \cap A) + \mu^*(X \cap \overline{A}).$$

♣

We can consolidate the above discussion as the statement of the following theorem. By $\mathcal{M}^*$, we shall denote to denote the set of all measurable sets.

Theorem 63 *The set of measurable sets* $\mathcal{M}^* \subsetneq \mathcal{P}(\mathbb{R})$ *(w.r.t.* μ^**) satisfies*

(i) $\emptyset \in \mathcal{M}^*$.

(ii) $A \in \mathcal{M}^* \Rightarrow \overline{A} \in \mathcal{M}^*$.

(iii) $A_i \in \mathcal{M}^*$ for every $i \in \mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} A_i$ and $\bigcap_{i=1}^{\infty} A_i \in \mathcal{M}^*$.

Proposition 53 *If* $\{A_i\}_{i\geq 1}$ *are good and pairwise disjoint, then*

$$\mu^*(\bigcup_i A_i) = \sum_i \mu^*(A_i)$$

Proof: Let $X = \bigcup_{i=1}^{\infty} A_i$. By the previous corollary, $\bigcup_{i=1}^{n} A_i$ is measurable for all natural n. Therefore,

$$\begin{aligned}
\mu^*(X) &= \mu^*\left(X \cap (A_1 \cup A_2 \cup ... \cup A_n)\right) + \mu^*\left(X \cap \overline{(A_1 \cup A_2 \cup ... \cup A_n)}\right) \\
&\geq \mu^*\left(X \cap (A_1 \cup A_2 \cup ... \cup A_n)\right) \\
&= \mu^*\left((X \cap A_1) \cup (X \cap A_2) ... \cup (X \cap A_n)\right) \\
&= \mu^*\left(A_1 \cup A_2 \cup ... \cup A_n\right) \\
&= \sum_{i=1}^{n} \mu^*(A_i).
\end{aligned}$$

Therefore,

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \geq \sum_{i=1}^{n} \mu^*(A_i) \quad \text{for all } n \in \mathbb{N}.$$

Taking the limit as n tends to infinity on the right hand side, we get

$$\mu^*\left(\bigcup_{i=1}^{\infty} A_i\right) \geq \sum_{i=1}^{\infty} \mu^*(A_i)$$

But, since the outer measure is sub-additive, we get the equality. ♣

Proposition 54 *If A_n are pairwise disjoint and measurable, then for any X,*

$$\mu^*\left(X \cap \left(\bigcup_{i=1}^{\infty} A_i\right)\right) = \sum_{i=1}^{\infty} \mu^*(X \cap A_i).$$

The previous proposition is the special case where $X = \bigcup_{i=1}^{\infty} A_i$.

Proof: It suffices to prove: If A_1 and A_2 are measurable, disjoint, then

$$\mu^*(X \cap (A_1 \cup A_2)) = \mu^*(X \cap A_1) + \mu^*(X \cap A_2).$$

This will yield finite additivity (induction) and passing on to the countable case is similar to the previous proposition. Since, A_1 is measurable, therefore,

$$\mu^*(X \cap (A_1 \cup A_2)) = \mu^*((X \cap (A_1 \cup A_2)) \cap A_1) + \mu^*((X \cap (A_1 \cup A_2)) \cap \overline{A_1}).$$

But, since A_1 and A_2 are disjoint, in the above equation, the first term is equal to $\mu^*(X \cap A_1)$ and the second term is equal to $\mu^*(X \cap A_2)$. ♣

Proposition 55 *If $\mathcal{Z}$ is a set of measure zero, then $\mathcal{Z}$ is measurable.*

Proof: Want to show: $\mu^*(X) = \mu^*(X \cap \mathcal{Z}) + \mu^*(X \backslash \mathcal{Z})$

By sub-additivity of μ^*, we have

$$\mu^*(X) \leq \mu^*(X \cap \mathcal{Z}) + \mu^*(X \backslash \mathcal{Z})$$

Since, $X \cap \mathcal{Z} \subseteq \mathcal{Z}$, $\mu^*(X \cap \mathcal{Z}) = 0$.

This implies $\mu^*(X) \leq \mu^*(X \backslash \mathcal{Z})$

Also,

$$X \backslash \mathcal{Z} \subseteq X \Rightarrow \mu^*(X \backslash \mathcal{Z}) \leq \mu^*(X)$$

Hence, $\mu^*(X) = \mu^*(X \setminus \mathcal{Z})$. As $\mu^*(X \cap \mathcal{Z}) = 0$, we are through. ♣

Proposition 56 $\mu^*((a,b)) = b - a$.

Proof: Since, (a, b) is an open cover for itself, therefore

$$\mu^*((a,b)) \leq b - a.$$

We now want to show: $\mu^*((a,b)) \geq b - a$.

We will instead show $\mu^*([a,b]) \geq b - a$ and then prove that $\mu^*((a,b)) = \mu^*([a,b])$.

$\{I_n\}_{n=1}^{\infty}$ is an open cover for $[a,b]$, then by compactness of $[a,b]$, there is a finite subcover, say $I_1, I_2, ..., I_n$. W.L.O.G., each I_i is a finite interval.

Consider $f_i(x) : \mathbb{R} \to \mathbb{R}$

$$\begin{aligned} f_i(x) &= 1 \text{ if } x \in I_i \\ &= 0 \text{ otherwise, for all } i = 1, 2, ..., n \\ f(x) &= 1 \text{ if } \ x \in [a,b] \\ &= 0 \text{ if } \ x \notin [a,b] \end{aligned}$$

Since f_i has only 2 points of discontinuity, $f_i \in R[-M, M]$ for a suitably large M. Similar is the case for f. Hence,

$$\int_{-M}^{M} f_i(x)dx = |I_i|$$

Since, $\{I_i\}_{i=1}^{n}$ cover $[a,b]$,

$$\sum_{i=1}^{n} f_i(x) \geq f(x) \ \text{ for all } x \in [-M, M].$$

Integrating both sides from $-M$ to M, we get

$$\sum_{i=1}^{n} |I_i| = \int_{-M}^{M} (\sum_{i=1}^{n} f_i) \geq \int_{-M}^{M} f(x)dx = b - a.$$

Therefore,

$$\mu^*([a,b]) \geq b - a.$$

Since, $(a,b) \subset [a,b]$,

$$\mu^*\left((a,b)\right) \leq \mu^*\left([a,b]\right)$$

Also, $[a,b] = (a,b) \cup \{a\} \cup \{b\}$, by sub-additivity we have,

$$\mu^*\left([a,b]\right) \leq \mu^*\left((a,b)\right)$$

using the fact that $\{a\}, \{b\}$ are measure zero sets. Hence, we get

$$\mu^*\left([a,b]\right) = \mu^*\left((a,b)\right) = b - a.$$

♣

Theorem 64

$$(a, \infty) \in \mathcal{M}^* \textit{ for all } a \in \mathbb{R}.$$

The following corollary is an immediate consequence.

Corollary 11 *Every open set and closed set in $\mathbb{R}$ is contained in $\mathcal{M}^*$.*

Proof: Want to show: $\mu^*(X) \geq \mu^*(X \cap (a, \infty)) + \mu^*(X \cap (-\infty, a])$
Let $\{I_n\}$ be a collection of open intervals covering X. Want to show,

$$\sum_n |I_n| \geq \mu^*(X \cap (a, \infty)) + \mu^*(X \cap (-\infty, a]).$$

W.L.O.G, assume $a = 0$.

We start with an observation: If $\mathcal{Z}$ is any set of measure zero,

$$\mu^*(X \backslash \mathcal{Z}) = \mu^*(X)$$

We shall assume, without loss of generality that $0 \notin X$ and $0 \notin I_n$ for any n.
Let $\mathfrak{I}^- = \{n \mid I_n \subseteq (-\infty, 0) \text{ and } X \cap I_n \neq \emptyset\}$.
and $\mathfrak{I}^+ = \{n \mid I_n \subseteq (0, \infty) \text{ and } X \cap I_n \neq \emptyset\}$.
In other words, if we write,

$$X^+ = X \cap (0, \infty),$$
$$X^- = X \cap (-\infty, 0),$$

we then have $X^+ \cup X^- = X$.
Since, $\{I_n\}_{n \in \mathfrak{I}^+}$ and $\{I_n\}_{n \in \mathfrak{I}^-}$ are open covers for X^+ and X^- respectively, we have

$$\sum_{n \in \mathfrak{I}^+} |I_n| \geq \mu^*(X^+)$$

$$\sum_{n \in \mathfrak{I}^-} |I_n| \geq \mu^*(X^-)$$

Therefore,

$$\begin{aligned} \sum_n |I_n| &= \sum_{n \in \mathfrak{I}^+} |I_n| + \sum_{n \in \mathfrak{I}^-} |I_n| \\ &\geq \mu^*(X^+) + \mu^*(X^-) \end{aligned}$$

Now, taking the infimum over all such open covers, we get

$$\mu^*(X) \geq \mu^*(X^+) + \mu^*(X^-).$$

♣

5.2 Sigma Algebras and the Borel Sigma Field

Definition 65 *A σ-algebra of $\mathbb{R}$ is a collection $\mathcal{A}$ satisfying*

(i) $\emptyset \in \mathcal{A}$.

(ii) $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$.

(iii) $\{A_i\}_{i=1}^{\infty} \subseteq \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

What we have shown is

(i) $\mathcal{M}^*$ is a σ-algebra.

(ii) $\mathcal{M}^*$ contains all open subsets of $\mathbb{R}$.

Proposition 57 *Suppose $\mathcal{A}_\alpha$ are σ-algebras where $\alpha \in \Lambda$ for some indexing set Λ, then $\bigcap_{\alpha \in \Lambda} \mathcal{A}_\alpha$ is also a σ-algebra.*

The proof is trivial, which we skip. The more important consequence is the following. Given any collection of sets $\mathcal{U}$ we may consider the *smallest* sigma algebra containing the members of $\mathcal{U}$.

This makes sense because, $\mathcal{U} \subset \mathcal{P}(\mathbb{R})$ and the latter is a sigma algebra. In particular, the collection of all σ-algebras containing all the members of $\mathcal{U}$ is non-empty. Hence by the preceding proposition, if we index by Λ the collection of all σ-algebras containing $\mathcal{U}$ then the intersection of all these σ-algebras is the *smallest* sigma algebra containing $\mathcal{U}$ since every σ-algebra containing all the members of $\mathcal{U}$ must necessarily contain the above intersection.

Definition 66 *$\mathcal{M}$ is the smallest σ-algebra containing all open sets.*

In the rest of our discussions, we shall restrict our attention only to the sigma field $\mathcal{M}$, which is also referred to as the **Borel Sigma Field**.

5.3 Lebesgue Integration

Definition 67 *For measurable sets $E_1, E_2, ..., E_n$ and constants $c_1, c_2, ..., c_n$ (≥ 0) satisfying $E_i \cap E_j = \phi$ for $i \neq j$, the function*

$$s(x) = \sum_{i=1}^{n} c_i 1_{E_i}(x) \ \forall \ x \in \mathbb{R},$$

where

$$
\begin{aligned}
1_E(x) &= 1 \text{ } \textit{if } x \in E \\
&= 0 \text{ } \textit{otherwise,}
\end{aligned}
$$

*is called a positive **Simple** function.*

Consider the notation:

$$
\begin{aligned}
\infty + \infty &= \infty \\
\infty . \infty &= \infty \\
\infty . 0 &= 0
\end{aligned}
$$

Define the Lebesgue Integral for a simple positive function,

$$\int s := \sum_{i=1}^{n} c_i \mu(E_i)$$

Example 19 *Suppose*

$$
\begin{aligned}
f(x) &= 1 \text{ } \textit{if } x \in [0,1] \cap (\mathbb{R} \setminus \mathbb{Q}) \\
&= 0 \text{ } \textit{otherwise.}
\end{aligned}
$$

f is simple and $\int f = \mu([0,1] \cap (\mathbb{R} \setminus \mathbb{Q})) \cdot 1 = 1$.

Recall the set $E \subseteq [0,1]$ obtained by picking a representative from each equivalence class of $\mathbb{R} \setminus \mathbb{Q}$:

$$
\begin{aligned}
f(x) &= 1 \text{ if } x \in E \\
&= 0 \text{ if } x \notin E
\end{aligned}
$$

Such functions cannot be assigned reasonable definition for a Lebesgue integral. This motivates the following:

Definition 68 *A function $f : \mathbb{R} \to \mathbb{R}$ is said to be **measurable** if $f^{-1}(E)$ is measurable if E is measurable.*

Suppose $E \in \mathcal{M}$. The following question is the most natural follow-up to the the definition from above. Is $s = 1_E$ measurable?

Ans: Yes. Consider $F \in \mathcal{M}$.

$$f^{-1}(F) := \{x \in \mathbb{R} \mid f(x) \in F\}.$$

1_E is measurable, since there are four possible choices for $f^{-1}(F)$, namely $\phi, E, \overline{E}, \mathbb{R}$, depending on

$$(i)\ 0,1 \notin F \qquad (ii)\ 1 \in F, 0 \notin F \qquad (iii)\ 0 \in F, 1 \notin F \qquad (iv)\ 0,1 \in F$$

respectively.

Similarly, we can prove simple functions are measurable by inducting on n. We skip the details.

The next proposition tells us that a'good' function (continuous) is measurable.

Proposition 58 *If f is continuous then it is measurable.*

Proof: Let $\mathcal{A} = \{E \in \mathcal{M} \mid f^{-1}(E) \text{ is measurable}\}$. We shall show that $\mathcal{A} = \mathcal{M}$.

Firstly, note that if E is open then $E \in \mathcal{A}$. This follows from the definition of continuity, since the inverse image of an open set (under a continuous function) is open, and open sets are members of $\mathcal{M}$. Thus, it suffices to show that $\mathcal{A}$ is a σ-algebra.

To see why, note that $\mathcal{M}$ is the smallest σ-algebra containing all open sets. So, if $\mathcal{A}$ is a σ-algebra and $\mathcal{A}$ contains all the open sets, then $\mathcal{A}$ must contain $\mathcal{M}$. But since, $\mathcal{A}$ is a subset of $\mathcal{M}$ by definition, we are through.

To check that $\mathcal{A}$ is a σ-algebra involves the following checks:

(i) $\emptyset \in \mathcal{A}$ This follows, since $f^{-1}(\phi) = \phi$ is measurable.

(ii) $A \in \mathcal{A} \Rightarrow f^{-1}(A) \in \mathcal{M} \Rightarrow \overline{f^{-1}(A)} \in \mathcal{M}$.

But, since $f^{-1}(\overline{A}) = \overline{f^{-1}(A)}$, this implies, $\overline{A} \in \mathcal{A}$. Thus, if $A \in \mathcal{A}$ then $\overline{A} \in \mathcal{A}$ as well.

(iii) $A_i \in \mathcal{A} \Rightarrow f^{-1}(A_i) \in \mathcal{M} \Rightarrow \displaystyle\bigcup_{i=1}^{\infty} f^{-1}(A) \in \mathcal{M}$

But, since $\displaystyle\bigcup_{i=1}^{\infty} f^{-1}(A) = f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right)$, we get

$$\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}.$$

♣

Note 9 *The $\mathcal{A}$ defined above is a σ-algebra regardless of whether f is a continuous function or not. To show that it contains all the open sets is the only part that requires the definition of continuity of f*

Note 10 *The proof above uses very useful and simple set-theoretic identities:*

(i) $f^{-1}(\overline{A}) = \overline{f^{-1}(A)}$

(ii) $\displaystyle\bigcup_{i=1}^{\infty} f^{-1}(A) = f^{-1}\left(\bigcup_{i=1}^{\infty} A_i\right)$

Definition 69 *Suppose we have a sequence* $\{f_n\}$, $f_n : \mathbb{R} \to \mathbb{R}$, *sup*$\{f_n\}$ *and limsup*$\{f_n\}$ *are the functions defined as,*

$$\sup\{f_n\}(x) := \sup_n\{f_n(x)\}$$
$$\limsup\{f_n\}(x) := \inf_m \ \sup_n\{f_{m+n}(x)\}$$

Proposition 59 *If* $\{f_n\}$ *are all measurable, then so are* $\sup\{f_n\}$ *and* $\limsup\{f_n\}$*.*

Proof:

$$\mathcal{A} = \{E \in \mathcal{M} \mid f^{-1}(E) \in \mathcal{M}\}$$

We will show that $\mathcal{A}$ is a σ-algebra and that $\mathcal{A}$ contains all the open sets.

(i) Let $f := \sup\{f_n\}$. Also,

$$f^{-1}(A) = \{x \mid f(x) \in A\}$$
$$f(x) \in A \text{ if and only if } \sup_n f_n(x) \in A$$

In particular, if $A = (a, \infty)$

$$f(x) > a \text{ if and only if } \sup_n f_n(x) > a$$

which implies,

$$x \in \bigcup_{n=1}^{\infty} f_n^{-1}((a, \infty)) = \bigcup_{n=1}^{\infty} f_n^{-1}(A).$$

So, we get, $f^{-1}(A)$ is measurable or equivalently,

$$(a, \infty) \in \mathcal{A} \ \forall \ a \in \mathbb{R}$$

We have proved earlier, that $\mathcal{A} = \{E \in \mathcal{M} \mid f^{-1}(E) \in \mathcal{M}\}$ is always a σ-algebra. Hence we are through.

(ii) Let $f = \limsup_n f_n = \inf_m\{\sup_n f_{n+m}\}$

By the same kind of reasoning, one can show that $\inf_n\{f_n\}$ is measurable if f_n are all measurable.

The only difference in the proof as in the previous case, is that this time we use $A = (-\infty, a)$ and show that $(-\infty, a) \in \mathcal{A}$ as defined there.

So, $\limsup f_n$ is also measurable if f_n's are all measurable. ♣

Definition 70 *Suppose f is a positive valued measurable function.*

$$\int f := \sup\left\{\int s \ \mid \ s(x) \leq f(x) \ \forall \ x \text{ and } s \text{ is simple}\right\}$$

Suppose, $f = \sum_{i=1}^{n} a_i 1_{E_i}$, we had defined earlier, then

$$\int f = \sum_{i=1}^{n} a_i \mu(E_i)$$

It is clear that if we use the definition of $\displaystyle\int f$, then $\displaystyle\int f \geq \sum_{i=1}^{n} a_i \mu(E_i)$.
This is because f can itself be used as a simple function underestimating the integral of f.

Suppose if possible that

$$\int f := \sup \left\{ \int s \ \mid \ 0 \leq s(x) \leq f(x) \ \forall \ x \ and \ s \ is \ simple \right\} > \sum_i a_i \mu(E_i)$$

Then in particular, there exists a simple function s, such that

(i) $0 \leq s(x) \leq f(x) \ \forall \ x$ *and,*

(ii) $\int s > \sum_i a_i \mu(E_i)$

In other words, we have the following question:

If s, t are simple functions and

$$0 \leq s(x) \leq t(x) \ \forall \ x$$

where

$$s(x) = \sum_{i=1}^{n} a_i 1_{E_i}(x)$$

$$t(x) = \sum_{i=1}^{m} b_i 1_{F_i}(x)$$

W.L.O.G., $m = n$ in the equations above.

Then, can it happen that

$$\sum_{i=1}^{n} a_i \mu(E_i) > \sum_{i=1}^{n} b_i \mu(F_i)$$

Answer: No.

Proof: We want to prove: If s, t are simple functions and

$$0 \leq s(x) \leq t(x) \ \forall \ x$$

then,

$$\int s \leq \int t$$

We prove by induction on n. For $n = 1$,

$$s(x) = a1_E(x)$$

$$t(x) = b1_F(x)$$

with $a, b > 0$ and $E, F \neq \phi$

Observe, if $x \in \overline{F}$, then

$$\begin{aligned} t(x) = 0 \Rightarrow s(x) = 0 \Rightarrow x \in \overline{E} \\ \Rightarrow \overline{F} \subseteq \overline{E} \\ \Rightarrow E \subseteq F \\ \Rightarrow \mu(E) \leq \mu(F) \end{aligned}$$

Also,

$$s(x) = a \Rightarrow x \in E \Rightarrow x \in F \Rightarrow t(x) = b$$

But, since $s(x) \leq t(x)\ \forall\ x$, implies

$$0 < a \leq b$$

Combining the above two inequalities, we get

$$a\mu(E) \leq b\mu(F)$$

or,

$$\int s \leq \int t$$

Let the hypothesis be true for $n = k$.

For n=k+1,

$$s = \sum_{i=1}^{k+1} a_i 1_{E_i} = \sum_{i=1}^{k} a_i 1_{E_i} + a_{k+1} 1_{E_{k+1}}$$

$$t = \sum_{i=1}^{k+1} b_i 1_{F_i} = \sum_{i=1}^{k+1} b_i 1_{F_i \setminus E_{k+1}} + \sum_{i=1}^{k+1} b_i 1_{F_i \cap E_{k+1}}$$

So, now we will prove the corresponding terms (for the integral) in the above two equations are less than or equal to each other. That will complete the proof.

Define

$$F = \bigcup_{i=1}^{k+1} F_i \ and\ E = \bigcup_{i=1}^{k+1} E_i$$

Now, if $x \in \overline{F}$, then $t(x) = 0 \Rightarrow s(x) = 0$, therefore, $x \in \overline{E}$

which implies that $\overline{F} \subseteq \overline{E}$ and therefore, $E \subseteq F$ and hence, we can write

$$E_{k+1} = \bigcup_{i=1}^{k+1} (F_i \cap E_{k+1})$$

Therefore, since F_i's are disjoint, we get

$$\mu(E_{k+1}) = \sum_{i=1}^{k+1} \mu\left(F_i \cap E_{k+1}\right)$$

Using the fact that $s(x) \le t(x) \; \forall \; x$,

$$a_{k+1}\mu(E_{k+1}) = \sum_{i=1}^{k+1} a_{k+1}\mu\left(F_i \cap E_{k+1}\right) \le \sum_{i=1}^{k+1} b_i\mu\left(F_i \cap E_{k+1}\right)$$

Note E_i's are disjoint, so that

$$\bigcup_{i=1}^{k+1}(F_i \backslash E_{k+1}) = \bigcup_{i=1}^{k}(E_i) \quad \cup (F \backslash E)$$

Let,

$$c_i = \min_{F_j \cap E_i \neq \phi} (b_j)$$

Observe that $a_i \le c_i$ for each i. So, by induction hypothesis,

$$\sum_{i=1}^{k} a_i\mu(E_i) \le \sum_{i=1}^{k} c_i\mu\left(E_i\right) \le \sum_{i=1}^{k+1} b_i\mu(F_i \backslash E_{k+1})$$

♣

This shows in particular, that the definition of Lebesgue integral given above for measurable functions, is consistent with the definition of the Lebesgue integral of simple functions as defined earlier. Question: Given a measurable function f, can we find "nice"simple functions $s \le f$?

Proposition 60 *Given $f : \mathbb{R}^+ \to \mathbb{R}^+$, measurable, there exists a sequence $\{s_n\}$ of simple functions such that*

(i) $s_n(x) \le s_{n+1}(x) \; \forall \; x$

(ii) $s_n(x) \uparrow f(x) \; \forall \; x$

Proof: Define

$$\begin{aligned} s_n(t) &= \min_{x \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right]} f(x) \; if \; t \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right] \; and \; 0 \le k \le n2^n - 1, \; k \in \mathbb{Z} \\ &= 0 \; if \; t > n \end{aligned}$$

Clearly, s_n is a simple function for all natural n. We claim that $\{s_n\}$ is the desired sequence.

(i) $s_n(t) \le s_{n+1}(t)\ \forall\ t, n$

if $t > n+1$, $s_n(t) = s_{n+1}(t) = 0$

else if, $t \in [n, n+1]$, then $s_{n+1}(t) \ge s_n(t) = 0$

else get $k < n2^n - 1$ such that, $t \in \left[\dfrac{k}{2^{n+1}}, \dfrac{k+1}{2^{n+1}}\right]$

then $\left[\dfrac{k}{2^{n+1}}, \dfrac{k+1}{2^{n+1}}\right] \subset \left[\dfrac{\lfloor \frac{k}{2} \rfloor}{2^n}, \dfrac{\lfloor \frac{k}{2} \rfloor + 1}{2^n}\right]$ and therefore,

$$\min_{x \in \left[\frac{k}{2^{n+1}}, \frac{k+1}{2^{n+1}}\right]} f(x) \ge \min_{x \in \left[\frac{\lfloor \frac{k}{2} \rfloor}{2^n}, \frac{\lfloor \frac{k}{2} \rfloor + 1}{2^n}\right]} f(x)$$

(ii) $s_n \to f$ as $n \to \infty$

Clear, since the width of the interval $\left[\dfrac{k}{2^n}, \dfrac{k+1}{2^n}\right] = \dfrac{1}{2^n} \to 0$ as $n \to \infty$ ♣

Note 11 *In the above few properties, we have that the function is non-negative everywhere. Everything above can be similarly extended to general functions in more or less the same way.*

Remark 36 *To check if a function is measurable, it may require more than just checking if $f^{-1}(\{x\})$ is measurable for each $x \in \mathbb{R}$. Indeed, consider $f : [0, 1] \to \mathbb{R}$*

$$\begin{aligned} f(x) &= 1 + x \quad if\ x \in E \\ &= -1 - x \quad if\ x \notin E \end{aligned}$$

where E is not a measurable set.

Note that $f^{-1}((0, \infty)) = E$ is not measurable, although $f^{-1}(\{x\})$ is measurable for every x.

Proposition 61 *Let s, t be simple functions.*

1. *Suppose $s(x) \ge 0$ for all $x \in \mathbb{R}$ and $\int s = 0$, then $s(x) = 0$ except on may be a set of measure zero.*
2. *$s \le t$, then $\int s \le \int t$.*
3. *$\int \alpha s = \alpha \int s$ for any $\alpha \in \mathbb{R}$.*
4. *$\int s + t = \int s + \int t$.*

Proof: We only prove (iv) since the others are practically trivial.
Suppose $s = \sum_{i=1}^{n} c_i 1_{E_i}$, $t = \sum_{j=1}^{m} d_j 1_{F_j}$.
Note that

$$\begin{aligned} s + t &= \sum_{i=1}^{n} \sum_{j=1}^{m} (c_i + d_j) 1_{E_i \cap F_j} \\ \int (s+t) &= \sum_{i=1}^{n} \left\{ \sum_{j=1}^{m} c_i \mu(E_i \cap F_j) + d_j \mu(E_i \cap F_j) \right\} \end{aligned}$$

Since, F_j's are pairwise disjoint $\sum_{j=1}^{m} \mu(E_i \cap F_j) = \mu\left(E_i \cap \bigcup_{j=1}^{m} F_j\right)$
Therefore,

$$\int (s+t) = \sum_{i=1}^{n} c_i \mu(E_i) + \sum_{j=1}^{m} d_j \mu(F_j) = \int s + \int t$$

♣

Recall: If $f : \Omega \to \mathbb{R}^+$ is measurable for some measurable $\Omega \subseteq \mathbb{R}$, then $\int f = \sup\{\int s | 0 \le s \le f$, s is simple$\}$

Proposition 62 *Suppose $f \geq 0$ for all $x \in \Omega$ and $\int f = 0$. Then $f(x) = 0$ "almost everywhere"(except on a set of measure zero).*

Proof: Suppose not; i.e., Suppose that $\{x | f(x) > 0\}$ is not of measure zero.

$$\{x|f(x) > 0\} = \bigcup_{n=1}^{\infty} \{x | f(x) > \frac{1}{n}\}$$

Let,

$$E_i = \{x | f(x) > \frac{1}{i}\}$$

Note that E_i's are measurable. Also, $E_1 \subseteq E_2 \subseteq E_3$...
If each $\mu(E_i) = 0$, then $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = 0$. Hence for some n, we must have $\mu(E_n) > 0$.

Consider

$$s_n(x) = \begin{cases} 1/n & : x \in E_n \\ 0 & : x \notin E_n \end{cases}$$

Observe that $s_n \leq f$. But then,

$$\int f = \frac{1}{n} \mu(E_n).$$

Therefore,

$$\int f \geq \int s_n = \frac{1}{n}\mu(E_n) > 0,$$

and this contradicts the hypothesis. ♣

Lemma 10 *1.* $0 \leq f \leq g$, *then* $\int f \leq \int g$. *Equality occurs iff* $f = g$ *almost everywhere.*

2. For $\alpha \geq 0$, $\int \alpha f = \alpha \int f$

3. $\int (f+g) \leq \int f + \int g$

Proof: We'll prove part 3; the others are easier to prove, and so we skip those proofs.

$$\int f = \sup\left\{\int s \mid 0 \leq s \leq f,\ s \textit{ is simple}\right\}.$$

$$\int g = \sup\left\{\int t \mid 0 \leq t \leq g,\ t \textit{ is simple}\right\}.$$

Hence,

$$\sup \int (s+t) \leq \sup \int s + \sup \int t.$$

or,

$$\int (f+g) \leq \int f + \int g$$

♣

5.4 Lebesgue's Monotone Convergence Theorem

The Monotone Convergence Theorem is really where the power of the Lebesgue Integral comes to the fore. The Riemann Integral's biggest weakness is its inability to provide simple proofs for statements of the type:

$$f_n \to f \Rightarrow \int f_n \to \int f.$$

The Riemann integral guarantees this under the stronger condition of uniform convergence. The Lebesgue integral gives us the same convergence under a much weaker condition.

Theorem 65 *Suppose f_i's are measurable and suppose $f_1 \leq f_2 \leq f_3 \leq \ldots$ i.e., $f_n(x) \leq f_{n+1}(x)$ for all x. Let $f = \lim_{n\to\infty} f_n(x)$. Then f is measurable and $\int f = \lim_{n\to\infty} \int f_n(x)$. To put it in other words,*

$$\int \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \int f_n(x).$$

Proof: $f_n(x) \leq f(x)$ for every x, so,

$$\int f_n \leq \int f$$

which implies

$$\lim_{n\to\infty} \int f_n(x) \leq \int f.$$

So, it suffices to show that $\lim_{n\to\infty} \int f_n(x) \geq \int f$.

We use the idea of using 'an epsilon of room' to prove this. Pick any $0 < \epsilon < 1$. It suffices to show that for any $0 < \epsilon < 1$,

$$\lim_{n\to\infty} \int f \geq (1-\epsilon) \int f = \int (1-\epsilon) f.$$

Since $0 < \epsilon < 1$, $(1-\epsilon)f < f$. Since $f_n \to f$ at each x, for all large n, we have

$$f_n(x) \geq (1-\epsilon) f(x).$$

Let $E_n = \{x | f_n(x) \geq (1-\epsilon) f(x)\}$. Since f_n's are increasing, we have

$$E_1 \subseteq E_2 \subseteq E_3 \ldots$$

Furthermore,

$$\bigcup_{i=1}^{\infty} E_i = \Omega.$$

We wish to show

$$\int f_n \geq \int (1-\epsilon) f$$

Take s simple and $s \leq f$ so that we have

$$\int_\Omega f_n \geq \int_\Omega f_n 1_{E_n} \geq \int_\Omega (1-\epsilon) f 1_{E_n} \geq (1-\epsilon) \int_\Omega f(1_{E_n}) \geq (1-\epsilon) \int_\Omega s 1_{E_n}$$

Since, s is simple, write $s = \sum_{i=1}^m c_i 1_{F_i}$ Then,

$$s 1_{E_n} = \sum_{i=1}^m c_i 1_{F_i \cap E_n}$$

Therefore,

$$\int_\Omega f_n \geq (1-\epsilon)\sum_{i=1}^{m} c_i\mu(F_i \cap E_n)$$

Note that for each fixed $1 \leq i \leq m$,

$$F_i \cap E_n \subseteq F_i \cap E_{n+1} \textit{ and}$$

$$\bigcup_i (F_i \cap E_n) = F_i$$

We know that if $A_1 \subseteq A_2 \subseteq \ldots$ are measurable, then

$$\mu(\cup_i A_i) = \lim_{n\to\infty} \mu(A_n)$$

Therefore,

$$\lim_{n\to\infty} \mu(F_i \cap E_n) = \mu(F_i)$$

So,

$$\lim_{n\to\infty}\int f_n \geq (1-\epsilon)\lim_{n\to\infty}\sum_{i=1}^{n} c_i\mu(F_i) = (1-\epsilon)\int s.$$

Therefore, for any $0 \leq s \leq f$,

$$\lim_{n\to\infty}\int f_n \geq (1-\epsilon)\int s.$$

Take supremum over $s \leq f$ to get

$$\lim_{n\to\infty}\int f_n \geq (1-\epsilon)\int f.$$

Take $\lim\limits_{\epsilon\to 0}$ to complete the proof.

♣

Corollary 12 $\int(f+g) = \int f + \int g$

Proof: We have $\displaystyle\int(f+g) \leq \int f + \int g$

Get sequences of simple functions $\{s_n\}$ and $\{t_n\}$ such that

$s_n \leq f$ such that $\displaystyle\int s_n \to \int f$ and sequence $\{s_n\}$ is increasing

$t_n \leq g$ such that $\displaystyle\int t_n \to \int g$ and sequence $\{t_n\}$ is increasing

Since $\{s_n + t_n\}$ is increasing and converging to $f+g$, by monotone convergence theorem, we have

$$\int \lim_{n\to\infty}(s_n+t_n) = \lim_{n\to\infty}\int(s_n+t_n) = \lim_{n\to\infty}\int s_n + \lim_{n\to\infty}\int t_n \qquad \textit{or,}$$

$$\int (f+g) = \int f + \int g$$

♣