# 21-355: Real Analysis 1

Carnegie Mellon University

Professor Ian Tice - Fall 2013

Project LaTeX'd by Ivan Wang

Last Updated: November 25, 2013

# Contents

# 1 The Number Systems

## 1.1 The Natural Numbers

**Theorem** (existence of $\mathbb{N}$): There exists a set $\mathbb{N}$ satisfying the following properties, known as the Peano Axioms:

**PA1** $0 \in \mathbb{N}$

**PA2** There exists a function $S : \mathbb{N} \to \mathbb{N}$ called the successor function. In particular, $S(n) \in \mathbb{N}$.

**PA3** $\forall n \in \mathbb{N}.\ S(n) \neq 0$

**PA4** $S(n) = S(m) \implies n = m$ ($S$ is injective, one-to-one)

**PA5** [Axiom of Induction] Let $P(n)$ be a property associated to each $n \in \mathbb{N}$.
If $P(0)$ is true, and $P(n) \implies P(S(n))$, then $P(n)$ is true $\forall n \in \mathbb{N}$.

**Definition**: **PA1** $\implies 0 \in \mathbb{N}$. **PA2** $\implies S(0) \in \mathbb{N}$.

Define $1 = S(0), 2 = S(1), 3 = S(2)$, etc.

**PA2** guarantees that $\{0, 1, 2, \cdots\} \subseteq \mathbb{N}$.

**PA3** prevents "wraparound": no successor can map to a "negative" number.

**PA4** prevents "stagnation": the cycle does not terminate.

**Theorem**: $\mathbb{N} = \{0, 1, 2, \cdots\}$

*Proof*: We know that $\{0, 1, 2, \cdots\} \subseteq \mathbb{N}$, so it suffices to prove that $\mathbb{N} \subseteq \{0, 1, 2, \cdots\}$.

Let $P(n)$ denote the proposition that $n \in \{0, 1, 2, \cdots\}$. Clearly $P(0)$ is true.

Suppose $P(n)$ is true; then $n \in \{0, 1, 2, \cdots\} \implies S(n) \in \{0, 1, 2, \cdots\}$ by construction. Hence, $P(S(n))$ is true. By induction, **PA5** guarantees that $P(n)$ is true $\forall n \in \mathbb{N}$.

It follows that $\mathbb{N} \subseteq \{0, 1, 2, \cdots\}$.

**Definition**: For any $m \in \mathbb{N}$, we define $0 + m = m$.
Then if $n + m$ is defined for $n \in \mathbb{N}$, we set $S(n) + m = S(n + m)$.

**Proposition** (Properties of Addition):

1. $\forall n \in \mathbb{N}.\ n + 0 = n$ (0 is the additive identity)
2. $\forall m, n \in \mathbb{N}.\ n + S(m) = S(n + m)$
3. $\forall m, n \in \mathbb{N}.\ m + n = n + m$ (commutativity)
4. $\forall k, m, n \in \mathbb{N}.\ k + (m + n) = (k + m) + n$ (associativity)
5. $\forall k, m, n \in \mathbb{N}.\ n + k = n + m \implies k = m$ (cancelation)

*Proof*:

1. Let $P(n)$ be $n + 0 = n$.
   $P(0)$ is true because $0 + 0 = 0$ by definition.
   Note $P(n) \implies S(n) + 0 = S(n + 0) = S(n)$, so $P(S(n))$ is true. By induction, (1) is true.

2. Fix $m \in \mathbb{N}$. Let $P(n)$ denote $n + S(m) = S(n + m)$.
   $P(0)$ is true because $0 + S(m) = S(m) = S(0 + m)$.
   $P(n) \implies S(n)+S(m) = S(n+S(m)) = S(S(n+m)) = S(S(n)+m)$, so $P(S(n))$ is true. By induction, since $m \in \mathbb{N}$ was arbitrary, (2) is true.
3. Let $m$ be fixed and $P(n)$ denote $n + m = m + n$.
   $P(0)$ is true since $0 + m = m$ by definition, and $m + 0 = m$ by 1, so $0 + m = m = m + 0$.
   Suppose $P(n)$; then $S(n) + m = S(n + m) = S(m + n) = m + S(n)$, so $P(S(n))$ is true. By induction and arbitrary choice of $m$, (3) is true.
4. Fix $k, m \in \mathbb{N}$ and let $P(n)$ denote $k + (m + n) = (k + m) + n$.
   $P(0)$ is true as $k + (m + 0) = k + m = (k + m) + 0$.
   Suppose $P(n)$; then $k+(m+S(n)) = k+S(m+n) = S(k+(m+n)) = S(k+m)+n = (k + m) + S(n)$ by (2). By induction and arbitrary choice, (4) is true.
5. Fix $m, n \in \mathbb{N}$ and let $P(k)$ denote proposition 5.
   $P(0)$ is true because $n + 0 = n = n + m \implies m = 0 \implies k = m$.
   Suppose $P(k)$; also, suppose $m + S(k) = n + S(k)$. Then $S(m + k) = m + S(k) = n + S(k) = S(n + k) \implies m + k = n + k \implies m = n$ (by 4). By the axiom of induction, (5) is true.

### 1.1.1 Positivity

**Definition**: We say that $n \in \mathbb{N}$ is *positive* if $n \neq 0$.

**Proposition** (Properties of Positivity):

1. $\forall n, m \in \mathbb{N}$, if $m$ is positive, then $m + n$ is positive.
2. $\forall n, m \in \mathbb{N}$, if $m + n = 0$, then $m = n = 0$.
3. $\forall n \in \mathbb{N}$, if $n$ is positive, then there exists a unique $m \in \mathbb{N}$ such that $n = S(m)$.

### 1.1.2 Order

**Definition**: For all $m, n \in \mathbb{N}$, $m \leq n$ or $n \geq m$ iff $n = m + p$ for some $p \in \mathbb{N}$.
$m < n$ or $n > m$ iff $m \leq n \wedge m \neq n$. The relation $\leq$ provides what is called an *order* on $\mathbb{N}$.

**Proposition** (Properties of Order):

Let $j, k, m, n \in \mathbb{N}$. Then:

1. $n \geq n$ (reflexitivity)
2. $m \leq n \wedge k \leq m \implies k \leq n$ (transitivity)
3. $m \geq n \wedge m \leq n \implies m = n$ (anti-symmetry)
4. $j \leq k \wedge m \leq n \implies j + m \leq k + n$ (order preservation)
5. $m < n \iff S(m) \leq n$
6. $m < n \iff n = m + p$ for some positive $p \in \mathbb{N}$.
7. $n \geq m \iff S(n) > m$
8. $n = 0 \oplus 0 < n$

**Theorem** (Trichotomy of Order): Let $m, n \in \mathbb{N}$. Then exactly one of the following is true:

$$m < n \quad \oplus \quad m = n \quad \oplus \quad m > n$$

*Proof*: Show that no two can be true simultaneously (by definition of $<$ and $>$), and then at least one must be true (by induction on $n$).

### 1.1.3 Multiplication

**Definition**: Fix $m \in \mathbb{N}$. Define $0 \cdot m = 0$. Now, if $n \cdot m$ is defined for some $n \in \mathbb{N}$, we define $S(n) \cdot m = n \cdot m + m$.

**Proposition** (Properties of Multiplication):

Fix $k, m, n \in \mathbb{N}$. Then:

1. $m \cdot n = n \cdot m$ (commutativity)
2. $m, n$ are positive $\implies mn$ is positive
3. $m \cdot n = 0 \iff m = 0 \vee n = 0$ (no zero divisors)
4. $k \cdot (m \cdot n) = (k \cdot m) \cdot n$ (associativity)
5. $k \cdot m = k \cdot n \wedge k$ is positive $\implies m = n$ (cancelation)
6. $k \cdot (m + n) = (m + n) \cdot k = k \cdot m + k \cdot n$ (distributivity)
7. $m < n \wedge k \leq l \wedge k, l$ are positive $\implies m \cdot k < n \cdot l$

## 1.2 The Integers

Consider the following relation on the set $\mathbb{N} \times \mathbb{N}$:

$$(m, n) \simeq (m', n') \iff m + n' = m' + n$$

**Lemma**: $\simeq$ is an equivalence relation.

*Proof*:

Reflexivity: $m + n = m + n \implies (m, n) \simeq (m, n)$

Symmetry: $(m, n) \simeq (m', n') \implies m + n' = m' + n \implies m' + n = m + n' \implies (m', n') \simeq (m, n)$

Transitivity: Suppose $(m, n) \simeq (m', n') \wedge (m', n') \simeq (m'', n'')$. Then:

$$\begin{aligned} & m + n' = m' + n \wedge m' + n'' = m'' + n' \\ \implies & m + n'' = m'' + n \\ \implies & (m, n) \simeq (m'', n'') \end{aligned}$$

**Definition**: Write the *equivalence class* of $(m, n)$ as $[(m, n)] = \{(p, q) \mid (p, q) \simeq (m, n)\}$. Define the *integers* $\mathbb{Z} = \{[(m, n)]\}$.

**Lemma**: Suppose $(m, n) \simeq (m', n'), (p, q) \simeq (p', q')$. Then:

1. $(m + p, n + q) \simeq (m' + p', n' + q')$

2. $(mp + nq, mq + np) \simeq (m'p' + n'q', m'q' + n'p')$

*Proof*: Consider equalities $(a) : m + n' = m' + n$ and $(b) : p + q' = p' + q$ (by definition of $\simeq$).

Using linear combinations of $(a)$ and $(b)$, we derive the two rules of the lemma:

1. $(a) + (b)$
2. $(a)(p' + q') + (b)(m + n)$

**Definition**: Let $[(m, n)], [(p, q)] \in \mathbb{Z}$. Then:

1. $[(m, n)] + [(p, q)] = [(m + p, n + q)]$ (addition of integers)
2. $[(m, n)] \cdot [(p, q)] = [(mp + nq, mq + np)]$ (multiplication of integers)

By the lemma, these are well-defined operations.

Note that for all $m, n \in \mathbb{N}$:

$$[(m, 0)] = [(n, 0)] \iff m + 0 = n + 0 \iff m = n$$
$$[(m, 0)] + [(n, 0)] = [(m + n, 0)]$$
$$[(m, 0)] \cdot [(n, 0)] = [(mn, 0)]$$

As such, the set $\{[(n, 0)] \mid n \in \mathbb{N}\} \subseteq \mathbb{Z}$ behaves exactly like a copy of $\mathbb{N}$.

**Definition**: For $n \in \mathbb{N}$ we set $n \in \mathbb{Z}$ to be $n := [(n, 0)]$.

For $x = [(m, n)] \in \mathbb{Z}$ we define $-x = [(n, m)]$.

### 1.2.1 Properties of Integers

(We can see that every integer $x \in \mathbb{Z}$ can be represented as $x := m - n$ where $x = [(m, n)]$.)

**Theorem**: Every $x \in \mathbb{Z}$ satisfies exactly one of the following:

1. $x = n$ for some $n \in \mathbb{N}\backslash\{0\}$
2. $x = 0$
3. $x = -n$ for some $n \in \mathbb{N}\backslash\{0\}$

*Proof*: Write $x = [(p, q)]$ for some $p, q \in \mathbb{N}$. By trichotomy of order on $\mathbb{N}$ we know that $p < q$ or $p = q$ or $p > q$. Each of these correlates to one of the three properties.

**Corollary**: $\mathbb{Z} = \{0, 1, 2, \ldots\} \cup \{-1, -2, -3, \ldots\}$

### 1.2.2 Algebraic Properties

**Proposition**: Let $x, y, z \in \mathbb{Z}$. Then the following hold:

1. $x + y = y + x$
2. $x + (y + z) = (x + y) + z$
3. $x + 0 = 0 + x = x$
4. $x + (-x) = (-x) + x = 0$
5. $xy = yx$

6. $(xy)z = x(yz)$
7. $x \cdot 1 = 1 \cdot x = x$
8. $x(y+z) = xy + xz$

**Definition**: Define $x - y = x + (-y)$. The usual properties hold.

**Definition**: For $x, y \in \mathbb{Z}$, we say $x \leq y$ or $y \geq x$ if $y - x = n$ for some $n \in \mathbb{N}$.
We say $x < y$ if $x \leq y \wedge x \neq y$.

## 1.3 The Rationals and Ordered Fields

Let a relation on $\mathbb{Z} \times (\mathbb{Z} \backslash \{0\})$ be given by $(m, n) \simeq (m', n') \iff mn' = m'n$.

**Lemma**: $\simeq$ is an equivalence relation. Proof follows from properties of $\mathbb{Z}$.

**Definition**: $\mathbb{Q} = \{[(m, n)]\}$

1. $[(m, n)] + [(p, q)] = [(mq + np, nq)]$ (addition)
2. $[(m, n)] \cdot [(p, q)] = [(mp, nq)]$ (multiplication)
3. $-[(m, n)] = [(-m, n)]$ (negation)
4. If $m \neq 0$ we set $[(m, n)]^{-1} = [(n, m)]$

Remark: the heuristic here is that $\frac{m}{n} = [(m, n)]$.

**Definition**: If $m \in \mathbb{Z}$, we write $m = [(m, 1)] \in \mathbb{Q}$; and thus $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}$.

1. For $x, y \in \mathbb{Q}$, we define $x - y = x + (-y) \in \mathbb{Q}$
2. For $x, y \in \mathbb{Q}, y \neq 0$ we define $\frac{x}{y} = x(y)^{-1}$. This is well defined because $y = 0 \iff y = [(0, n)]$.

**Proposition**: $\mathbb{Q} = \{\frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0\}$.

We define and propose the trichotomy of order on $\mathbb{Q}$, as per the integers.

### 1.3.1 Fields and Orders

**Definition**: A field is a set $\mathbb{F}$ endowed with two binary operations, $+, \cdot$, satisfying the following axioms:

(A1, M1) $\forall x, y \in \mathbb{F}.\ x + y \in \mathbb{F}, xy \in \mathbb{F}$ (closure)
(A2, M2) $\forall x, y \in \mathbb{F}.\ x + y = y + x,\ xy = yx$ (commutativity)
(A3, M3) $\forall x, y, z \in \mathbb{F}.\ x + (y + z) = (x + y) + z,\ x(yz) = (xy)z$ (associativity)
(A4, M4) $\exists (0, 1) \in \mathbb{F}.\ \forall x \in \mathbb{F}.\ 0 + x = x + 0 = x,\ 1 \cdot x = x \cdot 1 = x$ (identity)
(A5, M5) $\forall x \in \mathbb{F}.\ \exists (-x).\ x + (-x) = 0;\ \exists x^{-1} \in \mathbb{F}.\ xx^{-1} = x^{-1}x = 1$ (inverse)
(D1) $\forall x, y, z \in \mathbb{F}.\ x(y + z) = xy + xz$ (distributivity)

*Remark*: Field must have at least 2 elements (0, 1) by (A/M4). To prove field, must prove 5 properties of addition and multiplication (closure, commutativity, associativity, identity, inverse) as well as distributivity.

**Definition**: Let $E$ be a set; an *order* on $E$ is a relation $<$ satisfying the following:

1. $\forall x, y \in E$ exactly one of the following is true: $x < y$ or $x = y$ or $y < x$ (trichotomy)
2. $\forall x, y, z \in E$, $x < y \wedge y < z \implies x < z$ (transitivity)

**Definition**: Let $\mathbb{F}$ be a field. Then we define $x - y = x + (-y)$ and $\frac{x}{y} = xy^{-1}$ (for $y \neq 0$).

**Theorem**: $\mathbb{Q}$ is an ordered field with order $<$.

*Proof*: Follows from definitions and properties of $\mathbb{Z}$.

## 1.4 Problems with $\mathbb{Q}$

**Theorem**: There does not exist a $q \in \mathbb{Q}$ such that $q^2 = 2$.

*Proof*: Suppose not; i.e. there does exist such a $q \in \mathbb{Q}$.

Consider the set $S(q) = \{n \in \mathbb{N}^+ \mid q = \frac{m}{n} \text{ for some } m \in \mathbb{Z}\}$. Cleary $|S(q)| > 0$. Then the well-ordering principle implies that $\exists! n \in S(q).\ n = \min S(q)$.

Since $n \in S(q)$, we know that $q = \frac{m}{n}$ for some $m \in \mathbb{Z}$. Then $q^2 = (\frac{m}{n})^2 = \frac{m^2}{n^2} \implies m^2 = 2n^2 \implies m^2$ is even. We claim that $m$ is also even (proof is exercise to reader).

Then $\exists l \in \mathbb{Z}.\ m = 2l$. Then $4l^2 = (2l)^2 = m^2 = 2n^2 \implies n^2 = 2l^2 \implies n^2$ is even $\implies n$ is even $\implies n = 2p$ for some $p \in \mathbb{N}^+$.

Hence $q = \frac{m}{n} = \frac{2l}{2p} = \frac{l}{p} \implies p \in S(q)$. But clearly $p < n$, which contradicts the fact that $n$ is the minimal element. By contradiction, the theorem must be true.

### 1.4.1 Bounds (Infimum and Supremum)

Informally, $\mathbb{Q}$ has "holes":

**Definition**: Let $E$ be an ordered set with order $<$.

1. We say $A \subseteq E$ is bounded above iff $\exists x \in E.\ \forall a \in A.\ a \leq x$. We say $x$ is an upper bound of $A$.
2. We say $A \subseteq E$ is bounded below iff $\exists x \in E.\ \forall a \in A.\ x \leq a$. We say $x$ is a lower bound of $A$.
3. We say $A \subseteq E$ is bounded iff it's bounded above and below.
4. We say $x$ is a minimum of $A$ iff $x \in A$ and $x$ is a lower bound of $A$.
5. We say $x$ is a maximum of $A$ iff $x \in A$ and $x$ is an upper bound of $A$.

*Remark*: If a min or max exists, then it is unique.

**Definition**: Let $E$ be an ordered set and $A \subseteq E$.

1. We say $x \in E$ is the least upper bound (*supremum*) of $A$, written $x = \sup A$, iff $x$ is an upper bound of $A$ and $y \in E$ is an upper bound of $A \implies x \leq y$.
2. We say $x \in E$ is the greatest lower bound (*infimum*) of $A$, written $x = \inf A$, iff $x$ is a lower bound of $A$ and $y \in E$ is a lower bound of $A \implies y \leq x$.

*Remark*: If $x = \min(A)$, then $x = \inf(A)$. If $x = \max(A)$, then $x = \sup(A)$. But the converse is false; some sets have a supremum but no maximum, others a infimum but no minimum.

**Definition**: Let $\mathbb{F}$ be an ordered field. We say that $\mathbb{F}$ has the *least upper bound property* iff every $\varnothing \neq A \subseteq \mathbb{F}$ that is bounded above has a least upper bound.

**Theorem**: $\mathbb{Q}$ does not satisfy the least upper bound property.

*Proof*: Consider the set $A = \{x \in \mathbb{Q} \mid x > 0, x^2 \leq 2\}$.

Note that $0 < 1 = 1^2 \leq 2 \implies 1 \in A$, so $A$ is non-empty. Also, $2 \leq 4 = 2^2$ implies $(x \in A \implies 0 < x^2 < 2 < 2^2) \implies x < 2$. Then 2 is an upper bound of $A$.

Assume for sake of contradiction that $\mathbb{Q}$ has the least upper bound property. Then $A$ has a supremum. Let $x = \sup A \in \mathbb{Q}$ and write $x = \frac{p}{q}$ for $p, q \in \mathbb{Z}$.

By trichotomy, $x^2 < 2$ or $x^2 = 2$ or $x^2 > 2$. We know $x^2 \neq 2$.

**Case 1**: Suppose $x^2 < 2$. Then for any $n \in \mathbb{N}^+$ we have $(\frac{p}{q} + \frac{1}{n})^2 = \frac{p^2}{q^2} + \frac{2p}{qn} + \frac{1}{n^2} \leq \frac{p^2}{q^2} + \frac{1}{n}(\frac{2p+q}{q})$. From algebra, we derive $(\frac{p}{q} + \frac{1}{n})^2 < 2$ for some $n \in \mathbb{N}^+$.

Cleary $x > 0$ since otherwise $x \leq 0 < 1 \in A$. Hence $0 < x = \frac{p}{q} < \frac{p}{q} + \frac{1}{n} \in A$. But then $x$ is not an upper bound $\implies$ contradiction.

**Case 2**: Suppose $x^2 > 2$. Considering $(\frac{p}{q} - \frac{1}{n})^2 > 2$ and using the same logic as before, we can choose $n$ large enough such that $\frac{p}{q} - \frac{1}{n}$ is an upper bound of $A$. But $\frac{p}{q} - \frac{1}{n} < \frac{p}{q} = x$, which contradicts the fact that $x = \sup A$.

As all cases are false, we contradict trichotomy, and hence $\mathbb{Q}$ cannot have the least upper bound property.

## 1.5 The Real Numbers

We now construct an ordered field satisfying the least upper bound property using $\mathbb{Q}$.

**Definition**: We say $\mathbb{Q}$ is Archimedean iff $\forall(x \in \mathbb{Q}).\ x > 0 \implies \exists(n \in \mathbb{N}).\ x < n$.

**Lemma**: If $\mathbb{Q}$ is Archimedean, then $\forall(p < q \in \mathbb{Q}).\ \exists(r \in \mathbb{Q}).\ p < r < q$.
(Proofs in HW 2.)

### 1.5.1 Defining the Real Numbers: Dedekind Cuts

**Definition**: We say that $\mathcal{C} \in \mathcal{P}(\mathbb{Q})$ is a *cut* (Dedekind cut) iff the following hold:

(C1) $\varnothing \neq \mathcal{C}, \mathcal{C} \neq \mathbb{Q}$

(C2) If $p \in \mathcal{C}$ and $q \in \mathbb{Q}$ with $q < p$, then $q \in \mathcal{C}$.

(C3) If $p \in \mathcal{C}$, $\exists(r \in \mathbb{Q}).\ p < r \wedge r \in \mathcal{C}$.

**Lemma**: Suppose $\mathcal{C}$ is a cut. Then:

1. $p \in \mathcal{C}, q \notin \mathcal{C} \implies p < q$
2. $r \notin \mathcal{C}, r < s \implies s \notin \mathcal{C}$
3. $\mathcal{C}$ is bounded above

**Lemma**: Let $q \in \mathbb{Q}$. Then $\{p \in \mathbb{Q} \mid p < q\}$ is a cut.

*Proof*: Call the set $\mathcal{C}$. We prove the 3 properties of a cut:

(C1) $q - 1 \in \mathcal{C} \implies \mathcal{C} \neq \varnothing$; $q + 1 \notin \mathcal{C} \implies \mathcal{C} \neq \mathbb{Q}$.

(C2) If $p \in \mathcal{C}$ and $r \in \mathbb{Q}$ such that $r < p$, then $r < p < q \implies r < q \implies r \in \mathcal{C}$.

(C3) Let $p \in \mathcal{C}$ where $p < q$. Since $\mathbb{Q}$ is Archimedean, $\exists(r \in \mathbb{Q}).\ p < r < q \implies r \in \mathcal{C}$.

**Definition**: Given $q \in \mathbb{Q}$ we write $\mathcal{C}_q = \{p \in \mathbb{Q} \mid p < q\}$. By the above lemma, $\mathcal{C}_q$ is a cut.

**Definition**: We write $\mathbb{R} = \{\mathcal{C} \in \mathcal{P}(\mathbb{Q}) \mid \mathcal{C} \text{ is a cut}\} \neq \varnothing$.

**Lemma**: The following hold:

1. $\forall \mathcal{A}, \mathcal{B} \in \mathbb{R}$, exactly one of the following holds: $\mathcal{A} \subset \mathcal{B}, \mathcal{A} = \mathcal{B}, \mathcal{B} \subseteq \mathcal{A}$.
2. $\forall \mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{R}$, $\mathcal{A} \subset \mathcal{B} \wedge \mathcal{B} \subseteq \mathcal{C} \implies \mathcal{A} \subset \mathcal{C}$.

**Definition**: If $\mathcal{A}, \mathcal{B} \in \mathbb{R}$ we say that $\mathcal{A} < \mathcal{B} \iff \mathcal{A} \subset \mathcal{B}$, and $\mathcal{A} \leq \mathcal{B} \iff \mathcal{A} \subseteq \mathcal{B}$. This defines an order on $\mathbb{R}$ by the above lemma.

### 1.5.2 Defining the Real Numbers: The Least Upper Bound Property

**Lemma**: Suppose $\varnothing \neq E \subseteq \mathbb{R}$ is bounded above. Then $\mathcal{B} := \bigcup_{\mathcal{A} \in E} \mathcal{A} \in \mathbb{R}$.

**Theorem**: $\mathbb{R}$ satisfies the least upper bound property.

*Proof*: Let $\varnothing \neq E \subseteq \mathbb{R}$ be bounded above and set $\mathcal{B} = \bigcup_{\mathcal{A} \in E} \mathcal{A} \in \mathbb{R}$. We claim $\mathcal{B} = \sup E$.

First, we show that $\mathcal{B}$ is an upper bound of $E$. Let $\mathcal{A} \in E$. Then $\mathcal{A} \subseteq \mathcal{B} \implies \mathcal{A} \leq \mathcal{B}$ (by definition). This is true for all $\mathcal{A} \in E$, so $\mathcal{B}$ is an upper bound.

We claim that for $\mathcal{C} \in \mathbb{R}$. $\mathcal{C} < \mathcal{B} \implies \mathcal{C}$ is not an upper bound of $E$. If $\mathcal{C} < \mathcal{B}$, then $\mathcal{C} \subset \mathcal{B}$. This implies $\exists b \in \mathcal{B}.\ b \notin \mathcal{C} \implies \exists (\mathcal{A} \in E).\ b \in \mathcal{A} \wedge b \notin \mathcal{C}$. Then $\mathcal{A} > \mathcal{C}$ since otherwise $\mathcal{A} \subseteq \mathcal{C} \implies b \in \mathcal{C}, b \notin \mathcal{C}$. Hence $\mathcal{C} < \mathcal{A}$ and $\mathcal{C}$ is not an upper bound of $E$.

By the contrapositive: if $\mathcal{C}$ is an upper bound, $\mathcal{C} \geq \mathcal{B}$. Thus, $\mathcal{B}$ is the least upper bound, and the theorem holds.

### 1.5.3 Defining the Real Numbers: Addition

**Definition**: Given $\mathcal{A}, \mathcal{B} \in \mathbb{R}$, set $\mathcal{A} + \mathcal{B} = \{a + b \mid a \in \mathcal{A}, b \in \mathcal{B}\}$.

**Lemma**: If $\mathcal{A}, \mathcal{B} \in \mathbb{R}$, then $\mathcal{A} + \mathcal{B} \in \mathbb{R}$.

**Theorem**: Define $-\mathcal{A} = \{q \in \mathbb{Q} \mid \exists (p > q).\ -p \notin \mathcal{A}\}$. Then $\mathbb{R}, +, 0_{\mathbb{R}} = \mathcal{C}_0 = \{p \in \mathbb{Q} \mid p < 0\}$ satisfy the field axioms.

*Proof*:

(A1) $\mathcal{A} + \mathcal{B} \in \mathbb{R}$ by previous lemma.

(A2) $\mathcal{A} + \mathcal{B} = \{a + b\} = \{b + a\} = \mathcal{B} + \mathcal{A}$.

(A3) $\mathcal{A} + (\mathcal{B} + \mathcal{C}) = \{a + (b + c)\} = \{(a + b) + c\} = (\mathcal{A} + \mathcal{B}) + \mathcal{C}$.

(A4) Show $\forall \mathcal{A} \in \mathbb{R}.\ 0_{\mathbb{R}} + \mathcal{A} = \mathcal{A}$.

(A5) Show that $-\mathcal{A} \in \mathbb{R}$, then $\mathcal{A} + (-\mathcal{A}) = 0_{\mathbb{R}}$ using Archimedean property.

**Theorem** (Ordered Field): Let $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{R}$. If $\mathcal{A} < \mathcal{B}$ then $\mathcal{A} + \mathcal{C} < \mathcal{B} + \mathcal{C}$.

*Proof*: It's trivial to see that $\mathcal{A} \subseteq \mathcal{B} \implies \mathcal{A} + \mathcal{C} \subseteq \mathcal{B} + \mathcal{C} \implies \mathcal{A} + \mathcal{C} \leq \mathcal{B} + \mathcal{C}$.

If $\mathcal{A} + \mathcal{C} = \mathcal{B} + \mathcal{C}$, we can add $-\mathcal{C}$ to both sides and use the last theorem to see that $\mathcal{A} = \mathcal{B}$, a contradiction. Hence, $\mathcal{A} + \mathcal{C} < \mathcal{B} + \mathcal{C}$.

### 1.5.4 Defining the Real Numbers: Multiplication

**Lemma**: Let $\mathcal{A}, \mathcal{B} \in \mathbb{R}$, $\mathcal{A}, \mathcal{B} > 0_{\mathbb{R}}$. Then $\mathcal{C} = \{q \in \mathbb{Q} \mid q \leq 0\} \cup \{a \cdot b \mid a \in \mathcal{A}, b \in \mathcal{B}, a, b > 0\} \in \mathbb{R}$.

*Proof*:

(C1) $0 \in \mathcal{C} \implies \mathcal{C} \neq \varnothing$. $\mathcal{A}, \mathcal{B}$ are bounded above by, say $M_1, M_2$, so $M_1 \cdot M_2 + 1 \notin \mathcal{C}$ and $\mathcal{C} \neq \mathbb{Q}$.

(C2) Let $p \in \mathcal{C}$ and $q < p$. If $q \leq 0$ then $q \in \mathcal{C}$ by definition. If $q > 0$ then $0 < q < p$, but then $0 < p \implies p = a \cdot b$ for $a \in \mathcal{A}, b \in \mathcal{B}, a, b > 0$. Then $0 < q < a \cdot b \implies \frac{q}{a} < b \implies 0 < \frac{q}{a} \in \mathcal{B}$. Then $q = a(\frac{q}{a}) \in \mathcal{C}$.

(C3) Let $p \in \mathcal{C}$. If $p \leq 0$ then any $a \cdot b$ with $a \in \mathcal{A}, b \in \mathcal{B}, a, b > 0$ satisfies $p < a \cdot b \in \mathcal{C}$, so $r = a \cdot b$ is the desired element of $\mathcal{C}$. However, if $p > 0$, then $p = a \cdot b$ for $a \in \mathcal{A}, b \in \mathcal{B}, a, b > 0$. Choose $s \in \mathcal{A}$ such that $a < s, t \in \mathcal{B}$ such that $t > b$. Then $p = a \cdot b < s \cdot t \in \mathcal{S}$, so $r = s \cdot t$ proves the claim.

**Definition of Multiplication**: Let $\mathcal{A}, \mathcal{B} \in \mathbb{R}$.

1. If $\mathcal{A} > 0, \mathcal{B} > 0$ we set $\mathcal{A} \cdot \mathcal{B} = \{q \in \mathbb{Q} \mid q \leq 0\} \cup \{a \cdot b \mid a \in \mathcal{A}, b \in \mathcal{B}, a, b > 0\} \in \mathbb{R}$.
2. If $\mathcal{A} = 0$ or $\mathcal{B} = 0$, we set $\mathcal{A} \cdot \mathcal{B} = 0_{\mathbb{R}}$.
3. If $\mathcal{A} > 0$ and $\mathcal{B} < 0$, let $\mathcal{A} \cdot \mathcal{B} = -(\mathcal{A} \cdot (-\mathcal{B}))$.
4. If $\mathcal{A} < 0$ and $\mathcal{B} > 0$, let $\mathcal{A} \cdot \mathcal{B} = -((-\mathcal{A}) \cdot \mathcal{B})$.
5. If $\mathcal{A} < 0$ and $\mathcal{B} < 0$, let $\mathcal{A} \cdot \mathcal{B} = (-\mathcal{A}) \cdot (-\mathcal{B})$.

**Theorem**: $\mathbb{R}, \cdot$ satisfies (M1-M5) with $1_{\mathbb{R}} = \mathcal{C}_1$, and
$\mathcal{A} > 0 \implies \mathcal{A}^{-1} = \{q \in \mathbb{Q} \mid q \leq 0\} \cup \{q \in \mathbb{Q} \mid q > 0, \exists p > q.\ p^{-1} \notin \mathcal{A}\} \in \mathbb{R}$;
$\mathcal{A} < 0 \implies \mathcal{A}^{-1} = -(-\mathcal{A})^{-1}$.

*Proof*: HW3 (similar to addition).

**Theorem**: If $\mathcal{A}, \mathcal{B} > 0$, then $\mathcal{A} \cdot \mathcal{B} > 0$.

*Proof*: By definition $\mathcal{C}_0 \subseteq \mathcal{A} \cdot \mathcal{B} \implies 0 \leq \mathcal{A} \cdot \mathcal{B}$. Equality is impossible since $\mathcal{A}, \mathcal{B} > 0$.

### 1.5.5 Defining the Real Numbers: Distributivity

**Theorem**: Let $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathbb{R}$. Then $\mathcal{A} \cdot (\mathcal{B} + \mathcal{C}) = \mathcal{A} \cdot \mathcal{B} + \mathcal{A} \cdot \mathcal{C}$.

*Proof*: We prove the case where all are positive. The other cases are in HW.

Let $p \in \mathcal{A}(\mathcal{B} + \mathcal{C})$. If $p \leq 0$ then $p \in \mathcal{A} \cdot \mathcal{B} + \mathcal{A} \cdot \mathcal{B}$ is trivial (both products contain the interval less than 0).

If $p > 0$, $p = a(b + c)$ for $a \in \mathcal{A}, b \in \mathcal{B}, c \in \mathcal{C}$ for $a > 0, b + c > 0$.

Regardless of sign of $b$ or $c$, $a \cdot b \in \mathcal{A} \cdot \mathcal{B}, a \cdot c \in \mathcal{A} \cdot \mathcal{C}$. Hence $p = a(b + c) = a \cdot b + a \cdot c \in \mathcal{A} \cdot \mathcal{B} + \mathcal{A} \cdot \mathcal{C}$. So $\mathcal{A}(\mathcal{B} + \mathcal{C}) \subseteq \mathcal{A} \cdot \mathcal{B} + \mathcal{A} \cdot \mathcal{C}$.

Finally, we show the converse is true; let $p \in \mathcal{A} \cdot \mathcal{B} + \mathcal{A} \cdot \mathcal{C} \implies p = r + s$ for $r \in \mathcal{A} \cdot \mathcal{B}, s \in \mathcal{A} \cdot \mathcal{C}$. Case on positivity of $p, r, s$ to show $p \in \mathcal{A}(\mathcal{B} + \mathcal{C})$.

### 1.5.6 Defining the Real Numbers: Archimedean

**Theorem**: For $p, q \in \mathbb{Q}$, the following are true:

1. $\mathcal{C}_{p+q} = \mathcal{C}_p + \mathcal{C}_q$
2. $\mathcal{C}_{-p} = -\mathcal{C}_p$
3. $\mathcal{C}_{pq} = \mathcal{C}_p \mathcal{C}_q$
4. If $p \neq 0$ then $\mathcal{C}_{p^{-1}} = (\mathcal{C}_p)^{-1}$
5. $p < q \in \mathbb{Q} \iff \mathcal{C}_p < \mathcal{C}_q \in \mathbb{R}$

*Proof*: HW.

**Definition**: For $q \in \mathbb{Q}$ we say $\mathcal{C}_q \in \mathbb{R}$. Then $\mathbb{Q} \subseteq \mathbb{R}$.

**Theorem**: There exists an ordered field satisfying the least upper bound property; $\mathbb{R}$ is unique (for any ordered field $\mathbb{F}$ satisfying these properties, $\mathbb{F} = \mathbb{R}$ up to isomorphism; and $\mathbb{R}$ is Archimedean.

*Proof*: The basic assertion is Steps (0)-(4). Step (5) proves 1, Step (6) proves 3.

## 1.6 Properties of $\mathbb{R}$

Notation: think of $\mathbb{R}$ as numbers, not cut notation.

**Proposition**: $\mathbb{R}$ satisfies the following:

**Theorem**: For $p, q \in \mathbb{Q}$, the following are true:

1. $\mathbb{R}$ is Archimedean: $\forall x \in \mathbb{R}, x > 0.\ \exists n \in \mathbb{N}.\ x < n$
2. $\mathbb{N} \subset \mathbb{R}$ is not bounded above
3. $\inf\{\frac{1}{n} \mid n \in \mathbb{N}, n \geq 1\} = 0$
4. $\forall x \in \mathbb{R}$ the set $B(x) = \{m \in \mathbb{Z} \mid x < m\}$ has a minimum in $\mathbb{Z}$.
5. $\forall x, y \in \mathbb{R}, x < y.\ \exists q \in \mathbb{Q}.\ x < q < y$

*Remarks:*

1. (5) is interpreted as "the density of $\mathbb{Q} \subseteq \mathbb{R}$". Any element $x \in \mathbb{R}$ can be approximated to arbitrary accuracy by elements of $\mathbb{Q}$.
2. (4) allows us to define the integer part of any $x \in \mathbb{R}$. We can set $\lfloor x \rfloor = \min B(x) - 1 \in \mathbb{Z}$. Then $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

Next we show that $\mathbb{R}$ does not have the "holes" we saw in $\mathbb{Q}$.

**Theorem**: Let $x \in \mathbb{R}$ satisfy $x > 0$ and $n \in \mathbb{N}, n \geq 1$. Then $\exists! y \in \mathbb{R}.\ y > 0 \wedge y^n = x$.

*Proof*: The case $n = 1$ is trivial so assume $n \geq 2$.

Set $E = \{z \in \mathbb{R} \mid z > 0 \wedge z^n < x\}$. We want to show $E \neq \varnothing$ and is bounded above. Set $t = \frac{x}{1+x}$; then $0 < t < 1$ and $t < x$. Hence $0 < t^n < t < x$, and so $t \in E$ and $E \neq \varnothing$.

Set $s = 1 + x$. Then $1 < s \wedge x < s \implies x < s < s^n$; so if $z \in E$ then $z^n < x < s^n \implies z < s$. Then $s$ is an upper bound of $E$.

By least upper bound property, $\exists y \in \mathbb{R}.\ y = \sup E$. Since $t \in E$, $0 < t < y$, so $y > 0$. We claim that $y^n < x$ and $y^n > x$ are both impossible (proof is exercise), so $y^n = x$.

**Definition**: Let $n \geq 1$; for $x \in \mathbb{R}, x > 0$, we write $x^{\frac{1}{n}} = y$ where $y^n = x$. We set $0^{\frac{1}{n}} = 0$.

### 1.6.1 Absolute Value

For $x \in \mathbb{R}$, we define the function $|\cdot| : \mathbb{R} \to \{r \in \mathbb{R} \mid r \geq 0\}$:

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

**Proposition** (Properties of $|\cdot|$):

1. $\forall x \in \mathbb{R}.\ |x| \geq 0$ and $|x| = 0 \iff x = 0$
2. $\forall x, y \in \mathbb{R}.\ |x| < y \iff -y < x < y$
3. $\forall x, y \in \mathbb{R}.\ |xy| = |x||y|$
4. $\forall x, y \in \mathbb{R}.\ |x + y| \leq |x| + |y|$ (Triangle Inequality)
5. $\forall x, y \in \mathbb{R}.\ ||x| - |y|| \leq |x - y|$

# 2 Sequences

Let $E$ be a set. Then we may define a sequence $\{a_n\}_{n=l}^{\infty} \subseteq E$ as the set of values $a_n \equiv a(n)$ for some $l \in \mathbb{Z}$ and some function $a : \{n \in \mathbb{Z} \mid n \geq l\} \to E$.

## 2.1 Convergence and Bounds

**Definition**: We say a sequence $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$ converges to $a \in \mathbb{R}$, i.e. $a_n \to a$ as $n \to \infty$ or $\lim_{n\to\infty} a_n = a$, if for every $0 < \epsilon \in \mathbb{R}$, there exists $N \in \{m \in \mathbb{Z} \mid m \geq l\}$ such that $n \geq N \implies |a_n - a| < \epsilon$.

**Definition**: We say a sequence $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$ is bounded iff. $\exists M \in \mathbb{R}, M > 0.\ |a_n| < M\ (\forall n \geq l)$.

**Lemma**: If a sequence converges, then it is bounded.

**Definition**: Given $\{a_n\}, \{b_n\} \subseteq \mathbb{R}$ we define $\{a_n + b_n\} \subseteq \mathbb{R}$ to be the sequence whose elements are $a_n + b_n$. We similary define $\{ca_n\}$ for a fixed $c \in \mathbb{R}$, $\{a_n b_n\}$, and $\{a_n/b_n\}$ where $b_n \neq 0, n \geq l$.

**Theorem** (algebra of convergence): Let $\{a_n\}, \{b_n\} \subseteq \mathbb{R}, c \in \mathbb{R}$, and assume that $a_n \to a, b_n \to b$ as $n \to \infty$. Then the following hold:

1. $a_n + b_n \to a + b$ as $n \to \infty$
2. $ca_n \to ca$ as $n \to \infty$
3. $a_n b_n \to ab$ as $n \to \infty$
4. If $b_n \neq 0$ and $b \neq 0$, then $a_n/b_n \to a/b$ as $n \to \infty$.

*Proof*: (1), (2) are in next week's HW.

(3): Note that $|a_n b_n - ab| = |a_n b_n - ab_n + ab_n - ab| \leq |a_n b_n - ab_n| + |ab_n - ab| = |b_n||a_n - a| + |a||b_n - b|$. Since $b_n \to b$ we know that $\exists M > 0.\ |b_n| < M(\forall n \geq l)$.

Let $\epsilon > 0$. Since $a_n \to a$ and $b_n \to b$ we may choose $N_1$ such that $n \geq N_1 \implies |a_n - a| < \frac{\epsilon}{2M}$; and $N_2$ where $n \geq N_2 \implies |b_n - b| < \frac{\epsilon}{2(1+|a|)}$.

Then set $N = \max(N_1, N_2)$. So if $n \geq N$ we know that $|a_nb_n - ab| \leq |b_n||a_n - a| + |a||b_n - b| < M|a_n - a| + |a||b_n - b| < M \cdot \frac{\epsilon}{2M} + |a| \cdot \frac{\epsilon}{2(1+|a|)} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$.

Since $\epsilon$ was arbitrary, we deduce that $a_nb_n \to ab$.

(4): We know $|\frac{a_n}{b_n} - \frac{a}{b}| = |\frac{a_nb - ab_n}{b_nb}| = |\frac{a_nb - ab + ab - ab_n}{b_nb}| \leq \frac{|a_nb - ab|}{|b_n||b|} + \frac{|ab - ab_n|}{|b||b_n|} = \frac{|a_n - a|}{|b_n|} + \frac{|a|}{|b||b_n|}|b_n - b|$.

Let $\epsilon > 0$. Since $b_n \to b \neq 0$ we know that $\exists N_1$ such that $n \geq N_1 \implies |b_n - b| < \frac{|b|}{2}$. Then $n \geq N \implies 0 < |b| = |b - b_n + b_n| \leq |b - b_n| + |b_n| < \frac{|b|}{2} + |b_n| \implies 0 < \frac{|b|}{2} \leq |b_n| \implies 0 < \frac{1}{|b_n|} < \frac{2}{|b|}$.

Similarly, $a_n \to a \implies \exists N_2. (n \geq N_2 \implies |a_n - a| < \frac{\epsilon}{4}|b|$; and
$b_n \to b \implies \exists N_3. (n \geq N_3 \implies |b_n - b| < \frac{\epsilon|b|^2}{4(1+|a|)}$.

Set $N = \max(N_1, N_2, N_3)$. Then $n \geq N \implies |\frac{a_n}{b_n} - \frac{a}{b}| \leq \frac{|a_n - a|}{|b_n|} + \frac{|a|}{|b_n||b|}|b_n - b| < \frac{2}{|b||a_n - a|} + \frac{2|a|}{|b|^2}|b_n - b| < \frac{\epsilon}{2} + \frac{\epsilon}{2}\frac{|a|}{1+|a|} < \epsilon$.

Since $\epsilon > 0$ was arbitrary, we deduce $\frac{a_n}{b_n} \to \frac{a}{b}$ as $n \to \infty$.

**Lemma**: Let $\{a_n\}_{n=l}^{\infty}$ converge to $a \in \mathbb{R}$. Then $\forall \epsilon > 0. \exists N. m, n \geq N \implies |a_n - a_m| < \epsilon$.

**Definition**: We say $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$ is *Cauchy* iff $\forall \epsilon > 0. \exists N. m, n \geq N \implies |a_n - a_m| < \epsilon$.

**Lemma**: If $\{a_n\}$ is Cauchy, then it's bounded.

*Proof*: Let $\epsilon = 1$. Then $\exists N. m, n \geq N \implies |a_m - a_n| < 1$. Then $n \geq N \implies |a_n - a_N| < 1 \implies |a_n| < |a_n - a_N| + |a_N| < 1 + |a_N|$. Set $M = \max(1 + |a_N|, k)$, where $k = \max\{|a_l|, \ldots, |a_{N-1}|\}$. Then $|a_n| < M (\forall n \geq l)$, and $\{a_n\}$ is bounded.

**Theorem**: Let $\{a_n\} \subseteq \mathbb{R}$. Then $\{a_n\}$ converges $\iff$ $\{a_n\}$ is Cauchy.

*Proof*: $\implies$ is covered by 2nd-previous lemma. We show the converse:

Suppose $\{a_n\}$ is Cauchy. Then $|a_n| < M (\forall n \geq l)$ by the last lemma.

Set $E = \{x \in \mathbb{R} \mid \exists N. n \geq N \implies x < a_n\}$. Note that $-M < a_n (\forall n \geq l)$, and so $-M \in E$ and $E \neq \varnothing$.

Also, $x \in E \implies \exists N_x. n \geq N_x \implies x < a_n < M$, and so $M$ is an upper bound of $E$. By the least upper bound property of $\mathbb{R}$, $\exists a = \sup E \in \mathbb{R}$. We claim that $a_n \to a$ as $n \to \infty$.

Let $\epsilon > 0$. Then since $\{a_n\}$ is Cauchy, $\exists N. m, n \geq N \implies |a_n - a_m| < \frac{\epsilon}{2}$. In particular, $|a_n - a_N| < \frac{\epsilon}{2}$ when $n \geq N$. Then $n \geq N \implies a_N - \frac{\epsilon}{2} < a_n \implies a_N - \frac{\epsilon}{2} \in E \implies a_N - \frac{\epsilon}{2} \leq a$.

If $x \in E$, then $\exists E_x. (n \geq N_x \implies x < a_n < a_N + \frac{\epsilon}{2})$. Hence $a_N + \frac{\epsilon}{2}$ is an upper bound of $E \implies a \leq a_N + \frac{\epsilon}{2}$. Then $|a - a_N| < \frac{\epsilon}{2}$.

But if $n \geq N$, then $|a_n - a| \leq |a_n - a_N| + |a_N - a| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Hence $a_n \to a$.

### 2.1.1 Squeeze Lemma

**Lemma**: Let $\{a_n\}_{n=l}^{\infty}, \{b_n\}, \{c_n\} \subseteq \mathbb{R}$ and suppose that $a_n \to a, c_n \to a$ as $n \to \infty$. If $\exists k \geq l$ such that $a_n \leq b_n \leq c_n (\forall n \geq k)$, then $b_n \to a$ as $n \to \infty$.

Examples:

1. Suppose $a_n \to 0$ and $\{b_n\}$ is bounded, i.e. $|b_n| \leq M(\forall n \geq l)$. Then $|a_n b_n| = |a_n||b_n| \leq |a_n|M$. But $c_n \to 0 \iff |c_n| \to 0$. Then $0 \leq |a_n b_n| \leq |a_n|M$, both sides of which go to 0; and by the squeeze lemma, $|a_n b_n| \to 0 \implies a_n b_n \to 0$.
2. Fix $k \in \mathbb{N}$ with $k \geq 1$. Set $a_n = \frac{1}{n^k}, n \geq 1$. Then $0 \leq \frac{1}{n^k} \leq \frac{1}{n}$, and by squeeze lemma $\frac{1}{n^k} \to 0$.
3. Fix $k \in \mathbb{N}$ with $k \geq 2$. Let $a_n = \frac{1}{k^n}, n \geq 0$. We know $\forall n \in \mathbb{N}.\ n \leq k^n$ (proof by induction). Then $0 \leq \frac{1}{k^n} \leq \frac{1}{n}$, and by squeeze $\frac{1}{k^n} \to 0$.

## 2.2 Monotonicity and limsup, liminf

**Definition**: Let $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$. We say $\{a_n\}$ is:

1. increasing iff. $a_n < a_{n+1}(\forall n \geq l)$,
2. non-decreasing iff. $a_n \leq a_{n+1}(\forall n \geq l)$,
3. decreasing iff. $a_{n+1} < a_n(\forall n \geq l)$,
4. non-increasing iff. $a_{n+1} \leq a_n(\forall n \geq l)$.

We say $\{a_n\}$ is *monotone* iff. it is either non-increasing or non-decreasing.

*Remark*: increasing $\implies$ non-decreasing, decreasing $\implies$ non-increasing.

**Theorem**: Suppose that $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$ is monotone. Then $\{a_n\}$ is bounded iff $\{a_n\}$ is convergent.

*Proof*: $\impliedby$ is done in a previous lemma.

> $\implies$: We'll prove when the sequence is non-decreasing (other case handled by similar argument).
>
> Set $E = \{a_n \mid n \geq l\} \subseteq \mathbb{R}$. Clearly $E \neq \varnothing$. Also, since $\{a_n\}$ is bounded, $E$ is as well (in particular above). By least upper bound property of $\mathbb{R}$, $\exists a = \sup(E) \in \mathbb{R}$. We claim that $a = \lim_{n\to\infty} a_n$.
>
> Let $\epsilon > 0$. Since $a = \sup(E)$ we know that $a - \epsilon$ is not an upper bound of $E$; hence $\exists(N \geq l).\ a - \epsilon < a_N$. Also, since the sequence is non-decreasing, $a_n \leq a_{n+1}(\forall n \geq l)$, and so $n \geq N \implies a_N \leq a_n$. Then $n \geq N \implies a - \epsilon < a_N \leq a_n \leq a$ because $a$ is an upper bound of $E$.
>
> So $n \geq N \implies -\epsilon < a_n - a \leq 0 \implies |a_n - a| < \epsilon$. Since $\epsilon > 0$ was arbitrary, we deduce that $a_n \to a$ as $n \to \infty$.

**Lemma**: Suppose that $\{a_n\}$ is bounded. Set $S_m = \sup\{a_n \mid n \geq m\}$ and $I_m = \inf\{a_n \mid n \geq m\}$. Then $S_m, I_m \in \mathbb{R}$ are well-defined $\forall m \geq l$; $\{S_m\}$ is non-increasing; and $\{I_m\}$ is non-decreasing. Both sequences are bounded.

**Definition**: Suppose $\{a_n\} \subseteq \mathbb{R}$ is bounded. We set $\limsup_{n\to\infty} a_n = \lim_{m\to\infty} S_m \in \mathbb{R}$ and $\liminf_{n\to\infty} a_n = \lim_{m\to\infty} I_m \in \mathbb{R}$. Both limits exist by the lemma and previous theorem. We know that $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$ from HW.

## 2.3 Subsequences

**Definition**: Let $\phi : \{n \in \mathbb{Z} \mid n \geq l\} \to \{n \in \mathbb{Z} \mid n \geq l\}$ be order preserving (increasing), i.e. $m < n$ then $\phi(m) < \phi(n)$. Let $\{a_n\}_{l=k}^{\infty} \subseteq \mathbb{R}$ be a sequence. We say $\{a_{\phi(k)}\}_{k=l}^{\infty}$ is a *subsequence* of $\{a_n\}$.

*Remarks*:

1. $\phi(k) = k$ is order preserving, so every sequence is a subsequence of itself.
2. Not every $a_n$ has to be in the subsequence $\{a_{\phi(k)}\}$.
   For example, if $l = 0$ then $\phi(k) = 2k$ is order preserving. In this case $a_n, n$ odd does not appear in the subsequence $\{a_{\phi(k)}\}$.
3. We will often write $n_k = \phi(k)$ to simplify notation, so $\{a_{n_k}\}$ denotes a subsequence.
4. From HW1, we know $k \le \phi(k)$ $(\forall k \ge l)$.

*Proposition*: Suppose $\{a_n\}$ satisfies $a_n \to a \in \mathbb{R}$ as $n \to \infty$. Then any subsequence of $\{a_n\}$ also converges to $a$.

*Proof*:

Let $\{a_{\phi(k)}\}$ be a subsequence of $\{a_n\}$. Let $\epsilon > 0$. Since $a_n \to a$ as $n \to \infty$, we know $\exists N \ge l.\ n \ge N \implies |a_n - a| < \epsilon$. We claim $\exists K \ge l.\ k \ge K \implies \phi(k) \ge N$.

If not, then $\phi(k) < N(\forall k \ge l)$; but $k \le \phi(k) < N(\forall k \ge l)$ is a contradiction. Then the claim is true, and $k \ge K \implies \phi(k) \ge N \implies |a_{\phi(k)} - a| < \epsilon$. Since $\epsilon > 0$ was arbitrary, we deduce $\{a_{\phi(k)}\} \to a$ as $k \to \infty$.

*Remark*: Converse fails. Example: $a_n = (-1)^n$; $a_{2n} = +1 \to +1$, but $a_{2n+1} = -1 \to -1$.

### 2.3.1 Limsup Theorem

**Theorem**: Let $\{a_n\} \subseteq \mathbb{R}$ be bounded. The following hold:

1. Every subsequence of $\{a_n\}$ is bounded.
2. If $\{a_{n_k}\}$ is a subsequence, then $\limsup_{k\to\infty} a_{n_k} \le \limsup_{n\to\infty} a_n$.
3. If $\{a_{n_k}\}$ is a subsequence, then $\liminf_{n\to\infty} a_n \le \liminf_{k\to\infty} a_{n_k}$.
4. There exists a subsequence $\{a_{n_k}\}$ such that $\lim_{k\to\infty} a_{n_k} = \limsup_{n\to\infty} a_n$.
5. There exists a subsequence $\{a_{n_k}\}$ such that $\lim_{k\to\infty} a_{n_k} = \liminf_{n\to\infty} a_n$ $(\neq (4))$.

*Proof*:

1. Trivial.
2. Since $k \le \phi(k)$, $\{a_{\phi(n)} \mid n \ge k\} \subseteq \{a_n \mid n \ge k\}$ for every order-preserving $\phi$. Hence $S_k = \sup\{a_{\phi(n)}\} \mid n \ge k\} \subseteq \sup\{a_n \mid n \ge k\} = T_k$. But:
   $\limsup_{n\to\infty} a_{\phi(n)} = \limsup_{k\to\infty}\{a_{\phi(n)} \mid n \ge k\} \le \limsup_{k\to\infty}\{a_n \mid n \ge k\} = \limsup_{n\to\infty} a_n$.
3. Similar to (2); exercise to reader.
4. Too lazy to LaTeX; exercise to reader.
5. Exercise to reader.

**Theorem**: Suppose $\{a_n\} \subseteq \mathbb{R}$; the following are equivalent:

1. $a_n \to a$ as $n \to \infty$
2. $\{a_n\}$ is bounded, and every convergent subsequence converges to $a$.
3. $\{a_n\}$ is bounded, and $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n$.

*Proof*: (1) $\implies$ (2) proven already.

(2) $\implies$ (3)
Limsup theorem (4,5) $\implies \exists\{a_{\phi(k)}\},\{a_{\gamma(k)}\}$ subsequences such that $a_{\phi(k)} \to \limsup_{n\to\infty} a_n, a_{\gamma(k)} \to \liminf_{n\to\infty} a_n$ as $k \to \infty$. By (2) the limits must agree.

(3) $\implies$ (1)
Limsup theorem (1-3) $\implies \forall\{a_{\phi(k)}\}$. $\liminf_{n\to\infty} a_n \le \liminf_{k\to\infty} a_{\phi(k)} \le \limsup_{k\to\infty} a_{\phi(k)} \le \limsup_{n\to\infty} a_n$. As the first and last are equal, by transitivity it follows all subsequences satisfy $\liminf_{k\to\infty} a_{\phi(k)} = \limsup_{k\to\infty} a_{\phi(k)}$. As $a_n$ is a subsequence of itself, it therefore converges to some $a$ as $n \to \infty$.

**Theorem** (Bolzano-Weierstrass): If $\{a_n\} \subseteq \mathbb{R}$ is bounded then there exists a convergent subsequence. Proof from (4) or (5) of Limsup Theorem.

## 2.4 Special Sequences

**Definition**: Given $a_n \in \mathbb{R}$ for $0 \le k \le n, n \in \mathbb{N}$ we define $\sum_{k=0}^n a_n = a_0 + a_1 + \cdots + a_n$.

**Lemma** (Binomial Theorem): Let $x, y \in \mathbb{R}$ and $n \in \mathbb{N}$. Then $(x+y)^n = \sum\limits_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$, where $\binom{n}{k} := \frac{n!}{k!(n-k)!} \in \mathbb{N}$.

**Theorem**: In the following assuming that $n \ge 1$:

1. Let $x \in \mathbb{R}, x > 0$. Then $a_n = \frac{1}{n^x} \to 0$ as $n \to \infty$.
2. Let $x \in \mathbb{R}, x > 0$. Then $a_n = x^{1/n} \to 1$ as $n \to \infty$.
3. Let $a_n = n^{1/n}$; then $a_n \to 1$ as $n \to \infty$.
4. Let $a, x \in \mathbb{R}, x > 0$. Then $\frac{n^a}{(1+x)^a} \to 0$ as $n \to \infty$.
5. Let $x \in \mathbb{R}, |x| < 1$. Then $a_n = x^n \to 0$ as $n \to \infty$.

# 3 Series

**Definition**: Let $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$; for $p < q$ we write $\sum_{n=p}^{q} a_n = (a_p + \cdots + a_q)$.

1. We define, for each $n \ge l$, $S_n = \sum\limits_{k=l}^{n} a_k \in \mathbb{R}$ to be the $n^{\text{th}}$ partial sum of $\{a_n\}_{n=l}^{\infty}$.
2. If $\exists s \in \mathbb{R}$. $S_n \to s$ as $n \to \infty$, then $\sum_{n=l}^{\infty} a_n = s$. We say the "infinite series" $\sum_{n=l}^{\infty} a_n$ converges.
3. If the series does not converge, it diverges.

**Examples**

1. Let $a_n = x^n$ for $n \ge 0, x \in \mathbb{R}$. Then $S_n = \sum_{k=0}^n x^k$. Notice that $(1-x)S_n = \sum_{k=0}^n x^k - \sum_{k=0}^n x^{k+1} = \sum_{k=0}^n x^k - \sum_{k=1}^{n+1} x^k = 1 - x^{n+1}$.

   So $S_n = \sum_{k=0}^n x^k = (\frac{1-x^{n+1}}{1-x})$. If $|x| < 1$ then $S_n \to \frac{1}{1-x}$ by special seq (5).

2. Suppose $\{b_n\}_{n=0}^{\infty} \subseteq \mathbb{R}$ where $b_n \to b$ as $n \to \infty$. Set $a_n = b_{n+1} - b_n$ for $n \ge 0$. Then the series $\sum_{n=0}^{\infty} a_n$ converges and in fact $\sum_{n=0}^{\infty} = b - b_0$.

## 3.1 Convergence Results

We develop tools that will let us deduce the convergence of a series without knowing its value.

**Theorem**: Suppose $\sum_{n=l}^{\infty} a_n$ converges. Then $a_n \to 0$ as $n \to \infty$.

*Proof*: Notice that $a_n = S_n - S_{n-1}$ and so $\lim_{n\to\infty} a_n = \lim_{n\to\infty}(S_n - S_{n-1}) = S - S = 0$.

**Corollary**: $\sum_{n=0}^{\infty}(-1)^n$ and $\sum_{n=0}^{\infty} n$ diverge, as neither sequences converge to 0.

**Corollary**: The series $\sum_{n=0}^{\infty} x^n$ converges $\iff |x| < 1$.

*Proof*: $|x| \geq 1 \implies |x^n| = |x|^n \geq 1 (\forall n \in \mathbb{N})$. The converse was proved last time.

Next, we provide a characterization of convergence in terms of the size of the "tails" of the series.

**Theorem**: $\sum_{n=l}^{\infty} a_n$ converges $\iff \forall \epsilon > 0.\ \exists N \geq l.\ m \geq k \geq N \implies |\sum_{n=k}^{m} a_n| < \epsilon$.

*Proof*: $\sum_{n=l}^{\infty} a_n$ converges $\iff S_k = \sum_{n=l}^{k} a_n$ converges $\iff \{S_k\}$ is Cauchy.

This is useful in practice because we can guarantee a series converges without knowing its value.

**Theorem**:

1. If $\forall n \geq k.\ |a_n| \leq b_n$ for some $k \geq l$, and $\sum\limits_{n=l}^{\infty} b_n$ converges, then $\sum\limits_{n=l}^{\infty} a_n$ converges.
2. If $\forall n \geq k.\ 0 \leq a_n \leq b_n$ for some $k \geq l$, and $\sum\limits_{n=l}^{\infty} a_n$ diverges, then $\sum\limits_{n=l}^{\infty} b_n$ diverges.

*Proof*: (1) Let $\epsilon > 0$ and prove with previous theorem and induction on triangle inequality. (2) follows from contrapositive.

**Examples**:

1. $\sum_{n=0}^{\infty} \frac{(-1)^n}{2^n}$ converges because $|\frac{(-1)^n}{2^n}| = \frac{1}{2^n}$ and $\sum_{n=0}^{\infty} \frac{1}{2^n}$ converges ($\frac{1}{2} < 1$).
2. Suppose $\sum_{n=0}^{\infty} a_n$ converges and $a_n \geq 0\ \forall n \geq 0$. Let $\{b_n\} \subseteq \mathbb{R}$ be bounded, i.e. $|b_n| \leq M \forall n$. Then $|a_n b_n| = |a_n||b_n| \leq M a_n$. Then $MS_n = M \sum_{k=0}^{n} a_n = \sum_{k=0}^{n} M a_n$, so by the theorem, $\sum_{n=0}^{\infty} a_n b_n$ converges.
3. $\sum_{n=0}^{\infty} \frac{(-1)^n}{2^n} \cdot \frac{n!}{n^n} \cdot \frac{3n^2}{4n^2+2}$ converges because the product is bounded.

**Theorem**: Suppose $\forall n \geq l.\ a_n \geq 0$. Then $\sum_{n=l}^{\infty} a_n$ converges $\iff \{S_n\}_{n=l}^{\infty}$ is bounded.

*Proof*: Since $a_n \geq 0$, the sequence $S_n = \sum_{k=l}^{n} a_k$ is non-decreasing: $S_{n+1} = a_{n+1} + S_n \geq S_n$. Since $S_n$ is monotone and converges, it is bounded.

### 3.1.1 Cauchy Criterion Theorem

**Theorem**: Suppose that $\{a_n\}_{n=1}^{\infty} \subseteq \mathbb{R}$ satisfies $\forall n \geq l.\ a_n \geq 0$ and $\forall n \geq 1.\ a_{n+1} \leq a_n$. Then $\sum\limits_{n=1}^{\infty} a_n$ converges $\iff \sum\limits_{n=0}^{\infty} 2^n a_{2^n}$ converges.

*Proof*:

Let $S_n = \sum_{k=1}^{n} a_k$ and $T_n = \sum_{n=0}^{m} 2^n a_{2^n}$. Notice that if $m \leq 2^k$ then $S_m = a_1 + a_2 + \cdots + a_{2^k} \leq a_1 + (a_2 + a_3) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \leq a_1 + 2a_2 + \cdots + 2^k a_{2^k} = T_k$.

On the other hand, if $m \geq 2^k$, $S_m \geq a_1+\cdots+a_{2^k} = a_1+a_2+(a_3+a_4)+\cdots+(a_{2^{k-1}-1}+\cdots+a_{2^k}) \geq \frac{1}{2}a_1 + a_2 + \cdots + 2^{k-1}a_{2^k} = \frac{1}{2}T_k$.

Now, if $\sum_{n=0}^{\infty} 2^n a_{2^n}$ converges, then $T_n \to T$ as $n \to \infty$ and so $S_m \leq \lim_{n\to\infty} T_m = T$, which means $\{S_m\}$ is bounded and $\sum_{n=1}^{\infty} a_n$ converges.

Similarly, if $\sum_{n=1}^{\infty} a_n$ converges, then $T_k \leq 2\lim_{n\to\infty} S_n \implies \{T_k\}$ is bounded $\implies$ $\sum_{n=0}^{\infty} 2^n a_{2^n}$ converges.

**Theorem**: Let $p \in \mathbb{R}$. Then $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges $\iff p > 1$.

*Proof*:

If $p \leq 0$ the result is trivial since $\frac{1}{n^p} \geq 1$ (the sequences converges to 0). Assume that $p > 0$. Then $\frac{1}{(n+1)^p} \leq \frac{1}{n^p}$, so we can apply the Cauchy criterion:

$$\sum_{n=1}^{\infty} \frac{1}{n^p} \text{ converges } \iff \sum_{n=0}^{\infty} \frac{2^n}{(2^n)^p} \text{ converges.}$$

But $\sum_{n=0}^{\infty} \frac{2^n}{(2^n)^p} = \sum_{n=0}^{\infty} \frac{1}{(2^{p-1})^n}$, and this series converges $\iff \frac{1}{2^{p-1}} < 1 \iff p > 1$.

Notice $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent, but $\sum_{n=1}^{\infty} \frac{1}{n^{1+r}}$ converges $\forall r > 0$. To try to find intermediate series, we need the logarithm.

### 3.1.2 Logarithm

**Definition**: From Supplemental Reading 3, for every $1 < b \in \mathbb{R}$, we define a function $\log_b : \{x \in \mathbb{R} \mid x > 0\} \to \mathbb{R}$ such that

1. $b^{\log_b x} = x$ $(\forall x > 0)$
2. $\log_b(1) = 0,\ \log_b b = 1$
3. $0 < x < y \iff \log_b x < \log_b y$
4. $\log_b(x^z) = z\log_b(x)$ $(\forall x > 0, \forall z \in \mathbb{R})$
5. $\log_b$ is a bijection
6. $\lim\limits_{n\to\infty} \frac{\log_b n}{n^r} = 0$ $(\forall r \in \mathbb{R}, r > 0)$

Then from (6), for large $n$ and $p > 0$ we know:
$n \leq n(\log_b n)^p \leq n \cdot n^p = n^{1+p} \implies \frac{1}{n^{1+p}} \leq \frac{1}{n(\log_b n)^p} \leq \frac{1}{n}$.

So $\frac{1}{n(\log_b n)^p}$ is such an "intermediate series."

**Theorem**: Let $b > 1$. $\sum\limits_{n=2}^{\infty} \frac{1}{n(\log_b n)^p}$ converges $\iff p > 1$. $(n \geq 2 \implies \log_b n > 0)$

*Proof*:

$\sum\limits_{n=2}^{\infty} \frac{1}{n(\log_b n)^p}$ converges $\iff \sum\limits_{n=1}^{\infty} \frac{2^n}{2^n(\log_b 2^n)^p}$ converges by Cauchy criterion, but $\sum\limits_{n=1}^{\infty} \frac{1}{(\log_b 2)^p n^p} = \frac{1}{(\log_b 2)^p} \sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ converges $\iff p > 1$.

In particular, $\sum\limits_{n=2}^{\infty} \frac{1}{n \log_b n}$ is divergent.

## 3.2 The number $e$

**Lemma**: $\sum_{n=0}^{\infty} \frac{1}{n!}$ converges.

*Proof*: If $n \geq 2$ then:

$$\begin{aligned} S_n = \sum_{k=0}^{n} \frac{1}{k!} &= 1 + 1 + \frac{1}{2 \cdot 1} + \cdots + \frac{1}{n(n-1)\cdots 2 \cdot 1} \\ &\leq 1 + 1 + \frac{1}{2} + \frac{1}{2 \cdot 2} + \cdots + \frac{1}{2^{n-1}} \\ &\leq 1 + \sum_{k=0}^{\infty} \frac{1}{2^k} = 1 + 2 = 3 \end{aligned}$$

Since $S_n$ is increasing and bounded, we know that $\sum_{n=0}^{\infty} \frac{1}{n!}$ converges.

**Definition**: We set $e = \sum\limits_{n=0}^{\infty} \frac{1}{n!}$. Note that $e > 1$.

**Theorem**: $e = \lim_{n\to\infty}(1 + \frac{1}{n})^n$.

*Proof*: Let $S_n = \sum_{k=0}^{n} \frac{1}{k!}, T_n = (1 + \frac{1}{n})^n$. Then by the Binomial Theorem:

$$\begin{aligned} T_n = (1 + \frac{1}{n})^n &= \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \frac{1}{n^k} \\ &= 1 + 1 + \frac{1}{2!}\frac{n(n-1)}{n^2} + \cdots + \frac{1}{n!}\frac{n(n-1)\cdots 1}{n^n} \\ &= 1 + 1 + \frac{1}{2!}(1 - \frac{1}{n}) + \frac{1}{3!}(1 - \frac{1}{n})(1 - \frac{2}{n}) + \cdots + \frac{1}{n!}(1 - \frac{1}{n})\cdots(1 - \frac{n-1}{n}) \\ &\leq 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} = S_n \end{aligned}$$

Hence, $\limsup_{n\to\infty} T_n \leq \limsup_{n\to\infty} S_n = \lim_{n\to\infty} S_n = e$.

OTOH, fix $m \in \mathbb{N}$. Then for $n \geq m$:

$$T_n \geq 1 + 1 + \frac{1}{2!}(1 - \frac{1}{n}) + \cdots + \frac{1}{m!}(1 - \frac{1}{n})\cdots(1 - \frac{m-1}{n})$$

$$\implies \liminf_{n\to\infty} T_n \geq \liminf_{n\to\infty} \text{RHS} \geq 1 + 1 + \frac{1}{2!}\liminf_{n\to\infty}(1 - \frac{1}{n}) + \cdots + \frac{1}{m!}\liminf_{n\to\infty}(1 - \frac{1}{n}\cdots(1 - \frac{m-1}{n}) \quad = 1 + 1 +$$

Then, letting $m \to \infty$, $e = \lim_{m\to\infty} S_m \leq \liminf_{n\to\infty} T_n$.

Thus, $e \leq \liminf_{n\to\infty} T_n \leq \limsup_{n\to\infty} T_n \leq e \implies \lim_{n\to\infty} T_n = e$.

**Theorem**: $\forall n \geq 1$. $0 < e - S_n < \frac{1}{n \cdot n!}$. Also, $e \in \mathbb{R}\backslash\mathbb{Q}$ is irrational.

*Proof*: Since $S_n$ is increasing, $0 < e - S_n$ is clear. The other side can be seen from algebra.

Now, suppose $e \in \mathbb{Q}$; then $e = \frac{p}{q}$ for $p, q \in \mathbb{N}, p, q \geq 1$.

Then $0 < q!(e - S_q) < \frac{1}{q}$ $(\forall q \geq 1)$. Notice that $q!e = q!\frac{p}{q} = (q-1)!p \in \mathbb{N}$ and $q!(1 + \frac{1}{2!} + \cdots + \frac{1}{q!}) \in \mathbb{N}$.

Hence $q!(e - S_q) \in \mathbb{Z}$; but this yields an integer between 0 and 1, a contradiction. So $e$ is irrational.

*Remark*: In fact, $e$ is transcendental.

## 3.3 More Convergence Results

**Theorem (Root Test)**: Suppose $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$ and $\{|a_n|^{1/n}\}$ is bounded. Let $0 \leq \alpha = \limsup_{n\to\infty} |a_n|^{1/n}$. Then the following holds:

1. If $\alpha < 1$, then $\sum_{n=l}^{\infty} a_n$ converges.
2. If $\alpha > 1$, then $\sum_{n=l}^{\infty} a_n$ diverges.
3. if $\alpha = 1$, both convergence and divergence are possible.

**Theorem (Ratio Test)**: Let $\{a_n\}_{n=l}^{\infty} \subseteq \mathbb{R}$. Then $\sum_{n=l}^{\infty} a_n$:

1. converges if $\{|\frac{a_{n+1}}{a_n}|\}_{n=l}^{\infty}$ is bounded and $\limsup_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} < 1$.
2. diverges if $\exists k \geq l.\ |a_k| \neq 0$ and $|a_{n+1}| \geq |a_n| (\forall n \geq k)$.

**Lemma (Summation of Parts)**: Let $\{a_n\}_{n=0}^{\infty} \subseteq \mathbb{R}$ and define:

$$A_n = \begin{cases} \sum_{k=0}^{n} a_k & \text{if } n \geq 0 \\ 0 & \text{if } n = -1 \end{cases}$$

Then if $0 \leq p < q$:

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n(b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

**Theorem (Dirichlet Test)**: Suppose $\{a_n\}_{n=0}^{\infty}, \{b_n\}_{n=0}^{\infty} \subseteq \mathbb{R}$ satisfy:

1. The sequence $A_n = \sum_{k=0}^{n} a_k$ is bounded.
2. $0 \leq b_{n+1} \leq b_n (\forall n \in \mathbb{N})$
3. $\lim_{n\to\infty} b_n = 0$

Then $\sum_{n=0}^{\infty} a_n b_n$ converges.

**Corollary (Alternating Series)**: Suppose $0 \leq a_{n+1} \leq a_n, a_n \to 0$ as $n \to \infty$. Then $\sum_{n=l}^{\infty} (-1)^n a_n$ converges. Proof follows from Dirichlet Test.

**Corollary (Abel's Test)**: Suppose $\sum_{n=l}^{\infty} a_n$ converges, $b_{n+1} \leq b_n (\forall n \geq l)$ and $b_n \to b$ as $n \to \infty$. Then $\sum_{n=l}^{\infty} a_n b_n$ converges.

## 3.4 Algebra of Series

**Theorem**: If $A = \sum_{n=l}^{\infty} a_n, B = \sum_{n=l}^{\infty} B - N$, then

$$(1) A + B = \sum_{n=l}^{\infty} (a_n + b_n) \qquad (2) cA = \sum_{n=l}^{\infty} ca_n \ (\forall c \in \mathbb{R})$$

**Theorem**: Suppose $\{a_n\}_{n=0}^{\infty}, \{b_n\}_{n=0}^{\infty} \in \mathbb{R}$ satisfy:

$$(1) \sum_{n=0}^{\infty} |a_n| \text{ converges} \qquad (2) \sum_{n=0}^{\infty} b_n = B \qquad (3)\ c_n = \sum_{k=0}^{n} a_k b_{n-k} \text{ for } n \geq 0$$

Then $\sum_{n=0}^{\infty} c_n = A \cdot B$ converges.

**Definition**: The series $\sum\limits_{n=0}^{\infty} c_n$, where $c_n = \sum\limits_{k=0}^{n} a_k b_{n-k}$, is called the *Cauchy product* of the series $\sum\limits_{n=0}^{\infty} a_n$, $\sum\limits_{n=0}^{\infty} b_n$.

*Remark*: If $\sum a_n, \sum b_n$ converge, $\sum c_n$ does not necessarily converge if neither series has convergent absolute values.

## 3.5 Absolute Convergence and Rearrangements

**Proposition**: If $\sum_{n=l}^{\infty} |a_n|$ converges, then $\sum_{n=l}^{\infty} a_n$ converges. Proof is trivial.

**Definition**: Suppose $\sum_{n=l}^{\infty} a_n$ converges. If $\sum_{n=l}^{\infty} |a_n|$ converges, the series converges *absolutely*. If $\sum |a_n|$ diverges, the series is *conditionally convergent*.

*Example*: $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ is conditionally convergent, while $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2}$ is absolutely convergent.

Let's try to manipulate the series without being careful.

$$\begin{aligned}
\gamma &= \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots \\
&= \lim_{k\to\infty} (S_k = \sum_{n=0}^{k} \frac{(-1)^{n+1}}{n}) = \lim_{k\to\infty} (S_{2k} = \sum_{n=0}^{2k} \frac{(-1)^{n+1}}{n}) \\
\text{but: } S_{2k} &= (1 - \frac{1}{2}) + (\frac{1}{3} - \frac{1}{4} + \cdots + (\frac{1}{2k-1} - \frac{1}{2k}) > 0
\end{aligned}$$

Hence, $\gamma > 0$. But the next step is questionable:

$$\begin{aligned}
2\gamma &= \sum_{n=1}^{\infty} \frac{(2)(-1)^{n+1}}{n} \stackrel{?}{=} \sum_{k=0}^{\infty} \frac{2}{2k+1} - \sum_{k=1}^{\infty} \frac{2}{2k} \\
&\stackrel{?}{=} \sum_{k=0}^{\infty} \frac{2}{2k+1} - \sum_{k=1}^{\infty} \frac{1}{k} = \sum_{k=0}^{\infty} \frac{1}{2k+1} - \sum_{k=1}^{\infty} \frac{1}{2k} = \gamma \\
&\implies 2\gamma = \gamma \wedge \gamma > 0 \quad \text{a contradiction!}
\end{aligned}$$

Problem: rearrangement is a delicate issue.

**Definition**: Let $\gamma : \{m \in \mathbb{Z} \mid m \geq l\} \to \{m \in \mathbb{Z} \mid m \geq l\}$ be a bijection. The series $\sum_{n=l}^{\infty} a_{\gamma(n)}$ is called a rearrangement of $\sum_{n=l}^{\infty} a_n$.

**Theorem**: If $\sum_{n=l}^{\infty} a_n$ is absolutely convergent, then every rearrangement converges to $\sum_{n=l}^{\infty} a_n$.

*Proof*: Let $\epsilon > 0$.

Since $\sum_{n=l}^{\infty} a_n$ converges absolutely, $\exists N \geq l.\ k \geq m \geq N \implies \sum_{n=m}^{k} |a_n| < \frac{\epsilon}{2}$.
Let $k \to \infty : \sum_{n=m}^{\infty} |a_n| \leq \frac{\epsilon}{2} < \epsilon$.
Now choose $M \geq N$ such that $\{l, l+1, \ldots, N\} \subseteq \{\gamma(l), \gamma(l+1), \ldots, \gamma(M)\}$. Then $m \geq M \implies |\sum_{n=l}^{m} a_n - \sum_{n=l}^{m} a_{\gamma(n)}| \leq \sum_{n=N}^{\infty} |a_n| < \epsilon$.

Hence $\lim_{m\to\infty}(\sum_{n=l}^{m} a_n - \sum_{n=l}^{\infty} a_{\gamma(n)}) = 0$ and from this we deduce
$\lim_{m\to\infty} \sum_{n=l}^{m} a_{\gamma(n)} = \lim_{m\to\infty} \sum_{n=l}^{m} a_n = \sum_{n=l}^{\infty} a_n$.

When a series is only conditionally convergent, the situation is vastly worse.

**Theorem**: Suppose $\sum_{n=0}^{\infty} a_n$ is conditionally convergent. Let $c \in \mathbb{R}$.
There exists a rearrangement (bijection) $\gamma : \mathbb{N} \to \mathbb{N}$ such that $\sum_{n=0}^{\infty} a_{\gamma(n)} = c$.

**Lemma**: Suppose $\sum_{n=0}^{\infty} a_n$ is conditionally convergent and set:

$$b_n = \begin{cases} a_n & \text{if } a_n > 0 \\ 0 & \text{if } a_n \leq 0 \end{cases} \qquad c_n = \begin{cases} -a_n & \text{if } a_n < 0 \\ 0 & \text{if } a_n \geq 0 \end{cases}$$

Then $\sum_{n=0}^{\infty} b_n$ and $\sum_{n=0}^{\infty} c_n$ both diverge.

*Proof*: Suppose not; one of the series is convergent. If $\sum b_n$ converges, then $c_n = b_n - a_n \implies \sum c_n = \sum b_n - \sum a_n$; but $|a_n| = b_n + c_n$ and so $\sum |a_n| = \sum b_n + \sum c_n$ is convergent, a contradiction. A similar argument holds if $\sum c_n$ converges.

**Rearrangement Theorem Proof**:

Let $\{a_n^+\}_{n=0}^{\infty}$ denote the subsequence of $\{b_n \mid b_n > 0 \text{ or } b_n = 0 \wedge a_n = 0\}$. Let $\{a_n^-\}_{n=0}^{\infty}$ denote the subsequence of $\{c_n \mid c_n > 0\}$ (from last lemma). Note:

1. $a_n^+ \to 0, a_n^- \to 0$ since $a_n \to 0 \implies b_n \to 0, c_n \to 0$.
2. $\sum a_n^+$ and $\sum a_n^-$ both diverge because they differ by 0 from $\sum b_n, \sum c_n$ respectively.

Set $m_0 = n_0 = -1$. Since $\sum a_n^+$ diverges we may use the well-ordering principle: $\exists m_1 = \min\{k \in \mathbb{N} \mid \sum_{n=0}^{k} a_n^+ > c\}$. Similarly, $\exists n_1 = \min\{k \in \mathbb{N} \mid \sum_{n=0}^{m_1} a_n^+ - \sum_{n=0}^{k} a_n^- < c\}$.

Next, if $m_p$ and $n_p$ are known, we set:

$$m_{p+1} = \min \left\{ k \in \mathbb{N} \mid \sum_{l=0}^{p-1} \sum_{j=1+m_l}^{m_l} a_j^+ - \sum_{l=0}^{p-1} \sum_{j=1+n_l}^{n_l} a_j^- + \sum_{j=1+m_p}^{k} a_j^+ > c \right\}$$

$$n_{p+1} = \min \left\{ k \in \mathbb{N} \mid \sum_{l=0}^{p-1} \sum_{j=1+m_l}^{m_l} a_j^+ - \sum_{l=0}^{p-1} \sum_{j=1+n_l}^{n_l} a_j^- + \sum_{j=1+m_p}^{m_{p+1}} a_j^+ - \sum_{j=1+n_p}^{k} a_j^- < c \right\}$$

Consider the series $(a_1^+ + \cdots + a_{m_1}^+) - (a_1^- + \cdots + a_{n+1}^-) + (a_{1+m_1}^+ + \cdots + a_{m+2}^+) - (a_{1+n_1}^- + \cdots + a_{n_2}^-) + \cdots$. This is clearly a rearrangement of $\sum_{n=0}^{\infty} a_n$.

Write $A_p = \sum_{l=1+m_p}^{m_{p+1}} a_l^+, A_p^- = \sum_{l=1+n_p}^{n_{p+1}} a_l^-$, and let $S_j$ denote the $j^{\text{th}}$ partial sum of the rearrangement.

By construction, $\limsup_{j\to\infty} S_j = \limsup_{p\to\infty} (\sum_{l=0}^{p+1} A_l^+ - \sum_{l=0}^{p} A_l^-)$ and
$\liminf_{j\to\infty} S_j = \liminf_{p\to\infty} (\sum_{l=0}^{p} A_l^+ + \sum_{l=0}^{p} A_l^-)$.

Also, $c < \sum_{l=0}^{p+1} A_l^+ - \sum_{l=0}^{p} A_l^- < c + a_{m_{p+1}}^+$ and $c - a_{n_{p+1}}^- < \sum_{l=0}^{p+1} A_l^+ - \sum_{l=0}^{p+1} A_l^- < c$.

Thus, by the squeeze lemma, $\lim_{p\to\infty} (\sum_{l=0}^{p+1} A_l^+ - \sum_{l=0}^{p} A_l^-) = \lim_{p\to\infty} (\sum_{l=0}^{p} A_l^+ - \sum_{l=0}^{p} A_l^-) = c$, and so $\lim_{j\to\infty} S_j = c \implies \sum_{n=0}^{\infty} a_{\gamma(n)} = c$.

*Remark*: One can also rearrange such that $\sum a_{\gamma(n)} = \pm\infty$.

# 4 Topology of $\mathbb{R}$

Our goal in Section 4 is to develop some tools for understanding the "topology" of $\mathbb{R}$, which is a sort of generalized qualitative geometry.

## 4.1 Open and Closed Sets

### 4.1.1 Open Sets

**Definition**:

1. For $a, b \in \mathbb{R}$ with $a \leq b$, we define:
$$(a,b) = \{x \in \mathbb{R} \mid a < x < b\} \qquad [a,b) = \{x \in \mathbb{R} \mid a \leq x < b\}$$
$$(a,b] = \{x \in \mathbb{R} \mid a < x \leq b\} \qquad [a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$$
2. For $x \in \mathbb{R}$ and $\epsilon > 0$, we set $B(x,\epsilon) = (x-\epsilon, x+\epsilon)$ and $B[x,\epsilon] = [x-\epsilon, x+\epsilon]$. We call the set $B(x,\epsilon)$ a *neighborhood* of $x$ or a "ball of radius $\epsilon$ centered at $x$".
3. A set $E \subseteq \mathbb{R}$ is *open* if $\forall x \in E.\ \exists \epsilon > 0.\ B(x,\epsilon) \subseteq E$. In other words, every point in $E$ has a neighborhood contained in $E$.

*Examples*:

1. $\varnothing$ is vacuously open.
2. $\mathbb{R}$ is open because $\forall x \in \mathbb{R}.\ B(x,1) \subseteq \mathbb{R}$.
3. If $a < b$ then $(a,b)$ is open.
*Proof*: Fix $x \in (a,b)$ and let $\epsilon = \min\{x-a, b-x\} > 0$. Then $a \leq x - \epsilon < x < x + \epsilon \leq b$ by construction, and $B(x,\epsilon) \subseteq (a,b)$.
4. If $a < b$ then $[a,b)$ is not open.
*Proof*: For $x = a$ we know that $\forall \epsilon > 0.\ a - \epsilon \notin [a,b)$ and hence $B(a,\epsilon) \not\subseteq [a,b)$.
5. $[a,b]$ is not open, nor is $(a,b]$ by previous argument.
6. $E = \{a\}$ is not open.
7. $E = \{\frac{1}{n} \mid n \in \mathbb{N}, n \geq 1\}$ is not open: $\forall \epsilon > 0.\ B(1,\epsilon) \not\subseteq E$.

**Lemma**: If $E_\alpha \subseteq \mathbb{R}$ is open $\forall \alpha \in A$ (some index set), then $\bigcup_{\alpha \in A} E_\alpha$ is open.

*Proof*: Let $x \in \bigcup_{\alpha \in A} E_\alpha$. Then $x \in E_{\alpha_0}$ for some $\alpha_0 \in A$. Since $E_{\alpha_0}$ is open, $\exists \epsilon > 0.\ B(x,\epsilon) \subseteq E_{\alpha_0} \subseteq \bigcup_{\alpha \in A} E_\alpha$.

**Lemma**: If $E_i \subseteq \mathbb{R}$ is open for $i \in [n], n \in \mathbb{N}$, then $\bigcap_{i=1}^n E_i$ is open.

*Remark*: Infinite intersections of open sets need not be open. Let $E_n = (\frac{-1}{n}, \frac{1}{n}), n \geq 1$.
Then $\bigcap_{n=1}^\infty E_n = \{0\}$ which is closed.

### 4.1.2 Closed Sets

**Definition**: We say $E \subseteq \mathbb{R}$ is *closed* iff $E^c = \mathbb{R} \backslash E$ is open.

**Lemma**: $E$ is open $\iff$ $E^c$ is closed (by definition).

*Examples*:

1. $\varnothing$ is closed because $\varnothing^c = \mathbb{R}$ is open.
2. $\mathbb{R}$ is closed because $\mathbb{R}^c = \varnothing$ is open.
3. $[a, b]$ is closed because $[a, b]^c = (-\infty, a) \cup (b, \infty)$ is the union of open sets, and thus open.
4. $[a, b)$ and $(a, b]$ are not closed because $[a, b)^c = (-\infty, a) \cup [b, \infty)$ and $B(b, \epsilon) \not\subseteq [a, b)^c$ $(\forall \epsilon > 0)$.
5. $\{a\}$ is closed since $\{a\}^c = (-\infty, a) \cup (a, \infty)$, both open sets.
6. Suppose $E \subseteq \mathbb{R}$ is finite. Write $E = \{a_i \mid i \in [n]\}$ where $a_1 < a_2 < \ldots < a_n$. Then $E^c = (-\infty, a_1) \cup (a_1, a_2) \cup \cdots \cup (a_{n-1}, a_n) \cup (a_n, \infty)$, all of which are open.
7. $E = \{\frac{1}{n} \mid n \in \mathbb{N}, n \geq 1\}$ is not closed. $E^c = (-\infty, 0] \cup \bigcup_{n=1}^{\infty}(\frac{1}{n+1}, \frac{1}{n}) \cup (1, \infty)$ is not open because $B(0, \epsilon) \cap E = \{\frac{1}{n} \mid \frac{1}{\epsilon} < n\} \neq \varnothing \implies B(0, \epsilon) \notin E^c$.
8. $E = \{0\} \cup \{\frac{1}{n} \mid n \geq 1\}$ is closed, as $E^c = (-\infty, 0) \cup \bigcup_{n=1}^{\infty}(\frac{1}{n+1}, \frac{1}{n}) \cup (1, \infty)$ is open.

**Lemma**:

1. If $E_\alpha \subseteq \mathbb{R}$ is closed $\forall \alpha \in A$, then $\bigcap_{\alpha \in A} E_\alpha$ is closed.
2. If $E_i \subseteq \mathbb{R}$ is closed $\forall i \in [n]$ then $\bigcup_{i=1}^{n} E_i$ is closed.

*Proof*: The complement is the union of $E_\alpha^c$ (open by claim), which is open by previous lemma.

*Remark*: Example (7) shows that infinite unions of closed sets need not be closed.

### 4.1.3 Limit Points

**Definition**: Let $E \subseteq \mathbb{R}$.

1. A point $x \in \mathbb{R}$ is a *limit point* of $E$ iff. $\forall \epsilon > 0$. $(B(x, \epsilon) \cap E) \backslash \{x\} \neq \varnothing$.
2. A point $x \in E$ is called *isolated* if it is not a limit point.

*Example*: $E = \{\frac{1}{n} \mid n \geq 1\}$. 0 is a limit point, but $\frac{1}{n} \in E$ is isolated, since $B(\frac{1}{n}, \frac{1}{n(n+1)}) \cap E = \{\frac{1}{n}\}$.

**Theorem**: Let $E \subseteq \mathbb{R}$. $E$ is closed $\iff$ every limit point of $E$ is contained in $E$.

*Proof*:

$\implies$:
Assume $E$ is closed and $x \in \mathbb{R}$ is a limit point of $E$. If $x \in E^c$ then, since $E^c$ is open, $\exists \epsilon > 0$. $B(x, \epsilon) \subseteq E^c \implies B(x, \epsilon) \cap E = \varnothing$. But this contradicts the fact that $x$ is a limit point of $E$; thus $x \in E$.

$\impliedby$:
Suppose $E$ is not closed; then $E^c$ is not open and so $\forall \epsilon > 0$. $\exists x \in E^c$. $B(x, \epsilon) \cap E \neq \varnothing$. Since $x \in E^c$, $(B(x, \epsilon) \cap E) \backslash \{x\} = B(x, \epsilon) \cap E \neq \varnothing$ and hence $x$ is a limit point of $E$. Then $x \in E \cap E^c$, a contradiction; and $E$ is closed.

**Definition**: Let $\{x_n\}_{n=l}^{\infty} \subseteq S$ for some set $S$. We say $\{x_n\}$ is *eventually constant* if $\exists N \geq l$. $x_n = x_N$ $(\forall n \geq N)$.

**Proposition**: Let $E \subseteq \mathbb{R}$. Then $x$ is a limit point of $E \iff \exists \{x_n\}_{n=1}^{\infty} \subseteq E$ such that the sequence is not eventually constant and $x_n \to x$ as $n \to \infty$.

*Proof*:

$\implies$:
Suppose $x$ is a limit point of $E$, i.e. $\forall \epsilon > 0.\ (B(x,\epsilon) \cap E)\backslash\{x\} \neq \varnothing$. Set $r_1 = 1$ and choose $x_1 \in E$ such that $x_1 \in (B(x,r) \cap E)\backslash\{x\}$.
Set $r_n = \min(\frac{1}{n}, |x - x_{n-1}|)$ and choose $x_n \in (B(x_1, r_n) \cap E)\backslash\{x\}$.
Then $\forall n \geq 1.\ \{x_n\}_{n=1}^{\infty} \subseteq E$ and $|x - x_{n-1}| < |x - x_n|$ and $|x - x_n| < \frac{1}{n}$. It follows $\{x_n\}$ is not eventually constant, and $x_n \to x$ as $n \to \infty$.

$\impliedby$:
Let $\epsilon > 0$. $\exists N \geq 1.\ n \geq N \implies |x - x_n| < \epsilon$. Then $\{x_n \mid n \geq N\} \subseteq B(x,\epsilon) \cap E$. If $\{x_n \mid n \geq N\} = \{x\}$ then $\{x_n\}$ is eventually constant, a contradiction. Hence $\varnothing \neq \{x_n \mid n \geq N\}\backslash\{x\} \subseteq (B(x,\epsilon) \cap E)\backslash\{x\} \implies (B(x,\epsilon) \cap E)\backslash\{x\} \neq \varnothing$, and hence $x$ is a limit point.

*Corollary*: Let $E \in \mathbb{R}$. The following are equivalent (proof follows from last theorem):

1. $E$ is closed.
2. If $x \in \mathbb{R}$ is a limit point of $E$, $x \in E$.
3. If $\{x_n\}_{n=l}^{\infty} \subseteq E$ is such that $x_n \to x$ as $n \to \infty$, then $x \in E$.

*Corollary*: Let $E \subseteq \mathbb{R}$ and $E \neq \varnothing$. Suppose $E$ is closed.

1. If $E$ is bounded above, then $\sup E \in E$, i.e. $\sup E = \max E$.
2. If $E$ is bounded below, then $\inf E \in E$, i.e. $\inf E = \min E$.

### 4.1.4 Closure, Interior, and Boundary Sets

**Definition**: Let $E \subseteq \mathbb{R}$.

1. Let $\mathcal{O}(E) = \{V \subseteq \mathbb{R} \mid V \subseteq E \text{ and } V \text{ is open}\} \subseteq \mathcal{P}(\mathbb{R})$
   $\mathcal{C}(E) = \{C \subseteq \mathbb{R} \mid E \subseteq C \text{ and } C \text{ is closed}\} \subseteq \mathcal{P}(\mathbb{R})$.
   Note that $\varnothing \in \mathcal{O}(E)$ and $\mathbb{R} \in \mathcal{C}(E)$.
2. We define $E^0 = \bigcup_{V \in \mathcal{O}(E)} V$, and call this set the *interior* of $E$.
   We define $\bar{E} = \bigcap_{C \in \mathcal{C}(E)} C$, and call this set the *closure* of $E$.
3. We define $\partial E = E\backslash E^0$ to be the *boundary* of $E$.

**Theorem**: Let $E \subseteq \mathbb{R}$. The following hold:

1. $E^0 \subseteq E \subseteq \bar{E}$
2. $E^0$ is open and $\bar{E}, \partial E$ are closed.
3. For every $x \in E$, $x \in E^0 \oplus x \in \partial E$.
4. $\partial E = \{x \in \mathbb{R} \mid \forall \epsilon > 0.\ B(x,\epsilon) \cap E \neq \varnothing \text{ and } B(x,\epsilon) \cap E^c \neq \varnothing\}$.
5. $E$ is open $\iff E = E^0$, $E$ is closed $\iff E = \bar{E}$.

*Proof*:

1. Trivial.
2. $E^0$ is an arbitrary union of open sets and thus open; $\bar{E}$ is an arbitrary intersection of closed sets, so it's closed. $\partial E = \bar{E}\backslash E^0 = \bar{E} \cap (\mathbb{R}\backslash E^0)$ is the intersection of two closed sets, so it's closed.

3. Trivial.
4. Suppose $x \in \partial E$. Show the two properties of the set are satisfied via contradiction. Next, assume $x$ in the set, and show that $x \in \partial E$.
5. Trivial.

*Corollary*: Let $E \subseteq \mathbb{R}$. Then $E$ is closed $\iff \partial E \subseteq E$.

*Proof*: $E$ is closed $\implies E = \bar{E} \implies \partial E \subseteq \bar{E} \subseteq E$. On the other hand, if $\partial E \subseteq E$ then $E \subseteq \bar{E} = E^0 \cup \partial E \subseteq E$, so $E = \bar{E}$.

**Theorem (Bolzano-Weierstass, Part 2)**: Let $E \subseteq \mathbb{R}$ be infinite and bounded. Then $E$ has a limit point.

*Proof*: Since $E$ is infinite we may construct a non-eventually-constant sequence $\{x_n\}_{n=0}^{\infty} \subseteq E$. We do so by choosing $x_0 \in E$ arbitrarily, and $x_n \in E \backslash \{x_0, \ldots, x_{n-1}\}$ for any $n \in \mathbb{N}^+$. Since $E$ is bounded, the sequence is too, so B-W implies there exists a convergent subsequence $\{x_{n_k}\}_{k=0}^{\infty} \subseteq E$. This subsequence is not eventually constant by construction, so its limit is a limit point.

### 4.2 Compact Sets

**Definition**:

1. Let $A$ be some index set and assume $\forall \alpha \in A.\ V_\alpha \subseteq \mathbb{R}$. We write $\mathcal{V} = \{V_\alpha\}_{\alpha \in A}$ for the collection of all of these subsets.
2. If $E \subseteq \mathbb{R}$ and $E \subseteq \bigcup_{\alpha \in A} V_\alpha$, then we say $\mathcal{V}$ is a *cover* of $E$.
3. If $V_\alpha \subseteq \mathbb{R}$ is open $\forall \alpha \in A$ and $\mathcal{V}$ is a cover of $E$, we say $\mathcal{V}$ is an *open cover*.
4. Let $\mathcal{V}$ be a cover of $E$. We say $\mathcal{W} = \{V_\alpha\}_{\alpha \in A'}$ is a *subcover* of $E$ if $A' \subseteq A$ and $\mathcal{W}$ is a cover of $E$.
5. Let $\mathcal{V}$ be a cover of $E$. If $A$ is finite, then $\mathcal{W} = \{V_\alpha\}_{\alpha \in A'}$ is a *finite subcover* of $E$, if $\mathcal{W}$ is a subcover of $E$.

*Examples*:

1. Every $E \subseteq \mathbb{R}$ admits a cover: $E = \bigcup_{x \in E} \{x\}$.
2. Every $E \subseteq \mathbb{R}$ admits an open cover: $E \subseteq \bigcup_{x \in E} B(x, \epsilon)$ for $\epsilon > 0$.
3. If $E$ is finite and $\mathcal{V}$ is an open cover, we claim there is a finite open subcover. Indeed, write $E = \{a_i \mid 1 \leq i \leq n\}$ and choose $V_{\alpha_i}$ such that $a_i \in V_{\alpha_i}$. Then $E \subseteq \bigcup_{i=1}^{n} V_{\alpha_i}$ and $\{V_{\alpha_i}\}_{i=1}^{n} \subseteq \{V_\alpha\}_{\alpha \in A}$. Hence every open cover of a finite set admits a finite open subcover.
4. $E = \{\frac{1}{n} \mid n \geq 1\}$. $\mathcal{V} = \{B(\frac{1}{n}, \frac{1}{n(n+1)}\}_{n=1}^{\infty}$ is an open cover of $E$. Note that $B(\frac{1}{n}, \frac{1}{n(n+1)}) \cap E = \frac{1}{n}$, so there does not exist a finite subcover.
5. $E = \{0\} \cup \{\frac{1}{n} \mid n \geq 1\}$. Suppose $\mathcal{V}$ is an open cover of $E$. Since $0 \in E$, $\exists \alpha_0 \in A.\ 0 \in V_{\alpha_0}$. Since $V_{\alpha_0}$ is open, $\exists \epsilon > 0.\ B(0, \epsilon) \subseteq V_{\alpha_0}$. Then $B(0, \epsilon) \cap E = \{\frac{1}{n} \mid\mid; n \geq N\}$ where $N = \min\{n \in N \mid n \geq \frac{1}{\epsilon}\}$. Hence $E \backslash B(0, \epsilon) = \{\frac{1}{n} \mid 1 \leq n \leq N\}$. There exist $V_{\alpha_n}$ for $n \in [N]$ such that $\frac{1}{n} \in V_{\alpha_n}$. Then $E \subseteq \bigcup_{n=0}^{N} V_{\alpha_n}$ and $E$ has a finite subcover.
6. Let $a < b$ and $E = (a, b)$. Then $\mathcal{V} = \{(a + \frac{1}{n+1}, b - \frac{1}{n+1})\}_{n \in \mathbb{N}}$ is an open cover of $E$. Since these intervals are nested, there cannot be a finite subcover.

**Definition**: Let $E \subseteq \mathbb{R}$. We say that $E$ is *compact* if every open cover of $E$ admits a finite subcover.

*Examples*:

1. $\varnothing$ is trivially compact.
2. $\mathbb{R}$ is not compact because $\mathcal{V} = \{B(0,n)\}_{n\in\mathbb{N}}$ is an open cover that clearly does not admit a finite subcover of $\mathbb{R}$.
3. Any finite set $E \subseteq \mathbb{R}$ is compact.
4. $(a,b)$ for $a < b$ is not compact.
5. $\{\frac{1}{n} \mid n \geq 1\}$ is not compact.
6. $\{0\} \cup \{\frac{1}{n} \mid n \geq 1\}$ is compact.

Notice in each of our examples of compact sets that the set is closed and bounded.

### 4.2.1 Heine-Borel Theorem

**Theorem**: Let $K \subseteq \mathbb{R}$. Then $K$ is compact $\iff$ $K$ is closed and bounded.

*Proof*:

$\implies$ Suppose $K$ is compact.

Notice that $\bigcup_{n=1}^{\infty} B(0,n) = \mathbb{R}$ (since $\mathbb{R}$ is Archimedean) and so $K \subseteq \mathbb{R} = \bigcup_{n=1}^{\infty} B(0,n)$. Then $\{B(0,n)\}_{n=1}^{\infty}$ is an open cover of $K$. Since $K$ is compact, $\exists$ a finite subcover : $K \subseteq \bigcup_{i=1}^{m} B(0,n_i)$ for some $m \in \mathbb{N}$.

Set $r = \max_{i\in[m]} n_i$. Then $K \subseteq \bigcup_{i=1}^{m} B(0,n_i) \subseteq B(0,r) \implies K$ is bounded.

Now we show $K$ is closed. Let $x \in K^C$. For each $y \in K$ we set $r_y = \frac{1}{2}|x-y| > 0$. Then $B(y,r_y) \cap B(x,r_y) = \varnothing$ $(\forall y \in K)$. Also, $\{B(y,r_y)\}_{y\in K}$ is an open cover.

$K$ compact $\implies$ $\exists$ a finite subcover: $K \subseteq \bigcup_{i=1}^{n} B(y_i, r_{y_i})$. Set $r = \min_{i\in[n]} r_i > 0$ and notice that $B(y_i, r_{y_i}) \cap B(y,r) = \varnothing$. Hence $\bigcup_{i=1}^{n} B(y_i,r_{y_i}) \cap B(x,r) = \varnothing \implies K \cap B(x,r) = \varnothing \implies B(x,r) \subseteq K^C$. This means that $K^C$ is open and so $K$ is closed.

$\impliedby$ **(Heine-Borel)** Suppose $K$ is closed and bounded. If $K = \varnothing$ we're done, so suppose $K \neq \varnothing$.

Notice that $K$ bounded $\implies$ $\inf K, \sup K \in \mathbb{R}$, and $K$ closed $\implies$ $\inf K, \sup K \in K$. In particular, $\sup K = \max K, \inf = \min K$. Let $\mathcal{V}$ be an open cover of $K$.

Let $E = \{x \in K \mid \mathcal{V}$ admits a finite subcover of $K \cap [\inf K, x]\} \subseteq K$. Notice that $K \cap [\inf K, \inf K] = \{\inf K\}$ is a finite set and hence compact; thus $\mathcal{V}$ admits a finite subcover of $K \cap [\inf K, \inf K]$. Hence $\inf K \in E$, and so $E \neq \varnothing$. Clearly $E$ is bounded above by $\sup K$. By LUB property, $\exists \sup E \in \mathbb{R}$ and $\sup E \leq \sup K$.

We want to show $\sup E = \sup K = \max E$. Notice that $\forall n \geq 1$. $\exists x_n \in E \subseteq K$ such that $\sup E - \frac{1}{n} < x_n \leq \sup E$. Then $x_n \to \sup E$ as $n \to \infty$, and so $\sup E \in K$ (since $K$ is closed).

Write $\mathcal{V} = \{V_\alpha\}_{\alpha\in A}$. Since $\sup E \in K, \exists \alpha_0 \in A$ such that $\sup E \in V_{\alpha_0}$. But $V_{\alpha_0}$ is open so $\exists \epsilon > 0$. $B(\sup E, \epsilon) \subseteq V_{\alpha_0}$. By definition, $\exists x \in E$. $\sup E - \epsilon < x \leq \sup E$. Hence $\mathcal{V}$ admits a finite subcoverof $K \cap [\inf K, x]$, i.e. $K \cap [\inf K, x] \subseteq \bigcup_{i=1}^{n} V_{\alpha_i}$. Then $K \cap [\inf K, \sup E] \subseteq \bigcup_{i=0}^{n} V_{\alpha_i} \implies \sup E \in E \implies \sup E = \max E$.

Assume for sake of contradiction that $\max E < \max K$. Let $K' = K \setminus \bigcup_{i=0}^{n} V_{\alpha_i}$. $K'$ is closed since it's the intersection of closed sets. $K' \neq \varnothing$ since otherwise $K \subseteq \bigcup_{i=0}^{n} V_{\alpha_i} \implies \max E = \max K$.

Let $y = \inf K' = \min K'$ (since $K'$ is closed) and note that $y > \max E$. Then $K \cap [\inf K, y] = K \cap [\inf K, \min K'] \subseteq \bigcup_{i=0}^{n} V_{\alpha_i} \cup \{y\}$. But since $y \in K' \subseteq K$, $\exists V_{\alpha_{n+1}} \in \mathcal{V}$ such that $y \in V_{\alpha_{n+1}}$. Hence $K \cap [\inf K, y] \subseteq \bigcup_{i=0}^{n+1} V_{\alpha_i} \implies y \in E \implies \max E < y \leq \max E$, a contradiction. We then deduce that $\max E = \max K \implies K = K \cap [\min K, \max K]$ is covered by a finite subcover of $\mathcal{V}$; thus, $K$ is compact.

**Corollary**:

1. If $K \subseteq \mathbb{R}$ is compact and $E \subseteq \mathbb{R}$ is closed, then $E \cap K$ is compact.
2. If $K \subseteq \mathbb{R}$ is compact and $E \subseteq K$ is closed, then $E$ is compact.
3. If $K_i \subseteq \mathbb{R}$ is compact for $i \in [n]$, then $\bigcup_{i=1}^{n} K_i$ is compact.
4. If $K_\alpha \subseteq \mathbb{R}$ is compact $\forall \alpha \in A$, then $\bigcap_{\alpha \in A} K_\alpha$ is compact.

## 4.3 Connected Sets

**Definition**: We say two sets $A, B \subseteq \mathbb{R}$ are *separated* if $A \cap \bar{B} = \bar{A} \cap B = \varnothing$.
A set $E \subseteq \mathbb{R}$ is *disconnected* if $E = A \cup B$ such that $a \neq \varnothing, B \neq \varnothing$ and $A, B$ are separated.
If a set is $E \subseteq \mathbb{R}$ is not disconnected, we say it's *connected*.

*Examples*:

1. $(0,1)$ and $[1,2)$ are not separated, though they are disjoint, since $\overline{(0,1)} \cap [1,2) = [0,1] \cap [1,2) = \{1\} \neq \varnothing$.
2. $(a,b)$ and $(b,c)$ for $a < b < c$ are separated, since $\overline{(a,b)} \cap (b,c) = \varnothing = (a,b) \cap \overline{(b,c)}$.
   Then $(a,c) \setminus \{b\}$ is disconnected, since $(a,c) \setminus \{b\} = (a,b) \cup (b,c)$.
3. Similarly, $\forall a \in \mathbb{R}$. $(-\infty, a)$ an $(a, \infty)$ are separated.
   Then $\mathbb{R} \setminus \{a\} = (-\infty, a) \cup (a, \infty)$ is disconnected.

**Theorem**: Let $E \subseteq \mathbb{R}$. Then $E$ is connected $\iff$ ($x, y \in E$ and $x < z < y \implies z \in E$).

*Proof*:
$\neg 2 \implies \neg 1$:

If (2) is false then $\exists x, y \in E$ and $z \in (x,y)$ such that $z \notin E$. Then $E = L_z \cup R_z$ for $L_z = E \cap (-\infty, z)$ and $R_z = E \cap (z, \infty)$. Since $x \in L_z, y \in R_z$, and $L_z \subseteq (-\infty, z)$ and $R_z \subseteq (z, \infty)$, it follows that $L_z$ and $R_z$ are separated. Hence $E$ is disconnected.

$\neg 1 \implies \neg 2$

Suppose $E$ is disconnected. Write $E = A \cup B$ with $A, B \neq \varnothing$ and $\bar{A} \cap B = A \cap \bar{B} = \varnothing$. Let $x \in A$ and $y \in B$. Without loss of generality, we assume $x < y$.

Let $z = \sup(A \cap [x,y])$. Clearly $z \in \bar{A}$ and so $z \notin B \implies z \neq y \implies x \leq z \leq y$. If $z \notin A$ then $z \neq x \implies x < z < y$ and $z \notin A \cup B = E$. Otherwise, if $z \in A$, then $z \notin \bar{B}$. $\bar{B}$ is closed, so $\bar{B}^C$ is open; and hence we can find $w$ such that $z < w < y$, $w \notin B$, and $w \notin A$. Then $x < w < y$ and $w \notin A \cup B = E$. In all cases, then, $\neg 2$ is true.

*Corollary*: $\mathbb{R}, (-\infty, a), (-\infty, a], (a, \infty), [a, \infty), (a,b), (a,b], [a,b), and [a,b]$ are all connected.

# 5 Continuity

## 5.1 Limits of Functions

**Definition**: Let $E \subseteq \mathbb{R}$, $f : E \to \mathbb{R}$, and $p \in \mathbb{R}$ be a limit point. Let $q \in \mathbb{R}$.
We say $\lim_{x\to p} f(x) = q$ or $f(x) \to q$ as $x \to p$ iff $\forall \epsilon > 0.\ \exists \delta > 0.\ x \in E \wedge 0 < |x - p| < \delta \implies |f(x) - q| < \epsilon$.

*Examples*:

1. $E = [0, 1], f(x) = x$. Let $p = \frac{1}{2}$. $\lim_{x\to\frac{1}{2}} f(x) = \frac{1}{2}$.

   *Proof*: Let $\epsilon > 0$; choose $\delta = \epsilon > 0$. Then $x \in [0, 1]$ and $0 < |x - \frac{1}{2}| < S \implies |f(x) - \frac{1}{2}| < \epsilon$.
2. $E = [0, 1], f(x) = x$ (for $x \neq \frac{1}{2}$), $f(x) = 37$ (for $x = \frac{1}{2}$).
   By the proof of (1), the claim still holds.
3. $f(x) = x^n$ on $E = (0, 1)$ for $2 \leq n \in \mathbb{N}$. 0 is a limit point of $E$; we claim $\lim_{x\to 0} x^n = 0$.

   *Proof*: Let $\epsilon > 0$; choose $\delta = \epsilon^{1/n} > 0$. Then $x \in (0, 1)$ and $0 < |x - 0| < \delta \implies 0 < x < \delta \implies 0 < x^n < \delta^n = \epsilon \implies |f(x) - 0| = x^n < \epsilon$.
4. $\lim_{x\to p} x = p$ whenever $p$ is a limit point of $E$.
5. If $\forall x \in E.\ f(x) = 1$ then $\lim_{x\to p} f(x) = 1$ whenever $p$ is a limit point of $E$.
6. Let $E = \mathbb{R}$ and $f(x) = \cos(x)$. From HW6, $|\cos(x) - 1| \leq x^2 e^{x^2}$. We claim $\lim_{x\to 0} \cos(x) = 1$.

   *Proof*: Let $\epsilon > 0$. Choose $\delta = \min(1, \sqrt{\epsilon/e}) > 0$. Then for $x \in \mathbb{R}$, $0 < |x - 0| < \delta \implies |x| < \min(1, \sqrt{\epsilon/e}) \implies |\cos(x) - 1| \leq x^2 e^1$ (since $|x|^2 < \delta \leq 1 \implies e^{|x|^2} \leq e^1$) $\implies |\cos(x) - 1| < \delta^2 e \leq (\sqrt{\epsilon/e})^2 e = \epsilon$.
7. $E = \{\frac{1}{n} \mid n \geq 1\}, p = 0$. Let $f(x) = \frac{1}{x}$ for $x \in E$. We claim $\lim_{x\to 0} f(x)$ does not exist.

   *Proof*: Suppose not. Then for $\epsilon = 1.\ \exists \delta > 0.\ x \in E,\ 0 < |x - 0| < \delta \implies |f(x) - q| < 1$. But $x \in E, |x| < \delta \implies x = \frac{1}{n}, \frac{1}{\delta} < n$, and $|f(x) - q| = |\frac{1}{1/n} - q| = |n - q| < 1$, which is a contradiction.

**Definition**: Let $f : E \to \mathbb{R}$ for some $E \subseteq \mathbb{R}$. If $A \subseteq E$ we define $f(A) = \{f(x) \mid x \in A\} \subseteq \mathbb{R}$ as the *image* of $A$ under $f$. If $B \subseteq \mathbb{R}$ we define $f^{-1}(B) = \{x \in E \mid f(x) \in B\}$ as the *pre-image* of $B$ under $f$.

**Lemma**: Suppose $f : E \to \mathbb{R}$. Then $A \subseteq B \subseteq E \implies f(A) \subseteq f(B)$, and
$A \subseteq B \subseteq \mathbb{R} \implies f^{-1}(A) \subseteq f^{-1}(B) \subseteq E$.

### 5.1.1 Divergence Criteria

**Theorem (Divergence Criteria)**: Let $E \subseteq \mathbb{R}$, $f : E \to \mathbb{R}$, $p$ be a limit point of $E$, $q \in \mathbb{R}$. The following are equivalent:

1. $\lim_{x\to p} f(x) = q$
2. For every open set $V \subseteq \mathbb{R}$ such that $q \in V$, $\exists$ an open set $U \subseteq \mathbb{R}$ with $p \in U$ such that $f(U \cap E\backslash\{p\}) \subseteq V$. *(Topological characterization)*

3. If $\{x_n\}_{n=l}^{\infty} \subseteq E$ satisfies $x_n \neq p$ $(\forall n \geq l)$ and $x_n \to p$ as $n \to \infty$, the sequence $\{f(x_n)\}_{n=l}^{\infty} \subseteq \mathbb{R}$ converges and $f(x_n) \to q$ as $n \to \infty$. *(Sequential characterization)*

*Proof*:

(1) $\implies$ (2) :
Assume (1) and let $V \subseteq \mathbb{R}$ be open with $q \in V$. Since $V$ is open, $\exists \epsilon > 0.\ B(q, \epsilon) \subseteq V$. Since $\lim_{x\to p} f(x) = q$, $\exists \delta > 0.\ x \in E \wedge 0 < |x-p| < \delta \implies |f(x)-q| < \delta$. Let $U = B(p, \delta)$ (an open set). Then $x \in U \cap E\backslash\{p\} \implies x \in E \wedge |x-p| < \delta \implies |f(x)-q| < \epsilon \implies f(x) \in B(q, \epsilon) \subseteq V$. So $f(U \cap E\backslash\{p\}) \subseteq V$ as desired.

(2) $\implies$ (3):
Assume (2) and let $\{x_n\}_{n=l}^{\infty} \subseteq E$ satisfy $x_n \neq p, x_n \to p$. Let $\epsilon > 0$ and set $V = B(q, \epsilon)$ (open). From (2), $\exists$ open $U$ such that $f(U \cap E\backslash\{p\}) \subseteq V$ and $p \in U$. Since $U$ is open, $\exists \delta > 0.\ B(p, \delta) \subseteq U$. Since $x_n \to p$ as $n \to \infty$, $\exists N \geq l.\ n \geq N \implies ) < |x_n - p| < \delta \implies x_n \in U \cap E\backslash\{p\} \implies f(x_n) \in V = B(q, \epsilon)$. Hence $n \geq N \implies |f(x_n) - q| < \epsilon$, and $f(x) \to q$ as $n \to \infty$.

$\neg$(1) $\implies$ $\neg$(3):
Suppose (1) is false; then $\exists \epsilon > 0.\ \forall \delta > 0.\ \exists x \in E$ with $0 < |x-p| < \delta$ such that $|f(x) - q| \geq \epsilon$. For $n \in \mathbb{N}, n \geq 1$, set $\delta = \frac{1}{n}$ to find $x_n \in E$ such that $0 < |x_n - p| < \frac{1}{n}$ and $|f(x_n) - q| \geq \epsilon$. Clearly, $\{x_n\}_{n=1}^{\infty} \subseteq E$ satisfies $x_n \neq p$, $x_n \to p$. But $f(x_n)$ does not converge to $q$. Hence (3) fails.

**Corollary**: If $E \subseteq \mathbb{R}, f : E \to \mathbb{R}$, $p$ is a limit point of $E$, and $\lim_{x\to p} f(x) = q$, then $q$ is unique.

*Proof*: Limits of sequences are unique, so this follows from (3) in Divergent Criteria theorem.

**Corollary (Algebra of limits)**: Let $E \subseteq \mathbb{R}, f, g : E \to \mathbb{R}$, $p$ be a limit point of $E$. Assume $\lim_{x\to p} f(x) = q_1, \lim_{x\to p} g(x) = q_2$. The following hold:

1. If $\alpha, \beta \in \mathbb{R}$ then $\lim_{x\to p}(\alpha f(x) + \beta g(x)) = \alpha q_1 + \beta q_2$
2. $\lim_{x\to p} f(x)g(x) = q_1 q_2$
3. If $q_2 = \lim_{x\to p} g(x) \neq 0$, then $\frac{f}{g} : E\backslash g^{-1}(\{0\}) \to \mathbb{R}$ is well-defined, $p$ is a limit point of $E\backslash g^{-1}(\{0\})$, and $\lim_{x\to p} \frac{f(x)}{g(x)} = \frac{q_1}{q_2}$

*Proof*: All follow from the algebra of sequential limits and (3) in the Theorem.

As an application of this, we get a large class of limit examples.

**Corollary**: Let $P : E \to \mathbb{R}$ be a polynomial, i.e. $P(x) = a_0 + a_1 x + \cdots + a_n x^n$ for some $n \in \mathbb{N}, a_i \in \mathbb{R}$ for $i \in [n]$. If $p$ is a limit point of $E$, then $\lim_{x\to p} P(x) = P(p)$.

*Proof*: We know $\lim_{x\to p} 1 = 1$, $\lim_{x\to p} x = p$. Algebra of limits (2) and simple induction show $\lim_{x\to p} x^k = p^k$ $(\forall k \in \mathbb{N}^+)$. Then algebra of limits (1) and another induction argument prove $\lim_{x\to p} P(x) = \lim_{x\to p}(a_0 + a_1 x + \cdots + a_n x^n) = \lim_{x\to p}(a_0 + a_1 p + \cdots + a_n p^n) = P(p)$.

## 5.2 Continuous Functions

**Definition**: Let $E \subseteq \mathbb{R}, f : E \to \mathbb{R}, p \in E$. We say $f$ is *continuous* at $p$ iff:

$$\forall \epsilon > 0.\ \exists \delta > 0.\ x \in E \wedge |x-p| < \delta \implies |f(x) - f(p)| < \epsilon$$

If $f : E \to \mathbb{R}$ is continuous at each $p \in E$ we say $f$ is continuous on $E$.

*Remarks*:

1. In order to be continuous at $p \in E$, $f$ must be defined at $p$. Contrast this to $\lim_{x \to p} f(x)$, in which case $p$ need only be a limit point of $E$.
2. Informally one can think of continuous functions as those approximated well "near $p$" by $f(p)$, i.e. $f(x) \approx f(p)$ when $x \approx p$.
3. In the definition, the value of $\delta$ may depend on the point $p$. If a function is continuous on $E$ then for a given $\epsilon > 0$ the $\delta = \delta(p)$ may vary greatly as $p$ varies.
4. If $p \in E$ is isolated (not a limit point of $E$), then $f$ is vacuously continuous at $p$: $x \in E$, $|x - p| < \delta$ for $\delta$ small enough $\implies x = p$.

*Example*:

We saw last time that $\lim_{x \to p} P(x) = P(p)$ for all polynomials $P : \mathbb{R} \to \mathbb{R}$. Hence $\forall \epsilon > 0.\ \exists \delta > 0.\ x \in \mathbb{R}, 0 < |x - p| < \delta \implies |P(x) - P(p)| < \epsilon$. Hence $P$ is continuous at $p$.

**Theorem**: Let $E \subseteq \mathbb{R}, f : E \to \mathbb{R}, p \in E$ be a limit point of $E$. Then:

$$f \text{ is continuous at } p \iff \lim_{x \to p} f(x) = f(p)$$

**Corollary (Algebra of Continuity)**: Let $E \subseteq \mathbb{R}, f, g : E \to \mathbb{R}$, and $p \in E$. Assume that $f, g$ are continuous at $p$. Then the following hold:

1. If $\alpha, \beta \in \mathbb{R}$ then $\alpha f + \beta g$ is continuous at $p$.
2. $fg$ is continuous at $p$.
3. If $g(p) \neq 0$ then $\frac{f}{g} : E \backslash g^{-1}(\{0\}) \to \mathbb{R}$ is well-defined and continuous at $p$.

*Proof*: If $p$ is isolated, the claim is vacuously true. Assume $p$ is not isolated, i.e. $p$ is a limit point of $E$. Then the last theorem and algebra of limits gives the result.

**Corollary**: Let $E \in \mathbb{R}, f, g : E \to \mathbb{R}$. If $f, g$ are continuous on $E$, then:

1. If $\alpha, \beta \in \mathbb{R}$ then $\alpha f + \beta g$ is continuous on $E$.
2. $fg$ is continuous on $E$.
3. If $g(x) \neq 0$ ($\forall x \in E$), then $\frac{f}{g}$ is continuous on $E$.

**Theorem**: Let $E, F \subseteq \mathbb{R}$, $f : E \to \mathbb{R}, g : F \to \mathbb{R}$. Assume $f(E) \subseteq F$, $f$ is continuous at $p \in E$, and $g$ is continuous at $f(p) \in F$. Then $g \circ f : E \to \mathbb{R}$ (where $(g \circ f)(x) = g(f(x))$) is continuous at $p$. Moreover, if $f$ is continuous on $E$ and $g$ is continuous on $F$, then $g \circ f$ is continuous on $E$.

*Proof*: Let $\epsilon > 0$.

Since $g$ is continuous at $f(p)$, $\exists \eta > 0.\ y \in F$ and $|y - f(p)| < \eta \implies |g(y) - g(f(p))| < \epsilon$. Since $f$ is continuous at $p$, $\exists \delta > 0.\ x \in E, |x - p| < \delta \implies |f(x) - f(p)| < \eta$.

Since $f(E) \subseteq F$ we know that $x \in E, |x - p| < \delta \implies f(x) \in F, |f(x) - f(p)| < \eta \implies |g(f(x)) - g(f(p))| < \epsilon$. Hence, $g \circ f$ is continuous by definition.

*Examples*:

1. exp, cos, sin : $\mathbb{R} \to \mathbb{R}$ are continuous on $\mathbb{R}$ (proof in HW). YAlso, $\log : (0, \infty) \to \mathbb{R}$ is continuous on $(0, \infty)$.
2. Let $\alpha \in \mathbb{R}$ and set $f : (0, \infty) \to \mathbb{R}$ via $f(x) = x^\alpha$. Notice that $f(x) = \exp(\alpha \log x)$. Since log and exp are continuous, $f(x) = x^\alpha$ is continuous.

**Definition**: Let $E \subseteq \mathbb{R}$ and $A \subseteq E$. We say $A$ is *relatively open* in $E$ iff $A = U \cap E$ for some open set $U \subseteq \mathbb{R}$. Similarly, we say $A$ is *relatively closed* in $E$ iff $A = C \cap E$ for some closed $C \subseteq \mathbb{R}$.

**Proposition**: Let $A \subseteq E \subseteq \mathbb{R}$. The following hold:

1. $A$ is relatively open in $E \iff \forall x \in A.\ \exists \epsilon > 0.\ B(x, \epsilon) \cap A \subseteq E$.
2. $A$ is relatively closed in $E \iff A = B^C \cap E$ for some relatively open $B \subseteq E$.

**Theorem (Continuity Criteria)**: Let $E \subseteq \mathbb{R}, f : E \to \mathbb{R}$. The following are equivalent:

1. $f$ is continuous on $E$.
2. If $p \in E$ is a limit point of $E$, then $\lim_{x \to p} f(x) = f(p)$.
3. If $p \in E$ is a limit point of $E$ and $\{x_n\}_{n=l}^\infty \subseteq E$ satisfies $x_n \to p$ as $n \to \infty$, then $f(x_n) \to f(p)$ as $n \to \infty$.
4. If $V \subseteq \mathbb{R}$ is open, then $f^{-1}(V) \subseteq E$ is relatively open in $E$.
5. If $C \subseteq \mathbb{R}$ is closed, then $f^{-1}(C) \subseteq E$ is relatively closed in $E$.

*Proof*:

(1) $\iff$ (2) $\iff$ (3) follows from the sequential criterion of limits, previous theorem.
(4) $\iff$ (5) follows since $f^{-1}(V^C) = (f^{-1}(V))^C \cap E$.

(1) $\implies$ (4):
Let $V \subseteq \mathbb{R}$ be open and choose $p \in f^{-1}(V)$. Since $V$ is open, $\exists \epsilon > 0.\ B(f(p), \epsilon) \subseteq V$. It suffices to show, via previous proposition, that $\exists \delta > 0.\ B(p, \delta) \cap E \subseteq f^{-1}(V)$. Since $f$ is continuous on $E$, $\exists \delta > 0.\ x \in E, |x - p| < \delta \implies |f(x) - f(p)| < \epsilon$. That is, $x \in B(p, \delta) \cap E \implies |f(x) - f(p)| < \epsilon \implies f(x) \in B(f(p), \epsilon) \subseteq V$. Hence $B(p, \delta) \cap E \subseteq f^{-1}(V)$.

(4) $\implies$ (1):
Let $p \in E, \epsilon > 0$, and $V = B(f(p), \epsilon)$. Then $f^{-1}(B(f(p), \epsilon)) \subseteq E$ is relatively open in $E \implies$ (by previous proposition) $\exists \delta > 0.\ B(p, \delta) \cap E \subseteq f^{-1}(B(f(p), \epsilon))$. Then $x \in E$ and $|x - p| < \delta \implies f(x) \in B(f(p), \epsilon) \implies |f(x) - f(p)| < \epsilon$. Since $\epsilon, p$ were arbitrary, we deduce $f$ is continuous on $E$.

## 5.3 Compactness and Continuity

**Theorem**: Suppose $K \subseteq \mathbb{R}$ is compact and $f : K \to \mathbb{R}$ is continuous on $K$. Then $f(K)$ is compact.

*Proof*:

Note that for $E \subseteq \mathbb{R}$, $f(f^{-1}(E)) \subseteq E$ and $E \subseteq f^{-1}(f(E))$. Let $\{V_\alpha\}_{\alpha \in A}$ be an open cover of $f(K)$. Since $f$ is continuous and $V_\alpha$ is open, $f^{-1}(V_\alpha)$ is relatively open in $K \implies f^{-1}(V_\alpha) = U_\alpha \cap K$ for some open $U_\alpha \subseteq \mathbb{R}$.

Since $\{V_\alpha\}_{\alpha \in A}$ cover $f(K)$, we see that $\{f^{-1}(V_\alpha)\}_{\alpha \in A}$ is a cover of $K$. Then $\{U_\alpha\}_{\alpha \in A}$ is an open cover of $K$. Since $K$ is compact, there exists a finite subcover: $K \subseteq \bigcap_{i=1}^n U_{\alpha_i}$. Then $K \subseteq \bigcup_{i=1}^n U_{\alpha_i} \cap K = \bigcup_{i=1}^n f^{-1}(V_{\alpha_i}) \implies f(K) \subseteq \bigcup_{i=1}^n f(f^{-1}(V_{\alpha_i})) \subseteq \bigcup_{i=1}^n V_{\alpha_i}$. As we have extracted a finite open subcover of $f(K)$, $f(K)$ is compact.

**Extreme Value Theorem**: Let $K \subseteq \mathbb{R}$ be compact and $f : K \to \mathbb{R}$ be continuous. Then $\exists x_0, x_1 \in K$ such that $f(x_0) \leq f(x) \leq f(x_1)$ $(\forall x \in K)$. That is, $f(x_0) = \min_{x \in K} f(x) = \min f(K)$ and $f(x_1) = \max_{x \in K} f(x) = \max f(K)$.

*Proof*: From last theorem, we know $f(K)$ is compact, so it's closed and bounded. From a previous theorem, closed and bounded sets contain their infimum and supremum (and thus min, max).

**Definition**: Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$. We say $f$ is *uniformly continuous* on $E$ iff:

$$\forall \epsilon > 0.\ \exists \delta > 0.\ x, y \in E \wedge |x - y| < \delta \implies |f(x) - f(y)| < \epsilon$$

*Remarks*:

1. $f$ is uniformly continuous on $E \implies f$ is continuous on $E$.
2. The key difference is that for uniform continuity, $\delta > 0$ works for all points in $E$.

*Examples*:

1. Let $E = (0, 1)$ and $f(x) = \frac{1}{x}$. It's trivial that $f$ is continuous on $E$, but it is not uniformly continuous.

   *Proof*: Suppose it is; then for $\epsilon = \frac{1}{2}, \exists \delta > 0.\ x, y \in (0, 1) \wedge |x - y| < \delta \implies |f(x) - f(y)| < \frac{1}{2}$. Choose $N \in \mathbb{N}$ such that $\frac{1}{\sqrt{\delta}} < N$. Then $x = \frac{1}{n}, y = \frac{1}{n+1}$ satisfy $|x - y| = \frac{1}{n(n+1)} \leq \frac{1}{n^2} < \delta$ if $n \geq N$. Then $\frac{1}{2} > |f(x) - f(y)| = |n - (n+1)| = 1$, a contradiction.

**Definition**: A function $f : E \to \mathbb{R}$ is *Lipschitz* if $\forall x, y \in E.\ \exists k > 0.\ |f(x) - f(y)| \leq k|x - y|$.

**Claim**: If $f$ is Lipschitz, it is uniformly continuous. Proof: let $\delta = \frac{\epsilon}{k}$.

**Theorem**: Let $K \subseteq \mathbb{R}$ be compact and $f : K \to \mathbb{R}$ be continuous. Then $f$ is uniformly continuous on $K$.

## 5.4 Continuity and Connectedness

**Theorem**: Let $E \subseteq \mathbb{R}$ be connected and $f : E \to \mathbb{R}$ be continuous on $E$. If $X \subseteq E$ is connected, then $f(X)$ is connected.

**Intermediate Value Theorem**: Let $a < b \in \mathbb{R}$. Suppose $f : [a, b] \to \mathbb{R}$ is continuous. If $f(a) < c < f(b)$ or $f(b) > c > f(a)$ for some $c \in \mathbb{R}$, then $\exists x \in (a, b).\ f(x) = c$.

## 5.5 Discontinuities

**Lemma**: If $p$ is a limit point of $E \subseteq \mathbb{R}$ then $p$ is a limit point of $E_p^+ = E \cap (p, \infty)$ or $E_p^- = E \cap (-\infty, p)$.

**Definition**: Let $E \subseteq \mathbb{R}$, $f : E \to \mathbb{R}$, $p$ be a limit point of $E$, $q \in \mathbb{R}$.

1. If $p$ is a limit point of $E_p^-$, we say $\lim_{x \to p^-} f(x) = q \iff \forall \epsilon > 0.\ \exists \delta > 0.\ x \in E_p^-$, $0 < p - \delta \implies |f(x) - q| < \epsilon$.
2. If $p$ is a limit point of $E_p^+$, then $\lim_{x \to p^+} f(x) = q \iff \forall \epsilon > 0.\ \exists \delta > 0.\ x \in E_p^+$, $0 < x - p < \delta \implies |f(x) - q| < \epsilon$.

**Proposition**: If $p$ is not a limit point of $E_p^+$ then $\lim_{x\to p} f(x) = \lim_{x\to p^-} f(x)$. If $p$ is not a limit point of $E_p^-$ then $\lim_{x\to p} f(x) = \lim_{x\to p^+} f(x)$.

**Proposition**: If $p$ is both a limit point of either $E_p^+$ or $E_p^-$, then

$$\lim_{x\to p} f(x) = q \iff \lim_{x\to p^-} f(x) = \lim_{x\to p^+} f(x) = q$$

**Definition**: Suppose $E \subseteq \mathbb{R}$, $f : E \to \mathbb{R}$, $p \in E$ is a limit point of $E$. Suppose further that $p$ is not a point of continuity of $f$.

1. We say $f$ has a *simple discontinuity* of $p$ if
   $p$ is not a limit point of $E_p^+$ and $\lim_{x\to p^-} f(x)$ exists,
   $p$ is not a limit point of $E_p^-$ and $\lim_{x\to p^+} f(x)$ exists, or
   $p$ is a limit point of $E_p^+$ and $E_p^-$ and $\lim_{x\to p^+} f(x)$, $\lim_{x\to p^-} f(x)$ both exist.
2. Otherwise, we say $f$ has an *essential discontinuity* of $p$.

### 5.6 Monotone Functions

**Definition**: Let $E \subseteq \mathbb{R}$ and $f : E \to \mathbb{R}$. We say:
$f$ is *non-decreasing* (increasing) if $x, y \in E$ and $x < y \implies f(x) \leq f(y)$ $(f(x) < f(y))$, and
$f$ is *non-increasing* (decreasing) if $x, y \in E$ and $x < y \implies f(y) \leq f(x)$ $(f(y) < f(x))$.
If $f$ is non-increasing or non-decreasing, $f$ is *monotone*.

**Theorem**: Suppose $f : (a, b) \to \mathbb{R}$ is monotone, and let $p \in (a, b)$. Then $\lim_{x\to p^-} f(x)$ and $\lim_{x\to p^+} f(x)$ both exist. Moreover, if $f$ is non-decreasing, then

$$\lim_{x\to p^-} f(x) = \sup f((a, p)) \leq f(p) \leq \inf f((p, b)) = \lim_{x\to p^+} f(x)$$

**Corollary**: If $f : (a, b) \to \mathbb{R}$ is monotone, then $f$ has no essential discontinuities.

*Example*: $f(x) = \lfloor x \rfloor$ is non-decreasing and $f$ has countably many simple discontinuities.

**Theorem**: If $f : (a, b) \to \mathbb{R}$ is monotone, then $f$ has at most countably many simple discontinuities.

## 6 Differentiation

### 6.1 The Derivative

**Definition**: Assume $f : [a, b] \to \mathbb{R}$ for $a < b \in \mathbb{R}$. For all $x \in [a, b]$, the function $\phi : (a, b)\backslash\{x\} \to \mathbb{R}$ via $\phi(t) = \frac{f(t)-f(x)}{t-x}$ is well-defined, and $x$ is a limit point of $(a, b)\backslash\{x\}$. If $\lim_{t\to x} \phi(t)$ exists we write $f'(x) = \lim_{t\to x} \phi(t)$ and say that $f$ is *differentiable* at $x$.

We define $f' : \{x \in [a, b] \mid x \text{ is differentiable at } x\} \to \mathbb{R}$ to be the *derivative* of $f$. If $f$ is differentiable $\forall x \in E \subseteq [a, b]$, we say $f$ is differentiable on $E$.

**Definition (General)**: Let $E \subseteq \mathbb{R}$, $f : E \to \mathbb{R}$, and $x \in E$ be a limit point of $E$. Define $\phi : E\backslash\{x\} \to \mathbb{R}$ by $\phi(t) = \frac{f(t)-f(x)}{t-x}$. If $\lim_{t\to x} \phi(t)$ exists we say $f$ is differentiable at $x$, and write $f'(x) = \lim_{t\to x} \phi(t)$.

**Proposition (locality of derivative)**: Suppose $f : E \to \mathbb{R}$, $g : F \to \mathbb{R}$, $x \in E \cap F$ is a limiit point of $E \cap F$, and that $f$ and $g$ are differentiable at $x$. If $f = g$ on $E \cap F$ then $f'(x) = g'(x)$. This shows that $f'(x)$ only depends on the value of $f$ "near $x$".

**Proposition (Newtonian approximation)**: Let $f : E \to \mathbb{R}$, $x \in E$ be a limit point of $E$, and $L \in \mathbb{R}$. Then the following are equivalent:

1. $f$ is differentiable at $x$ and $f'(x) = L$
2. $\forall \epsilon > 0.\ \exists \delta > 0.\ t \in E \wedge |x - t| < \delta \implies |f(t) - (f(x) + L(t - x))| < \epsilon|t - x|$

Proof follows from definition of $\lim_{t\to x} \phi(t)$. Newton's approximation says differentiable functions are those that can be "well-approximated" by affine functions $\alpha+\beta x$. Continuous functions are those well-approximated by constants, while differentiable functions are well-approximated by the "next" simplest function.

**Theorem**: Suppose $f : E \to \mathbb{R}$, $x \in E$ is a limit point of $E$, and $f$ is differentiable at $x$. Then $f$ is continuous at $x$.

*Proof*: By definition, if $t \in E\backslash\{x\}$ then $f(t)-f(x) = \phi(t)(t-x)$. Then $f(t) = f(x)+\phi(t)(t-x)$ and hence $\lim_{t\to x} f(t) = f(x) + \lim_{t\to x} \phi(t)(t - x) = f(x) + f'(x)0 = f(x)$. By the limit chracterization of continuity, we deduce that $f$ is continuous at $x$.

*Remark*: The converse fails. Let $f(x) = |x|$ on $\mathbb{R}$. Since $||x| - |y|| \leq |x - y|$, $f$ is Lipschitz and hence uniformly continuous. However, for $x = 0$, $t > 0 \implies \phi(t) = \frac{|t|-0}{t-0} = 1$ and $t < 0 \implies \phi(t) = \frac{-t-0}{t-0} = -1$. Then $\lim_{t\to 0^-} \phi(t) = -1 \neq \lim_{t\to 0^+} \phi(t) = 1$, so $f'(0)$ does not exist.

**Theorem (Algebra of Derivatives)**: Let $f, g : E \to \mathbb{R}$ be differentiable at $x \in E$. Then:

1. $f + g : E \to \mathbb{R}$ is differentiable at $x$ and $(f + g)'(x) = f'(x) + g'(x)$
2. $fg : E \to \mathbb{R}$ is differentiable at $x$ and $(fg)'(x) = f(x)g'(x) + f'(x)g(x)$
3. If $g(x) \neq 0$ then $\frac{f}{g} : E\backslash g^{-1}(\{0\}) \to \mathbb{R}$ is differentiable at $x$ and $(\frac{f}{g})'(x) = \frac{g(x)f'(x)-f(x)g'(x)}{g(x)^2}$

*Examples*:

1. $f(x) = \alpha + \beta x$ on $\mathbb{R} \implies f'(x) = \lim_{t\to x} \frac{f(t)-f(x)}{t-x} = \beta$ ($\forall x \in \mathbb{R}$).
2. $f(x) = x^n$ for $n \in \mathbb{N} \implies f'(x) = nx^{n-1}$. Proof by induction.
3. Every polynomial $P(x) = \sum_{n=0}^{N} a_n x^n$ is differentiable, and $P'(x) = \sum_{n=0}^{N} na_n x^{n-1}$.
4. $R(x) = \frac{P(x)}{Q(x)}$ is dfiferentiable when $P, Q$ are polynomials at points $p \in \mathbb{R}$ where $Q(p) \neq 0$.

**Theorem (Chain Rule)**: Suppose $f : E \to \mathbb{R}$ is differentiable at $x \in E$, $f(E) \subseteq F$, and $g : F \to \mathbb{R}$ is differentiable at $f(x) \in F$. Then $g\circ f : E \to \mathbb{R}$ is differentiable at $x$ and $(g\circ f)'(x) = g'(f(x))f'(x)$.

## 6.2 Mean Value Theorems

**Definition**: Let $f : E \to \mathbb{R}$. We say that $f$ has a *local maximum* at $x \in E$ if $\exists \delta > 0.\ t \in E$ and $|x - t| < \delta \implies f(t) \leq f(x)$. We say $f$ has a *local minimum* at $x \in E$ if $-f$ has a local maximum.

If $f$ has either a local max or min at $x \in E$, we say $f$ has a *local extremum* at $x$.

**Theorem**: Suppose $f : E \to \mathbb{R}$ is differentiable at $x \in E$ and $x$ is a limit point of both $E_x^+$ and $E_x^-$. If $f$ has a local extremum at $x$, then $f'(x) = 0$.

*Proof*: It suffices to assume that $f$ has a local max at $x$. Let $\delta > 0$ such that $t \in E$ and $|x-t| < \delta \implies f(t) \leq f(x)$. Then $t \in E, 0 < x - t < \delta \implies \frac{f(t)-f(x)}{t-x} \geq 0$ and $0 < t - x < \delta \implies \frac{f(t)-f(x)}{t-x} \leq 0$. So $\lim_{t \to x^-} \frac{f(t)-f(x)}{t-x} = f'(x) \geq 0$, $\lim_{t \to x^-} \frac{f(t)-f(x)}{t-x} = f'(x) \leq 0$, and thus $f'(x) = 0$.

*Remark*: The result is false if $x$ is not a limit point of either $E_x^+$ or $E_x^-$. Consider $f(x) = x$ on $E = [0,1]$; $f$ has a local min at $x = 0$, local max at $x = 1$, but $f'(x) = 1 \forall x \in [0,1]$.

**Theorem (Monotonicity part 1)**: Let $f : E \to \mathbb{R}$ be differentiable at $x \in E$.

1. If $f$ is non-decreasing on $E$, then $f'(x) \geq 0$.
2. If $f$ is non-increasing on $E$, then $f'(x) \leq 0$.

**Cauchy's Mean Value Theorem**: Suppose that $f, g : [a,b] \to \mathbb{R}$ are continuous on $[a,b]$, differentiable on $(a,b)$. Then $\exists x \in (a,b)$. $(g(b) - g(a))f'(x) = (f(b) - f(a))g'(x)$.

*Proof*:

Consider $h : [a,b] \to \mathbb{R}$ via $h(x) = (g(b) - g(a))f(x) - (f(b) - f(a))g(x)$. To prove the result, it suffices to find $x \in (a,b)$ such that $h'(x) = 0$ (since $h$ is cont, diff on $[a,b]$ and $(a,b)$).

Notice that $h(a) = g(b)f(a) - g(a)f(b) = h(b)$.

If $h$ is constant, then $h'(x) = 0$ trivially and we're done. Assume $h$ is not constant; then $\exists t \in (a,b)$. $h(t) > h(a)$ or $h(t) < h(a)$.

If $h(t) \geq h(a)$, then Extreme Value Theorem guarantees that $\exists x \in (a,b)$. $h(x) = \max h([a,b])$ and Local Extremum Theorem $\implies h'(x) = 0$.

If $h(t) < h(a)$ then EVT guarantees $\exists x \in (a,b)$. $h(x) = \min h([a,b])$ and LET $\implies h'(x) = 0$.

**Corollary (Mean Value Theorem)**: If $f : [a,b] \to \mathbb{R}$ is cont and diff on $[a,b], (a,b)$ then $\exists x \in (a,b)$. $f(b) - f(a) = f'(x)(b-a)$. *Proof*: Set $g(x) = x$.

**Corollary (Monotonicity part 2)**: Suppose $f : (a,b) \to \mathbb{R}$ is differentiable on $(a,b)$. The following hold:

1. $\forall x \in (a,b)$. $f'(x) > 0 \implies f$ is increasing.
2. $\forall x \in (a,b)$. $f'(x) \geq 0 \implies f$ is non-decreasing.
3. $\forall x \in (a,b)$. $f'(x) = 0 \implies f$ is constant.
4. $\forall x \in (a,b)$. $f'(x) \leq 0 \implies f$ is non-increasing.
5. $\forall x \in (a,b)$. $f'(x) < 0 \implies f$ is decreasing.

*Proof*: by MVT, if $a < x_1 < x_2 < b$, then $f(x_2) - f(x_1) = f'(x)(x_2 - x_1)$.

## 6.3 Darboux's Theorem

**Definition**: We say a function $g : \mathbb{R} \to \mathbb{R}$ is periodic with period $p > 0$ if $g(x+p) = g(x)$ $(\forall x \in \mathbb{R})$.

**Theorem (Darboux)**: Suppose $f : [a,b] \to \mathbb{R}$ is differentiable on $[a,b]$ and $f'(a) < \gamma < f'(b)$. Then $\exists x \in (a,b)$. $f'(x) = \gamma$.

**Corollary**: If $f : [a,b] \to \mathbb{R}$ is differentiable on $[a,b]$, then $f'$ has no simple discontinuities.

## 6.4 L'Hôpital's Rule

**Theorem**: Suppose $f, g : [a,b] \to \mathbb{R}$ are continuous on $[a,b]$, differentiable on $(a,b)$, and $g'(x) \neq 0$ $(\forall x \in (a,b))$. Assume that $\lim_{x\to a} \frac{f'(x)}{g'(x)} = L$. If $f(a) = g(a) = 0$, then $\lim_{x\to a} \frac{f(x)}{g(x)} = L$.

*Proof*:

We claim first that $g(x) \neq 0$ for $x \in (a,b]$. Otherwise, $g(x) = 0$ for some $x \in (a,b] \implies 0 = \frac{g(x)-g(a)}{x-a} = g'(z)$ for some $z \in (a,x)$, a contradiction. So $\frac{f}{g} : (a,b] \to \mathbb{R}$ is well-defined.

Let $\{x_n\}_{n=l}^{\infty} \subseteq (a,b]$ satisfy $x_n \to a$ as $n \to \infty$. We claim that $\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = L$. Once this is established, the sequential characterization of limits yields the desired result.

To prove the claim, we apply Cauchy's Mean Value Theorem on $[a, x_n]$: $\exists y_n \in (a, x_n)$ such that $f'(y_n)g(x_n) = f'(y_n)(g(x_n) - g(a)) = g'(x_n)(f(x_n) - f(a)) = g'(x_n)f(x_n)$. Then $\forall n \geq l$. $\frac{f(x_n)}{g(x_n)} = \frac{f'(x_n)}{g'(x_n)}$. Since $a < y_n < x_n$, the squeeze lemma implies $y_n \to a$. Hence $\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = \lim_{n\to\infty} \frac{f'(x_n)}{g'(x_n)} = \lim_{x\to a} \frac{f'(x)}{g'(x)} = L$.

**Remarks**:

1. The theorem is also true if we take limits at $t$.
2. If $f, g : (a,b] \to \mathbb{R}$ and $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$, then the theorem still works.

## 6.5 Higher Derivatives and Taylor's Theorem

**Definition**: Suppose $f : E \to \mathbb{R}$ is differentiable at $x \in E$, and $x$ is a limit point of $\{y \in E \mid f'(x) \text{ exists}\}$. We say $f$ is twice differentiable at $x$ if $f' : \{y \in E \mid f'(y) \text{ exists}\} \to \mathbb{R}$ is differentiable at $x$; and $f''(x) = f^{(2)}(x) = (f')'(x)$. Similarly, for $n \in \mathbb{N}$ with $n > 2$, we say $f$ is $n$-times differentiable at $x$ if $x$ is a limit point of $\{y \in E \mid f^{(n-1)}(y) \text{ exists}\}$ and $f^{(n-1)}$ is differentiable at $x$, in which case $f^{(n)}(x) = (f^{(n-1)})'(x)$.

If $f^{(n)}$ exists $\forall n \in \mathbb{N}$, $n \geq 1$ we say $f$ is infinitely differentiable at $x$.

**Theorem (Taylor)**: Suppose $f : [a,b] \to \mathbb{R}$. Assume $f^{(n-1)}$ is continuous on $[a,b]$ and $f^(n)$ exists on $(a,b)$. Let $x, y \in [a,b]$ with $x \neq y$. Then $\exists z \in (\min\{x,y\}, \max\{x,y\})$ such that

$$f(y) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!}(y-x)^k + \frac{f^{(n)}(z)}{n!}(y-x)^n$$

(called the Taylor polynomial or Taylor approximation).

*Proof*:

Suppose $x < y$ ($y < x$ is handled without loss of generality). Let $P(t) = \sum_{k=0}^{n-1} \frac{f^{(k)}(x)}{k!}(t-x)^k$, and set $M = \frac{f(y)-P(y)}{(y-x)^n}$. It suffices to prove that $M = \frac{f^{(n)}(z)}{n!}$ for some $z \in (x,y)$.

Define $g(t) = f(t) - P(t) - M(t-x)^n$, and notice that $g^{(n)}(t) = f^{(n)}(t) - n!M$. As such, it suffices to show that $g^{(n)}(z) = 0$ for some $z \in (x,y)$.

By construction, $g^{(k)}(x) = 0$ $(\forall k = 0, \ldots, n-1)$, and $g(y) = 0$ (by choice of $M$). By Mean Value Theorem, $\exists x_i \in (x,y)$. $g'(x_1) = \frac{g(y)-g(x)}{y-x} = 0$. Similarly, $\exists x_2 \in (x, x_1)$. $g''(x_2) = \frac{g'(x_1)-g'(x)}{x-1-x} = 0$. Iterating, we eventually find $x_{n-1} \in (x,y)$. $g^{(n-1)}(x_{n-1}) = 0$. Then $0 = g^{(n)}(z) = \frac{g^{(n-1)}(x_{n-1})-g^{(n-1)}(x)}{x_{n-1}-x} = 0$ for some $z \in (x, x_{n-1})$.

# 7 Riemann-Stieltjes Integration

## 7.1 The R-S Integral

**Definition**: Let $a, b \in \mathbb{R}$ with $a \leq b$. A partition of $[a, b]$ is a finite ordered set $P = \{x_0, \ldots, x_n\}$ such that $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$. Write $\Pi[a, b] = \{P \mid P \text{ is a partition of } [a, b]\}$. For brevity we'll write $\Pi = \Pi[a, b]$.

**Universal Assumptions**: Throughout §7 we will always assume that:

1. $f : [a, b] \to \mathbb{R}$ is bounded: $\forall x \in [a, b].\ m \leq f(x) \leq M$, where $m = \inf f([a, b]), M = \sup f([a, b])$
2. $\alpha : [a, b] \to \mathbb{R}$ (the integrator or weight function) is non-decreasing (in particular, $\alpha$ is also bounded)

**Definition**: For each $P \in \Pi[a, b]$ we associate to $f$ the following quantities ($P = \{x_0, \ldots, x_n\}$):

1. $m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$ for $i \in [n]$
2. $M_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}$ for $i \in [n]$
3. $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) \geq 0$ for $i \in [n]$

We write $U(P, f, \alpha) = \sum_{i=1}^n M_i \Delta\alpha_i$, $L(P, f, \alpha) = \sum_{i=1}^n m_i \Delta\alpha_i$.
$U$ is the upper Riemann-Stieltjes sum, and $L$ is the lower R-S sum.

*Remark*: Clearly $m(\alpha(b) - \alpha(a)) = \sum_{i=1}^n m(\alpha(x_i) - \alpha(x_{i-1})) = \sum_{i=1}^n m\Delta\alpha_i \leq \sum_{i=1}^n M_i \Delta\alpha_i \leq M \sum_{i=1}^n \Delta\alpha_i = M(\alpha(b) - \alpha(a))$. Hence $\forall P \in \Pi[a, b].\ m(\alpha(b) - \alpha(a)) \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq M(\alpha(b) - \alpha(a))$.

**Definition of Integral**: We define

$\underline{\int_a^b} f d\alpha = \sup\{L(P, f, \alpha) \mid P \in \Pi[a, b]\}$, and $\overline{\int_a^b} f d\alpha = \inf\{U(P, f, \alpha) \mid P \in \Pi[a, b]\}$. Both are well-defined by the remark.

If $\underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha$ then we say $f$ is R-S integrable with respect to $\alpha$, and write
$\int_a^b f d\alpha = \underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha$.

We write $\mathcal{R}([a, b]; \alpha) = \{f : [a, b] \to \mathbb{R} \mid f \text{ is bounded, } f \text{ is R-S integrable with respect to } \alpha\}$. When $\alpha(x) = x$, $\int_a^b f dx$ is the Riemann integral and we write $\mathcal{R}([a, b])$.

*Heuristics*: The function $\alpha$ assigns different weights to different points in $[a, b]$. The intuition is that $\int_a^b f d\alpha$ is a "weighted Riemann integral". If $\alpha$ is continuous then we have a geometric interpretation of $\int_a^b f d\alpha$: consider the curve in $\mathbb{R}^2$ parameterized by $(x(t), y(t)) = (\alpha(t), f(t))$; $\int_a^b f d\alpha =$ area under this curve.

**Lemma**: Let $f(x) = C$ ($\forall x \in [a, b]$). Then $f \in \mathcal{R}([a, b]; \alpha)$ and $\int_a^b f d\alpha = C(\alpha(b) - \alpha(a))$.

*Proof*: For any $P \in \Pi[a, b]$ we have $m_i = M_i = C$. Hence $U(P, f, \alpha) = L(P, f, \alpha) = \sum_{i=1}^n C\Delta\alpha_i = C(\alpha(b) - \alpha(a))$. So $\underline{\int_a^b} f d\alpha = \sup\{(L(P, f, \alpha)\} = C(\alpha(b) - \alpha(a)) = \inf\{U(P, f, \alpha)\} = \overline{\int_a^b} f d\alpha$.

**Definition**: If $P, P' \in \Pi[a, b]$ and every point in $P$ is in $P'$, we say $P'$ is a *refinement* of $P$.
If $P_1, P_2 \in \Pi[a, b]$ we define the *common refinement* $P_1 \# P_2 \in \Pi[a, b]$ by $P_1 \# P_2 = P_1 \cup P_2$, ordered appropriately.

**Proposition**: If $P, P' \in \Pi[a,b]$ and $P'$ is a refinement of $P$, then $L(P,f,\alpha) \leq L(P', f\alpha) \leq U(P', f, \alpha) \leq U(P, f, \alpha)$.

**Theorem**: $\underline{\int_a^b} f d\alpha \subseteq \overline{\int_a^b} f d\alpha$.

*Proof*: Let $P_1, P_2 \in \Pi[a,b]$. Then $L(P_1, f, \alpha) \leq L(P_1 \# P_2, f, \alpha) \leq U(P_1 \# P_2, f, \alpha) \leq U(P_2, f, \alpha)$ by last proposition. Hence $\underline{\int_a^b} f d\alpha = \sup\{L(P_1, f, \alpha) \mid P_1 \in \Pi[a,b]\} \leq U(P_2, f, \alpha)$. Then $\underline{\int_a^b} f d\alpha \leq \inf\{U(P_2, f, \alpha) \mid P_2 \in \Pi[a,b]\} = \overline{\int_a^b} f d\alpha$.

## 7.2 Integrability Criteria

**Theorem (Riemann)***: $f \in \mathcal{R}([a,b];\alpha) \iff \forall \epsilon > 0.\ \exists P \in \Pi[a,b].\ U(P,f,\alpha) - L(P,f,\alpha) < \epsilon$.

*Proof*:

(1) $\implies$ (2):
Let $\epsilon > 0$. By definition, $\exists P_1, P_2 \in \Pi[a,b]$. $\underline{\int_a^b} f d\alpha - \frac{\epsilon}{2} < L(P_1, f, \alpha)$, and $U(P_2, f, \alpha) \leq \overline{\int_a^b} f d\alpha + \frac{\epsilon}{2}$. Let $P = P_1 \# P_2$. Then $U(P_1, f, \alpha) \leq U(P_2, f, \alpha) < \overline{\int_a^b} f d\alpha + \frac{\epsilon}{2}$. Hence $U(P, f, \alpha) - L(P, f, \alpha) < \epsilon$.

(2) $\implies$ (1):
For any partition $P \in \Pi[a,b]$ we know $\overline{\int_a^b} f d\alpha \leq U(P_1, f, \alpha)$ and $L(P, f, \alpha) \leq \underline{\int_a^b} f d\alpha$. We also know $\underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha$. Then (2) implies that for $\epsilon > 0$ we have $P \in \Pi[a,b]$. $U(P_1, f, \alpha) - L(P_1, f, \alpha) < \epsilon$. But then $0 \leq \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha < \epsilon$ $(\forall \epsilon > 0)$ and so $\underline{\int_a^b} f d\alpha = \overline{\int_a^b} f d\alpha \implies f \in \mathcal{R}([a,b];\alpha)$.

**Lemma**: Let $P \in \Pi[a,b]$. The following are true:

1. If $P'$ is a refinement of $P$, then $U(P', f, \alpha) - L(P', f, \alpha) \leq U(P, f, \alpha) - L(P, f, \alpha)$.
2. If $s_i, t_i \in [x_{i-1}, x_i]$ $(\forall i \in [n])$, then $0 \leq \sum_{i=1}^n |f(s_i) - f(t_i)| \Delta\alpha_i \leq U(P, f, \alpha) - L(P, f, \alpha)$.
3. If $f \in \mathcal{R}([a,b];\alpha)$, then $t_i \in [x_{i-1}, x_i] \implies |\sum_{i=1}^n f(t_i)\Delta\alpha_i - \int_a^b f d\alpha| \leq U(P, f, \alpha) - L(P, f, \alpha)$.

**Theorem**: If $f$ is continuous on $[a,b]$, then $f \in \mathcal{R}([a,b];\alpha)$.

*Proof*:

Note that EVT implies $f$ is bounded. Let $\epsilon > 0$ and choose $k > 0$ so $k(\alpha(b) - \alpha(a)) < \epsilon$. Since $f$ is cont on compact $[a,b]$, $f$ is uniformly continuous. Then $\exists \delta > 0$. $x, y \in [a,b]$ and $|x - y| < \delta \implies |f(x) - f(y)| < k$.

Choose a partition $P \in \Pi[a,b]$ such that $x_i - x_{i-1} < \delta$ $(\forall i \in [n])$. Then by EVT, $\exists s_i, t_i \in [x_{i-1}, x_i]$ such that $m_i = f(s_i), M_i = f(t_i)$. Then $U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^n (M_i - m_i)\Delta\alpha_i = \sum_{i=1}^n (f(t_i) - f(s_i))\Delta\alpha_i \leq \sum_{i=1}^n k\Delta\alpha_i = k\sum_{k=1}^n \Delta\alpha_i = k(\alpha(b) - \alpha(a)) < \epsilon$. Since $\epsilon > 0$ was arbitrary, we deduce via Riemann's Theorem that $f \in \mathcal{R}([a,b];\alpha)$.

**Theorem**: Suppose that $f$ is monotone and $\alpha$ is continuous on $[a,b]$; then $f \in \mathcal{R}([a,b];\alpha)$.

*Proof*:

Since $\alpha$ is continuous, $\alpha([a,b])$ is connected. Then for every $n \in \mathbb{N}$ with $n \geq 1$, we know $\alpha(a) + \frac{i}{n}(\alpha(b) - \alpha(a)) \in \alpha([a,b])$ for $i \in [n]$. In particular, $\exists x_i \in [a,b]$ such that $\alpha(x_i) = \alpha(a) + \frac{i}{n}(\alpha(b) - \alpha(a))$. Since $\alpha$ is non-decreasing, we may choose these $x_i$ such that $a = x_0 \leq x_1 \leq \cdots \leq x_n = b$. Then $P_n = \{x_0, x_1, \ldots, x_n\}$ is a partition of $[a,b]$.

Assume $f$ is non-decreasing (the case of $f$ non-increasing is similar). Let $\epsilon > 0$.

Since $f$ is non-decreasing, $f(x_i) = M_i$, $f(x_{i-1}) = m_i$ for $i \in [n]$. Then

$$\begin{aligned}
U(P_n, f, \alpha) - L(P_n, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\
&= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\Delta\alpha_i \\
&= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1}))\frac{(\alpha(b) - \alpha(a))}{n} \\
&= \frac{\alpha(b) - \alpha(a)}{n}(f(b) - f(a))
\end{aligned}$$

Choose $n$ such that $n > \frac{(\alpha(b)-\alpha(a))(f(b)-f(a))}{\epsilon}$. Then $U(P_n, f, \alpha) - L(P_n, f, \alpha) < \epsilon$.

*Remarks*:

1. If $\alpha = x$, then all monotone functions $f$ are in $\mathcal{R}([a,b])$, i.e. all monotone functions are Riemann integrable.
2. Monotone functions don't have to be continuous, so $\exists f$ with simple discontinuities in $\mathcal{R}([a,b]; \alpha)$ when $\alpha$ is continuous.
3. Monotone functions can have countably infinite sets of discontinuity; R-S integrals can handle infinite discontinuities.

Next, we show that R-S integration can handle more general discontinuities.

**Theorem**: Suppose that $f$ is continuous on $[a,b] \backslash E$, where $E \subseteq [a,b]$ is finite. Assume that $\alpha$ is continuous on $E$; then $f \in \mathcal{R}([a,b]; \alpha)$.

*Remark*: There is a much more powerful version known as Lebesgue's Theorem, which says that:

$$f \in \mathcal{R}([a,b]; \alpha) \iff (E = \{x \in [a,b] \mid f \text{ is not continuous at } x\} \implies E \text{ is } \alpha\text{-null})$$

Here, a set $E$ is $\alpha$-null if $\forall \epsilon > 0$. $\exists$ intervals $(u_j, v_j) \subseteq \mathbb{R}$ for $j \geq 1 \in \mathbb{N}$ such that $E \subseteq \bigcup_{j=1}^{\infty}(u_j, v_j)$ and $\sum_{j=1}^{\infty} \alpha(v_j) - \alpha(u_j) < \epsilon$.

In particular, if $\alpha(x) = x$, then all countable sets are $\alpha$-null, and hence if $f$ has discontinuities only on a countable set, then $f \in \mathcal{R}([a,b])$.

**Theorem**: Let $I \in \mathbb{R}$. Then the following are equivalent:

1. $f \in \mathcal{R}([a,b]; \alpha)$ and $\int_a^b f d\alpha = I$
2. $\forall \epsilon > 0$. $\exists P \in \Pi[a,b]$. $(\forall i \in [n].\ t_i \in [x_{i-1}, x_i] \implies |\sum_{i=1}^{n} f(t_i)\Delta\alpha_i - I| < \epsilon)$.

## 7.3 Properties of $\mathcal{R}([a,b];\alpha)$

**Theorem**: If $f_1, f_2 \in \mathcal{R}([a,b];\alpha)$ and $c_1, c_2 \in \mathbb{R}$, then $c_1 f_1 + c_2 f_2 \in \mathcal{R}([a,b];\alpha)$ and $\int_a^b (c_1 f_1 + c_2 f_2) d\alpha = c_1 \int_a^b f_1 d\alpha + c_2 \int_a^b f_2 d\alpha$.

*Remark*: This theorem gives us two important consequences:

1. $\mathcal{R}([a,b];\alpha)$ is a linear space;
2. The integral is a linear function, i.e. $\int_a^b d\alpha : \mathcal{R}([a,b];\alpha) \to \mathbb{R}$ given by $f \mapsto \int_a^b f d\alpha$ is a linear map.

**Theorem**: Suppose $f \in \mathcal{R}([a,b];\alpha_i)$ for $i \in \{1,2\}$. Let $c_1, c_2 \geq 0$. Then $f \in \mathcal{R}([a,b]; c_1\alpha_1 + c_2\alpha_2)$ and $\int_a^b f d(c_1\alpha_1 + c_2\alpha_2) = c_1 \int_a^b f d\alpha_1 + c_2 \int_a^b f d\alpha_2$.

**Theorem**: Suppose $f \in \mathcal{R}([a,b];\alpha)$. Let $c \in (a,b)$; then $f \in \mathcal{R}([a,c];\alpha)$ and $f \in \mathcal{R}([c,b];\alpha)$; i.e. $\int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha$.

**Theorem**: Suppose $f \in \mathcal{R}([a,b];\alpha)$ and $g : [m = \inf f, M = \sup f] \to \mathbb{R}$ is continuous on $[m, M]$. Then $g \circ f \in \mathcal{R}([a,b];\alpha)$.

**Theorem**: Suppose $f, g \in \mathcal{R}([a,b];\alpha)$. Then $fg \in \mathcal{R}([a,b]\ \alpha)$.

*Proof*: The function $\psi : \mathbb{R} \to \mathbb{R}$ where $\phi(x) = x^2$ is cont. on $\mathbb{R}$. By the last theorem, if $h \in \mathcal{R}([a,b];\alpha)$ then $h^2 \in \mathcal{R}([a,b];\alpha)$. Note that $fg = \frac{1}{4}(f+g)^2 - \frac{1}{4}(fg)^2 \in \mathcal{R}([a,b];\alpha)$ since $f \pm g \in \mathcal{R}([a,b];\alpha)$.

This result shows that $\mathcal{R}([a,b];\alpha)$ has even more structure than a linear space; it is an algebra, i.e. a vector space closed under product.

## 7.4 Integration and Order

**Theorem**: Suppose $f_1, f_2 \in \mathcal{R}([a,b];\alpha)$ and $\forall x \in [a,b]$. $f_1(x) \leq f_2(x)$. Then $\int_a^b f_1 d\alpha \leq \int_a^b f_2 d\alpha$.

**Theorem**: Let $f \in \mathcal{R}([a,b];\alpha)$. Then $|f| \in \mathcal{R}([a,b];\alpha)$ and $|\int_a^b f d\alpha| \leq \int_a^b |f| d\alpha$.

*Proof*: $\psi : \mathbb{R} \to \mathbb{R}$ where $\psi(x) = |x|$ is continuous on $\mathbb{R}$, so the composition theorem guarantess that $|f| \in \mathcal{R}([a,b];\alpha)$. Let $c = 1$ if $\int_a^b f d\alpha \geq 0$ and $-1$ otherwise. Then $|\int_a^b f d\alpha| = c \int_a^b f d\alpha = \int_a^b c f d\alpha$, but $cf \leq |cf| = |c||f| = |f|$ so the last theorem gives the desired result.

**Definition**: Given a function $f : E \to \mathbb{R}$ ($E \subseteq \mathbb{R}$ we define the *positive part* of $f$ as $f^+ : E \to [0, \infty)$ given by $f^+(x) = \max\{f(x), 0\}$. Similarly, $f^- : E \to [0,\infty)$ via $f^-(x) = \max\{-f(x), 0\} = (-f)^+(x)$.

**Lemma**: $f = f^+ - f^-$ and $|f| = f^+ + f^-$. Proof is trivial.

**Theorem**: Let $f \in \mathcal{R}([a,b];\alpha)$. Then $f^+, f^- \in \mathcal{R}([a,b];\alpha)$.

*Proof*: $f^+ = \frac{f+|f|}{2}$, $f^- = \frac{f-|f|}{2}$.

This leads to a theorem $f, g \in \mathcal{R}([a,b];\alpha) \implies f \wedge g, f \vee g \in \mathcal{R}([a,b];\alpha)$, where $(f \wedge g)(x) = \min\{f(x), g(x)\}$ and $(f \vee g)(x) = \max\{f(x), g(x)\}$. (Proof on HW.)

## 7.5 Fundamental Theorem of Calculus

**Lemma**: Suppose $f \in \mathcal{R}([a,b];\alpha)$. Define $F : [a,b] \to \mathbb{R}$ via $F(x) = \int_a^x f d\alpha$. Let $M = \sup\{|f(x)| \mid x \in [a,b]\}$. Then $|F(x) - F(y)| \leq M|\alpha(x) - \alpha(y)|$ ($\forall x, y \in [a,b]$).

**Theorem**: Suppose $f \in \mathcal{R}([a,b];\alpha)$ and define $F(x) = \int_a^x f d\alpha$. The following hold:

1. If $\alpha$ is continuous at a point $x \in [a,b]$, then $F$ is continuous at $x$.
2. If $\alpha$ is continuous on $[a,b]$, then $F$ is uniformly continuous at $[a,b]$.
3. If $\alpha$ is Lipschitz on $[a,b]$ $(|\alpha(x) - \alpha(y)| \leq K|x-y|.\ \forall x, y \in [a,b])$, then $F$ is Lipschitz.

**Theorem (FTC1)**: Assume that $f$ is continuous at $x \in [a,b]$ and that $\alpha$ is differentiable at $x$. Let $F(s) = \int_a^b f d\alpha$. Then $F$ is differentiable at $x$ and $F'(x) = f(x)\alpha'(x)$.

*Proof*:

For $t \in [a,b]\backslash\{x\}$ set $u = \min\{t,x\}$, $v = \max\{t,x\}$. Then:

$$\frac{F(t)-F(x)}{t-x} - f(x)\left(\frac{\alpha(t)-\alpha(x)}{t-x}\right) = \frac{1}{v-u}\int_u^v (f - f(x))d\alpha$$

Hence

$$\begin{aligned}\left|\frac{F(t)-F(x)}{t-x} - f(x)\alpha'(x)\right| &\leq \frac{1}{v-u}\left|\int_u^v (f-f(x))d\alpha\right| + \left|f(x)(\frac{\alpha(t)-\alpha(x)}{t-x}) - f(x)\alpha'(x)\right| \\ &\leq \frac{1}{v-u}\int_u^v |f-f(x)|d\alpha + |f(x)|\left|\frac{\alpha(t)-\alpha(x)}{t-x} - \alpha'(x)\right|\end{aligned}$$

Let $\epsilon > 0$. Since $f$ is cont. at $x$ and $\alpha$ is diff. at $x$, we can find $\delta > 0$ such that

$$|x-y| < \delta \implies |f(y)-f(x)| < \frac{\epsilon}{2(1+|\alpha'(x)|)} \wedge \left|\frac{\alpha(y)-\alpha(x)}{y-x} - \alpha'(x)\right| < \min\{1, \frac{\epsilon}{2(1+|f(x)|)}\}$$

Then, if $|x-t| < \delta$ we can deduce:

$$\begin{aligned}\left|\frac{F(t)-F(x)}{t-x} - f(x)\alpha'(x)\right| &\leq \frac{1}{v-u}\int_u^v \frac{\epsilon}{2(1+|\alpha'(x)|)}d\alpha + |f(x)|\frac{\epsilon}{2(1+|f(x)|)} \\ &= \frac{\epsilon}{2(1+|\alpha'(x)|)}|\frac{\alpha(t)-\alpha(x)}{t-x}| + \frac{\epsilon}{2} \\ &\leq \frac{\epsilon}{2(1+|\alpha'(x)|)}(1+|\alpha'(x)|) + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

Hence, $F$ is differentiable at $x$ and $F'(x) = f(x)\alpha'(x)$.

**Theorem (FTC2)**: If $f \in \mathcal{R}([a,b])$ and $\exists F : [a,b] \to \mathbb{R}$ that is continuous on $[a,b]$ and differentiable on $(a,b)$ and $\forall x \in (a,b)$. $F'(x) = f(x)$, then $\int_a^b f dx = F(b) - F(a)$.

*Proof*:

Let $\epsilon > 0$. Then $\exists P \in \Pi[a,b]$ such that $|\sum_{i=1}^n f(t_1)\Delta x_i - \int_a^b f dx| < \epsilon$ for any $t_i \in [x_{i-1}, x_i]$ for $i \in [n]$.

The Mean Value Theorem guarantees that $\exists t_i \in [x_{i-1}, x_i]$ such that $F(x_i) - F(x_{i-1}) = (x_i - x_{i-1})F'(t_i) = f(t_i)\Delta x_i$. Hence:

$$\left|F(b) - F(a) - \int_a^b f dx\right| = \left|\sum_{i=1}^n F(x_i) - F(x_{i-1}) - \int_a^b f dx\right| = \left|\sum_{i=1}^n f(t_i)\Delta x_i - \int_a^b f dx\right| < \epsilon$$

Since this is true $\forall \epsilon > 0$, we deduce $\int_a^b f dx = F(b) - F(a)$.

**Theorem (Integration by Parts)**: Suppose $f, g : [a, b] \to \mathbb{R}$ are differentiable on $[a, b]$, and $f', g' \in \mathcal{R}([a, b])$. Then $\int_a^b f' g dx = f(b)g(b) - f(a)g(a) - \int_a^b f g' dx$.

*Proof*: Let $h(x) = f(x)g(x)$. Then $h'(x) = f(x)g'(x) + f'(x)g(x)$, and $h' \in \mathcal{R}([a, b])$ (since $f, g$ diff. $\implies f, g$ cont). Applying FTC2 to $h$ yields the desired result.

## 7.6 Advanced Results in R-S Integration

**Theorem**: Suppose $\alpha$ is diff. on $[a, b]$ and $\alpha' \in \mathcal{R}([a, b])$. Then:

$$f \in \mathcal{R}([a, b]; \alpha) \iff fa' \in \mathcal{R}([a, b]) \text{ and } \int_a^b f dx = \int_a^b f \alpha' dx$$

*Proof*: Write $K = \sup |f|$. For any $P \in \Pi[a, b]$ we know for $t_i, s_i \in [x_{i-1}, x_i]$ for $i \in [n]$,

$$\sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)| \Delta x_i \leq U(P, \alpha', x) - L(P, \alpha', x)$$

(since $\alpha' \in \mathcal{R}([a, b])$). The Mean Value Theorem allows us to find $s_i \in (x_{i-1}, x_i)$ such that $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \alpha'(s_i)(x_i - x_{i-1}) = \alpha'(s_i)\Delta x_i$. Then for any choice of $t_i \in [x_{i-1}, x_i]$, we may estimate

$$\begin{aligned}
\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \sum_{i=1}^{n} f(t_i)\alpha'(t_i)\Delta x_i \right| &= \left| \sum_{i=1}^{n} f(t_i)\alpha'(s_i)\Delta x_i - \sum_{i=1}^{n} f(t_i)\alpha'(t_i)\Delta x_i \right| \\
&= \left| \sum_{i=1}^{n} f(t_i)(\alpha'(s_i) - \alpha'(t_i))\Delta x_i \right| \\
&\leq K \sum_{i=1}^{n} |\alpha'(s_i) - \alpha'(t_i)| \Delta x_i \\
&\leq K(U(P, \alpha', x) - L(P, \alpha', x))
\end{aligned}$$

Now, (1) $\implies$ (2):
Note $f \in \mathcal{R}([a, b]; \alpha)$ and $\alpha' \in \mathcal{R}([a, b]) \implies \exists P_1, P_2.\ U(P_1, f, \alpha) - L(P_1, f, \alpha) < \frac{\epsilon}{2}$ and $U(P_2, \alpha', x) - L(P_2, \alpha', x) < \frac{\epsilon}{2(1+K)}$. Then

$$\begin{aligned}
\left| \sum_{i=1}^{n} f(t_i)\alpha'(t_i)\Delta x_i - \int_a^b f d\alpha \right| &\leq \left| \sum_{i=1}^{n} f(t_i)\alpha'(t_i)\Delta x_i - \sum_{i=1}^{n} f(t_i)\Delta\alpha_i \right| + \left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f d\alpha \right| \\
&\leq K(U(P, \alpha', x) - L(P, \alpha', x)) + U(P, f, \alpha) - L(P, f, \alpha) \\
&< K\left(\frac{\epsilon}{2(K+1)}\right) + \frac{\epsilon}{2} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

and the last theorem in §7.2 guarantees that $f\alpha' \in \mathcal{R}([a, b])$ and $\int_a^b f\alpha' dx = \int_a^b f d\alpha$.
The proof that (2) $\implies$ (1) is similar (exercise to reader).