Notes for Math 451

Advanced Calculus I

Albon Wu

May 16, 2024

Contents

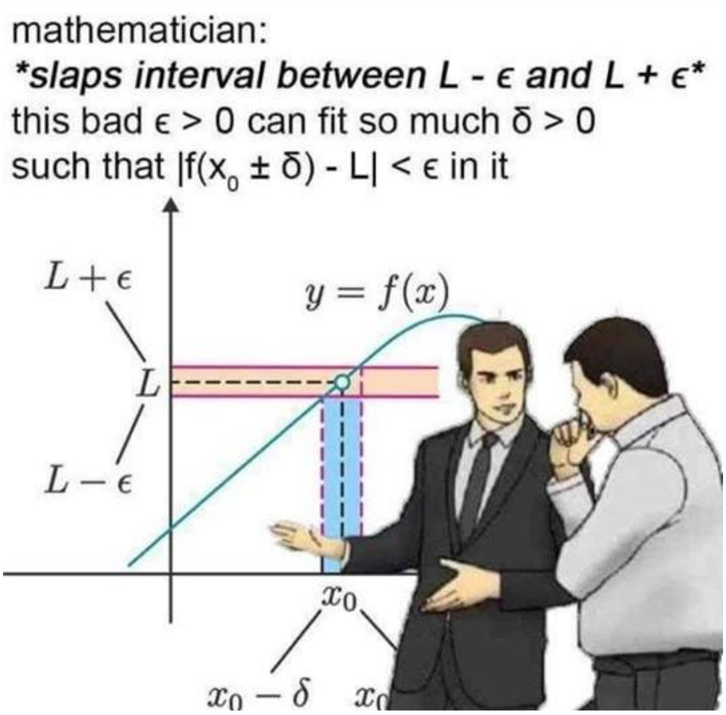
*Additional measure theory content.

0 Introduction

Elementary analysis—sequences, differentiation, and integration, plus some basic algebra and topology.

Professor: Tasho Kaletha.

Textbook: *Elementary Analysis: The Theory of Calculus* by Ross. Additional measure theory notes use *Measures, Integrals and Martingales* by Schilling.



Source: “Measure 0 Memes for Lebesgue Integrable Teens”

1 Sets

1.1 The Natural Numbers $\mathbb{N}$

$\mathbb{N} = \{0, 1, 2, 3, \ldots\} \subset \mathbb{Z}$ is the set of natural numbers (some authors exclude 0). Below are some properties of $\mathbb{N}$:

N1) $\mathbb{N}$ is not empty.

N2) $\mathbb{N}$ has a smallest element.

N3) Every $n \in \mathbb{N}$ has a successor $n + 1 \in \mathbb{N}$.

N4) If $X \subset \mathbb{N}$ is such that $0 \in X$ and $n \in X \longrightarrow n + 1 \in X$, then $X = \mathbb{N}$.

The last property says that $\mathbb{N}$ is the smallest set with the first three properties. It is then natural to conjecture the following:

> **Claim.** The four properties above uniquely characterize $\mathbb{N}$.

We will formulate this with more precise language later. First, we introduce some familiar technology.

> **Theorem 1.1.1** (Induction)**.** *Let $P(n)$ be a logical statement with parameter $n \in \mathbb{N}$. Assume $P(0)$ is true and $P(n) \longrightarrow P(n + 1)$. Then $\forall n \in \mathbb{N}$, $P(n)$ is true.*

Proof. Define $X := \{n \in \mathbb{N} \mid P(n) \text{ is true}\} \subset \mathbb{N}$. Then, since $0 \in X$ and $n \in X \longrightarrow n + 1 \in X$, we know $X = \mathbb{N}$. □

We also introduce the concept of recursion. To construct a collection $(S_n)_{n\in\mathbb{N}}$ of sets or maps, it is enough to define S_0 and S_{n+1} given S_n. For instance, to construct $f : \mathbb{N} \longrightarrow S$, it is enough to specify $f(0) \in S$ and $f(n + 1) \in S$ given $f(n) \in S$.

> **Lemma 1.1.1.** Let $(S_n)_{n\in\mathbb{N}}$ and $(S'_n)_{n\in\mathbb{N}}$ be collections of sets. Assume $S_0 = S'_0$ and $S_n = S'_n \longrightarrow S_{n+1} = S'_{n+1}$. Then $S_n = S'_n$ for $n \in \mathbb{N}$.

Proof. Follows directly from induction on n. □

> **Definition 1.1.1.** A **Peano triple** (P, e, s) consists of:
>
> - a set P
> - an element $e \in P$
> - an injective map $s : P \longrightarrow P$
>
> such that
>
> P1) $e \notin S(P)$
>
> P2) If $X \subset P$ is such that $e \in X$ and $S(X) \subset X$, then $X = P$.

Peano triples are essentially abstractions of the properties of $\mathbb{N}$ we stated above. For instance, it is easy to show that for a Peano triple (P, e, s), we have $P = \{e\} \cup S(P)$ using P2). If we use the successor function for s, the result mirrors property N4).

Now we can address our conjecture from earlier.

Theorem 1.1.2. *Let (P, e, s) be a Peano triple. There exists a unique bijection $f : \mathbb{N} \longrightarrow P$ such that $f(0) = e$ and $f(n+1) = s(f(n))$.*

This result says that, for any Peano triple (P, e, s), we can map every natural number n to one element in P whose successor is the image of $n + 1$. That is, all Peano triples are equivalent to $\mathbb{N}$ up to bijections.

Proof. f is recursively defined, meaning it is unique by Lemma 1.1.1. It suffices to show it is bijective.

For injectivity, define the logical statement $T(n) := (\forall m \in \mathbb{N} : f(n) = f(m) \implies n = m)$. We induce on n. First, consider $T(0)$ and take $m \in \mathbb{N}$. If $m = n = 0$, we are done. Otherwise, we can write $m = m' + 1$ for $m' \in \mathbb{N}$, and we write

$$f(m) = f(m' + 1) = S(f(m')) \in S(P).$$

By definition, $e \notin S(P)$, so $f(m) \neq e = f(n)$. This shows the contrapositive of $T(0)$.

Now suppose $T(n)$. Again, we will show the contrapositive. Take $m \in \mathbb{N}$ such that $m \neq n + 1$. If $m = 0$, we are done by $T(0)$. Otherwise, $m = m' + 1$ for $m' \in \mathbb{N}$. So $m' + 1 \neq n + 1 \implies m' \neq n$, and $f(m') \neq f(n)$ by assumption. It follows that

$$f(n+1) = s(f(n)) \neq s(f(m')) = f(m' + 1)$$

since s is injective.

It remains to show surjectivity. Let $X = f(\mathbb{N}) \subseteq P$. Now $e = f(0) \in f(\mathbb{N}) = X$ and $s(X) = s(f(\mathbb{N})) = f(\mathbb{N} + 1) \subseteq f(\mathbb{N}) = X$. By P2), $X = P$, completing the proof. □

1.2 The Integers $\mathbb{Z}$

$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ is the set of integers. Although the construction of $\mathbb{Z}$ as “the natural numbers and their negatives” is intuitive, it would be nice to define $\mathbb{Z}$ in a way that only uses $\mathbb{N}$ and its axioms without resorting to ad-hoc definitions like “negative” and their behavior with arithmetic.

One construction represents each integer as a difference of natural numbers. Since $-5 = 0 - 5$, we would represent -5 as $(0, 5)$. We would also need a new notion of equality since $(0, 5)$ and $(1, 6)$ represent the same number; try $(a, b) \equiv (a', b') \iff a - b = a' - b'$.

Now we introduce an abstraction for $\mathbb{Z}$ analogous to Peano triples for $\mathbb{N}$:

Definition 1.2.1. An **abelian group** $(G, e, +)$ consists of

- a set G
- an element $e \in G$
- a map $+ : G \times G \longrightarrow G$

such that $\forall a, b, c \in G$:

A1) $(a + b) + c = a + (b + c)$

A2) $a + b = b + a$

A3) $a + e = a$

A4) There exists $a' \in G$ where $a + a' = e$.

Lemma 1.2.1. Let $(G, e, +)$ be an abelian group. Then $a + c = b + c \Longrightarrow a + b$.

Proof. Exercise, or see 412 notes. □

Our goal is to characterize $\mathbb{Z}$ with respect to an arbitrary abelian group. We will start with the following map:

Lemma 1.2.2. Let $(G, e, +)$ be an abelian group. Take $g \in G$. There exists a unique map $f : \mathbb{Z} \longrightarrow G$ such that $f(1) = g$ and $f(a + b) = f(a) + f(b)$.

Proof. Pretend we had f. What properties would it have?

1. $f(0) = f(0 + 0) = f(0) + f(0) \Longrightarrow f(0) = e$
2. For $n \in \mathbb{N}$, we have $f(n + 1) = f(n) + f(1) = f(n) + g$
3. $e = f(0) = f(n + (-n)) = f(n) + f(-n) \Longrightarrow -f(n) = -f(n)$

Properties 1) and 2) give a unique $f : \mathbb{N} \longrightarrow G$ (by the recursive uniqueness lemma from earlier), and 3) extends the domain to $\mathbb{Z}$.

We then get the first stated property of f for free: $f(1) = f(0 + 1) = f(0) + g = e + g = g$. It remains to show that f respects addition. Consider $f(a + b)$; we have a few cases.

<u>Case 1:</u> $a, b \in \mathbb{N}$. We induce on $a + b$. In the base case, we have $a + b = 0 \Longrightarrow a = 0 = b \Longrightarrow f(a + b) = f(0) = e = e + e = f(a) + f(b)$.

Now suppose the result holds for $a + b > 0$. WLOG, let $b > 0$. Then $f(a + b) = f(a + (b - 1) + 1) = f(a + (b - 1)) + g$. By assumption, we can simplify to $f(a) + f(b - 1) + g = f(a) + f(b)$.

<u>Case 2:</u> $a \in \mathbb{N}, b \in -\mathbb{N}$, $a + b \in \mathbb{N}$. Since $-b \in \mathbb{N}$, use Case 1: $f(a + b) + f(-b) + f(b) = f(a + b + (-b)) + f(b) = f(a) + f(b)$.

<u>Case 3:</u> $a \in \mathbb{N}, b \in -\mathbb{N}$, $a + b \in \mathbb{N}$. Set $a' = -b$ and $b' = -a$. Then $a' \in \mathbb{N}$ and $b' \in -\mathbb{N}$; moreover, $a' + b' \in \mathbb{N}$ since $a' + b' = -(a + b)$. So $f(a + b) = f(-(a' + b')) = -f(a' + b') = -(f(a') + f(b')) = -(f(-b) + f(-a)) = f(b) + f(a)$.

<u>Case 4</u>: $a \in -\mathbb{N}, b \in -\mathbb{N}$. Let $a' = -b$ and $b' = -a$. Adapt the proof from Case 3. □

So if we have an abelian group and a map f from $\mathbb{Z}$ respecting $+$ and e, then f will have these properties. But this map is not always meaningful; for instance, choosing $g = e$ collapses $\mathbb{Z}$ into $\{e\}$. We need a more precise characterization.

Definition 1.2.2. Let G be abelian. $g \in G$ is a **generator** if, for all $h \in G$, we have $h = g + g + \cdots + g$ or $-h = g + g + \cdots + g$.

g is **free** if $g + \overset{n \text{ times}}{\cdots} + g \neq e$ when $n > 0$.

Theorem 1.2.1. *Let G be abelian and take $g \in G$. The map $f : \mathbb{Z} \longrightarrow G$ of the previous lemma is injective if g is free, and surjective if it is a generator.*

Proof. First, we show injectivity. WLOG, let $a - b \in \mathbb{N}$. Then if $a \neq b$, we have $a - b \neq 0$, so $f(a-b) = g + \overset{a-b \text{ times}}{\cdots} + g \neq e \implies f(a) \neq f(b)$.

For surjectivity, it suffices to show $f(\mathbb{Z}) = G$ if g is a generator. Note that

$$f(\mathbb{Z}) = f(\mathbb{N}) \cup f(-\mathbb{N}) = \{g + \overset{n \text{ times}}{\cdots} + g \mid n \in \mathbb{N}\} \cup \{-(g + \overset{n \text{ times}}{\cdots} + g) \mid n \in \mathbb{N}\}.$$

The result follows since g is a generator. $\square$

So we have just shown a condition on g for when f is bijective. That is, up to bijections, $\mathbb{Z}$ is the unique abelian group with a free generator. Nice!

1.3 The Rational Numbers $\mathbb{Q}$

$\mathbb{Q} := \{\frac{m}{n} \mid m, n \in \mathbb{Z}, n \neq 0\}$ is the set of rational numbers. Take (a, b) as a name for $\frac{a}{b}$. Then $(a, b) \equiv (a', b')$ when $ab' = a'b$.

For $\mathbb{N}$, our abstraction was the Peano triple. For $\mathbb{Z}$, it was the abelian group. The underlying structure of $\mathbb{Q}$ is the ordered field.

Definition 1.3.1. A **field** is a tuple $(F, 0, 1, +, \cdot)$ consisting of

- a set F
- elements $0, 1 \in F$ such that $0 \neq 1$
- maps $+, \cdot$ from $F \times F$ to F

subject to

F1) $(F, 0, +)$ is abelian

F2) $(F^\times, 1, \cdot)$ is abelian

F3) $a(b + c) = ab + ac$

Remark. This definition is an equivalent formulation of "commutative, nontrivial ring where every nonzero element is a unit."

Definition 1.3.2. Let S be a set. A **total order** on S is a relation $\leq$ with the following properties for all $a, b, c \in S$:

O1) $a \leq b \lor b \leq a$

O2) $a \leq b \land b \leq a \implies a = b$

O3) $a \leq b \land b \leq c \implies a \leq c$

Definition 1.3.3. An **ordered field** is a tuple $(F, 0, 1, +, \cdot, \leq)$ where $(F, 0, 1, +, \cdot)$ is a field and $\leq$ is a total order on F subject to

O4) $a \leq b \implies a + c \leq b + c$

O5) $a \leq b \implies ac \leq bc$ if $0 \leq c$

Remark.

1. $(F, 0, +)$ uses additive notation, while $(F^\times, 1, \cdot)$ uses multiplicative notation.
2. $\cdot$ is defined on all of $F \times F$, so $a \cdot 0$ makes sense.
3. Many more properties follow. For instance, $a \leq b \implies a + (-a - b) \leq b + (-a - b) \implies -b \leq -a$.

Theorem 1.3.1. *Let F be any ordered field. There exists a unique map $f : \mathbb{Q} \longrightarrow F$ such that:*

- $f(a + b) = f(a) + f(b)$
- $f(ab) = f(a)f(b)$
- $a \leq b \implies f(a) \leq f(b)$

Moreover, f is automatically injective.

Proof. We start with the map $f : \mathbb{Z} \longrightarrow F$ whose existence is given from a previous lemma, where F is the abelian group $(F, 0, +)$ with generator 1.

Moreover, we claim that f is injective and satisfies $f(n) > 0$ for all $n > 0$. We can quickly check that $1_F > 0_F$, so by induction it follows that $1_F + \cdots + 1_F > 0_F$. We then express positive $n \in \mathbb{N}$ as the sum of n ones; since f respects addition and sends 1 to 1_F, we conclude $f(n) > 0$.

For injectivity, take $a, b \in \mathbb{Z}$ where $f(a) = f(b)$. WLOG, let $a \geq b$. Note that $a - b \in \mathbb{N}$, so it directly follows from the above result that $f(a) - f(b) = 0 \implies f(a - b) = 0 \implies a = b$.

We also claim that f respects multiplication of $a, b \in \mathbb{Z}$. Suppose $b \in \mathbb{N}$; then we induce on b. In the base case $b = 1$, we simply have $f(a \cdot 1) = f(a) = f(a) \cdot 1 = f(a) \cdot f(1)$.

Now suppose the property holds for arbitrary b. We write

$$f(a \cdot (b + 1)) = f(a \cdot b + a) = f(a \cdot b) + f(a) = f(a) \cdot f(b) + f(a)$$
$$= f(a) \cdot (f(b) + 1) = f(a) \cdot f(b + 1).$$

$b = 0$ is an exercise. Now if $b \in -\mathbb{N}$, we write

$$f(a \cdot b) = f(a \cdot (-(-b))) = f((-a) \cdot (-b)) = f(-a) \cdot f(-b)$$
$$= (-f(a)) \cdot (-f(b)) = f(a) \cdot f(b).$$

This f is useful because it leads naturally to a map $f : \mathbb{Q} \longrightarrow F$. For $a, b \in \mathbb{Z}$ and $b \neq 0$, define $f(a/b) := f(a) \cdot f(b)^{-1}$. Note that since we previously showed f is injective when its domain is $\mathbb{Z}$, we have $b \neq 0 \implies f(b) \neq 0 \implies f(b) \in F^\times$. So $f(b)^{-1}$ exists.

It is also worth checking that f is well defined; that is, equivalent representations of the same rational number map to the same image. Suppose $a/b = a'/b'$ so that $ab' = a'b$. Then $f(a) \cdot f(b') = f(a') \cdot f(b) \implies f(a) \cdot f(b)^{-1} = f(a') \cdot f(b')^{-1}$, as desired.

We must also show that this extended f is a ring homomorphism. The proof is fairly routine, so we leave it as an exercise.

It remains to prove that if $r < s$, then $f(r) < f(s)$ for all $r, s \in \mathbb{Q}$. Write $r = a/b$ and $s = a'/b$, where we force a common denominator $b > 0$. Then

$$r < s \implies a < a' \implies a' - a > 0 \implies f(a' - a) > 0 \implies f(a') - f(a) > 0$$

$$\Longrightarrow f(a') > f(a) \Longrightarrow f(a')f(b)^{-1} > f(a)f(b)^{-1}$$

where the last step is possible since $f(b)^{-1} > 0$. So $f(s) > f(r)$, and in fact we get injectivity for free because $r \neq s \Longrightarrow (r < s) \lor (s < r) \Longrightarrow f(s) \neq f(r)$. □

Theorem 1.3.1 tells us that $\mathbb{Q}$ is the smallest ordered field. This is because there exists an injection from $\mathbb{Q}$ to any ordered field F, meaning that $\mathbb{Q}$ is bijective to some subset of F. Pretty cool!

For bonus points, here are some miscellaneous properties about fields.

Theorem 1.3.2. *For a field F and $a, b, c \in F$:*

1. $a + c = b + c \Longrightarrow a = b$
2. $a \cdot 0 = 0$
3. $a(-b) = -ab$
4. $(-a)(-b) = ab$
5. *If* $c \neq 0$, *then* $ac = bc \Longrightarrow a = b$
6. $ab = 0$ *implies* $a = 0$ *or* $b = 0$

Proof.

1. Follows from right addition of $-c$ to both sides.
2. See 412 notes.
3. $ab + a(-b) = a(b + (-b)) = a \cdot 0 = 0$. So $a(-b)$ is the additive inverse of ab, as desired.
4. $(-a)(-b) + a(-b) = (a + (-a))(-b) = 0 \cdot (-b) = 0$. By the previous part, $(-a)(-b)$ then equals ab, the additive inverse of $-ab$.
5. Follows from right multiplication by c^{-1} on both sides.
6. Suppose $b \neq 0$ and $ab = 0$. Then $0 = ab(b^{-1}) = a$. Otherwise, done.

□

We can also prove some results using the ordered field axioms:

Theorem 1.3.3. *For a field F and $a, b, c \in F$:*

1. $a \leq b \Longrightarrow -b \leq -a$
2. $a \leq b$ *and* $c \leq 0$ *implies* $bc \leq ac$
3. $0 \leq a$ *and* $0 \leq b$ *implies* $0 \leq ab$
4. $0 \leq a^2$ *for all* a
5. $0 < 1$
6. $0 < a$ *implies* $0 < a^{-1}$
7. $0 < a < b$ *implies* $0 < b^{-1} < a^{-1}$

Proof.

1. $a \leq b$ implies $a + (-a + (-b)) \leq b + (-a + (-b))$, so $-b \leq -a$.
2. By the previous part, $0 \leq -c$, so $-ac \leq -bc$ and $bc \leq ac$.
3. $0 \cdot a \leq ba \Longrightarrow 0 \leq ab$.
4. $0 \leq a$ is straightforward. If $a \leq 0$, we have $0 \leq a \cdot a = a^2$ by (1).
5. Suppose $1 \leq 0$. Then $0 \cdot 1 \leq 1 \cdot 1 \Longrightarrow 0 \leq 1$, a contradiction.
6. Suppose $0 < a$ but $a^{-1} < 0$. Then $0 \cdot a^{-1} > aa^{-1} \Longrightarrow 0 > 1$, a contradiction.
7. Adapt the proof of (1) using multiplicative inverses to obtain $b^{-1} < a^{-1}$. Then $0 < b^{-1}$ follows from (5).

□

1.4 The Real Numbers $\mathbb{R}$

So far, we have $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{Q}$. But this is not enough; it is well-known that $\sqrt{2} \notin \mathbb{Q}$. So our goal is to "fill up" $\mathbb{Q}$ with the missing numbers.

Let F be an ordered field.

Definition 1.4.1. For $a \in F$, the **absolute value** of a, denoted $|a|$, is the following function:

$$|x| = \begin{cases} x & x \geq 0 \\ -x & x \leq 0 \end{cases}$$

Definition 1.4.2. For $a, b \in F$, the **distance** between a and b, denoted $\text{dist}(a, b)$, is defined as $\text{dist}(a, b) = |a - b|$.

Theorem 1.4.1. *Take $a, b \in F$. Then the following properties hold:*

1. $|a| \geq 0$
2. $|ab| = |a| \cdot |b|$
3. $|a + b| \leq |a| + |b|$

Proof.

1. Follows by definition.
2. It is straightforward to check that if a and b have the same sign, $|ab| = |a| \cdot |b| = ab$. Otherwise, $|ab| = |a| \cdot |b| = -ab$.
3. By definition, $-|a| \leq a \leq |a|$ and $-|b| \leq b \leq |b|$. So $-|a| - |b| \leq a + b \leq |a| + |b|$. This implies $\pm(a + b) \leq |a| + |b|$, so $|a + b| \leq |a| + |b|$.

□

The last result is also called the **Triangle Inequality** because for $x, y, z \in F$, we can substitute $a = x - y$ and $b = y - z$ to obtain $|x - z| \leq |x - y| + |y - z| \Longrightarrow \text{dist}(x, z) \leq \text{dist}(x, y) + \text{dist}(y, z)$.

Geometrically, this is analogous to the statement that the combined length of any two sides of a triangle is greater than the length of the third.

We will introduce some important definitions shortly, but here is a motivating example for the new terminology.

Claim. If $a \in \mathbb{Q}$, $a > 0$, $a^2 < 2$, there exists $b \in \mathbb{Q}$ where $a < b$, $b^2 < 2$.

Proof. Define $b := a - \frac{a^2-2}{a+2} \in \mathbb{Q}$. Since $a > 0$, we have $a < b$. Then we write

$$\begin{aligned} b^2 - 2 &= \left(\frac{2a+2}{a+2}\right)^2 - 2 = \frac{4a^2 + 8a + 4}{a^2 + 4a + 4} - 2 \\ &= \frac{2a^2 - 4}{a^2 + 4a + 4} = 2\left(\frac{a^2 - 2}{a^2 + 4a + 4}\right). \end{aligned}$$

$a^2 - 2$ is negative, so $b^2 < 2$. □

This implies that the set $S = \{a \in \mathbb{Q} \mid a^2 \leq 2\}$ has no largest element. Even so, it seems like the elements of S tend toward a number $\sqrt{2}$ without actually reaching it.

Definition 1.4.3. Let $(S, \leq)$ be an ordered set. An element $s_0 \in S$ is a **maximum** if $s \leq s_0$ for all $s \in S$.

Definition 1.4.4. Let $(S, \leq)$ be an ordered set with $T \subseteq S$.

1. An element $s_0 \in S$ is an **upper bound** for T if $t \leq s_0$ for all $t \in T$.
2. If T has an upper bound, it is **bounded above**.
3. An $s_0 \in S$ is called a least upper bound if
   - s_0 is an upper bound for T
   - if s_1 is an upper bound for T, we have $s_0 < s_1$

The following properties are immediate from these definitions:

Claim. Let S be an ordered set with $T \subseteq S$. Then the following hold:

- If T has a max, it is unique.
- If T has a least upper bound, it is unique.
- A max is a least upper bound.
- A least upper bound for T is a max iff it lies in T.

Definition 1.4.5. An ordered set S has the **least upper bound property (LUBP)** if any subset that is non-empty and bounded above has a least upper bound.

Definition 1.4.6. If $T \subseteq S$ has a max/min, denote it by $\max T$ or $\min T$, respectively. If $T \subseteq S$ has a least upper bound, denote it by $\sup T$—the **supremum** of T. The greatest lower bound is called the **infimum**, inf for short.

The LUBP is significant because it formulates the notion of "missing numbers" solely in terms of order. So, for instance, $\mathbb{Q}$ does not have the LUBP because $S \subset \mathbb{Q}$ does not have a least upper bound.

From now on, let F be an ordered field with LUBP.

> **Theorem 1.4.2** (Archimedian Property). *Let $x, y \in F$ and $x > 0$. There exists $n \in \mathbb{N}$ such that $nx > y$.*

Proof. If $y < 0$, pick $n = 1$ and we are done.

Otherwise, suppose the contrary. Then y is an upper bound for the set $S = \{nx \mid n \in \mathbb{N}\}$. By LUBP, $z = \sup S \in F$. Then $z - x$ is not an upper bound for S, so there exists $n \in \mathbb{N}$ such that $nx > z - x$. But we win because this implies $(n+1)x > z$, a contradiction. □

> **Remark.** When $x = 1$, the Archimedian Property implies that $\mathbb{N}$ is unbounded in F.

> **Theorem 1.4.3** (Density of $\mathbb{Q}$ in F). *Given $z, w \in F$ where $z < w$, there exists $q \in \mathbb{Q}$ with $z < q < w$.*

Proof. Our approach is to "zoom in" on z and w.

Apply the Archimedian Property to $x = w - z$ and $y = 1$ to get $n \in \mathbb{N}$ such that $n(w - z) > 1$. This gives us $nw > nz + 1$. Now consider $S = \{m \in \mathbb{Z} \mid nz < m\}$. We claim that S is non-empty and bounded below (by an integer).

Before proving this, let's admit the claim and let $m_0 := \min S$. Then $m_0 - 1 \leq nz < m_0$, hence $nz < m_0 \leq nz + 1 < nw$ and $z < \frac{m_0}{n} < w$, as desired.

Now we show the claim. We apply the Archimedian Property to $x = 1, y = nz$ and $x = 1, y = -nz$ to get $m_1, m_2 \in \mathbb{N}$ with $m_1 > nz, m_2 > -nz$. So $m_1 \in S$ by definition, so S is non-empty, and m_2 is a lower bound for S, as desired. □

As a sanity check, we should make sure the LUBP actually guarantees the existence of elements we know are not in $\mathbb{Q}$.

> **Proposition 1.4.1.** $\sqrt{2} \in F$. *That is, there exists an element in F that equals $1_F + 1_F$ when multiplied by itself.*

Proof. Consider $S = \{q \in \mathbb{Q} \mid q^2 < 2\} \subset \mathbb{Q} \subset F$. Clearly, S is non-empty and bounded above. By LUBP, $x := \sup S \in F$. We will show $x^2 = 2$.

First, we claim $x^2 \geq 2$. Suppose the contrary: that $x^2 < 2$. We "nudge" x toward 2 and apply density for a contradiction. Take $0 < k < \frac{2-x^2}{2(x+1)}$. Since $x > 0$, we have $0 < k < 1$. So $(x+k)^2 - x^2 = 2xk + k^2$.

We then weaken the equality to obtain $(x+k)^2 - x^2 \leq 2xk + 2k$, which Kaletha calls "the key to analysis—being generous at the right moment." We write

$$(x+k)^2 - x^2 \leq 2xk + 2k = 2k(x+1) < 2 - x^2,$$

where the last step follows from construction of k. Hence $(x+k)^2 < 2$. By density of $\mathbb{Q}$, there exists $q \in \mathbb{Q}$ such that $x < q < x+k$. Then $q^2 < (x+k)^2 < 2$, so $q \in S$. But $x < q$ by construction, which contradicts the definition of x.

Now we show that $x^2 \leq 2$. Suppose $x^2 > 2$; then take $0 < k = \frac{x^2}{2x} < x$. Given $t \geq x - k$, we have

$$\begin{aligned} x^2 - t^2 &\leq x^2 - (x-k)^2 = 2kx - k^2 \\ &\leq 2kx < x^2 - 2 \end{aligned}$$

Thus, $t^2 > x^2 - x^2 + 2 \Longrightarrow t^2 > 2$, implying that $t \notin S$. Then $x - k$ is an upper bound for S, a contradiction. □

So we know that if such a field F with LUBP existed, it would have these fun properties. Now we actually construct such an F.

Definition 1.4.7. A **Dedekind cut** is a subset $S \subsetneq \mathbb{Q}$ such that

1. $S \neq \emptyset$.
2. If $a \in S$ and $b \in \mathbb{Q}$ such that $b < a$, then $b \in S$. This is called **downward closure**.
3. If $a \in S$, there exists $b \in S$ where $b > a$. In other words, $\max S$ does not exist.

So $\{x \in \mathbb{Q} \mid x^2 < 2 \text{ or } x < 0\}$ is a Dedekind cut because we showed previously that it has no upper bound.

Lemma 1.4.1. A Dedekind cut S is bounded above.

Proof. Assume not. Then for any $q \in \mathbb{Q}$, there exists $s \in S$ where $s > q$. By downward closure, $q \in S$. Thus, $S = \mathbb{Q}$, a contradiction. □

Theorem 1.4.4. *The maps $\alpha : DC \longrightarrow F$ and $\beta : F \longrightarrow DC$, where DC is the set of Dedekind cuts in F and*

$$\begin{aligned} \alpha(S) &= \sup S \in F \\ \beta(f) &= \{q \in \mathbb{Q} \mid q < f\}, \end{aligned}$$

satisfy $\alpha = \beta^{-1}$.

Proof. We prove a series of claims.

Claim. Let $f \in F$ and define $S := \beta(f)$. Let $f_1 = \alpha(S)$. Then $f_1 = f$.

We have $S := \{q \in \mathbb{Q} \mid q < f\}$. By construction, f is an upper bound for S, so $f_1 \leq f$. If $f_1 < f$, density gives $x \in \mathbb{Q}$ such that $f_1 < x < f$.

But since $x < f$, we know $x \in F$, contradicting the construction of f_1.

Claim. Let $S \in DC$ and $f = \alpha(S)$. If $T = \beta(f)$, then $S = T$.

First, we show $S \subseteq T$. We have $T = \{x \in \mathbb{Q} \mid x < f\}$. Take $x \in S$. Then $x \leq f$ since $f = \sup S$. If $x = f$, then $f \in S$ and thus S always has a max, which is untrue. So $x < f$ and $x \in T$.

For the opposite inclusion, let $x \in T$. So $x < f = \sup S$. Then x is not an upper bound of s, and there exists a witness $s \in S$ where $s > x$. By downward closure, $x \in S$. □

So we know that there is a correspondence between Dedekind cuts and the elements of an arbitrary field. This leads to the following important result:

Theorem 1.4.5. *There exists an ordered field with the LUBP.*

Proof sketch. Take DC as the underlying set and endow it with order by inclusion. Then properties O2) and O3) are immediate.

For O1), let $S, T \in DC$ where $S \neq T$. Assume WLOG that there exists $s \in T$. Then downward closure for T implies $t < s$ for all $t \in T$, so $t \in S$. Hence $T \subseteq S$.

Now we show that $(DC, \leq)$ satisfies LUBP.

Take non-empty $X \subseteq DC$ and let $S_1 \in DC$ such that $S \leq S_1$ for all $S \in X$. Define $S_0 = \bigcup_{S \in X} S \subseteq \mathbb{Q}$. We will show $S_0 \in DC$. Since S_0 subsumes $S \in X$, we have $S \subset S_0 \subset S_1$; thus, S_0 dominates S. This implies S_0 is an upper bound, but it is dominated by any upper bound S_1. So S_0 is the least upper bound of X.

But we still have to check $S_0 \in DC$! Since $X \neq \emptyset$, there exists $S \in X$. And since $S \in DC$, we have $\emptyset \neq S \subset S_0$, so $S_0 \neq \emptyset$. Moreover, since $S \subset S_1$, then for all $S \in X$, we have $S_0 \subset S_1 \in DC$, so $S_0 \neq \mathbb{Q}$.

To show S_0 has no max, let $a \in S_0$. Then there exists $S \in X$ such that $a \in S$. Since $S \in DC$, there exists $b \in S \subset S_0$ where $b > a$. Hence S_0 has no maximum and it is indeed a Dedekind cut.

Now we turn $F := DC$ into a field. For the identities, we define

$$0_F = \{x \in \mathbb{Q} \mid x < 0\}$$
$$1_F = \{x \in \mathbb{Q} \mid x < 1\}.$$

Given $S, T \in DC$, define

$$S + T = \{s + t \mid s \in S, t \in T\}$$
$$-S = \{-s \mid \exists r > 0 \,:\, s - r \notin S\}.$$

The idea behind the additive inverse is to reflect the "ray" representing the Dedekind cut about the point 0. This breaks downward closure, so we then take the complement and omit any resulting maximum.

Now assume $S > 0$ and $T > 0$ using our definition of order. Then define

$$S \cdot T := \{s \cdot t \mid s \in S, t \in T \text{ for } s > 0 \text{ or } t > 0\}.$$

If we didn't restrict the sign of s, t, then $S \cdot T$ would equal $\mathbb{Q}$ by downward closure.

So we have some operations for DC; now we show that they satisfy the field axioms.

Claim. $(F, 0, +)$ is abelian.

Proof. A1) and A2) follow from $\mathbb{Q}$. For A3), take $S \in DC$. We want to prove $S + 0_F \subset S$, so take $s \in S$ and $t \in 0_F$. Then $t < 0$, so $s + t < s \Longrightarrow s + t \in S$.

Now take $s \in S$. Then there exists $t \in S$ such that $t > s$, and then $s = t + (s - t)$. Since $t \in S$ and $s - t < 0$, we have $s \in S + 0_F$. So $S = S + 0_F$.

For A4), we want to show $S + (-S) \subset 0_F$. Take $s \in S$ and $t \in -S$. By definition, $-t \notin S$, so $s < -t$ by downward closure. Therefore, $s + t < 0$, so $s + t \in 0_F$.

Now take $v \in 0_F$. Let $w = -v/2 > 0$. By a "blowing up" argument like the one from the Archimedian Property proof, there exists $n \in \mathbb{Z}$ such that $n \cdot w \in S$ and $(n + w) \cdot w \notin S$. Then $-(n + 2)w \in -S$, which gives $n \cdot w + (-(n + 2)w) = -2w = v$, and we are done. $\square$

Claim. $(F^{\times}, 1, \cdot)$ is abelian.

Proof. A1), A2) follow from $\mathbb{Q}$. For A3), take $s \in S$ and $t \in 1_F$. If $s < 0$ or $t < 0$, then $s \cdot t < 0$ and thus $s \cdot t \in S$ (a ray pointing downward and starting at a positive number absorbs all negative numbers).

If $s, t > 0$, then since $t < 1$, we have $s \cdot t < s$, so $s \cdot t \in S$ by downward closure. This shows $S \cdot 1_F \subset S$.

To show the opposite inclusion, take $s \in S$. Since S has no max, we can choose $t \in S$ such that $t > s$. Then $s = t \cdot \frac{s}{t}$. The result follows because $\frac{s}{t} \in 1_F$. $\square$

We leave A4), F3), O4), and O5) as exercises. $\square$

Theorem 1.4.6. *If F_1, F_2 are ordered fields with LUBP, there exists a unique bijection $f : F_1 \longrightarrow F_2$ satisfying*

$$
\begin{aligned}
f(a + b) &= f(a) + f(b) \\
f(ab) &= f(a)f(b) \\
a < b &\Longrightarrow f(a) < f(b).
\end{aligned}
$$

Proof. We define $f : F_1 \longrightarrow F_2$ such that $f = \alpha_2 \circ \beta_1$ for $\alpha_2 : DC \longrightarrow F_2$ and $\beta_1 : F_1 \longrightarrow DC$ as defined previously.

Now take $n_1, n_2 \in F_1$. We have

$$
\begin{aligned}
(\alpha_2 \circ \beta_1)(n_1 + n_2) &= \sup\{q \in \mathbb{Q} \mid q < n_1 + n_2\} \\
&= n_1 + n_2 \\
&= \sup\{q \in \mathbb{Q} \mid q < n_1\} + \sup\{q \in \mathbb{Q} \mid q < n_2\} \\
&= (\alpha_2 \circ \beta_1)(n_1) + (\alpha_2 \circ \beta_1)(n_2)
\end{aligned}
$$

So f respects addition. Similarly, we can show it respects multiplication:

$$
\begin{aligned}
(\alpha_2 \circ \beta_1) \cdot (n_1 n_2) &= \sup\{q \in \mathbb{Q} \mid q < n_1 n_2\} \\
&= n_1 n_2 \\
&= \sup\{q \in \mathbb{Q} \mid q < n_1\} \cdot \sup\{q \in \mathbb{Q} \mid q < n_2\} \\
&= (\alpha_2 \circ \beta_1)(n_1) \cdot (\alpha_2 \circ \beta_1)(n_2).
\end{aligned}
$$

Now suppose $n_1 < n_2$. Then we have $\beta_1(n_1) = \{q \in \mathbb{Q} \mid q < n_1\} \subset \{q \in \mathbb{Q} \mid q < n_2\} = \beta_1(n_2)$, which implies $\beta_1(n_1) < \beta_1(n_2)$, so β_1 respects the order. And by a property of sup, $\beta_1(n_1) \subset \beta_1(n_2)$ implies

$$\alpha_2(\beta_1(n_1)) = \sup(\beta_1(n_1)) < \sup(\beta_1(n_2)) = \alpha_2(\beta_1(n_2))$$

so α_2 also respects the order. Therefore, f respects order.

It remains to show uniqueness. Take $q \in S_x$. Then, by assumption, $i_1(q) < x$. We then have $f(i_1(q)) < x$. But since it is the composition of injections, $f \circ i_1$ is itself an injection. Hence $f \circ i_1 = i_2$ since i_2 is the unique injection $\mathbb{Q} \longrightarrow F_2$, so $i_2(q) = f(i_1(q)) < f(x)$. Therefore, $q \in S_{f(x)}$ and so $S_x \subset S_{f(x)}$.

And since f is a bijection, we can use the existence of the injection f^{-1} to show the opposite inclusion. Consider $q \in \mathbb{Q}$ so that $i_2(q) < f(q)$. Then $f^{-1}(i_2(q)) < f^{-1}(f(q)) = q$, and by uniqueness of i_1, we write $i_1(q) = f^{-1}(i_2(q)) < q$. This shows $S_x = S_{f(x)}$.

So the Dedekind cut in F_1 induced by $x \in F_1$, which we obtain from $\beta_1(x)$, corresponds to the Dedekind cut in F_2 induced by $f(x)$, namely $S := \{x_2 \in F_2 \mid x_2 < f(x)\}$. Therefore, $\alpha_2(S) = \sup S = f(x)$, so we conclude that $(\alpha_2 \circ \beta_1)(x) = f(x)$. □

Definition 1.4.8. The unique ordered field with LUBP is called the field of **real numbers** $\mathbb{R}$.

Note that, by Theorem 1.4.6, any two fields with the LUBP are isomorphic. This means that there is no "canonical real number;" $\sqrt{2}$, $S \in DC$, or a Cauchy sequence (more on this later) are all *presentations* of real numbers: elements of the unique field with the LUBP.

Nevertheless, $\mathbb{R}$ still has the properties we would expect it to have in the familiar sense (as the rationals plus the irrationals).

Lemma 1.4.2. Let $a \in \mathbb{R}$ and $a \geq 0$. Assume $a < q$ for all $q \in \mathbb{Q}$ and $q > 0$. Then $a = 0$.

Proof. Assume not. Then $a > 0$. By density, there is $q \in \mathbb{Q}$ where $0 < q < a$, contradiction. □

2 Sequences and Series

2.1 The Basics

Definition 2.1.1. A **sequence** of real numbers is a map $x : \mathbb{N} \longrightarrow \mathbb{R}$.

Remark.

1. We may think of x as the enumeration of its image: $x_0, x_1, x_2, \dots$. This is not to be confused with the set $\{x_0, x_1, x_2, \dots\}$ since sets cannot have duplicate elements.
2. We can shift the index as we wish. So $x_{-3}, x_{-2}, \dots, x_0, \dots$ is an equally valid sequence. In general, if we want to move the index, we can just choose a Peano triple with a different starting element.
3. We sometimes denote a sequence by $(x_n)_{n=0}^{\infty}$.

Example 2.1.1. The following are sequences:

1. $x_n = n^2$
2. $x_n = (-1)^n$
3. $x_n = \frac{1}{n}$.
4. $x_1 = 1$, $x_2 = 1$, $x_{n+2} = x_{n+1} + x_n$.

Definition 2.1.2. A sequence x **converges** to $L \in \mathbb{R}$ if:

$$\forall \epsilon > 0 \,:\, \exists N \in \mathbb{N} \,:\, \forall n > N \,:\, |x_n - L| < \epsilon.$$

We write $x_n \longrightarrow L$.

This definition says that the sequence eventually falls within the margin of error ϵ, which can be arbitrarily small (but it doesn't *always* have to, only eventually—which is a metaphor for life or something).

Lemma 2.1.1.

$$\frac{1}{n} \longrightarrow 0.$$

Proof. We start with a "once upon a time:" let $\epsilon > 0$.

By the Archimedian Property, there exists $N \in \mathbb{N}$ where $N > 1/\epsilon$. Then for $n > N$, we have $|x_N - L| = |1/n - 0| = 1/n < 1/N < \epsilon$. □

Lemma 2.1.2. If $x_n \longrightarrow L_1$ and $x_n \longrightarrow L_2$, then $L_1 = L_2$.

Proof. Take $\epsilon > 0$ and choose the lower bound N as in Definition 2.1.2. We can do this by choosing $\max(N_{L_1}, N_{L_2})$. For $n > N$, we have

$$|L_1 - L_2| = |L_1 - x_n + x_n - L_2| \leq |L_1 - x_n| + |L_2 - x_n| < \epsilon + \epsilon = 2\epsilon.$$

Since ϵ can be arbitrarily small, $|L_1 - L_2| = 0$. □

Definition 2.1.3. A sequence (x_n) is **bounded above** or **below** if

$$S = \{x_n \mid n \in \mathbb{N}\} \subseteq \mathbb{R}$$

is so.

Proposition 2.1.1. *A convergent sequence is bounded.*

Proof. Suppose $x_n \longrightarrow L$. Then there exists some $N \in \mathbb{N}$ such that $n > N$ implies $|x_n - L| < \epsilon$; pick $\epsilon = 1$.

Now define $M = \max(\{|x_0 - L|, |x_1 - L|, \ldots, |x_n - L|, 1\})$. M is therefore the maximum deviation of x_n from L; that is, for all $n \in \mathbb{N}$, we have $|x_n - L| \leq M$. This implies $-M \leq x_n - L \leq M$, or $L - M \leq x_n \leq L + M$, which bounds (x_n). □

Proposition 2.1.2. *If $x_n \longrightarrow L$, then $|x_n| \longrightarrow |L|$.*

Proof. Let $\epsilon > 0$. Then we apply Definition 2.1.2 to get N such that $n > N \implies |x_n - L| < \epsilon$. By a previous result, we write

$$||x_n| - |L|| \leq |x_n - L| < \epsilon,$$

and the result follows. □

Proposition 2.1.3. *If $x_n \longrightarrow L$ and $k \in \mathbb{R}$, then $kx_n \longrightarrow kL$.*

Proof. If $k = 0$, we are done. Otherwise, apply Definition 2.1.2 on (x_n) and $\epsilon/|k|$ to obtain N such that $n > N \implies |k \cdot x_n - k \cdot L| = |k| \cdot |x_n - L| < |k| \cdot \epsilon/|k| = \epsilon$. The result follows from this N. □

Proposition 2.1.4. *If $x_n \longrightarrow L$ and $y_n \longrightarrow M$, then $(x_n + y_n) \longrightarrow L + M$.*

Proof. Let $\epsilon > 0$. For (x_n), pick N such that $n > N$ implies $|x_n - L| < \epsilon/2$ and $|y_n - M| < \epsilon/2$. Then

$$|(x_n + y_n) - (L + M)| = |x_n - L + y_n - M| \leq |x_n - L| + |y_n - M| = \epsilon/2 + \epsilon/2 = \epsilon.$$

□

Proposition 2.1.5. *If $x_n \longrightarrow L$ and $y_n \longrightarrow M$, then $x_n y_m \longrightarrow LM$.*

Proof. Since (x_n) is convergent, it is bounded by some $k > 0$ so that $|x_n| < k$ for all $n \in \mathbb{N}$. Now choose N such that $|x_n - L| < \frac{\epsilon}{2 \cdot |M|}$ and $|y_n - M| < \frac{\epsilon}{2k}$. This lets us write

$$\begin{aligned} |x_n y_n - LM| &= |x_n y_n + x_n M - x_n M - LM| \\ &\leq |x_n| \cdot |(y_n - M)| + |M| \cdot |(x_n - L)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

□

Proposition 2.1.6. *If $x_n \longrightarrow L$, and $L > 0$, there exists $N \in \mathbb{N}$ and $r > 0$ such that for all $n > N$, we have $x_n > r$. That is, there exists a positive eventual lower bound for (x_n).*

Proof. Take $\epsilon = \frac{L}{2}$. Then there exists $N \in \mathbb{N}$ such that $n > N \implies |x_n - L| < \frac{L}{2}$, implying that $-\frac{L}{2} < x_n - L < \frac{L}{2}$. So $\frac{L}{2} < x_n < \frac{3L}{2}$. Pick $r = \frac{L}{2}$. □

Proposition 2.1.7. *If $x_n \longrightarrow L$ and $L \neq 0$, then $\frac{1}{x_n} \longrightarrow \frac{1}{L}$.*

Proof. Let $\epsilon > 0$. By Proposition 2.1.2, $|x_n| \longrightarrow |L|$, and by Proposition 2.1.6, there exists $r > 0$ and $N_1 \in \mathbb{N}$ such that $|x_n| > r$ for $n > N_1$.

Choose $N > N_1$ such that for all $n > N$, we have $|x_n - L| < \epsilon \cdot |L| \cdot r$. Then

$$\left|\frac{1}{x_n} - \frac{1}{L}\right| = \left|\frac{x_n - L}{x_n \cdot L}\right| < \frac{\epsilon \cdot |L| \cdot r}{|L| \cdot r} = \epsilon.$$

□

Proposition 2.1.8 (Squeeze theorem)**.** *Let $(x_n), (y_n), (z_n)$ be sequences. Assume:*

1. *$x_n \longrightarrow L$ and $z_n \longrightarrow L$*
2. *There exists N such that for all $n > N$, we have $x_n \leq y_n \leq z_n$.*

Then $y_n \longrightarrow L$.

Proof. Let $\epsilon > 0$. Choose $N' > N$ such that $|x_n - L| < \epsilon$ and $|z_n - L| < \epsilon$ for $n > N'$. Then we can write

$$-\epsilon < x_n - L \leq y_n - L \leq z_n - L < \epsilon,$$

which gives $|y_n - L| < \epsilon$. □

The squeeze theorem gives a general strategy for determining if an abstract sequence is convergent.

Definition 2.1.4. A sequence (x_n) is **non-decreasing** if $x_n \geq x_m$ when $n > m$, **non-increasing** if $x_n \leq x_m$ when $n > m$, and **monotonic** if either.

Proposition 2.1.9. *A bounded, monotonic sequence converges to some $L \in \mathbb{R}$.*

Proof. This result should make sense. Think about a monotonically increasing sequence that never goes past its "maximum threshold" L. Since the sequence never moves away from L, it should converge to it.

Concretely, we write the following. Replacing (x_n) with $(-x_n)$ if necessary, we can assume (x_n) is non-decreasing. Now consider $S = \{x_n \mid n \in \mathbb{N}\}$. We claim $x_n \longrightarrow L$, where $L = \sup S$.

Take $\epsilon > 0$. Since L is an upper bound of S, $L - \epsilon$ is not an upper bound. Hence there exists a witness $x_n \in S$ such that $x_n > L - \epsilon$. This gives $L - \epsilon < x_n \leq L < L + \epsilon$, so $|x_n - L| < \epsilon$. □

Now we answer the question: *what should $x_n \longrightarrow \infty$ mean*?

Definition 2.1.5. Let (x_n) be a sequence. We say $x_n \longrightarrow +\infty$ if

$$\forall M \in \mathbb{R} \,:\, \exists N \in \mathbb{N} \,:\, \forall n > N \,:\, x_n > M,$$

and $x_n \longrightarrow -\infty$ if

$$\forall M \in \mathbb{R} \,:\, \exists N \in \mathbb{N} \,:\, \forall n > N \,:\, x_n < M.$$

Proposition 2.1.10. *Let $x_n \longrightarrow +\infty$ and y_n have a positive eventual lower bound. Then $x_n \cdot y_n \longrightarrow +\infty$.*

Proof. Take $M \in \mathbb{R}$. Then there exists $N \in \mathbb{N}$ and $r > 0$ such that $n > N \implies x_n > M/r$ and $y_n > r$. It follows immediately that $x_n y_n > M$. $\square$

Example 2.1.2. Consider the series $x_n = \frac{n^3-7}{n+1}$. We have

$$\frac{n^3 - 7}{n + 1} = \frac{n^3(1 - \frac{7}{n^3})}{n(1 + \frac{1}{n})} = n^2 \cdot \frac{1 - \frac{7}{n^3}}{1 + \frac{1}{n}}.$$

We have showed already that $\frac{1}{n} \longrightarrow 0$. It follows from the above properties that $-\frac{7}{n^3} \longrightarrow 0$.

By addition, $1 - \frac{7}{n^3} \longrightarrow 1$ and $1 + \frac{1}{n} \longrightarrow 1$. By division, their quotient also approaches 1. And since $1 > 0$, this fraction has an eventual positive lower bound. Therefore, $n^2 \cdot \frac{1-\frac{7}{n^3}}{1+\frac{1}{n}}$ diverges because $n^2 \longrightarrow +\infty$.

Proposition 2.1.11. *Any monotonic sequence converges in $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$. More precisely, for a monotonic sequence (x_n), the following cases are possible:*

1. *(x_n) is bounded. Then it converges to $L \in \mathbb{R}$.*
2. *(x_n) is unbounded above. Then $x_n \longrightarrow \infty$.*
3. *(x_n) is unbounded below. Then $x_n \longrightarrow -\infty$.*

Proof. For 1), see Proposition 2.1.9.

Assume (x_n) is unbounded. Replacing (x_n) with $(-x_n)$ if necessary, we may assume (x_n) is monotonically increasing. Thus, it is unbounded above.

Now take $M \in \mathbb{R}$. Since M is not an upper bound, there exists $N \in \mathbb{N}$ such that $x_N > M$. And by monotonicity of (x_n), we know $n > N \implies x_n \geq x_N > M$. Hence $x_n \longrightarrow \infty$. $\square$

2.2 Subsequences

Definition 2.2.1. Let (x_n) be a sequence. A **subsequence** of (x_n) is a sequence (y_n) for which there exists a strictly increasing map $f : \mathbb{N} \longrightarrow \mathbb{N}$ such that $y_n = x_{f(n)}$.

Example 2.2.1. $2, 4, 6, \ldots$ is a subsequence of $1, 2, 3, \ldots$. A subsequence doesn't need to have a pattern, so $1, 2, 5, 7, \ldots$ is another example.

Proposition 2.2.1. *If $x_n \longrightarrow L$ and (y_n) is a subsequence of (x_n), then $y_n \longrightarrow L$.*

Proof. Let $\epsilon > 0$. Choose N such that $|x_n - L| < \epsilon$ for $n > N$. Since f is strictly increasing, $f(n) \geq n$, so $|y_n - L| = |x_{f(n)} - L| < \epsilon$. $\square$

Corollary 2.2.1. *A sequence (x_n) converges to L iff every subsequence of (x_n) converges to L.*

Proposition 2.2.2. *Every sequence has a monotonic subsequence.*

Proof. Let (x_n) be a sequence. Call x_{n_0} **dominant** if $x_{n_0} \geq x_m$ for all $m > n_0$. Then we have two cases:

<u>Case 1:</u> There are infinitely many dominant terms. Then take the terms and form a monotonically decreasing subsequence.

<u>Case 2:</u> There are finitely many dominant terms. Let N be such that for $n \geq N$, x_n is not dominant. Define (x_{n_k}) recursively as follows:

- $x_{n_0} = x_N$
- If x_{n_k} is selected, since it is not dominant, we have $x_{n'_k} > x_{n_k}$ for some $n'_k > n_k$. Define $x_{n_{k+1}} = x_{n'_k}$.

(x_{n_k}) is a monotonically increasing subsequence. □

Theorem 2.2.1 (Bolzano-Weierstrass). *Every bounded sequence has a subsequence that converges in* $\mathbb{R}$. *If we allow* $\pm\infty$ *as limits, then every sequence has a convergent subsequence.*

Proof. Follows from Proposition 2.1.9 and Proposition 2.2.2. □

Definition 2.2.2. The **extended real number line** is given by $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$.

Moreover, declare $-\infty < a < +\infty$ for all $a \in \mathbb{R}$. We write:

- $a + (\pm\infty) = \pm\infty$
- $\frac{a}{\pm\infty} = 0$ if $a \neq 0$
- $a \cdot (\pm\infty) = \pm\infty$ if $a > 0$
- $a \cdot (\pm\infty) = \mp\infty$ if $a < 0$.

We don't define $(+\infty) + (-\infty)$, $\frac{0}{+\infty}$, or $\frac{0}{-\infty}$. $\overline{\mathbb{R}}$ is not a field.

Also, for $S \subset \mathbb{R}$, if $S = \emptyset$, then $\sup S = -\infty$. If S is not bounded above, $\sup S = +\infty$.

Definition 2.2.3. Let (x_n) be a sequence. An element $L \in \overline{\mathbb{R}}$ is called a **subsequential limit** if there is a subsequence of (x_n) converging to L.

Proposition 2.2.3. *Let* (x_n) *be a sequence and denote its set of subsequential limits by* $S \subset \overline{R}$. *Then* $S \neq \emptyset$ *and* $\sup S \in S$.

Proof. By Bolzano-Weierstrass, $S \neq \emptyset$. It remains to show that S has a maximum.

<u>Case 1:</u> $\sup S \in \mathbb{R}$. Write $L = \sup S$. For any $k \in \mathbb{N}$, $L - \frac{1}{2k}$ is not an upper bound for S, so there exists $L_k \in S$ such that $L - \frac{1}{2k} < L_k \leq L$. And since L_k is a subsequential limit, there exists a subsequence of (x_n) converging to it; hence there is some (x_{n_k}) such that $|x_{n_k} - L_k| < \frac{1}{2k}$ eventually.

Now we write

$$|L - x_{n_k}| = |L - L_k + L_k - x_{n_k}| \leq |L - L_k| + |L_k - x_{n_k}| < \frac{1}{k}.$$

Thus (x_{n_k}) is a subsequence of (x_n) converging to L, so $L \in S$.

Case 2: $\sup S = +\infty$. For $M \in \mathbb{N}$, M is not an upper bound for S, so there exists $L_m \in S$ where $L_m > M$. Since L_m is a subsequential limit, there exists (x_{n_m}) such that $|x_{n_m} - L_m| < L_m - M$ eventually, so $x_{n_m} > M$.

Thus (x_{n_m}) is a subsequence going to $+\infty$, so $+\infty \in S$.

Case 3: $\sup S = -\infty$. Then $S = \{-\infty\}$. □

Definition 2.2.4. Let (x_n) be a sequence. Then

$$\limsup x_n := \lim_{n\to\infty} \sup\{x_m \mid m \geq n\}$$

$$\liminf x_n := \lim_{n\to\infty} \inf\{x_m \mid m \geq n\}.$$

That is, lim sup is the limit of the suprema of the tails of (x_n), and analogously for lim inf.

Remark. The sequence $v_n := \sup\{x_m \mid m \geq n\}$ is non-increasing, so $\lim_{n\to\infty} v_n$ exists, and analogously for $w_n := \inf\{x_m \mid m \geq n\}$.

Example 2.2.2.

1. $x_n = 1, 2, 3, \ldots$

   $v_n = +\infty$, so $\limsup x_n = \lim v_n = +\infty$

   $w_n = x_n$, so $\liminf x_n = \lim w_n = +\infty$

2. $x_n = \frac{1}{n}$

   $v_n = x_n$, so $\limsup x_n = 0$

   $w_n = 0$, so $\liminf x_n = 0$

3. $x_n = (-1)^n$

   $v_n = 1$, so $\limsup x_n = 1$

   $w_n = -1$, so $\liminf x_n = -1$.

It's not a coincidence that the first two sequences satisfy $\limsup x_n = \liminf x_n$ and are convergent:

Theorem 2.2.2. *Let (x_n) be a sequence. Denote its set of subsequential limits by $S \subset \overline{\mathbb{R}}$. Then the following hold:*

1. $\limsup x_n = \sup S \in S$

   $\liminf x_n = \inf S \in S$

2. *There exists a monotonic subsequence converging to* $\limsup x_n$ *and one to* $\liminf x_n$.

3. *TFAE:*

   (a) *(x_n) converges*

   (b) *S has one element*

(c) $\limsup x_n = \liminf x_n$

Moreover, in this case, $\limsup x_n = \liminf x_n = \lim x_n$ *is the unique element of* S*.*

Proof.

1. By the previous result, it suffices to show that $\limsup x_n = \sup S$.

   We claim that $\limsup x_n \geq \sup S$. Note that there exists a subsequence (x_{n_k}) converging to $\sup S$. For any n, $v_n = \sup\{x_m \mid m \geq n\} \geq x_{n_k}$ for sufficiently large k. Thus $v_n \geq \lim_{k\to\infty}(x_{n_k}) = \sup S$ and $\limsup x_n \geq \sup S$.

   Now we will show $\limsup x_n \leq \sup S$. Toward a contradiction, suppose the contrary and denote $x = \limsup x_n$ and $y = \sup S$. We will construct a subsequence (x_{n_k}) such that $x_{n_k} \geq \frac{x+y}{2}$.

   After doing so, we may assume (x_{n_k}) is monotonic (otherwise, take some monotonic subsequence of it), hence convergent to some $z \geq \frac{x+y}{2} > y$, contradicting the definition of y. Now to construct the actual sequence (x_{n_k}), iterate the following procedure:

   Write $\epsilon = \frac{x-y}{4}$. Since $v_n \longrightarrow x$, there exists v_{n_0} such that $|v_{n_0} - x| < \epsilon$. And because $v_{n_0} = \sup\{x_m \mid m \geq n_0\}$, we can find x_{m_0} in the sequence such that $v_{n_0} \geq x_{m_0} \geq v_{n_0} - \epsilon$; that is, $|x_{m_0} - v_{n_0}| < \epsilon$. Thus, $|x - x_{m_0}| < 2\epsilon = \frac{x-y}{2}$, which yields $x_{m_0} > \frac{x+y}{2}$, as desired.

2. By the previous part, there exists a subsequence converging to $\limsup x_n$. We can extract from it a monotonic subsequence using Proposition 2.2.2.

3. $(a) \implies (b)$ follows from Proposition 2.2.1. $(b) \implies (c)$ follows from part 1. $(c) \implies (a)$ follows from the squeeze theorem on $w_n \leq x_n \leq v_n$.

□

So we know S can have one element, in which case (x_n) converges. But what else can S look like?

Definition 2.2.5. A set S is called:

- **finite** if there is a bijection $S \longrightarrow \{m \in \mathbb{N} \mid m \leq n\}$ for some $n \in \mathbb{N}$
- **countable** if there is a bijection $S \longrightarrow \mathbb{N}$
- **uncountable** otherwise.

Proposition 2.2.4. $\mathbb{Q}$ *is countable.*

Proof. Exercise! The TeX is too much work. □

Proposition 2.2.5. $\mathbb{R}$ *is uncountable.*

Proof. Diagonal argument. Note that we assume the existence of decimal expansions, which we will prove later. □

It turns out that S can be any of these:

- Finite, e.g., $1, 2, 3, 1, 2, 3, \ldots \implies S = \{1, 2, 3\}$

- Countable, e.g., $1, 1, 2, 1, 2, 3, 1, 2, 3, 4, \ldots \implies S = \mathbb{N}$.
- Uncountable: pick a bijection $\lambda : \mathbb{N} \longrightarrow \mathbb{Q}$. By density of $\mathbb{Q}$, any real number is approached by some sequence of rational numbers. So $S = \mathbb{R}$.
- Any non-empty, closed subset of $\mathbb{R}$.

2.3 Cauchy Sequences

Definition 2.3.1. A sequence (x_n) is **Cauchy** if

$$\forall \epsilon > 0 \; : \; \exists N \in \mathbb{N} \; : \; \forall n, m > N \; : \; |x_n - x_m| < \epsilon.$$

Recall that $x_n \longrightarrow L$ if

$$\forall \epsilon > 0 \; : \; \exists N \in \mathbb{N} \; : \; \forall n > N \; : \; |x_n - L| < \epsilon.$$

This requires an explicit limit. Cauchy-ness, on the other hand, is an intrinsic property of (x_n). This is especially useful due to the following result:

Theorem 2.3.1. *A sequence of real numbers converges in $\mathbb{R}$ iff it is Cauchy.*

Proof. ($\Longrightarrow$): Let L be the limit and choose N such that $n > N \implies |x_n - L| < \epsilon/2$. Then for any $n, m > N$, we have

$$|x_n - x_m| = |x_n - L + L - x_m| \le |x_n - L| + |x_m - L| < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

($\Longleftarrow$): Let $\epsilon > 0$ and choose N such that $n, m > N \implies |x_n - x_m| < \epsilon$. WLOG, suppose $x_n - x_m < \epsilon$ for all $n, m > N$. This gives us $\sup\{x_n \mid n \ge k\} \le \epsilon + x_m$ because $\epsilon + x_m$ is an upper bound for $\{x_n \mid n \ge k\}$, so $\limsup(x_n) \le \epsilon + x_m$.

Since m is arbitrary, we use the same argument to obtain $\limsup(x_n) - \epsilon \le \liminf(x_n)$. So $\limsup(x_n) - \liminf(x_n) \le \epsilon$. And since ϵ can be arbitrarily small, we conclude $\limsup(x_n) = \liminf(x_n)$, so (x_n) converges by Theorem 2.2.2. $\square$

Therefore, checking whether a sequence is Cauchy lets us check for convergence without a candidate limit.

It turns out that convergent implies Cauchy for any ordered field—we didn't use LUBP to show the forward direction of Theorem 2.3.1.

The backward direction, though, is actually *equivalent* to LUBP; the result doesn't hold without it. Why? Take any sequence in $\mathbb{Q}$ that converges to an irrational number. Then it is Cauchy (by density) but not convergent.

Indeed, our proof of the backward direction assumes the existence of $\sup\{x_n \mid n \ge k\}$.

Proposition 2.3.1.

1. $\lim_{n\to\infty} a^n = 0$ *if* $|a| < 1$
2. $\lim_{n\to\infty} n^{1/n} = 1$
3. $\lim_{n\to\infty} a^{1/n} = 1$ *if* $a > 0$

Proof.

1. If $|a| < 1$, then $|\frac{1}{a}| > 1$, so $|\frac{1}{a}| = 1 + b$ for $b > 0$. By the binomial expansion, $(1 + b)^n \geq n \cdot b$. Thus $0 \leq |a|^n \leq \frac{1}{nb} \implies |a|^n \longrightarrow 0$ by the squeeze theorem.
2. It suffices to show $n^{1/n} - 1 \longrightarrow 0$. Denote the LHS by x_n. Then, using the binomial expansion,
$$n^{1/n} = 1 + x_n \implies n = (1 + x_n)^n \geq \frac{n(n-1)}{2} \cdot x_n^2 \implies 1 \geq \frac{n-1}{2} \cdot x_n^2.$$
So $0 \leq x_n^2 \leq \frac{2}{n-1} \implies x_n^2 \longrightarrow 0 \implies x_n^2 \longrightarrow 0$.
3. If $a > 1$, then $1 < a^{1/n} < n^{1/n}$ when $n > a$ (which it is, since $n \longrightarrow +\infty$). Then the previous part implies $a^{1/n} \longrightarrow 1$.

   If $a \leq 1$, then $\frac{1}{a} \geq 1$, so $(\frac{1}{a})^{1/n} \longrightarrow 1 \implies \frac{1}{a^{1/n}} \longrightarrow 1 \implies a^{1/n} \longrightarrow 1$.

□

Proposition 2.3.2. *Let (a_n) be a sequence of nonzero real numbers. Then*
$$\liminf \left|\frac{a_{n+1}}{a_n}\right| \leq \liminf |a_n|^{1/n} \leq \limsup |a_n|^{1/n} \leq \limsup \left|\frac{a_{n+1}}{a_n}\right|.$$

Proof. We will prove the rightmost inequality. Firstly, denote
$$\alpha = \limsup |a_n|^{1/n}$$
$$\beta = \limsup \left|\frac{a_{n+1}}{a_n}\right|.$$

It suffices to show $\alpha \leq \beta_1$ for all $\beta_1 > \beta$. Since $\beta_1 > \limsup |\frac{a_{n+1}}{a_n}|$, there exists N such that $n > N \implies |\frac{a_{n+1}}{a_n}| < \beta_1$. Now take some $n > N$ and note that
$$|a_n| = \frac{|a_n|}{|a_{n-1}|} \cdot \frac{|a_{n-1}|}{|a_{n-2}|} \cdot \ldots \cdot \frac{|a_{N+1}|}{|a_N|} \cdot |a_N|.$$

Each term is less than β_1 by construction, so $|a_n| < \beta_1^{n-N} \cdot |a_n| = \beta_1^n \cdot \frac{|a_N|}{\beta_1^N}$. This gives
$$|a_n|^{1/n} < \beta_1 \cdot \left(\frac{|a_N|}{\beta_1^N}\right)^{1/n}$$
$$\implies \limsup |a_n|^{1/n} \leq \beta_1 \cdot \limsup \left(\frac{|a_N|}{\beta_1^N}\right)^{1/n} \leq \beta_1.$$

□

2.4 Series

Definition 2.4.1. Let (a_n) be a sequence.

1. The **n-th partial sum** of (a_n) is defined recursively by $s_0 = a_0$ and $s_{n+1} = s_n + a_{n+1}$.
2. The **infinite series** associated with (a_n) is the sequence (s_n). We write $\sum_{k=0}^{n} a_k = s_n$ and $\sum a_k = (s_n)$ for the whole series.

   If (s_n) converges, we write $\sum_{k=0}^{\infty} a_k$ for its limit.

Proposition 2.4.1 (Cauchy Test). *A series $\sum a_n$ converges iff for all $\epsilon > 0$, there exists N such that $n, m > N$ implies*

$$|s_n - s_m| = \left|\sum_{k=n}^{m} s_n\right| < \epsilon.$$

Proof. Apply Theorem 2.3.1 to (s_n). □

Proposition 2.4.2 (Silly Test). *If $\sum a_n$ converges, then $a_n \longrightarrow 0$.*

Proof. Apply the Cauchy Test with $n = m$. □

Example 2.4.1. Consider the following series:

$$\sum \frac{n^2 + 1}{300n^2 + 1000000n + 35 \text{ gazillion}}.$$

It does not converge since its underlying sequence approaches 1/300, which fails the contrapositive of the Silly Test.

Theorem 2.4.1 (Geometric Series). *Let $a \in \mathbb{R}$. The series $\sum_{n=0}^{\infty} a^n$*

1. *doesn't converge if $|a| \geq 1$*
2. *converges to $\frac{1}{1-a}$ if $|a| < 1$.*

Proof. If $|a| \geq 1$, then a^n doesn't converge to 0. We win by Silly Test.

Otherwise, we compute $s_n = 1 + a + a^2 + \cdots + a^n$ as follows:

$$(1 - a)(1 + a + a^2 + \cdots + a^n) = 1 + a + a^2 + \cdots + a^n - (a + a^2 + a^3 + \cdots + a^{n+1}) = 1 - a^{n+1}.$$

So $s_n = \frac{1-a^{n+1}}{1-a}$. Since $a < 1$, this converges to $\frac{1}{1-a}$, as desired. □

Definition 2.4.2. A series $\sum a_n$ **converges absolutely** if $\sum |a_n|$ converges.

Proposition 2.4.3. *If $\sum a_n$ converges absolutely, then it converges.*

Proof. By the Triangle Inequality,

$$\left|\sum_{k=n}^{m} a_k\right| \leq \sum_{k=n}^{m} |a_k|$$

The result follows immediately from Cauchy. $\square$

Proposition 2.4.4. *Let (a_n) be a sequence of non-negative reals. Then $\sum a_n$ converges iff its sequence of partial sums is bounded.*

Proof. Observe that the sequence of partial sums is monotonic, then apply Proposition 2.1.11. $\square$

Theorem 2.4.2 (Comparison Test)**.** *Let (a_n) and (b_n) be sequences of non-negative reals. Assume $a_n \leq b_n$ eventually. Then (a_n) converges if (b_n) does.*

Proof. Let N be such that $a_n \leq b_n$ for $n > N$. Then, for $m > n > N$, we have

$$\sum_{i=n}^{m} a_i \leq \sum_{i=n}^{m} b_i.$$

Apply Cauchy. $\square$

Example 2.4.2. Famous example: the harmonic series diverges by the Comparison Test.

Proposition 2.4.5 (Limit Comparison Test)**.** *Let (a_n) and (b_n) be sequences of non-negative reals. Assume* $\limsup \frac{a_n}{b_n} \neq +\infty$. *Then $\sum a_n$ converges if $\sum b_n$ does.*

Proof. Let $c = \limsup \frac{a_n}{b_n}$. Let $\epsilon > 0$. There is N such that, for $n > N$, we have $\frac{a_n}{b_n} < c + \epsilon$. Then $a_n < (c + \epsilon)b_n$, and apply Comparison Test. $\square$

Corollary 2.4.1. *Let (a_n) and (b_n) be sequences of non-negative reals. Assume* $\lim \frac{a_n}{b_n}$ *exists and is a positive real number. Then $\sum a_n$ converges iff $\sum b_n$ does.*

Proof. Consider $\lim \frac{a_n}{b_n}$ and $\lim \frac{b_n}{a_n}$. $\square$

Theorem 2.4.3 (Root Test)**.** *Let (a_n) be any sequence. Denote* $\alpha = \limsup |a_n|^{1/n}$.

1. *If $\alpha < 1$, then $\sum a_n$ converges absolutely.*
2. *If $\alpha > 1$, then $\sum a_n$ diverges.*
3. *If $\alpha = 1$, the test is inconclusive.*

Proof. Assume $\alpha < 1$. Then take $\alpha_1 \in \mathbb{R}$ where $\alpha < \alpha_1 < 1$. It is easy to show that there exists $N \in \mathbb{N}$ such that for $n > N$, we have $|a_n|^{1/n} < \alpha_1$. Thus $|a_n| < \alpha_1^n$. Apply the Comparison Test and convergence of geometric series.

Now assume $\alpha > 1$. By a previous result, there is a subsequence $|a_{n_k}|^{1/n_k}$ converging to α. So there is K such that for $k > K$, $|a_{n_k}| > 1$, so $|a_{n_k}| > 1$. Thus $a_n \not\to 0$. Apply Silly Test. $\square$

Theorem 2.4.4 (Ratio Test)**.** *Let (a_n) be a sequence.*

1. *If* $\limsup |\frac{a_{n+1}}{a_n}| < 1$, *then* $\sum a_n$ *converges absolutely.*
2. *If* $\limsup |\frac{a_{n+1}}{a_n}| > 1$, *then* $\sum a_n$ *diverges.*
3. *Otherwise, the test is inconclusive.*

Proof. By a previous result, we have

$$\liminf \left|\frac{a_{n+1}}{a_n}\right| \leq \limsup |a_n|^{1/n} \leq \limsup \left|\frac{a_{n+1}}{a_n}\right|.$$

Apply the Root Test. □

Lemma 2.4.1. Let $a_1 \geq a_2 \geq \cdots \geq a_n \geq \cdots$ be a non-increasing sequence of non-negative numbers. Then $\sum a_n$ converges iff $\sum 2^n \cdot a_{2^n}$ does.

Proof. Consider

$$s_n = a_1 + \cdots + a_n$$
$$t_n = a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_k.$$

We claim that if (t_k) is bounded, then so is (s_n). To show this, let $n < 2^k$. Then

$$\begin{aligned} s_n &= a_1 + \cdots + a_n \\ &\leq a_1 + (a_2 + a_3) + (a_4 + \cdots + a_7) + \cdots + (a_{2^k} + \cdots + a_{2^{k+1}-1}) \\ &\leq a_1 + 2a_2 + 4a_4 + \cdots + 2^k a_{2^k} \\ &= t_k. \end{aligned}$$

We also claim the converse is true:

$$\begin{aligned} s_n &= a_1 + \cdots + a_n \\ &\geq a_1 + a_2 + (a_3 + a_4) + (a_5 + \cdots + a_8) + (a_{2^{k-1}-1} + \cdots + a^{2k}) \\ &\geq \frac{1}{2}a_1 + a_2 + 2a_4 + 4a_8 + \cdots + 2^{k-1} a_{2^k} \\ &= \frac{1}{2} t_k. \end{aligned}$$

Now apply Proposition 2.4.4. □

Theorem 2.4.5 (p-series Test)**.**

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges iff $p > 1$.

Proof. If $p \leq 0$, then $\frac{1}{n^p} \not\to 0$, and we are done by Silly Test. So assume $p > 0$. Apply the previous lemma and consider

$$\sum 2^n \frac{1}{(2^n)^p} = \sum 2^n \cdot 2^{-np} = \sum (2^{(1-p)})^n.$$

This is the geometric series and converges iff $2^{1-p} < 1$; that is, if $1 - p < 0$, or $p > 1$. □

Example 2.4.3. Here is an illustration of why the root and ratio tests can be inconclusive:

$\sum \frac{1}{n}$	diverges	$\lim \left(\frac{1}{n}\right)^{1/n} = 1$	$\lim \frac{1/(n+1)}{1/n} = 1$
$\sum \frac{1}{n^2}$	converges	$\lim \left(\frac{1}{n^2}\right)^{1/n} = 1$	$\lim \frac{1/(n+1)^2}{1/n^2} = 1$

It is also natural to ask about how we multiply series.

Proposition 2.4.6. *Let $\sum a_n$ and $\sum b_n$ be convergent series, at least one absolutely. Define*

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{\substack{i,j \\ i+j=n}} a_i b_j.$$

Then $\sum c_n$ converges and $\sum c_n = (\sum a_n)(\sum b_n)$.

Proof. WLOG, suppose $\sum a_n$ converges absolutely. Set $A_n = \sum_{i=0}^n a_i$, $B_n = \sum_{i=0}^n b_i$, $C_n = \sum_{i=0}^n c_i$ and write $B = \lim B_n$ and $\beta_n = B - B_n$. Then

$$\begin{aligned} C_n &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \cdots + (a_0 b_n + a_1 b_{n-1} + \cdots + a_n b_0) \\ &= a_0 B_n + a_1 B_{n-1} + \cdots a_n B_0 \\ &= a_0(B + \beta_n) + a_1(B + \beta_{n-1}) + \cdots + a_n(B + \beta_0) \\ &= A_n \cdot B + a_0 \beta_n + \cdots a_n \beta_0. \end{aligned}$$

Claim. $\gamma_n := a_0 \beta_n + \cdots a_n \beta_0 \longrightarrow 0.$

Admitting the claim, we see that $C_n = A_n \cdot B + \gamma \longrightarrow A \cdot B$, as desired.

To prove it, write $\alpha = \sum_{n=0}^{\infty} |a_n|$ and choose $\epsilon > 0$. Since $\beta_n \longrightarrow 0$, there is N such that $|\beta_n| < \epsilon$ for $n > N$. Then we partition γ_n at N:

$$\begin{aligned} |\gamma_n| &= |\beta_0 a_n + \cdots + \beta_n a_0| \\ &\leq |\beta_0 a_n + \cdots + \beta_N a_{n-N}| + |\beta_{N+1} a_{n-N-1} + \cdots + \beta_n a_0|. \end{aligned}$$

Note that the rightmost term is bounded above by $\epsilon \cdot \alpha$. Since $a_n \longrightarrow 0$, there is M such that $|a_n| < \epsilon/(|\beta_0| + \cdots + |\beta_n|$ for $n > M$. Thus, for $n > N + M$, we have

$$|\gamma_n| \leq |\beta_0||a_n| + \cdots + |\beta_N||a_{n-N}| + \epsilon \cdot \alpha \leq \epsilon(1 + \alpha).$$

□

Definition 2.4.3. $\sum a_n$ **converges conditionally** if it converges, but not absolutely.

Theorem 2.4.6 (Alternating Series Test). *Let (a_n) be a sequence such that:*

- $a_n \geq a_{n+1} \geq \cdots \geq 0$

- $a_n \longrightarrow 0$.

Then $\sum(-1)^n a_n$ *converges.*

Proof. Note that

$$s_{2k} = s_{2k-2} + (a_{2k} - a_{2k-1}) \leq s_{2k-2}$$
$$s_{2k-1} = s_{2k-3} + (a_{2k-2} - a_{2k-1}) \geq s_{2k-3}$$

and that $s_{2k} = s_{2k-1} + a_{2k} \geq s_{2k-1}$. So s_{2k} is non-increasing and bounded below by s_1. Similarly, s_{2k-1} is non-decreasing and bounded above by it, so both converge (to the same limit L, since $s_{2k} - s_{2k-1} = a_{2k} \longrightarrow 0$).

Let $\epsilon > 0$, and choose K such that for $k > K$, we have $|s_{2k} - L| < \epsilon$. Then, for $n > 2k$, $|s_n - L| < \epsilon$. □

Example 2.4.4. $\sum \frac{(-1)^n}{n}$ converges conditionally.

In fact, a family of conditionally convergent series follows from the p-series.

Definition 2.4.4. A **rearrangement** of $\sum a_n$ is a series of the form $\sum a_{f(n)}$ for some bijection $f : \mathbb{N} \longrightarrow \mathbb{N}$.

Theorem 2.4.7. *If* $\sum a_n$ *converges absolutely, so does any rearrangement, to the same value.*

Proof. Let L be the limit and take $\epsilon > 0$. Choose N such that for $m \geq n \geq N$,

$$\left|\sum_{k=0}^{n} a_k - L\right| < \epsilon/2 \qquad \sum_{k=n}^{m} |a_k| < \epsilon/2.$$

Then for $M = \max(f^{-1}(0), \dots, f^{-1}(N))$ and $m > M$, so

$$\begin{aligned}\left|\sum_{k=0}^{n} a_k - L\right| &= \left|\sum_{k=0}^{m} a_{f(k)} - \sum_{k=0}^{N} a_k + \sum_{k=0}^{N} a_k - L\right| \\ &\leq \left|\sum_{k=0}^{m} a_{f(k)} - \sum_{k=0}^{N} a_k\right| + \left|\sum_{k=0}^{N} a_k - L\right| \\ &< \sum_{k=N+1}^{\infty} |a_k| + \epsilon/2 \leq \epsilon.\end{aligned}$$

The last step follows from the fact that $f(m) > N$, which implies that $\sum_{k=0}^{N} a_k$ is contained entirely within $\sum_{k=0}^{m} a_{f(k)}$. Hence the difference is the sum from $N+1$ to m, which is bound by sending m to infinity. □

Theorem 2.4.8 (Riemann). *Let* $\sum a_n$ *be conditionally convergent. Then there is a rearrangement which:*

- *converges to any desired element of* $\overline{\mathbb{R}}$
- *oscillates and doesn't converge.*

Proof sketch. Let (a_n^+) and (a_n^-) be the subsequences of (a_n) consisting of positive and negative terms, respectively. Then both approach 0 by the Silly Test.

Moreover, note that $\sum a_n^+ = +\infty$ and $\sum a_n^- = -\infty$. We argue as follows. Suppose toward a contradiction that both limits, denoted L^+, L^-, respectively, are in $\mathbb{R}$. Then $\sum |a_n| = L^+ - L^- \in \mathbb{R}$ and (a_n) converges absolutely, a contradiction. If either L^+ or L^- diverges, their sum $\sum a_n$ does too, another contradiction.

Now let $L \in \mathbb{R}$. Sum just enough of a_n^+ until the sum exceeds L, then just enough until the sum is less than L, and repeat. We can do this due to the claim shown above. Since $a_n^+ \longrightarrow 0$ and $a_n^- \longrightarrow 0$, the distance between the sum and L goes to zero with each step.

To obtain a non-convergent rearrangement, choose $L^+ > L^- \in \mathbb{R}$ and select terms so that the positive ones always exceed the negative ones, or vice versa. $\square$

3 Continuity and Differentiation

3.1 Continuity

Definition 3.1.1. Let $D \subseteq \mathbb{R}$ and take $f : D \longrightarrow \mathbb{R}$ and $a \in D$. Then f is **continuous at** a if

$$\forall \epsilon > 0 : \exists \delta > 0 : \forall x \in D : |x - a| < \delta \Longrightarrow |f(x) - f(a)| < \epsilon.$$

This definition says that we can arbitrarily bound variation in f around $f(a)$ by bounding its argument. So, for instance, constant functions are continuous because their variation is always bounded.

Proposition 3.1.1. *TFAE:*

1. *f is continuous at a*
2. *If $x_n \in D$ is a sequence converging to $a \in D$, then $f(x_n)$ converges to $f(a)$.*

Proof. ($\Longrightarrow$): Let $x_n \in D$ converge to a and take $\epsilon > 0$. Let δ be as in the definition of continuity and take N such that $n > N \Longrightarrow |x_n - a| < \delta$. Then $|f(x_n) - f(a)| < \epsilon$ by continuity of f at a.

$\neg$($\Longrightarrow$): Take $\epsilon > 0$ as above. For each n, apply with $\delta = \frac{1}{n}$ to get $x_n \in D$ with $|x_n - a| < \frac{1}{n}$ and $|f(x_n) - f(a)| \geq \epsilon$. Then $x_n \longrightarrow a$ and $f(x_n) \not\rightarrow f(a)$. □

Corollary 3.1.1. *If $f, g : D \longrightarrow \mathbb{R}$ are both continuous at $a \in D$, so are $f + g$, $f - g$, fg, and f/g when $g \neq 0$.*

Proof. Follows immediately from the limit theorems and the previous proposition. □

Corollary 3.1.2. *If $f : S \longrightarrow T$ is continuous at $a \in S$ and $g : T \longrightarrow \mathbb{R}$ is continuous at $b = f(a) \in T$, then $g \circ f$ is continuous at a.*

Proof. Let $x_n \in D$ and $x_n \longrightarrow a$. Then $f(x_n) \longrightarrow f(a)$, so $(g \circ f)(x_n) \longrightarrow (g \circ f)(a)$. □

Definition 3.1.2. $f : D \longrightarrow \mathbb{R}$ is **continuous** if it is continuous at a for all $a \in D$.

Theorem 3.1.1. *A continuous function $f : [a, b] \longrightarrow \mathbb{R}$ attains a minimum and maximum. In other words, there exist $x_-, x_+ \in [a, b]$ such that $f(x_-) \leq f(x) \leq f(x_+)$ for all $x \in [a, b]$.*

Proof. We will prove a series of claims—firstly, that f is bounded; i.e., $f([a, b]) \subseteq \mathbb{R}$ is bounded.

Assume not. Then for each $n \in \mathbb{N}$, there is $x_n \in [a, b]$ such that $|f(x_n)| > n$, implying that $\lim_{n\to\infty} |f(x_n)| = +\infty$. By Bolzano-Weierstrass, there is a subsequence (x_{n_k}) of (x_n) which converges. Let c be its limit. Then $|f(x_{n_k})| \longrightarrow |f(c)|$, contradiction.

So $M_+ := \sup f([a, b]) \in \mathbb{R}$ and $M_- := \inf f([a, b]) \in \mathbb{R}$. It remains to show that there exists $x_+ \in [a, b]$ with $f(x_+) = M_+$.

For $u \in \mathbb{N}\backslash\{0\}$, we know $M_+ - 1/u$ is not an upper bound for $f([a, b])$, so there is $x_n \in [a, b]$ with $M_+ \geq f(x_n) \geq M_+ - 1/u$. By Bolzano-Weierstrass, there is a subsequence (x_{n_k}) converging to some $x_+ \in [a, b]$. Then $f(x_+) = \lim_{k\to\infty} f(x_{n_k})$, and $M_+ \geq f(x_{n_k}) \geq M_+ - 1/u_k$ implies $f(x_+) = M_+$.

An analogous argument on $-f(x)$ gives x_-. $\square$

Note that this theorem does not hold if the domain is open! Take $f(x) = x$ on $(0, 1)$. So where in the proof did we assume that the domain is closed?

The answer is when we claimed that $x_+ \in [a, b]$. That's because x_+ is an adherent point of $\{x_n \mid n \in \mathbb{N}\} \subset [a, b]$ and we need to use that $[a, b]$ is closed.

Definition 3.1.3. Let $D \subset \mathbb{R}$ and $a \in \mathbb{R}$. Then a is called:

- an **adherent point** of D if $\forall \epsilon > 0 : \exists x \in D : |a - x| < \epsilon$.
- a **limit point** of D if $\forall \epsilon > 0 : \exists x \in D : 0 < |a - x| < \epsilon$.
- an **isolated point** of D if $\exists \epsilon > 0 : \forall x \in D : x \neq a \implies |a - x| > \epsilon$.
- an **interior point** of D if $\exists \epsilon > 0 : \forall x \in \mathbb{R} : |a - x| < \epsilon \implies x \in D$.

Example 3.1.1.

- 0 is a limit point and adherent point of $\{1/n \mid n \in \mathbb{N}\}$, but not isolated or interior.
- 1 is an adherent and isolated point of $\mathbb{N}$, but not limit or interior.
- 1/2 is a limit, adherent, and interior point of $(0, 1)$, but not isolated.

Definition 3.1.4. A subset $D \subset \mathbb{R}$ is called

- **open** if every $a \in D$ is interior to D.
- **closed** if every adherent $a \in R$ lies in D.

Proposition 3.1.2. *Let $D \subset \mathbb{R}$. Then D open $\iff$ $\mathbb{R}\backslash D$ closed.*

Proof. Fix $a \in D$.

$$\begin{aligned} &a \text{ interior to } D \\ \iff &\exists \epsilon > 0 : (a - \epsilon, a + \epsilon) \subset D \\ \iff &\exists \epsilon > 0 : (a - \epsilon, a + \epsilon) \cap (\mathbb{R}\backslash D) = \emptyset \\ \iff &a \text{ not adherent to } \mathbb{R}\backslash D \end{aligned}$$

$\square$

Definition 3.1.5. Take $D \subset \mathbb{R}$. Then $f : D \longrightarrow \mathbb{R}$ is **continuous** if

$$\forall y \in D : \forall \epsilon > 0 : \exists \delta > 0 : \forall x \in D : |x - y| < \delta \implies |f(x) - f(y)| < \epsilon.$$

Definition 3.1.6. Take $D \subset \mathbb{R}$. Then $f : D \longrightarrow \mathbb{R}$ is **uniformly continuous** if

$$\forall \epsilon > 0 : \exists \delta > 0 : \forall y \in D : \forall x \in D : |x - y| < \delta \implies |f(x) - f(y)| < \epsilon.$$

Proposition 3.1.3. *Let $f : [0, 1] \longrightarrow [0, 1]$ be continuous. Then f has a fixed point; i.e., there is $x_0 \in [0, 1]$ such that $f(x_0) = x_0$.*

Proof. If $f(0) = 0$ or $f(1) = 1$, then we are done. Otherwise, we have $f(0) > 0$ and $f(1) < 1$. Now define $g(x) = f(x) - x$. We are now looking for a point at which $g(x) = 0$, which we can do using IVT by noting that $g(0) > 0$ and $g(1) < 0$. The result follows. □

Remark. This is the one-dimensional analog of the fixed-point problem in topology: informally, no matter how you stir a cup of coffee, there will always be one molecule that has not moved.

We previously showed that $f(x) = 1/x$ is continuous but not uniformly because it "explodes" at $x = 0$. It turns out that some nice functions also do not qualify as uniformly continuous:

Claim. $f(x) = x^2$ is continuous but not uniformly.

Proof. Note that $|f(x) - f(y)| = |x^2 - y^2| = |x - y| \cdot |x + y|$ and suppose that $|x - y| < \delta$. Then $|x + y|$ can still explode if x, y can be arbitrarily large, as in the definition of uniform continuity, so f is not uniformly continuous.

On the other hand, if we fix y and then choose $|x - y| < \delta$, like in the definition of continuity, we can make $|x - y| \cdot |x + y| < \delta|x + y|$ arbitrarily small since $x + y \longrightarrow 2y$ as $x \longrightarrow y$. Hence f is continuous. □

Proposition 3.1.4. *Let $f : D \longrightarrow \mathbb{R}$ be uniformly continuous. If (x_n) is a Cauchy sequence in D, then $f(x_n)$ is also a Cauchy sequence.*

Proof. Let $\epsilon > 0$. Take $\delta > 0$ such that $|x - y| \Longrightarrow |f(x) - f(y)| < \epsilon$. Take N such that for $n, m > N$, $|x_n - x_m| < \delta$. Then $|f(x_n) - f(x_m)| < \epsilon$. □

Remark. This result does NOT imply the following:

$$f \text{ uniformly continuous} \Longleftrightarrow (x_n) \text{ Cauchy} \Longleftrightarrow (x_n) \text{ continuous} \Longleftrightarrow f \text{ continuous}$$

because the Cauchy-continuous biconditional only holds in $\mathbb{R}$.

Take, for instance, $f(x) = 1/x$ on $D = (0, 1)$ and $x_n = 1/n$. We see that $1/n$ is Cauchy, but $f(x_n) = n$, which is obviously neither Cauchy nor convergent.

Moreover, the opposite direction of this result also fails. Consider $f(x) = x^2$ on $D = \mathbb{R}$, which sends Cauchy sequences to Cauchy sequences (since it sends sequences convergent in $\mathbb{R}$ to sequences convergent in $\mathbb{R}$). But we showed previously that it is not uniformly continuous.

Definition 3.1.7. Let $D \subset \tilde{D} \subset \mathbb{R}$.

1. Given $\tilde{f} : \tilde{D} \longrightarrow \mathbb{R}$, the **restriction** of $\tilde{f}$ to D is the function $f : D \longrightarrow \mathbb{R}$ defined by $f(x) = \tilde{f}(x)$ for all $x \in D$.

   We write $f = \tilde{f}|_D$.

2. Given $f : D \longrightarrow \mathbb{R}$, an **extension** of f is a function $\tilde{f} : \tilde{D} \longrightarrow \mathbb{R}$ such that $\tilde{f}|_D = f$.

Note that we refer to "the" restriction of $\tilde{f}$ to D and "an" extension of f because the restriction of $\tilde{f}$ to a given domain is uniquely determined, while many functions can extend f.

Example 3.1.2. Let $D = (0,1)$ and $\tilde{D} = [0,1]$. Take $f(x) = x$ and $\tilde{f}_1(x) = x$. Consider also $\tilde{f}_2$ given by $\tilde{f}_2(x) = x$ for $0 < x < 1$ and $\tilde{f}_2(x) = 1/2$ for $x = 0, 1$. Both $\tilde{f}_1, \tilde{f}_2$ are perfectly valid extensions of f.

So extensions are not uniquely determined, but it seems like they are if we restrict ourselves to continuous extensions.

Definition 3.1.8. Let $D \subset \tilde{D} \subset \mathbb{R}$. We say D is **dense** in $\tilde{D}$ if every point in $\tilde{D}$ is an adherent point of D.

The point of this definition is to characterize D that might not have *all* the points in $\tilde{D}$, but are such that we can approach any point in $\tilde{D}$ from within D. In some sense, D "knows" about all the points in $\tilde{D}$.

We can check that $\mathbb{Q}$ is dense in $\mathbb{R}$—in the topological sense, but it turns out that this agrees with our earlier definition of density. This is not necessarily the case with any pair of sets:

Example 3.1.3.

- $(0,1)$ is dense in $[0,1]$
- $\mathbb{Z}$ is dense in $\mathbb{Z}$
- $\mathbb{Z}$ is not dense in $\mathbb{R}$
- $\mathbb{Z}$ is not dense in $\frac{1}{2}\mathbb{Z}$ because we cannot approach any element in $\frac{1}{2}\mathbb{Z}$ using a sequence of integers. But this disagrees with our earlier definition of density: "between any two real numbers, you can find a rational number."

Lemma 3.1.1. Let $D \subset \tilde{D} \in \mathbb{R}$. Let $f : D \longrightarrow \mathbb{R}$ be a function. If D is dense in $\tilde{D}$, then there is at most one continuous extension of f to $\tilde{D}$.

Proof. Let $\tilde{f}_1, \tilde{f}_2 : \tilde{D} \longrightarrow \mathbb{R}$ be continuous extensions of f. Then $\tilde{f}_1(x) = f(x) = \tilde{f}_2(x)$ for all $x \in D$.

Let $x_0 \in \tilde{D} \backslash D$. Then x_0 is a limit point of D i.e., there is a sequence (x_n) in D with $x_n \longrightarrow x_0$. Then

$$\begin{aligned} \tilde{f}_1(x_0) &= \lim_{n\longrightarrow\infty} \tilde{f}_1(x_n) = \lim_{n\longrightarrow\infty} f(x) \\ \tilde{f}_2(x_0) &= \lim_{n\longrightarrow\infty} \tilde{f}_2(x_n) = \lim_{n\longrightarrow\infty} f(x), \end{aligned}$$

which are equal. □

Remark. A continuous extension need not exist. Take $D = \mathbb{R}\backslash\{0\}, \tilde{D} = \mathbb{R}, f(x) = \sin(1/x)$. Graphically, f is a sinusoidal function whose oscillations increase in frequency near 0.

So a continuous extension of f does not exist, since we can pick an arbitrary value for $f(0)$, but the values of $f(x)$ do not tend toward it (nor anything) as $x \longrightarrow 0$.

It is then natural to ask if we can easily identify whether a function has a continuous extension.

Proposition 3.1.5. *A function $f : (a,b) \longrightarrow \mathbb{R}$ has a continuous extension to $[a,b]$ iff it is uniformly continuous.*

Proof. If a continuous extension $\tilde{f} : [a, b] \longrightarrow \mathbb{R}$ exists, then $\tilde{f}$ is uniformly continuous, so f is also.

Conversely, assume f is uniformly continuous. If $\tilde{f}$ existed, then for any sequence (x_n) where $x_n \in (a, b)$ and $x_n \longrightarrow a$, we would have $\tilde{f}(a) = \lim_{n\to\infty} \tilde{f}(x_n) = \lim_{n\to\infty} f(x_n)$.

Now let (x_n) be a sequence in (a, b) with $x_n \longrightarrow a$. Then we claim $\lim_{n\to\infty} f(x_n)$ exists. To show this, note that since (x_n) converges in $\mathbb{R}$, it is Cauchy. But because uniform continuity preserves Cauchy-ness, $(f(x_n))$ is Cauchy, hence convergent in $\mathbb{R}$.

Moreover, in this situation, we further clam that $\lim_{n\to\infty} f(x_n)$ is independent of (x_n). Take $(x_n), (y_n)$ such that $x_n \longrightarrow a$ and $y_n \longrightarrow a$. Then construct (z_n) as $z_{2n} = x_n, z_{2n+1} = y_n$; this gives $z_n \longrightarrow a$, so $f(z_n) \longrightarrow a$. But $(f(x_n)), (f(y_n))$ are subsequences of $(f(z_n))$, so they are equal.

Define $\tilde{f}(a) = \lim_{n\to\infty} f(x_n)$ for any sequence (x_n) with $x_n \in (a, b)$ and $x_n \longrightarrow a$. Then $\tilde{f}(a)$ is well-defined; it remains to show $\tilde{f}$ is continuous at a.

Assume not. Then there exists (x_n) with $x_n \in (a, b)$ and $x_n \longrightarrow a$ but $\tilde{f}(x_n) \not\longrightarrow \tilde{f}(a)$. That is, we can find $\epsilon > 0$ such that for all N, there exist $n > N$ with $|\tilde{f}(x_n) - \tilde{f}(a)| \geq \epsilon$. So there is a subsequence (x_{nk}) such that $|\tilde{f}(x_{nk}) - \tilde{f}(a)| \geq \epsilon$. In particular, $\tilde{f}(x_{nk}) = f(x_{nk}) \not\longrightarrow \tilde{f}(a)$, a contradiction. $\square$

3.2 Limits of Functions

So far, we have discussed limits of sequences $x : \mathbb{N} \longrightarrow \mathbb{R}$. Now we want to discuss functions $f : D \longrightarrow \mathbb{R}$ for any $D \subset \mathbb{R}$, and consider limits toward any $a \in \mathbb{R}$.

Definition 3.2.1. Let $f : D \longrightarrow \mathbb{R}$ be a function and take a limit point $a \in \mathbb{R}$ of D. Let $L \in \mathbb{R}$. We define $\lim_{x\to a} f(x) = L$ to mean:

$$\forall \epsilon > 0 : \exists \delta > 0 : \forall x \in D : 0 < |x - a| < \delta \Longrightarrow |f(x) - L| < \epsilon.$$

Note that a need not be in D and we don't allow $x = a$.

Example 3.2.1. Define the following:

$$f(x) = \begin{cases} 0 & x \neq 0 \\ 1 & x = 0. \end{cases}$$

Then $\lim_{x\to 0} f(x) = 0$. Alternatively, consider

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0. \end{cases}$$

Now the limit at 0 does not exist.

Using the fact that the definition of the limit resembles that of continuity, we see immediately that if $f : D \longrightarrow \mathbb{R}$ is a function and $a \in D$ is a limit point, then TFAE:

- f is continuous at a
- $\lim_{x\to a} f(x) = f(a)$

Proposition 3.2.1. *TFAE:*

1. $\lim_{x \to a} f(x) = L$
2. *For any sequence* (x_n) *in* $D \backslash \{a\}$ *and* $x_n \to a$*, we have* $f(x_n) \to L$

Proof. (1) $\Longrightarrow$ (2): Let $\epsilon > 0$. Take $\delta > 0$ as in the definition of limit. Take N such that for $n > N$, we have $|x_n - a| < \delta$. Then $|f(x_n) - L| < \epsilon$.

$\neg(1) \Longrightarrow \neg(2)$: There exists $\epsilon > 0$ such that for all $\delta > 0$, there is $x \in D$ such that $0 < |x - a| < \delta$ and $|f(x) - L| \geq \epsilon$. Choose $\delta = 1/n$ for $0 \neq n \in \mathbb{N}$ to get a sequence $(x_n), x_n \in D \backslash \{a\}$ where $x_n \to a$ and $f(x_n) \not\to L$. □

Corollary 3.2.1. *Let* $f_1, f_2 : D \to \mathbb{R}$, $a \in \mathbb{R}$ *a limit point of* D*, and* $L_i = \lim_{x \to a} f_i(x)$ *for* $i = 1, 2$*. Then*

$$\lim_{x \to a} (f_1(x) + f_2(x)) = L_1 + L_2$$
$$\lim_{x \to a} (f_1(x) - f_2(x)) = L_1 - L_2$$
$$\lim_{x \to a} (f_1(x) \cdot f_2(x)) = L_1 \cdot L_2$$
$$\lim_{x \to a} (f_1(x)/f_2(x)) = L_1/L_2 \text{ if } L_2 \neq 0.$$

Proof. Apply Proposition 3.2.1 and sequence limit theorems. □

Corollary 3.2.2. *Let* $f : D \to E \subset \mathbb{R}$ *and* $g : E \to \mathbb{R}$*. If* $a \in \mathbb{R}$ *is a limit point of* D *and* $\lim_{x \to a} f(x) = b \in E$*, then*

$$\lim_{x \to a} (g \circ f)(x) = g(b).$$

Proof. Apply Proposition 3.2.1 and the analogous result for sequence limits. □

Moreover, based on our proof of Proposition 3.1.5, if we take $f : D \to \mathbb{R}$ and a limit point $a \in \mathbb{R} \backslash D$ of D, then f has a continuous extension $\tilde{f}$ to $\tilde{D} := D \cup \{a\}$ iff $\lim_{x \to a} f(x) = L$ exists and $\tilde{f}(a) = L$.

Definition 3.2.2. Let $f : D \to \mathbb{R}$, $a \in \mathbb{R}$ being a limit point of D. We say

$$\lim_{x \uparrow a} f(x) = L \Longleftrightarrow \forall \epsilon > 0 : \exists \delta > 0 : \forall x \in D : a - \delta < x < a \Longrightarrow |f(x) - L| < \epsilon$$

$$\lim_{x \downarrow a} f(x) = L \Longleftrightarrow \forall \epsilon > 0 : \exists \delta > 0 : \forall x \in D : a < x < a + \delta \Longrightarrow |f(x) - L| < \epsilon.$$

Combining the constraints on x, we get

$$\lim_{x \to a} = L \Longleftrightarrow \lim_{x \uparrow a} f(x) = L = \lim_{x \downarrow a} f(x).$$

Example 3.2.2. Recall this example:

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0. \end{cases}$$

We have $\lim_{x \uparrow 0} = 0 \neq 1 = \lim_{x \downarrow 0}$, so the limit at 0 does not exist.

3.3 Differentiation

The motivation for differentiation begins with us considering the most basic type of function:

Definition 3.3.1. An **affine linear function** is of the form $f(x) = ax + b$ for $a, b \in \mathbb{R}$.

It is very easy to understand—it contains exactly two pieces of information: its rate of change and its value at 0.

The point of differentiation is any function looks like an affine linear function if we look "closely enough" around a given point. Intuitively, we say that a function $f : (a, b) \longrightarrow \mathbb{R}$ is differentiable at $c \in (a, b)$ if there exists an affine linear function $L : \mathbb{R} \longrightarrow \mathbb{R}$ such that $f(x) \approx L(x)$ for x close to c.

That is, $f(x) = L(x) + e(x)$, where $e : (a, b) \longrightarrow \mathbb{R}$ satisfies

$$e(c) = 0 \qquad \lim_{x \to c} \frac{e(x)}{x - c} = 0.$$

We say that e vanishes at c **superlinearly**. $(x - c)^2$ vanishes quadratically at c, and $(x - c)^n$ vanishes to the order of n at c.

In other words, the linear behavior of f at c is fully captured at L. The rate of change of L is called the infinitesimal rate of change of f at c.

Now it is natural to ask:

Question. How can we check if an arbitrary f has this property?

Proposition 3.3.1. *Let $f : (a, b) \longrightarrow \mathbb{R}$ and $c \in (a, b)$. TFAE:*

1. *There exists an affine linear function $L : \mathbb{R} \longrightarrow \mathbb{R}$ such that $e(x) = f(x) - L(x)$ satisfies $e(c) = 0$ and e vanishes superlinearly.*
2. $\lim_{x \to c} \frac{f(x) - f(c)}{x - c} =: f'(c)$ *exists.*

In that case, L is uniquely determined by $L(x) = f(c) + f'(c)(x - c)$.

Proof. (1) $\Longrightarrow$ (2):

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c} \frac{(L(x) + e(x)) - (L(c) + e(c))}{x - c} = \lim_{x \to c} \frac{L(x) - L(c)}{x - c} + \lim_{x \to c} \frac{e(x)}{x - c}.$$

The first term is constant in x, and the second term is 0 by construction.

(2) $\Longrightarrow$ (1): Define $L(x) = f(x) + f'(c)(x - c)$. Then $e(x) = f(x) - L(x)$ satisfies

$$e(c) = f(c) - L(c) = f(c) - (f(c) + f'(c)(c - c)) = 0$$

and

$$\lim_{x \to c} \frac{e(x)}{x - c} = \lim_{x \to c} \frac{f(x) - f(c) - f'(c)(x - c)}{x - c} = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} - \lim_{x \to c} f'(c) \frac{x - c}{x - c}.$$

The first term is, by definition, $f'(c)$, and the second term is $\lim_{x \to c} f'(c) = f'(c)$. So the whole expression is 0.

To prove uniqueness of L, let $L : \mathbb{R} \longrightarrow \mathbb{R}$ be an arbitrary linear function satisfying (1). Then $L(c) = f(c) - e(c) = f(c)$. We write

$$\begin{aligned}
\frac{L(x) - L(c)}{x - c} = \lim_{x \to c} \frac{L(x) - L(c)}{x - c} &= \lim_{x \to c} \frac{(f(x) - e(x)) - (f(c) - e(c))}{x - c} \\
&= \lim_{x \to c} \frac{f(x) - f(c)}{x - c} - \lim_{x \to c} \frac{e(x)}{x - c} \\
&= f'(c).
\end{aligned}$$

□

Definition 3.3.2. A function $f : (a, b) \longrightarrow \mathbb{R}$ is **differentiable at** $c \in (a, b)$ if it satisfies the equivalent conditions above.

In other words, f is differentiable at c iff $f(x) = f(c) + f'(c)(x - c) + e(x)$ and $e(x)$ vanishes superlinearly at c.

Proposition 3.3.2. *If f is differentiable at c, then f is continuous at c.*

Definition 3.3.3.

$$\lim_{x \to c}(f(x) - f(c)) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}(x - c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \cdot \lim_{x \to c}(x - c) = f'(c) \cdot 0 = 0.$$

So f is continuous at c.

Proposition 3.3.3. *Let $f, g : (a, b) \longrightarrow \mathbb{R}$ be differentiable at c. Then so are $f + g, fg, f/g$ if $g \neq 0$. Moreover,*

1. $(f + g)'(c) = f'(c) + g'(c)$
2. $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$
3. $(f/g)(c) = \frac{f'(c)g(c) - f(c)g'(c)}{g(c)^2}$

Proof.

1. $$\begin{aligned}
\lim_{x \to c} \frac{(f + g)(x) - (f + g)(c)}{x - c} &= \lim_{x \to c} \frac{f(x) - f(c) + g(x) - g(c)}{x - c} \\
&= \lim_{x \to c} \frac{f(x) - f(c)}{x - c} + \lim_{x \to c} \frac{g(x) - g(c)}{x - c}.
\end{aligned}$$

2. $$\begin{aligned}
\lim_{x \to c} \frac{(fg)(x) - (fg)(c)}{x - c} &= \lim_{x \to c} \frac{f(x)g(x) - f(c)g(x) + f(c)g(x) + f(c)g(c)}{x - c} \\
&\lim_{x \to c} \frac{f(x) - f(c)}{x - c} g(x) + \lim_{x \to c} \frac{g(x) - g(c)}{x - c} f(x).
\end{aligned}$$

   Note that $\lim_{x \to c} g(x) = g(c)$ because g is continuous (since it is differentiable). The result follows immediately.

3. By (2), we can assume $f(x) = 1$. Then

$$\lim_{x \to c} \frac{1/g(x) - 1/g(c)}{x - c} = \lim_{x \to c} \frac{g(c) - g(x)}{g(x)g(c)(x - c)} = -\frac{g'(c)}{g(c)^2}.$$

$\square$

Corollary 3.3.1. *A polynomial function* $f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ *for* $a_n, \ldots, a_0 \in \mathbb{R}$ *is differentiable at any* $c \in \mathbb{R}$ *and* $f'(c) = a_n \cdot nc^{n-1} + a_{n-1} \cdot (n-1)c^{n-2} + \cdots + a_1$.

Proof. Since the derivative respects arithmetic, it is enough to treat the special cases $f(x) = a_0$ and $f(x) = x^n$. In the first case, we simply have $f'(x) = 0$.

To show the second case, we induce on n. For $n = 1$, we have $\lim_{x \to c} \frac{x-c}{x-c} = 1$. For general n, write $f(x) = x \cdot x^{n-1}$ so $f'(c) = 1 \cdot c^{n-1} + c \cdot (n-1) \cdot c^{n-1} = c^{n-1} + (n-1)c^{n-1} = n \cdot c^{n-1}$ by the product rule. $\square$

Proposition 3.3.4 (Chain rule). *Let* $f_1 : (a_1, b_1) \longrightarrow (a_2, b_2) \subset \mathbb{R}$ *and* $f_2 : (a_2, b_2) \longrightarrow \mathbb{R}$. *Assume* f_1 *is differentiable at* $c_1 \in (a_1, b_1)$ *and* f_2 *is differentiable at* $c_2 = f_1(c_1)$. *Then* $f_2 \circ f_1$ *is differentiable at* c_1 *and*

$$\begin{aligned}(f_2 \circ f_1)'(c) &= f_2'(c_2) \cdot f_1'(c_1) \\ &= f_2'(f_1(c_1)) \cdot f_1'(c_1).\end{aligned}$$

Proof. Define

$$g(y) = \begin{cases} \frac{f_2(y) - f_2(c_2)}{y - c_2} & y \neq c_2 \\ f'(c_2) & y = c_2 \end{cases}.$$

Then, by the definition of the derivative, $\lim_{y \to c_2} g(y) = g(c_2)$, so g is continuous at c_2. Now we write

$$\frac{f_2(f_1(x)) - f_2(f_1(c_1))}{x - c_1} = g(f_1(x)) \cdot \frac{f_1(x) - f_1(c_1)}{x - c_1}.$$

Why is this true? Suppose we pick x such that $f_1(x) \neq f_1(c_1)$. Then $g(f_1(x))$ takes the first case and we have

$$\frac{f_2(f_1(x)) - f_2(f_1(c_1))}{x - c_1} = \frac{f_2(y) - f_2(c_2)}{y - c_2} \cdot \frac{f_1(x) - f_1(c_1)}{x - c_1} = \frac{f_2(y) - f_2(c_2)}{x - c_1}.$$

If $f_1(x) = f_1(c_1)$, then the equality collapses to $0 = 0$.

Now we take the limit as $x \longrightarrow c_1$ on the left to obtain $(f_2 \circ f_1)'(c)$ and do the same on the right:

$$\begin{aligned}&\lim_{x \to c_1} \left(g(f_1(x)) \cdot \frac{f_1(x) - f_1(c_1)}{x - c_1} \right) \\ &= \lim_{x \to c_1} g(f_1(x)) \cdot \lim_{x \to c_1} \frac{f_1(x) - f_1(c_1)}{x - c_1}.\end{aligned}$$

The left limit exists because it equals $g(f_1(c_1))$ by continuity of g, f_1 at c_1, and the right limit exists by differentiability of f_1 at c_1.

But then the product simply reduces to $g(f_1(c_1))f_1'(c_1)$, as desired. $\square$

3.4 Properties of the Derivative

Before we proceed, let's recall some intuition for the derivative.

Analytically, we can think of $f'(c)$ as the infinitesimal rate of change of f at c; that is, if $y = f(x)$, $x - c = \Delta x$, and $f(x) - f(x) = \Delta y$, then $f'(c) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$.

Geometrically, we can fix some x and compute the difference quotient $\frac{f(x)-f(c)}{x-c}$ as the slope of the secant line $x - c$ through $(c, f(c))$ and $(x, f(x))$. As we take the limit $x \longrightarrow c$, the secant lines approach the tangent line, so $f'(c)$ is the slope of the tangent line.

We start with the following question:

Question. We know that if $f : (a, b) \longrightarrow \mathbb{R}$ is constant, then $f'(c) = 0$ for all $c \in (a, b)$. Does the converse hold?

Definition 3.4.1. Let $f : (a, b) \longrightarrow \mathbb{R}$ and $c \in (a, b)$. Then c is:

- a **local minimum** if $\exists \epsilon > 0 : \forall x \in (c - \epsilon, c + \epsilon) : f(x) \geq f(c)$
- a **local maximum** if $\exists \epsilon > 0 : \forall x \in (c - \epsilon, c + \epsilon) : f(x) \leq f(c)$
- a **local extremum** if either.

Proposition 3.4.1. *Assume $f : (a, b) \longrightarrow \mathbb{R}$ is differentiable at c. If c is a local extremum, then $f'(c) = 0$.*

Proof. Assume not. Say WLOG that $f'(c) > 0$. Then choose $0 < \epsilon < f'(c)$ and let $\delta > 0$ be such that for $0 < |x - c| < \delta$,

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

or, equivalently,

$$0 < f'(c) - \epsilon < \frac{f(x) - f(c)}{x - c} < f'(c) + \epsilon.$$

Thus, for $c < x < c + \delta$, we have

$$f(x) > f(c) + (f'(c) - \epsilon)(x - c) > f(c),$$

because $f'(c) - \epsilon > 0$ and $x - c > 0$. For $c - \delta < x < c$, we have

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

because $f'(c) - \epsilon > 0$ again but $x - c < 0$. Hence c is neither a local min nor a local max, contradiction. □

Definition 3.4.2. A function $f : (a, b) \longrightarrow \mathbb{R}$ is called **differentiable** if it is differentiable at every $c \in (a, b)$.

This gives us a new function $f' : (a, b) \longrightarrow \mathbb{R}$!

Corollary 3.4.1 (Rolle's theorem)**.** *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous, and suppose f is differentiable on (a, b). If $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. f attains its min and max at $x_-, x_+ \in [a, b]$. If $\{x_-, x_+\} \subseteq \{a, b\}$, then $f(x_-) = f(x_+)$, so f is constant and $f' = 0$. Otherwise, let $c = x_-$ or $c = x_+$, whichever lies in (a, b), and apply the previous result. □

Theorem 3.4.1 (Mean value theorem). *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous, and suppose it is differentiable on (a, b). There exists $c \in (a, b)$ such that $\frac{f(b)-f(a)}{b-a} = f'(c)$.*

Proof. We tilt our heads and apply Rolle. More precisely, let $L : \mathbb{R} \longrightarrow \mathbb{R}$ be an affine linear function whose graph is the secant line for a, b; that is,

$$L(x) = \frac{f(b) - f(a)}{b - a}(x - a) + f(a)$$

and let $h(x) = f(x) - L(x)$. Then

$$h(a) = f(a) - L(a) = 0$$

$$h(b) = f(b) - L(b) = 0.$$

By Rolle, there is $c \in (a, b)$ such that

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

□

Corollary 3.4.2. *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous, differentiable on (a, b).*

1. *If $f'(c) > 0$ for all $c \in (a, b)$, then f is strictly increasing.*
2. *If $f'(c) < 0$ for all $c \in (a, b)$, then f is strictly decreasing.*
3. *If $f'(c) = 0$ for all $c \in (a, b)$, then f is constant.*

Proof. For any $x_1, x_2 \in [a, b]$ where $x_1 < x_2$, there exists by MVT some $x_1 < c < x_2$ such that $f(x_2) - f(x_1) = f'(c)(x_2 - x_1)$. □

Definition 3.4.3. Let $f : I \longrightarrow \mathbb{R}$ be a function defined on an open interval. An **antiderivative** of f is a differentiable function $F : I \longrightarrow \mathbb{R}$ such that $F' = f$.

Proposition 3.4.2. *Let $F_1, F_2 : I \longrightarrow \mathbb{R}$ be antiderivatives of the same $f : I \longrightarrow \mathbb{R}$. Then there exists $C \in \mathbb{R}$ such that $F_2(x) = F_1(x) + C$ for all $x \in I$.*

Proof. By construction of F_1, F_2, we can write $F_1' - F_2' = 0 \implies (F_1 - F_2)' = 0$, so $F_1 - F_2$ is constant. □

This means that, if it exists, an antiderivative is almost completely uniquely determined.

Corollary 3.4.3. *Let I be an open interval and $f : I \longrightarrow \mathbb{R}$ differentiable. If $f' : I \longrightarrow \mathbb{R}$ is bounded, then f is uniformly continuous.*

Proof. Let $M \in \mathbb{R}$ be such that $|f'(x)| \leq M$ for all $x \in I$. By MVT, for any $x, y \in I$, there is $z \in I$ such that $|f(x) - f(y)| = |f'(z)| \cdot |x - y| \leq M \cdot |x - y|$. Thus, for any $\epsilon > 0$, we can take $\delta = \epsilon/M$, and we are done. □

Example 3.4.1. Consider $f(x) = 1/x^2$, which has $f'(x) = -2/x$. The derivative is bounded on $(1, \infty)$, so f is uniformly continuous on the same interval.

Note that the converse is *not* true; there are uniformly continuous functions with unbounded derivatives.

4 Integration

4.1 The Riemann Integral

We want a conceptual understanding of area, and we get there by a conceptual understanding of the area "under the graph of a function." This leads us to the following questions:

Question.

1. How do we make this precise?
2. For what functions does this work?
3. What properties does this procedure have?

Denote the area under a function f from a to b by $\int_a^b f(x)dx$. Let's start with some immediate observations.

1. If $f(x) = k \in \mathbb{R}$ is constant, then
$$\int_a^b f(x)dx = (b-a)k.$$
2. For any f_1, we have
$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

So we know how to integrate constant and step (piecewise constant) functions.

Definition 4.1.1.

1. A **partition** of $[a, b]$ is a finite sequence P where $a = a_0^P \leq a_1^P \leq \cdots \leq a_n^P = b$ for $n = |P|$.
2. A partition Q **refines** P if
$$\{a_0^P, \ldots, a_{|P|}^P\} \subseteq \{a_0^Q, \ldots, a_{|Q|}^Q\}.$$
3. A **step function adapted to** P is a function $h : [a, b] \longrightarrow \mathbb{R}$ such that $h|_{(a_i^P, a_{i+1}^P)}$ is constant for all $i = 0, \ldots, |P| - 1$.
4. A **step function** is a function $h : [a, b] \longrightarrow \mathbb{R}$ such that there exists a partition P satisfying 3).

We specify an open interval for 3) because it's arbitrary which "step" we pick on the boundary of a partition.

Definition 4.1.2. If h is a step function adapted to P, then
$$\int_P h(x)dx := \sum_{i=0}^{|P|-1} (a_{i+1}^P - a_i^P) \cdot h\left(\frac{a_i^P + a_{i+1}^P}{2}\right).$$

Remark. The values of h at a_i^P are irrelevant because of the point above and because always we sample h between two partition boundaries.

It follows immediately that if Q is a refinement of P, then

$$\int_P h(x)dx = \int_Q h(x)dx.$$

Note also that any two partitions have a common refinement—simply take their union. This gives us the following:

Corollary 4.1.1. *Let h be a step function adapted to two partitions P_1, P_2. Then*

$$\int_{P_1} h(x)dx = \int_{P_2} h(x)dx.$$

Proof. See above. □

Definition 4.1.3. Let $h : [a,b] \longrightarrow \mathbb{R}$ be a step function. Define

$$\int_a^b h(x)dx = \int_P h(x)dx$$

where P is any partition to which h is adapted.

Lemma 4.1.1. If $h_1, h_2 : [a,b] \longrightarrow \mathbb{R}$ are step functions and $h_1 < h_2$ (that is, $h_1(x) \leq h_2(x)$) for all $x \in [a,b]$, then

$$\int_a^b h_1(x)dx \leq \int_a^b h_2(x)dx.$$

Proof. Take partitions for h_1, h_2 and find a common refinement. Apply Definition 4.1.2 and Definition 4.1.3 □

Now we try to approximate general functions using refinement of partitions.

Definition 4.1.4. Let $f : [a,b] \longrightarrow \mathbb{R}$ be bounded.

1. $US(f) = \{h : [a,b] \longrightarrow \mathbb{R} \mid h \text{ step and } h \geq f\}$
2. $LS(f) = \{h : [a,b] \longrightarrow \mathbb{R} \mid h \text{ step and } h \leq f\}$
3. $$\overline{\int_a^b} f(x)dx := \inf\left\{\int_a^b h(x)dx \mid h \in US(f)\right\}.$$ $$\underline{\int_a^b} f(x)dx := \sup\left\{\int_a^b h(x)dx \mid h \in LS(f)\right\}.$$

   We call these the **upper** and **lower** integrals of f, respectively.

Lemma 4.1.2.

$$+\infty \overset{(1)}{>} (b-1)\sup f([a,b]) \overset{(2)}{\geq} \overline{\int_a^b} f(x)dx$$
$$\overset{(3)}{\geq} \underline{\int_a^b} f(x)dx \overset{(4)}{\geq} (b-a)\inf f([a,b]) \overset{(5)}{>} -\infty.$$

Proof. By assumption, f is bounded, so (1) and (5) follow.

Also note that

$$\begin{aligned} h_2(x) &:= \sup f([a,b]) \in US(f) \\ h_1(x) &:= \sup f([a,b]) \in LS(f), \end{aligned}$$

which gives (2) and (4).

Now define

$$\begin{aligned} us(f) &= \left\{ \int h(x)dx \mid h \in US(f) \right\} \\ ls(f) &= \left\{ \int h(x)dx \mid h \in LS(f) \right\}. \end{aligned}$$

Note that $h_2 \in US$ and $h_1 \in LS$, so $h_1 \leq f \leq h_2$, so $ls(f) \leq us(f)$. It is easy to show that, for $A, B \subset \mathbb{R}$, we have $A \leq B \implies \sup A \leq \inf B$. Hence

$$\overline{\int_a^b} f(x)dx = \inf us(f) \geq \sup ls(f) = \underline{\int_a^b} f(x)dx.$$

□

Definition 4.1.5. A bounded $f : [a,b] \longrightarrow \mathbb{R}$ is called **Riemann integrable** if

$$\overline{\int_a^b} f(x)dx = \underline{\int_a^b} f(x)dx.$$

We denote this value by $\int_a^b f(x)dx$.

At this point in the class, Kaletha paused to check in on us:

> How do you feel? I will give you some options: "I feel happy, sad, lost, distraught, confused, tormented, devastated..."

4.2 Properties of the Riemann Integral

Now we are probably inclined to ask:

Question. Which functions are Riemann integrable?

Lemma 4.2.1. Any step function is Riemann integrable; moreover, Definition 4.1.2 and Definition 4.1.5 agree.

Proof. Let h be a step function. Then $\max LS(h) = h = \min US(h)$. So

$$\overline{\int_a^b} h(x)dx = \int_a^b h(x)dx = \underline{\int_a^b} h(x)dx.$$

□

Example 4.2.1. Consider the function

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

on $[0, 1]$. Note that $\min US(f) = 1 \neq 0 = \max LS(f)$. So f is not Riemann integrable.

Lemma 4.2.2. For bounded $f : [a, b] \longrightarrow \mathbb{R}$, TFAE:

1. f is Riemann integrable
2. For $\epsilon > 0$, there exist step functions $h_1 \leq f \leq h_2$ such that

$$\int_a^b h_2(x)dx - \int_a^b h_1(x)dx < \epsilon.$$

Proof. 1) $\implies$ 2): Define $I := \int_a^b f(x)dx \in \mathbb{R}$. Then $I - \epsilon/2$ is not an upper bound for $ls(f)$. Thus there exists $h_1 \in LS(f)$ such that

$$\int_a^b h_1(x)dx > I - \epsilon/2.$$

Similarly, $I + \epsilon/2$ is not a lower bound for $us(f)$, so there exists $h_2 \in US(f)$ such that

$$\int_a^b h_2(x)dx < I + \epsilon/2.$$

Now, because $h_1 \leq f \leq h_2$,

$$\int_a^b h_2(x)dx - \int_a^b h_1(x)dx < (I + \epsilon/2) - (I - \epsilon/2) = \epsilon.$$

2) $\implies$ 1): Fix ϵ. Let $h_1 \leq f \leq h_2$; then

$$0 \leq \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx \leq \int_a^b h_2(x)dx - \int_a^b h_1(x)dx < \epsilon.$$

□

Theorem 4.2.1. *Any continuous $f : [a, b] \longrightarrow \mathbb{R}$ is Riemann integrable.*

Proof. Fix $\epsilon > 0$. By a previous result, f is uniformly continuous. Now choose $\delta > 0$ such that for $|x - y| < \epsilon$, we have $|f(x) - f(y)| < \epsilon/(b - a)$.

By Archimedian Property, there is $n \in \mathbb{N}$ where $n > 1/\delta$. Define $P : a_i^P = a + i \cdot \frac{b-a}{n}$. Then

$$h_2|_{(a_i^P, a_{i+1}^P)} = \sup f([a_i^P, a_{i+1}^P]) = f(M_i)$$

$$h_1|_{(a_i^P, a_{i+1}^P)} = \inf f([a_i^P, a_{i+1}^P]) = f(m_i)$$

for $m_i, M_i \in [a_i^P, a_{i+1}^P]$. We then write

$$\int_a^b h_2(x)dx - \int_a^b h_1(x)dx = \sum_{i=1}^{|P|-1} (a_{i+1}^P - a_i^P)(f(M_i) - f(m_i)) < \epsilon/(b-a) < \epsilon.$$

□

Theorem 4.2.2. *If $f : [a, b] \longrightarrow \mathbb{R}$ is bounded and monotonic, then f is integrable.*

Proof. WLOG, assume f is non-decreasing. Let $\epsilon > 0$ and $n \in \mathbb{N}$ such that

$$(f(b) - f(a)) \cdot (b - a)/\epsilon < n$$

by Archimedian property. Let P be the equi-spaced partition $a_i^P = a + i \cdot \frac{b-a}{n}$. Define step functions

$$h_1|_{(a_i^P, a_{i+1}^P)} = f(a_i^P)$$

$$h_2|_{(a_i^P, a_{i+1}^P)} = f(a_{i+1}^P).$$

Then $h_1 \leq f \leq h_2$ and

$$\begin{aligned} &\int_a^b h_2(x)dx - \int_a^b h_1(x)dx \\ &= \sum_{i=1}^{n-1} (a_{i+1}^P - a_i^P)(f(a_{i+1}^P) - f(a_i^P)) \\ &= \frac{b-a}{n} \sum_{i=1}^{n-1} (f(a_{i+1}^P) - f(a_i^P)) = \frac{b-a}{n}(f(b) - f(a)) < \epsilon. \end{aligned}$$

□

Theorem 4.2.3. *Let $f, g : [a, b] \longrightarrow \mathbb{R}$ be integrable. Then*

1. *$f + g$ is also integrable and*

$$\int_a^b f(x) + g(x)dx = \int_a^b f(x)dx + \int_a^b g(x)dx.$$

2. *If $k \in \mathbb{R}$ then $k \cdot f$ is integrable and*

$$\int_a^b k \cdot f(x)dx = k \cdot \int_a^b f(x)dx.$$

3. *If $f \leq g$ then*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx.$$

4. *If $c \in (a,b)$, then*

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

5. *If $m \leq f \leq M$ with $m, M \in \mathbb{R}$, then*

$$(b-a)m \leq \int_a^b f(x)dx \leq (b-a)M.$$

Proof.

1. Let $\epsilon > 0$. There exist step functions h_1, h_2, h_3, h_4 such that $h_1 \leq f \leq h_2$ and $h_3 \leq g \leq h_4$ and $\int h_2 - \int h_1 < \epsilon/2$ and $\int h_4 - \int h_3 < \epsilon/2$.

   Then $h_1 + h_3$ and $h_2 + h_4$ are step functions (exercise) and $h_1 + h_3 \leq f + g \leq h_2 + h_4$. We can write

   $$\int h_4 + h_2 - \int h_3 + h_1 = \int h_4 - \int h_3 + \int h_2 - \int h_1 < \epsilon/2 + \epsilon/2 = \epsilon,$$

   which is not circular because $h_4 + h_2$ and $h_3 + h_1$ are step functions. So $f + g$ is integrable by Lemma 4.2.2.

   Now to show the addition property, it is easy to see $LS(f) + LS(g) \subseteq LS(f+g)$. This implies $ls(f) + ls(g) \subseteq ls(f+g)$. Using basic properties of sup, we have $\sup ls(f) + \sup ls(g) \leq \sup ls(f+g)$. Hence

   $$\underline{\int_a^b} f(x)dx + \underline{\int_a^b} g(x)dx \leq \underline{\int_a^b} f(x) + g(x)dx.$$

   Since $f, g, f+g$ are integrable, we conclude

   $$\int_a^b f(x)dx + \int_a^b g(x)dx \leq \int_a^b (f(x) + g(x))dx.$$

   Now, using US and $\overline{\int}$, we get the opposite inequality.

2. "Similar, but easier."

3. By 1) and 2), it is enough to show

   $$f \geq 0 \implies \int_a^b f(x)dx \geq 0.$$

   But $f \geq 0 \implies 0 \in LS(f)$, so

   $$0 \leq \underline{\int_a^b} f(x)dx = \int_a^b f(x)dx.$$

4. Note $LS(f) = LS(f|_{[a,c]}) + LS(f|_{[c,b]})$, which implies $ls(f) = ls(f|_{[a,c]}) + ls(f|_{[c,b]})$ and $\sup ls(f) = \sup ls(f|_{[a,c]}) + \sup ls(f|_{[c,b]})$. Hence

   $$\underline{\int_a^b} f(x)dx = \underline{\int_a^c} f(x)dx + \underline{\int_c^b} f(x)dx,$$

   and the result follows by integrability.

5. Apply 3) to $m \le f \le M$.

$\square$

Remark. In 1), we subtly use the fact that the integral respects addition of step functions. This is immediate if we take a common refinement of the partitions of both integrals.

Lemma 4.2.3. If $f : [a, b] \longrightarrow \mathbb{R}$ is continuous, then

$$\left| \int_a^b f(x) dx \right| \le \int_a^b |f(x)| dx.$$

Proof. Note that $-|f| \le f \le |f|$. Apply 3) from Theorem 4.2.3.

If we accept without proof that the composition of a continuous and integrable function is itself integrable, then the result further holds whenever f is integrable. $\square$

4.3 The Fundamental Theorem of Calculus

Question. We already know that $f : [a, b] \longrightarrow \mathbb{R}$ is integrable. Can we compute $\int_a^b f(x) dx$ efficiently?

Theorem 4.3.1 (Fundamental theorem of calculus). *Let $f : [a, b] \longrightarrow \mathbb{R}$ be integrable. Define $F(x) = \int_a^x f(t) dt$.*

Then F is uniformly continuous. If f is continuous at $c \in (a, b)$, then F is differentiable at c and $F'(c) = f(c)$.

Proof. Let $x, y \in [a, b]$ and $x < y$. Then

$$|F(y) - F(x)| = \left| \int_x^y f(t) dt \right| \le \int_x^y |f(t)| dt \le |y - x| \cdot \sup |f([a, b])|.$$

So F is uniformly continuous.

Now assume f is continuous at c. Let $\epsilon > 0$. Choose $\delta > 0$ such that $|t - c| < \delta \implies |f(t) - f(c)| < \epsilon$. Then

$$\begin{aligned}
\left| \frac{F(x) - F(c)}{x - c} - f(c) \right| &= \left| \frac{1}{x - c} \int_c^x f(t) dt - f(c) \right| \\
&= \left| \frac{1}{x - c} \int_c^x (f(t) - f(c)) dt \right| \\
&\le \frac{1}{|x - c|} \int_c^x |f(t) - f(c)| dt \\
&\le \frac{1}{|x - c|} \int_c^x \epsilon dt = \epsilon
\end{aligned}$$

when $|x - c| < \delta$. $\square$

Recall that G is an antiderivative of f if $G' = f$. We showed already that G is unique up to the addition of a constant. And if f is continuous, FTC shows that an antiderivative exists.

Importantly, the concept of the antiderivative is, *a priori*, totally independent of integration. Integrals compute area; the antiderivative asks if a given function came from differentiating another function.

This is why the FTC is significant. Integrals aren't inherently antiderivatives, like we are taught in high school—the FTC shows in hindsight that the two are related. The following example shows this:

Example 4.3.1. Consider the function

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

on $[-1, 1]$. We can write

$$F(x) = \int_{-1}^{x} f(t)dt = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases}.$$

This isn't differentiable, and it turns out that f does not have an antiderivative despite being integrable.

Corollary 4.3.1. *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous and let G be any antiderivative of f. Then*

$$\int_{a}^{b} f(x)dx = G(b) - G(a) =: [G]_{a}^{b}.$$

Proof. Consider F from FTC. Then

$$\int_{a}^{b} f(x)dx = F(b) = F(b) - F(a) = G(b) - G(a),$$

where the last equality follows from the antiderivative uniqueness property. □

This gives is the opposite insight as FTC; we can compute an integral given any antiderivative (even one provided by "my friend Joe, who may have gotten the antiderivative from the black market").

Moreover, it holds even if we only assume f is integrable.

Corollary 4.3.2. *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous, differentiable on (a, b), with f' uniformly continuous. Then*

$$f(b) - f(a) = \int_{a}^{b} f'(x)dx.$$

Proof. Apply the previous result with $G = f$. Now, since it is uniformly continuous, we can extend f' to the closed interval. □

Proposition 4.3.1 (Integration by parts)**.** *Let $f, g : [a, b] \longrightarrow \mathbb{R}$ be continuous and differentiable on (a, b). Suppose f', g' are uniformly continuous. Then*

$$\int_{a}^{b} f(x)g'(x)dx = f(b)g(b) - f(a)g(a) - \int_{a}^{b} g'(x)g(x)dx.$$

Proof. Let $h(x) = f(x)g(x)$. Then h is continuous and differentiable on (a, b) and $h'(x) = f(x)g'(x) + f'(x)g(x)$ is uniformly continuous. Apply the previous result. $\square$

Proposition 4.3.2 (Substitution)**.** *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous, and let $u : [c, d] \longrightarrow [a, b]$ be continuous and differentiable on (c, a) with u' uniformly continuous. Then*

$$\int_c^d f(u(x))u'(x)dx = \int_{u(c)}^{u(d)} f(y)dy.$$

Proof. Let F be an antiderivative of f, which exists by FTC. Then $(F \circ u)'(x) = F'(u(x))u'(x) = f(u(x)) \cdot u'(x)$. So

$$\int_c^d f(u(x))u'(x)dx = [F \circ u]_c^d = [F]_{u(c)}^{u(d)} = \int_{u(c)}^{u(d)} f(y)dy.$$

$\square$

We can remember this by thinking of x as an independent variable and $y = u(x)$ as dependent on x.

5 Sequences and Series of Functions

5.1 The Basics

Definition 5.1.1. Let $D \subset \mathbb{R}$.

1. Assume given, for each $n \in \mathbb{N}$, a function $f_n : D \longrightarrow \mathbb{R}$. We call (f_n) a **sequence of functions**. Alternatively, we can write $f : \mathbb{N} \times D \longrightarrow \mathbb{R}$.
2. Let $f : D \longrightarrow \mathbb{R}$. We say $f_n \longrightarrow f$ **pointwise** if

$$\forall x \in D : \lim_{n \to \infty} f_n(x) = f(x).$$

Question.

1. If each f_n is continuous, is f also continuous?
2. If each f_n is differentiable at $c \in (a, b)$, is f also differentiable at c? If so, is

$$f'(c) = \lim_{n \to \infty} f_n'(c)?$$

3. If each f_n is integrable, is f also integrable? If so, is

$$\int_a^b f(x)dx = \lim_{n \to \infty} \int_a^b f_n(x)dx?$$

To answer the second question (the reasoning extends analogously to the other two), it suffices to answer the following:

$$\lim_{x \to c} \lim_{n \to \infty} \frac{f_n(x) - f_n(c)}{x - c} \stackrel{?}{=} \lim_{n \to \infty} \lim_{x \to c} \frac{f_n(x) - f_n(c)}{x - c}.$$

This gives the more general question:

Question. Can we commute two limits arbitrarily?

Example 5.1.1. Consider $a_{n,m} = \frac{n}{n+m}$. Then

$$\lim_{n \to \infty} \lim_{m \to \infty} a_{n,m} = 0$$

$$\lim_{m \to \infty} \lim_{n \to \infty} a_{n,m} = 1.$$

So it appears not. It seems that pointwise convergence doesn't play nice with our operations. Here's an alternative definition of convergence:

Definition 5.1.2. Let $D \subseteq \mathbb{R}$ and take a sequence f_n of functions $D \longrightarrow \mathbb{R}$. Then $f_n \longrightarrow f$ **uniformly** for $f : D \longrightarrow \mathbb{R}$ if

$$\forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall n > N : \forall x \in D : |f_n(x) - f(x)| < \epsilon.$$

If we expand the definition of pointwise convergence, we can compare it with uniform convergence:

$$\exists x \in D \,:\, \forall \epsilon > 0 \,:\, \exists N \in \mathbb{N} \,:\, \forall n > N \,:\, |f_n(x) - f(x)| < \epsilon.$$

They only differ in the placement of the quantifier $\forall x$, which resembles our discussion of regular/uniform continuity. The "uniformity" of "uniform convergence" refers to the fact that the $N \in \mathbb{N}$ applies uniformly across all x.

Example 5.1.2. The sequence $f_n = x^n$ with $D = [0, 1]$ converges non-uniformly to a function that is 0 everywhere except at 1. This is a sequence of continuous functions that converges to a non-continuous function.

Proposition 5.1.1. *Let $f_n \longrightarrow f$ uniformly. If each f_n is (uniformly) continuous, then f is (uniformly) continuous.*

Proof. Let $\epsilon > 0$, $a \in D$. Choose N as in the definition of uniform convergence. Choose $\delta > 0$ such that

$$|x - a| < \delta \implies |f_{N+1}(x) - f_{N+1}(a)| < \epsilon.$$

Then

$$\begin{aligned} |f(x) - f(a)| &= |f(x) - f_{N+1}(x) + f_{N+1}(x) - f_{N+1}(a) + f_{N+1}(a) - f(a)| \\ &\leq |f(x) - f_{N+1}(x)| + |f_{N+1}(x) - f_{N+1}(a)| + |f_{N+1}(a) - f(a)| \\ &< 3\epsilon. \end{aligned}$$

□

Proposition 5.1.2. *Let $f_n \longrightarrow f$ uniformly on $D = [a, b]$. If each f_n is integrable, so is f, and*

$$\int_a^b f(x)dx = \lim_{n\longrightarrow\infty} \int_a^b f_n(x)dx.$$

Proof. Let $\epsilon > 0$. Choose N as in the definition of uniform convergence. Then for $n > N$:

$$\forall x \in [a, b] \,:\, f_n(x) - \epsilon < f(x) < f_n(x) + \epsilon.$$

Thus

$$\underline{\int_a^b} (f_n(x) - \epsilon)dx \leq \underline{\int_a^b} f(x)dx \leq \overline{\int_a^b} f(x)dx \leq \overline{\int_a^b} (f_n(x) + \epsilon)dx \quad (*).$$

But we also have

$$\underline{\int_a^b} (f_n(x) - \epsilon)dx = \int_a^b f_n(x)dx - (b - a) \cdot \epsilon$$

$$\overline{\int_a^b} (f_n(x) + \epsilon)dx = \int_a^b f_n(x)dx + (b - a) \cdot \epsilon,$$

so we write

$$0 \leq \overline{\int_a^b} f(x)dx - \underline{\int_a^b} f(x)dx \leq 2(b - a)\epsilon.$$

This is true for all ϵ, so f is integrable. Use $(*)$ again to see

$$\left|\int_a^b f(x)dx - \int_a^b f_n(x)dx\right| < (b-a)\epsilon.$$

□

Proposition 5.1.3. *Let f_n be a sequence of functions on $[a, b]$ that are continuous, differentiable on (a, b), with f_n' uniformly continuous. Assume $f_n' \longrightarrow g$ uniformly, and there is some $c \in [a, b]$ such that the sequence of real numbers $(f_n(c))$ converges.*

Then $f_n \longrightarrow f$, f is differentiable, and $f' = g$.

Proof. Define

$$f(x) = \lim_{n\to\infty} f_n(c) + \int_c^x g(t)dt.$$

Note that g is integrable because it is uniformly continuous on (a, b), so it has a continuous extension to $[a, b]$.

By FTC, f is differentiable and $f' = g$. Moreover, using FTC for f_n',

$$\begin{aligned}|f(x) - f_n(x)| &= \left|\lim_{k\to\infty} f_k(c) + \int_c^x g(t)dt - \left(f_n(c) + \int_c^x f_n'(t)dt\right)\right| \\ &\leq |\lim_{k\to\infty} f_k(c) - f_n(c)| + \int_c^x |g(t) - f_n'(t)|dt.\end{aligned}$$

Both terms can be made arbitrarily small. □

Definition 5.1.3. For bounded $f : D \longrightarrow \mathbb{R}$, define $\|f\|_\infty := \sup\{|f(x)| : x \in D\}$. We call this the **sup norm** of f.

Note that $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$ by properties of the sup. Moreover, $\|f\|_\infty = 0 \Longrightarrow f = 0$.

Consider the set $C_b(D) = \{f : D \longrightarrow \mathbb{R} \text{ bounded}\}$. If we endow it with our notions of arithmetic, $C_b(D)$ forms a ring (since multiplicative inverses are not guaranteed).

Note also that $f_n \longrightarrow f$ uniformly iff $\|f_n - f\|_\infty \longrightarrow 0$.

Proposition 5.1.4 (Cauchy criterion). *$f_n \longrightarrow f$ uniformly iff (f_n) is Cauchy for the sup norm. That is,*

$$\forall\epsilon > 0 : \exists N \in \mathbb{N} : \forall n, m > N : \|f_n - f_m\|_\infty < \epsilon.$$

Proof. For the forward direction, if $f_n \longrightarrow f$ uniformly, let $\epsilon > 0$. Choose N as in the definition of uniform convergence. Then, for $n, m > N$,

$$\|f_n - f_m\|_\infty = \|f_n - f + f - f_m\|_\infty \leq \|f_n - f\|_\infty + \|f_m - f\|_\infty < 2\epsilon.$$

Now we show the backward direction. For $a \in D$, $|f_n(a) - f_m(a)| \leq \|f_n - f_m\|_\infty$ by construction of the sup norm. Thus $(f_n(a))$ is Cauchy, hence convergent. Let $f(a)$ be the limit of $(f_n(a))$. This defines a function $f : D \in \mathbb{R}$ and $f_n \longrightarrow f$ pointwise.

To check uniform convergence, let $\epsilon > 0$. Let N be such that $\|f_n - f_m\|_\infty < \epsilon$ for $n, m > N$. For $a \in D$, we have $|f_n(a) - f_m(a)| \le \|f_n - f_m\|_\infty < \epsilon$. Choose $m > N$ such that $|f_m(a) - f(a)| < \epsilon$; then $|f_n(a) - f(a)| < 2\epsilon$, which is true for all $a \in D$. □

Proposition 5.1.5 (Weierstrass M-test)**.** *Let (f_n) be a sequence of function $D \longrightarrow \mathbb{R}$. If $\sum \|f_n\|_\infty$ converges, then $\sum f_n$ converges uniformly.*

Proof.

$$\sum \|f_n\|_\infty \text{ converges} \iff \forall \epsilon > 0 : \exists N \in \mathbb{N} : \forall m > n > N : \left\| \sum_{k=n}^{m} f_n \right\|_\infty \le \sum_{k=n}^{m} \|f_k\|_\infty < \epsilon.$$

Hence the sequence of partial sums of $\sum f_n$ is Cauchy for the sup norm. □

Remark. We call this the "M test" because we can write

$$\begin{aligned} &\sum \|f_n\|_\infty \text{ converges} \\ \iff &\forall n : \exists M_n : \forall x \in D : |f_n(x)| \le M_n \text{ and } \sum M_n \text{ converges} \end{aligned}$$

Why do we care about this property? We started with the idea that "most" functions look like a line if you zoom in enough. But Weierstrass showed that there exist functions that are continuous everywhere and differentiable nowhere (and, in fact, that these form the majority of continuous functions).

The lesson is that our intuition is often misleading.

Example 5.1.3. Here, we construct an everywhere-continuous but nowhere-differentiable function.

Consider $\phi(x)$ obtained from $|x|$ on $D = [-1, 1]$ by periodic extension. This function is not differentiable at any integer; moreover, it has these properties:

0. $\|\phi\|_\infty = 1$
1. $\phi(x) = \phi(y)$ if $x - y$ is even
2. $|\phi(x) - \phi(x)| = |x - y|$ if there is no integer between x, y
3. $|\phi(x) - \phi(y)| \le |x - y|$ for any x, y

Now define

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

We have

$$\sum \left\| \left(\frac{3}{4}\right)^n \phi(4^n \cdot x) \right\| = \sum \left(\frac{3}{4}\right)^n \text{ converges by geometric series.}$$

Therefore, by the M-test,

$$\sum \left(\frac{3}{4}\right)^n \phi(4^n x) \text{ converges uniformly.}$$

By a previous result, f is continuous.

Claim. f is nowhere differentiable.

Proof. Let $a \in \mathbb{R}$. We want to show
$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$
does not exist. It is enough to find a sequence $\delta_m \longrightarrow 0$ such that $\lim_{m \to \infty} \frac{f(a+\delta_m) - f(a)}{\delta_m}$ also does not exist.

For any m, let $\alpha_m \in \{-1/2, 1/2\}$ be such that there is no integer between $4^m a + \alpha_m$ and $4^m a$. Define $\delta_m = \alpha_m / 4^m \longrightarrow 0$. Now we write
$$\frac{f(a+\delta_m) - f(a)}{\delta_m} = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \frac{\phi(4^n(a+\delta_m)) - \phi(4^n a)}{\delta_m}.$$
Call the rightmost sum term $\gamma_{n,m}$. Suppose $n > m$. Then $4^n \cdot \delta_m = 4^{n-m} \cdot \alpha_m$ is even, so $\gamma_{n,m} = 0$ by Property 1.

If $n = m$, then $4^n \delta_m = \alpha_m$. By Property 2, $|\gamma_{n,m}| = 4^n$.

If $n < m$, we use Property 3 to write $|\gamma_{n,m}| \leq 4^n$. Now we can write
$$\left|\frac{f(a+\delta_m) - f(a)}{\delta_m}\right| = \left|\sum_{n=0}^{m} \left(\frac{3}{4}\right) \gamma_{n,m}\right|$$
$$= | \pm 3^m \pm (\leq 3^{m-1}) \pm (\leq 3^{m-2}) + \cdots \pm (\leq 3^0)| \geq 3^m - 3^{m-1} - 3^{m-2} - \cdots$$
$$= \frac{1}{2}(3^m + 1) \longrightarrow \infty.$$
$\square$

The idea is that as we add more slopes, the function at our δ_m will get steeper, thus making the slope approach infinity.

Definition 5.1.4. Let $I \subseteq \mathbb{R}$ be an open interval, $f : I \longrightarrow \mathbb{R}$. Define $f^(n)$ for all (possible) $n \in \mathbb{N}$ recursively by $f^{(0)} = f$ and $f^{(n+1)} = (f^n)'$ if possible.

We say f is ***n*-times differentiable** if $f^{(n)}$ exists. f is **smooth** if $f^{(n)}$ exists for all $n \in \mathbb{N}$.

Define
$$C^n(I) = \{f : I \longrightarrow \mathbb{R} \mid f^{(n)} \text{ exists and is continuous}\}$$
$$C^\infty(I) = \{f : I \longrightarrow \mathbb{R} \mid f \text{ smooth}\}.$$
These are vector spaces; moreover, we can write
$$C^0(I) \supset C^1(I) \supset C^2(I) \supset \cdots$$
$$C^\infty(I) = \bigcup_{n=0}^{\infty} C^n(I).$$

They relate to the derivative as such:
$$C^0(I) \xleftarrow{d} C^1(I) \xleftarrow{d} C^2(I) \xleftarrow{d} \cdots$$
$$C^\infty(I) \text{ self loop.}$$

But what if we want to extend the chain to the left of $C^0(I)$?

5.2 Distributions

Not to be confused with probability distributions.

Definition 5.2.1. A function $f : \mathbb{R} \longrightarrow \mathbb{R}$ has **compact support** if there exists $k > 0$ such that $f(x) = 0$ for $x \notin [-k, k]$.

Define the vector space

$$C_C^\infty(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f \text{ smooth and has compact support}\}.$$

We call this the "space of smooth functions with compact support." Take $f \in C^0(\mathbb{R})$. Define

$$\begin{aligned} \ell_f : C_C^\infty(\mathbb{R}) &\longrightarrow \mathbb{R} \\ \varphi &\longmapsto \int_{-\infty}^{\infty} f(x) \cdot \phi(x) dx. \end{aligned}$$

We can write some properties about ℓ_f:

1. It is linear.
2. If φ_k is a sequence in $C_C^\infty(\mathbb{R})$ and $\varphi \in C_C^\infty(\mathbb{R})$, then for all $n \in \mathbb{N}$, $\varphi_k^{(n)} \longrightarrow \varphi^{(n)}$ uniformly implies $\ell_f(\varphi_k) \longrightarrow \ell_f(\varphi)$.
3. $\ell_f = 0 \implies f = 0$

Definition 5.2.2. A **distribution** is a linear map $\ell : C_C^\infty(\mathbb{R}) \longrightarrow \mathbb{R}$ which is continuous in the sense of point 2 above.

Denote the vector space of distributions by $\mathcal{D}(R)$. The map $f \longmapsto \ell_f$ is a linear map $C^0(\mathbb{R}) \longrightarrow \mathcal{D}(\mathbb{R})$. Let $f \in C'(\mathbb{R})$. Then $f' \in C^0(\mathbb{R})$. We write

$$\begin{aligned} \ell_{f'}(\varphi) &= \int_{-\infty}^{+\infty} f'(x) \cdot \varphi(x) dx \\ [f \cdot \varphi]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} f(x) \cdot \varphi'(x) dx &= -\ell_f(\varphi'). \end{aligned}$$

So we have just generalized the derivative to non-differentiable functions!

Definition 5.2.3. For any $\ell \in \mathcal{D}(\mathbb{R})$, define $\ell' \in \mathcal{D}(\mathbb{R})$ by $\ell'(\varphi) = -\ell(\varphi')$.

We have $(\ell_f)' = \ell_{f'}$. This gives a linear map $d : \mathcal{D}(\mathbb{R}) \longrightarrow \mathcal{D}(\mathbb{R})$ generalizing function differentiation.

Example 5.2.1. Consider the Dirac delta distribution:

$$\delta \in \mathcal{D}(\mathbb{R}), \delta(\varphi) = \varphi(0).$$

This is a distribution that is not a function. As an exercise, check that we can produce δ via the following chain:

$$f(x) = \begin{cases} 0 & x < 0 \\ x^2/2 & x \geq 0 \end{cases}$$
$$f'(x) = \begin{cases} 0 & x < 0 \\ x & \geq 0 \end{cases}$$
$$f''(x) = \begin{cases} 0 & x < 0 \\ 1 & x > 0 \end{cases}$$
$$f'''(x) = \delta$$

5.3 Power Series

Recall that a polynomial is a formal expression $a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ for $a_i \in \mathbb{R}$. It provides a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ for $c \in \mathbb{R}$ given by evaluation at c.

Polynomials are advantageous in that they are nice to work with—they are smooth and we can compute their derivatives easily. But many useful functions (e.g., trig) do not come from polynomials.

But many such functions are "almost polynomials."

Definition 5.3.1. A **power series** is a formal expression of the form $a_0 + a_1 x + a_2 x^2 + \cdots = \sum a_n x^n$.

Remark. The info in a power series is the sequence (a_n) of coefficients. *A priori*, we don't get a function.

Lemma 5.3.1. Let $\sum a_n x^n$ be a power series.

1. There exists a unique $R \in \mathbb{R}_{\geq 0} \cup \{+\infty\}$ such that for all $c \in \mathbb{R}$:

$$|c| < R \Longrightarrow \sum a_n c^n \text{ converges absolutely}$$

$$|c| > R \Longrightarrow \sum a_n c^n \text{ diverges.}$$

2. $R = (\limsup \sqrt[n]{|a_n|})^{-1} \in \mathbb{R}_{\geq 0} \cup \{+\infty\}$.

Proof. Uniqueness is immediate. Existence and 2) follow from the Root Test:

$$\limsup \sqrt[n]{|a_n c^n|} = \limsup |c| \cdot \sqrt[n]{|a_n|} = |c|/R.$$

□

Definition 5.3.2. The above R is called the **radius of convergence**.

If $R > 0$, we get a function $f : (-R, R) \longrightarrow \mathbb{R}$ given by $c \mapsto \sum_{n=0}^{\infty} a_n c^n$. If $R = 0$, we don't get a function. $R = 0$ can in fact happen: consider $\sum n^n \cdot x^n$.

We can think of $\sum a_n x^n$ as a series of functions. By definition, $\sum a_n x^n \longrightarrow f$ pointwise on $(-R, R)$.

Example 5.3.1. The convergence for $|c| = R$ is subtle. $\sum x^n$ converges on $(-1, 1)$, $\sum \frac{1}{n} x^n$ converges on $[-1, 1)$, and $\frac{1}{n^2} x^n$ converges on $[-1, 1]$. All have $R = 1$.

Lemma 5.3.2. For any $0 < r < R$,

$$\sum a_n x^n \longrightarrow f \text{ uniformly on } [-r, r].$$

Proof. By the previous lemma,

$$\sum |a_n r^n| \text{ converges } \iff \sum \|a_n x^n|_{[-r,r]}\|_\infty.$$

Then Weierstrass M-test implies that $\sum a_n x^n$ converges uniformly on $[-r, r]$. $\square$

Example 5.3.2. Convergence on $(-R, R)$ need not be uniform. If $\sum x^n$ converged uniformly on $(-1, 1)$, then by Silly Test, $1 = \|x^n|_{(-1,1)}\|_\infty \longrightarrow 0$, contradiction.

Proposition 5.3.1. *Let $\sum a_n x^n$ be a power series with radius of convergence $R > 0$. Then $f : (-R, R) \longrightarrow \mathbb{R}$ is differentiable and f' also comes from a power series, namely $\sum n a_n x^{n-1}$ whose radius of convergence is still R.*

Proof. The domain of convergence of $\sum n a_n x^{n-1}$ is the same as for $x \cdot \sum n a_n x^{n-1} = \sum n a_n x^n$. Its radius of convergence is

$$(\limsup \sqrt[n]{|n a_n|})^{-1} = (\limsup \sqrt[n]{n})^{-1} \cdot (\limsup \sqrt[n]{|a_n|})^{-1} = 1 \cdot R = R.$$

Take $0 < r < R$. We know that $\sum n a_n x^{n-1} \longrightarrow g$ uniformly on $(-r, r)$ and $\sum a_n x^n \longrightarrow f$ on $(-r, r)$. By a previous result, f is differentiable and $f' = g$ on $(-r, r)$. Since $0 < r < R$ is arbitrarily, the same holds on $(-R, R)$. $\square$

So functions coming from power series behave as nicely as functions coming from polynomials—we can differentiate them freely while preserving the radius of convergence.

Corollary 5.3.1. *If f is associated to $\sum a_n x^n$, then f is smooth and $a_n = \frac{f^{(n)}(0)}{n!}$.*

Proof. Induce on n. If $n = 0$, done. For $n + 1$, we have $f^{(n+1)}(0) = (f)'^{(n)}(0) = n! \cdot b_n$, where

$$\sum b_n x^n = f' = n!(n+1)a_{n+1} = (n+1)! \cdot a_{n+1}.$$

$\square$

Now we reverse directions.

Question. Given a smooth $f : I \longrightarrow \mathbb{R}$, where I is an open interval around 0, does f come from a power series?

Corollary 5.3.1 shows that the only power series f could come from is

$$\sum \frac{f^{(n)}(0)}{n!} x^n.$$

So we give it a name:

Definition 5.3.3.

1. The power series $\sum \frac{f^{(n)}(0)}{n!}x^n$ is called the **Taylor series of** f.
2. $$\sum_{n=0}^{N} \frac{f^{(n)}(0)}{n!}x^n$$
   is called the **Taylor polynomial of degree** N of f.
3. $$R_N(x) = f(x) - \sum_{n=0}^{N-1} \frac{f^{(n)}(0)}{n!}x^n$$
   is called the N**-th Taylor remainder**.

The answer to our question is as follows:

$$\begin{aligned} &f \text{ comes from a power series} \\ \iff &f \text{ is represented by its Taylor series} \\ \iff &R_N \longrightarrow 0 \text{ on } I \subset (-R, R) \end{aligned}$$

Example 5.3.3. The following is a smooth function not represented by its Taylor series.

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0. \end{cases}$$

We claim f is smooth. If $x \neq 0$,

$$f'(x) = e^{-1/x^2} \cdot \frac{1}{x^3} \longrightarrow 0$$

$$f''(x) = e^{-1/x^2} \cdot \left(\frac{1}{x^6} - \frac{1}{x^4}\right) \longrightarrow 0.$$

This goes forever, so the Taylor series is 0. The Taylor remainder carries the whole value of the function.

$R = 0$ can happen; take

$$f(x) = \sum e^{-n} \cdot \cos(n^2 x).$$

Theorem 5.3.1 (Taylor's theorem). *For any $a \in (-R, R)$, there exists b between 0 and a such that:*

$$R_N(a) = \frac{f^{(N)}(b)}{N!} \cdot a^N.$$

Proof. Let M be the unique real number such that

$$f(a) = \sum_{n=0}^{N-1} \frac{f^{(n)}(0)}{n!}a^n + \frac{M}{N!}a^N.$$

It is enough to show $M = f^{(N)}(b)$ for some b. Define

$$g(x) = \sum_{n=0}^{N-1} \frac{f^{(n)}(0)}{n!}x^n + \frac{M}{N!}x^N - f(x).$$

Now $g(0) = g(a) = 0$, so by MVT, there is b_1 between 0 and a such that $g'(b_1) = 0$.

So $g'(0) = g'(b_1) = 0$, so we can find b_2 between 0 and b_1 such that $g''(b_2) = 0$. Repeat this process until we have $g^{(N-1)}(0) = g^{(N-1)}(b_{N-1}) = 0$, which gives b_N between 0 and b_{N-1} such that

$$0 = g^{(N)}(b_N) = M - f^{(N)}(b_N).$$

□

Corollary 5.3.2. *If $I \subset (-R, R)$ and there is $c > 0$ such that $\|f^{(c)}\|_\infty < c$ for all a, then* $\sum \frac{f^{(n)}(0)}{n!} x^n \longrightarrow f$.

Proof. By the previous result, we have $R_N(x) \leq c \cdot \frac{a^N}{N!} \longrightarrow 0$. □

Proposition 5.3.2 (Integral form of remainder)**.**

$$R_N(a) = \int_0^a \frac{(a-t)^{N-1}}{(N-1)!} f^{(N)}(t) dt.$$

Proof. We induce on N. For $N = 1$, we have

$$R_1(a) = f(a) - f(0) = \int_0^a f'(t) dt$$

by FTC. For $N \longrightarrow N + 1$, we apply the induction hypothesis

$$R_N(a) = \int_0^a \frac{(a-1)^{N-1}}{(N-1)!} f^{(N)}(t) dt$$

and integrate by parts. Let $u'(t) = \frac{(a-1)^{N-1}}{(N-1)!}$ and $v(t) = f^{(N)}(t)$; then

$$u(a)v(a) - u(0)v(0) - \int_0^a u(t)v'(t) dt$$

$$= \frac{a^N}{N!} f^{(N)}(0) + \int_0^a \frac{(a-t)^N}{N!} f^{(N+1)}(t) dt.$$

Thus we write

$$R_{N+1}(a) = R_N(a) - \frac{a^N}{N!} f^{(N)}(0) = \int_0^a \frac{(a-t)^N}{N!} f^{(N+1)}(t) dt.$$

□

Corollary 5.3.3 (Cauchy remainder)**.** *For $a \in (-R, R)$, there is b between $0, a$ such that*

$$R_N(a) = \frac{(a-b)^{N-1}}{(N-1)!} f^{(N)}(b) \cdot a.$$

Proof. Firstly, note that for any continuous $g : [z, w] \longrightarrow \mathbb{R}$, there is $s \in [z, w]$ such that $\int_z^w g(t) dt = (w - z) g(s)$. We can see this by considering the maximum and minimum values of g, denoted m, M respectively. Then write

$$m(w - z) \leq \int_z^w g(t) dt \leq M(w - z)$$

and apply IVT. Now

$$R_N(a) = \int_0^a \frac{(a-t)^{N-1}}{(N-1)!} f^{(N)}(t)dt = a \cdot \frac{(a-b)^{N-1}}{(N-1)!} f^{(N)}(b).$$

□

What if $f : I \longrightarrow \mathbb{R}$ is smooth but $0 \neq I$? Just shift everything by c for some $c \in I$.

Definition 5.3.4. Let $I \subset \mathbb{R}$ be an open interval and $c \in I$ and $f : I \longrightarrow \mathbb{R}$ be a smooth function.

0. A **power series centered at** c is $\sum a_u(x-c)^u$.
1. The **Taylor series of** f **centered at** c is $\sum \frac{f^{(u)}(c)}{u!}(x-c)^u$.
2. The **Taylor polynomial of** f **at** c is $\sum_{u=0}^N \frac{f^{(u)}(c)}{u!}(x-c)^u$.
3. The **Taylor remainder of** f **at** c **of degree** N is $R_n^c(x) = f(x) - \sum_{u=0}^{N-1} \frac{f^{(u)}(c)}{u!}(x-c)^u$.

Theorem 5.3.2.

1. *If* $f(x) = \sum_{u=0}^\infty a_u(x-c)^u$*, then* $a_n = \frac{f^{(u)}(c)}{u!}$
2. *For any* $a \in I$*, there is* b *between* c, a *such that* $R_N^c(a) = \frac{f^{(N)}(b)}{N!}(a-c)^N$ *(Lagrange form)*
3. *For any* $a \in I$, $R_N^c(a) = \int_c^a \frac{(a-t)^{N-1}}{(N-1)!} f^{(N)}(t)dt$ *(integral form)*
4. *For any* $a \in I$ *there is* b *between* c, a *such that* $R_N^c(a) = \frac{f^{(N)}(b)}{(N-1)!}(a-b)^{N-1}(a-c)$ *(Cauchy form)*

Proof. Apply the previous results to $g(x) = f(x-c)$. □

Definition 5.3.5. Let $I \subset \mathbb{R}$ be an open interval, $f : I \longrightarrow \mathbb{R}$ smooth. We call f **analytic at** $a \in I$ if there is $\epsilon > 0$ such that $f|_{(a-\epsilon, a+\epsilon)}$ is represented by its Taylor series at a. We call f **analytic** if it is analytic at all $a \in I$.

Proposition 5.3.3. *Let* $\sum a_n x^n$ *be a power series with radius of convergence* $R > 0$ *and let* $f : (-R, R) \longrightarrow \mathbb{R}$ *be the associated function. Then* f *is analytic.*

Proof. Let $a \in (-R, R)$. Then

$$\sum_{u=0}^\infty a_u x^u = \sum_{u=0}^\infty a_u(x-a+a)^u = \sum_{u=0}^\infty a_u \sum_{k=0}^u \binom{u}{k}(x-a)^k a^{u-k}.$$

The sum runs over $\{(u,k) \mid k \leq u\}$. If we could interchange summands, we would have

$$\sum_{k=0}^\infty \left(\sum_{u=k}^\infty \binom{u}{k} a^{u-k} a_u \right) \cdot (x-a)^k = \sum_{k=0}^\infty b_k \cdot (x-a)^k.$$

We first claim that

$$\sum_{u=0}^{\infty} \sum_{k=0}^{u} |a_u| \binom{u}{k} |x-a|^k |a|^{u-k}$$

converges. Indeed, this is just $\sum_{k=0}^{u} |a_n|(|x-a|+|a|)^u$ and converges for $|x-a|+|a| < R$.

Moreover, we claim that if s_{ij} is a double sequence and $t_i = \sum_{j=0}^{\infty} |s_{ij}|$ converges and $\sum_{i=0}^{\infty} t_i$ converges, then

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} s_{ij} = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} s_{ij}.$$

Define $D = \{1/u \mid u \in \mathbb{N}\} \cup \{0\}$ and a sequence $f_i : D \longrightarrow \mathbb{R}$ by

$$f_i\left(\frac{1}{u}\right) = \sum_{j=0}^{u} s_{ij} \qquad f_k(0) = \sum_{j=0}^{\infty} s_{ij}.$$

Then f_i is continuous. Since $|f_i(x)| \leq t_i$ and $\sum t_i < \infty$, Weierstrass-M test shows $\sum f_i$ converges uniform to a function g which is then continuous. Now

$$\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} s_{ij} = \sum_{i=0}^{\infty} f_i(0) = g(0) = \lim_{u\to\infty} g\left(\frac{1}{u}\right)$$

$$= \lim_{u\to\infty} \sum_{i=0}^{\infty} f_i\left(\frac{1}{u}\right) = \lim_{u\to\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{u} s_{ij}$$

$$= \lim_{u\to\infty} \sum_{j=0}^{u} \sum_{i=0}^{\infty} s_{ij} = \sum_{j=0}^{\infty} \sum_{i=0}^{\infty} s_{ij}.$$

□

5.4 Fourier Series

Consider a function $f : [-\pi, \pi] \longrightarrow \mathbb{R}$, which we can think of as a function $f : \mathbb{R} \longrightarrow \mathbb{R}$ that is periodic with period 2π. Assume f is continuous (integrable is enough).

Definition 5.4.1. The series

$$\sum_{n=-\infty}^{\infty} a_n \cos(nx) + b_n \sin(nx)$$

is the **Fourier series** of f, where

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx$$

$$b_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx.$$

Does the Fourier series recover f?

Recall the following properties of sin, cos:

- sin, cos are smooth, $\sin' = \cos$, $\cos' = -\sin$

- $\sin(-x) = -\sin(x)$, $\cos(-x) = \cos(x)$
- $\sin(0) = 0$, $\cos(0) = 1$
- $\sin(x)^2 + \cos(x)^2 = 1$
- $|\sin(x)| \leq 1$, $|\cos(x)| \leq 1$
- $\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y)$
- $\cos(x)\sin(y) = \frac{1}{2}(\sin(x + y) - \sin(x - y))$

Lemma 5.4.1 (Riemann-Lebesgue)**.** If $f : [a, b] \longrightarrow \mathbb{R}$ is integrable, then as $u \longrightarrow \infty$,

$$\int_a^b f(x)\sin(ux)dx \longrightarrow 0$$

$$\int_a^b f(x)\cos(ux)dx \longrightarrow 0.$$

Proof. Let $\epsilon > 0$. Choose step functions $h_1 \leq f \leq h_2$ with $\int_a^b h_2(x) - h_1(x)dx < \epsilon$.

Let $P : a = a_0 \leq a_1 \leq \cdots \leq a_k = b$ be a partition to which h_1, h_2 are adapted. Then

$$\left|\int_a^b f(x)\cos(ux)dx\right| \leq \left|\int_a^b h_1(x)\cos(ux)dx\right| + \left|\int_a^b (f(x) - h_1(x))\cos(ux)dx\right|.$$

Now note that

$$\left|\int_a^b (f(x) - h_1(x))\cos(ux)dx\right| \leq \int_a^b |f(x) - h_1(x)|dx \leq \int_a^b (h_2(x) - h_1(x))dx < \epsilon$$

and write

$$\left|\int_a^b h_1(x)\cos(ux)dx\right| = \left|\sum_{i=1}^k \int_{a_{i-1}}^{a_i} h_1(x)\cos(ux)dx\right| = \left|\sum_{i=1}^k h_1\Big(\frac{a_{i-1} + a_i}{2}\Big)\int_{a_{i-1}}^{a_i} \cos(ux)dx\right| \longrightarrow 0.$$

□

Definition 5.4.2. The N**-th partial Fourier sum** for f is

$$S_N, f(x) = \sum_{n=-N}^{N} a_n \cos(nx) + b_n \sin(nx).$$

Remark. $a_n = \int f(x)\cos(nx)dx \Longrightarrow a_{-n} = a_n$, and analogously for b_n. So

$$S_N, f(x) = a_0 + 2\sum_{n=1}^{N} a_n \cos(nx) + b_n \sin(nx).$$

Definition 5.4.3. The N**-th Dirichlet kernel** is

$$D_N(x) = \frac{\sin((N + \frac{1}{2})x)}{\sin(\frac{1}{2}x)} \qquad x \neq 0.$$

Lemma 5.4.2. $D_N(x)$ extends continuously to 0 with $D_N(0) = 2N + 1$.

Proof. L'Hopital. □

Exercise: graph $D_N(x)$ and see that it approximates the Dirac delta function.

Proposition 5.4.1.

$$D_N(x) = 1 + 2\sum_{n=1}^{N} \cos(nx).$$

Proof. We have

$$\begin{aligned}\sin(x/2)\left(1 + 2\sum_{n=1}^{N}\cos(nx)\right) &= \sin(x/2) + 2\sum_{n=1}^{N}\sin(x/2)\cos(nx)\\ &= \sin(x/2) + \sum_{n=1}^{N}\left(\sin\left(n+\frac{1}{2}\right)x - \sin\left(\left(n-\frac{1}{2}\right)x\right)\right)\\ &= \sin\left(\left(N+\frac{1}{2}\right)x\right).\end{aligned}$$

□

Corollary 5.4.1.

$$S_N, f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)D_N(x-t)dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+t)D_N(t)dt.$$

Proof. The RHS equals

$$\begin{aligned}&\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\left(1 + 2\sum_{n=1}^{N}\cos(nx-nt)\right)dt\\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt + 2\sum_{n=1}^{N}\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)(\cos(nx)\cos(nt) + \sin(nx) + \sin(nt))dt\\ &= a_0 + 2\sum_{n=1}^{N} a_n\cos(nx) + b_n\sin(nx),\end{aligned}$$

which is exactly the LHS. □

Corollary 5.4.2.

$$\int_{-\pi}^{0} D_N(x)dx = \pi = \int_{0}^{\pi} D_N(x)dx.$$

Proof. It is enough to show that

$$\int_{-\pi}^{0}\cos(nx)dx = 0 = \int_{0}^{\pi}\cos(nx)dx.$$

Use FTC. □

Definition 5.4.4. A function $f : [-\pi, \pi] \longrightarrow \mathbb{R}$ is **piecewise differentiable** if it is differentiable at all $x \in [-\pi, \pi]$ with possibly finitely many exceptions $a_1, \dots, a_k \in [-\pi, \pi]$, where the following limits exist (denote $a = a_i$):

$$\begin{aligned} f(a_+) &:= \lim_{x \downarrow a} f(x) \\ f(a_-) &:= \lim_{x \uparrow a} f(x) \\ f'(a_+) &:= \lim_{x \downarrow a} \frac{f(x) - f(a_+)}{x - a} \\ f'(a_-) &:= \lim_{x \uparrow a} \frac{f(x) - f(a_-)}{x - a} \end{aligned}$$

Theorem 5.4.1. *Let $f : [-\pi, \pi] \longrightarrow \mathbb{R}$ be piecewise differentiable. Then*

$$S_N, f(x) \longrightarrow \frac{f(x_+) + f(x_-)}{2},$$

which equals $f(x)$ iff f is continuous at x.

Proof. We have shown already that

$$S_N, f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+t) D_N(t) dt.$$

It is enough to show

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{0} f(x+t) D_N(t) dt &\longrightarrow \frac{f(x_-)}{2} \\ \frac{1}{2\pi} \int_{0}^{\pi} f(x+t) D_N(t) dt &\longrightarrow \frac{f(x_+)}{2}. \end{aligned}$$

To prove this, we claim that

$$\frac{1}{2\pi} \int_{0}^{\pi} f(x+t) D_N(t) dt - \frac{f(x_+)}{2} \longrightarrow 0.$$

This is equivalent to

$$\begin{aligned} &\frac{1}{2\pi} \int_{0}^{\pi} (f(x+t) - f(x_+)) D_N(t) dt \\ = &\frac{1}{2\pi} \int_{0}^{\pi} \frac{f(x+t) - f(x_+)}{\sin(\frac{1}{2}t)} \sin\left(\left(N + \frac{1}{2}\right)t\right) dt. \end{aligned}$$

Claim. $g(t)$ is piecewise continuous.

Admitting this, g is integrable, so by Riemann-Lebesgue, the above expression approaches 0. Then we are done.

Proof of claim. By definition, g is piecewise differentiable at all $t \neq 0$. It suffices to check that the following limit exists:

$$\lim_{t \downarrow 0} g(t) = \lim_{t \downarrow 0} \frac{f(x+t) - f(x_+)}{t} \cdot \frac{t}{\sin(\frac{1}{2}t)}.$$

The left term's limit exists since f is piecewise differentiable. The right term's limit exists by L'Hopital. □

$\square$

Theorem 5.4.2.
$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

Proof. Consider $f(x) = |x|$, which is piecewise differentiable. Then

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)dx = \frac{1}{2\pi} \cdot \frac{2\pi^2}{2} = \frac{\pi}{2}$$

$$\begin{aligned}
a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} |x| \cos(nx)dx &= \frac{1}{2\pi} \int_{-\pi}^{0} (-x) \cos(nx)dx + \frac{1}{2\pi} \int_{0}^{\pi} x \cos(nx)dx \\
= \frac{1}{\pi} \int_{0}^{\pi} x \cos(nx)dx &= \frac{1}{\pi} \left(\left[\frac{x \cdot \sin(nx)}{n} \right]_0^{\pi} - \int_0^{\pi} \frac{\sin(nx)}{n} dx \right) \\
&= \frac{1}{\pi} \left(0 + \left[\frac{\cos(nx)}{n^2} \right]_0^{\pi}) \right) \\
&= \frac{1}{\pi \cdot n^2} (\cos(n\pi) - 1),
\end{aligned}$$

which is 0 when n is even and $-\frac{2}{n^2\pi}$ when n is odd. Also,

$$b_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} |x| \cdot \sin(nx)dx = 0.$$

Therefore,

$$\begin{aligned}
0 = f(0) &= \lim_{N \to \infty} S_N, f(0) \\
& a_0 + 2 \sum_{n=1}^{\infty} a_n \\
&= \frac{\pi}{2} - \sum_{n>0} \frac{4}{n^2\pi} \\
&\Rightarrow \sum_{n>0} \frac{1}{n^2} = \frac{\pi^2}{8}
\end{aligned}$$

if n is odd. If n is even, we have the following trick:

$$\begin{aligned}
\frac{3}{4} \sum_{n>0} \frac{1}{n^2} &= \sum_{n>0} \frac{1}{n^2} - \frac{1}{4} \sum_{n>0} \frac{1}{n^2} \\
& \sum_{n>0} \frac{1}{n^2} - \sum_{n>0} \frac{1}{4n^2} \\
&= \sum_{n>0 \text{ for odd } n} \frac{1}{n^2} \\
\sum_{n>0} \frac{1}{n^2} &= \frac{4}{3} \cdot \frac{\pi^2}{8} = \frac{\pi^2}{6}.
\end{aligned}$$

$\square$

6 σ-algebras*

6.1 The Basics

Given a set X and $A \subset X$, denote $A^c = X \backslash A$.

Definition 6.1.1. A **σ-algebra** $\mathcal{A}$ on a set X is a family of subsets of X with the following properties:

1. $X \in \mathcal{A}$
2. $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
3. $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A} \implies \bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$

A set $A \in \mathcal{A}$ is said to be **measurable** or **$\mathcal{A}$-measurable**.

The third requirement says that the union of countably many subsets of $\mathcal{A}$ must also be a subset of $\mathcal{A}$.

Theorem 6.1.1. *Consider a σ-algebra $\mathcal{A}$. Then the following properties hold:*

- $\emptyset \in \mathcal{A}$
- $A, B \in \mathcal{A} \implies A \cup B \in \mathcal{A}$
- $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A} \implies \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

Proof.

- Since $X \in \mathcal{A}$, we write $\emptyset = X^c \in \mathcal{A}$ by properties 1 and 2.
- Follows directly from property 3.
- By property 2, we write $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A} \implies (A_n^c)_{n \in \mathbb{N}} \subset \mathcal{A}$. Then we use DeMorgan to obtain

$$\left(\bigcap_{n \in \mathbb{N}} A_n \right)^c = \bigcup_{n \in \mathbb{N}} A_n^c \in \mathcal{A} \implies \bigcap_{n \in \mathbb{N}} A_n \in \mathcal{A}.$$

□

Example 6.1.1. Denote the cardinality of a set A by $\#A$. Show that the following set is a σ-algebra:

$$\mathcal{A} := \{A \subset X : \#A \leq \mathbb{N} \text{ or } \#A^c \leq \mathbb{N}\}.$$

That is, $\mathcal{A}$ is the set of countable subsets of X and their complements.

Proof. We show that $\mathcal{A}$ satisfies the three properties of a σ-algebra.

1. $X^c = \emptyset$, which is countable. Hence $X \in \mathcal{A}$.
2. If $A \in \mathcal{A}$, then $A^c \in \mathcal{A}$ because $(A^c)^c = A$.

3. Fix a set of (A_n) and suppose all are countable. Then $\bigcup_{n\in\mathbb{N}} A_n$ is the countable union of countable sets; hence, is is countable.

   Now suppose some $A_i \in \mathcal{A}$ is uncountable. Then A_i^c must be countable, so we write

$$\left(\bigcup_{n\in\mathbb{N}} A_n\right)^c = \bigcap_{n\in\mathbb{N}} A_n^c \subset A_i^c.$$

   So the leftmost expression is countable, and thus its complement is in $\mathcal{A}$.

□

Theorem 6.1.2 (Existence of generators)**.** *For every system of sets $\mathcal{G} \subset \mathcal{P}(X)$ there exists a smallest σ-algebra containing $\mathcal{G}$.*

Proof. Consider the union of all σ-algebras containing $\mathcal{G}$:

$$\mathcal{A} := \bigcap_{\mathcal{F}\supset\mathcal{G}} \mathcal{F}, \quad \text{where } \mathcal{F} \text{ is a } \sigma\text{-algebra.}$$

We claim that $\mathcal{A}$ is the minimal family in question. Using Definition 3.1.1, it is easy to check that the intersection of arbitrarily many σ-algebras is itself a σ-algebra.

But, by definition, if $\mathcal{G} \subset \mathcal{A}'$ for a σ-algebra $\mathcal{A}'$, then $\mathcal{A} \subset \mathcal{A}'$, so $|\mathcal{A}| \leq |\mathcal{A}'|$. So $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{G}$. □

6.2 Borel σ-algebras