Math 115

Functions of a Real Variable

Jia Xiaodong

May 20, 2020

1 Introduction

This is an introductory course in real analysis. The goal of this class is to prove the Fundamental Theorem of Calculus. A worthy and noble cause. Let us begin.

2 Real numbers

The real numbers ($\mathbb{R}$) are difficult to construct, hence we take a synthetic approach from axioms. We say that:

Definition 2.1. $\mathbb{R}$ is a *complete ordered field.* $\square$

We shall explore all three of those terms.

2.1 Fields

A *field* is a set of numbers F that possesses two binary operations (implies closure) $+, \cdot$, and two special constants, $0, 1$.

The field axioms:

- (NT) Non-triviality: $0 \neq 1$.
- (A1) Associativity: $\forall x, y, z \in F, (x + y) + z = x + (y + z) = x + y + x$
- (A2) Commutativity: $\forall x, y \in F, x + y = y + x$
- (A3) Additive Identity: $\forall x \in F, x + 0 = x$
- (A4) Additive Inverse: $\forall x \in F, \exists y \in F, x + y = 0$

- (M1) $\forall x, y, z \in F, (x \cdot y) \cdot z = x \cdot (y \cdot z) = x \cdot y \cdot x$
- (M2) $\forall x, y \in F, x \cdot y = y \cdot x$
- (M3) $\forall x \in F, x \cdot 1 = x$
- (M4) $\forall x \in F, x \neq 0 \implies \exists y \in F, x \cdot y = 1$
- (DL) $\forall x, y, z \in F, x \cdot (y + z) = x \cdot y + x \cdot z$

Henceforth we also drop the dot i.e. $x \cdot y = xy$ for convenience's sake.

2.2 Ordered Fields

An *ordered field* is a field with an *order*. An order is a relation $\leq$.

- (O1) $\forall x, y \in F, x \leq y \vee y \leq x$
- (O2) Anti-symmetry: $\forall x, y \in F, (x \leq y \wedge y \leq x) \implies x = y$
- (O3) Transitivity: $\forall x, y, z \in F, (x \leq y \wedge y \leq z) \implies x \leq z$
- (O4) $\forall x, y, z \in F, x \leq y \implies x + z \leq y + z$
- (O5) $\forall x, y, z \in F, (x \leq y \wedge 0 \leq z) \implies xz \leq yz$

Afterwards we might use the symbol $\geq$ as well for convenience: $x \leq y \implies y \geq x$.

2.3 Consequences of the Field Axioms

A field has an unique additive and multiplicative identity. If $x \in F$, the additive inverse of x is unique. If $x \in F \setminus \{0\}$, then the multiplicative inverse of x is unique.

Proposition 2.1. *Suppose $z \in F$ where F is a field. If it satisfies*

$$\forall x \in F, x + z = x$$

then $z = 0$.

Proof.

$$\begin{aligned} z &= z + 0 && \text{(by A3)} \\ &= 0 + z && \text{(by A2)} \\ &= 0 && \text{(by construction)} \end{aligned}$$

■

Similarly, let $x \in F$, suppose $x + y = x + y' = 0$. Then,

$$\begin{aligned} y' &= y' + 0 \\ &= y' + (x + y) \\ &= (y' + x) + y \\ &= (x + y') + y \\ &= 0 + y \\ &= y \end{aligned}$$

Hence now we will write the unique additive inverse for x as $-x$, and x^{-1} will serve as the unique multiplicative inverse [1].

[1] The proof for the uniqueness of the multiplicative inverse is left out as homework

Theorem 2.1. *Let F be a field. Then $\forall a, b, c \in F$:*

i. $a + c = b + c \implies a = b$

ii. $a \cdot 0 = 0$

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

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

v. $(-a) \cdot (-b) = ab$

vi. $(ac = bc) \wedge c \neq 0 \implies a = b$

vii. $ab = 0 \implies (a = 0) \vee (b = 0)$

We will prove a few of them:

Proof.

i. Suppose $a + c = b + c$.

$$\begin{gathered} (a + c) + (-c) = (b + c) + (-c) \\ \vdots \\ a = b \end{gathered}$$

ii. Take $a \cdot 0 = a(0 + 0) = a \cdot 0 + a \cdot 0$. Then it follows that $a \cdot 0 = 0$.

iii.

$$\begin{aligned} ab + (-a)b &= (a + (-a))b \\ &= 0 \cdot b \\ &= 0 \end{aligned}$$

Hence $(-a)b$ serves as the additive inverse for ab.

vi. Suppose $a \neq 0 \wedge b \neq 0$. Consider ab :

$$a^{-1}(ab) = (a^{-1}a)b = 1 \cdot b = b \neq 0$$

We know a^{-1} is not 0. ab cannot be 0 either.

∎

Theorem 2.2. *Let F be a field. $\forall a, b, c \in F$:*

i. $a \leq b \implies -b \leq -a$

ii. $(a \leq b) \wedge (c \leq 0) \implies bc \leq ac$

iii. $(a \geq 0) \wedge (b \geq 0) \implies ab \geq 0$

iv. $a^2 \geq 0$

v. $0 < 1$

Proof.

i.

$$\begin{aligned} a &\leq b \\ a + [(-a) + (-b)] &\leq b + [(-a) + (-b)] \\ &\vdots \\ -b &\leq -a \end{aligned}$$

■

2.4 Absolute Value and Distance in $\mathbb{R}$

Let $x \in \mathbb{R}$. We say

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

and $\text{dist}(x, y) = |x - y|$.

Theorem 2.3. *Let $a, b, c \in \mathbb{R}$.*

i. $|a| \geq 0$

ii. $|ab| = |a| \cdot |b|$

iii. Triangle inequality: $|a + b| \leq |a| + |b|$

Proof.

iii. We know that:

$$\begin{aligned} -|a| \leq\, &a \leq |a| \\ -|b| \leq\, &b \leq |b| \\ &\vdots \\ -(|a| + |b|) \leq a + &b \leq |a| + |b| \end{aligned} \qquad \begin{aligned} -|a| &\leq -a \leq |a| \\ -|b| &\leq -b \leq |b| \\ &\vdots \\ -(|a| + |b|) &\leq -(a + b) \leq |a| + |b| \end{aligned}$$

Therefore $|a + b| \leq |a| + |b|$

■

Corollary 2.3.1. $\forall x, y, z \in \mathbb{R}$, $\text{dist}(x, z) \leq \text{dist}(x, y) + \text{dist}(y, z)$.

Proof.

$$\begin{aligned} |x - z| &= |x - y + y - z| \\ &= |x - y| + |y - z| \end{aligned}$$

■

2.5 The Completeness Axiom

We finally get to the last term that defines $\mathbb{R}$.

Let $S \in R, S \neq \emptyset$.

- S has a maximum if $\exists M \in S, \forall x \in S, x \leq m$.
- S has a minimum if $\exists M \in S, \forall x \in S, m \leq x$.

Let S be a non-empty subset of R.

- $B \in R$ is an upper bound on S if $\forall x \in S, x \leq B$.
- The *best* upper bound is the minimum of the set of all upper bounds. We call such a least upper bound the *supremum* of S ($\sup S$).

Example 2.1. $\sup(\{0, 1\}) = 1$.

- 1 is an upper bound by inspection.
- Let $x < 1$. Then x cannot bound 1. Hence 1 is also the *least* upper bound.

□

Definition 2.2 (The Completeness Axiom)**.** Let S be a non-empty subset of $\mathbb{R}$. If S is bounded above then S has a least upper bound. □

Remark. $\mathbb{Q}$ is not complete.

Take for example: $S = \{r \in \mathbb{Q} \mid r^2 \leq 2\} = \{r \in \mathbb{Q} \mid 0 \leq r \leq \sqrt{2}\}$. $\sup S = \sqrt{2} \neq \mathbb{Q}$.

2.6 Consequences of completeness

A brief aside to recap our construction. Let us check this fact: $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$.

We take $\mathbb{Q} = \{\frac{a}{b} \mid a, b, \in \mathbb{Z} \wedge b \neq 0\}$, where $\frac{a}{b} = \frac{c}{d} \iff ad = bc$. Then $\mathbb{Z}$ is in $\mathbb{Q}$ since we can say $\mathbb{Z} = \{\frac{a}{1} \mid a \in \mathbb{Z}\}$.

$0, 1 \in \mathbb{R}$, and also $1 + 1 \in \mathbb{R}$, $1 + 1 + 1 \in \mathbb{R}$, and so on and $1 + 1 + \ldots \neq 0$ since $\mathbb{R}$ is ordered. Then $\mathbb{N} \subseteq \mathbb{R}$, and $-\mathbb{N} \subseteq \mathbb{R}$ as well. Then $\mathbb{Z} \subseteq \mathbb{R}$, and we can see now that $\mathbb{Q} \in \mathbb{R}$ as well. It is interesting to note that $\mathbb{Q}$ is the smallest *field* such that $0 \neq 1 \neq 1 + 1 \neq \ldots$.

***

Any non-empty set bounded below also has a greatest lower bound. We call it the *infimum* of S ($\inf S$).

Proof. If S is non-empty and bounded below, then $-S$ is non-empty and bounded above. Hence $\inf(S) = -\sup(-S)$. ■

Some brief notes on convention:

- If S is not bounded above then we write $\sup S = \infty$.
- If S is empty then all $M \in \mathbb{R}$ are upper bounds for it, since the statement becomes vacuously true. We say $\sup(\emptyset) = -\infty$ and $\inf(\emptyset) = +\infty$.
- If S is not bounded below then $\inf S = -\infty$.

***

Theorem 2.4 (Archimedean property). *If $a, b \in \mathbb{R}^+$, then $\exists n \in \mathbb{N}, na > b$.*

Proof. Suppose not. Then $\exists a, b > 0, \forall n \in \mathbb{N}, na \leq b$.

Let $S = \{na \mid n \in \mathbb{N}\}$. S is obviously not empty. It is also bounded from above by b. Let $s^* \in \mathbb{R}$ be $\sup S$ by completeness. s^* is an upper bound on S, so $\forall \varepsilon > 0, s^* - \varepsilon$ in not an upper bound. In particular, $s^* - a$ is not an upper bound. Then $\exists y \in S, y > s^* - a$. Then $y + a > s^*$. This means that we have $y + a \in S$ but it exceeds the upper bound of S, contradiction. ■

Since $\mathbb{Q} \in \mathbb{R}$, $\mathbb{Q}$ also possesses the Archimedean property. However, the proof for AP in $\mathbb{Q}$ does not require completeness.

Claim. *If $S \subseteq \mathbb{N} (\subseteq \mathbb{R})$ and $S \neq \emptyset$, then S has a smallest element.*

Proof. Take the contrapositive: if S has no smallest element, then $S \neq \emptyset$. We know $0 \notin S$, since $\forall n \in \mathbb{N}, n \geq 0$, then 0 would be $\min S$. Suppose then $0 \notin S \wedge \ldots \wedge n \notin S$, then $n + 1 \notin S$ since otherwise that would become the minimum element. Thus S is empty. ■

Similarly, if $S \subseteq \mathbb{Z}$ and S is bounded below, then S has a smallest element.

Lemma 2.5. *Suppose $a, b \in \mathbb{R}, a < b$. If $b - a > 1$, then $\exists m \in \mathbb{Z}$, $a < m < b$.*

Proof. Let $S = \{k \in \mathbb{Z} \mid k > a\}$. S is bounded below by a. S is not empty: if $a \leq 0$ then $0 \in S$. If $a > 0$, then $\exists k \in \mathbb{N}$, $k \cdot 1 > a$ by the Archimedean property, so $k \in S$. So now let $m = \min S$. Then,

$$\begin{aligned} m &\geq a \\ m - 1 &< a \\ m &< a + 1 < a + (b - a) = b \end{aligned}$$

■

Theorem 2.6 (Density of $\mathbb{Q}$ in $\mathbb{R}$). $\forall a, b \in \mathbb{R}, a < b \implies \exists r \in \mathbb{Q}, a < r < b$.

Proof. $b - a > 0$. By the Archimedean Property, $\exists n \in \mathbb{N}, n(b - a) > 1$, or $nb - ba > 1$. By the lemma above, $\exists m \in \mathbb{Z}, na < m < nb$. Then $a \leq \frac{m}{n} < b$. ■

2.7 Summary

Ordered Field: Just the 15 axioms, fairly simple.

Completeness: If $S \subseteq \mathbb{R}$ and $S \neq \emptyset$ and $\exists M \in \mathbb{R}, \forall x \in S, x \leq M$, then $\exists s^* \in \mathbb{R}, (\forall x \in S \mid x \leq s^*) \wedge (\forall \varepsilon > 0, \exists y \in S \mid y > s^* - \varepsilon)$. A convenient way of saying "upper bound" and "least upper bound". Similarly, every non-empty set bounded below also has an infimum, the greatest lower bound $\in \mathbb{R}$.

Archimedean property (AP): $\forall a, b > 0, \exists n \in \mathbb{N}, na > b$.

("Really obvious") Lemma: $\forall a, b \in \mathbb{R}, b - a > 1 \implies \exists m \in \mathbb{Z}, a < m < b$.

Density of $\mathbb{Q}$ in $\mathbb{R}$: $\forall a, b \in \mathbb{R}, a < b \implies (\exists r \in \mathbb{Q}, a < r < b)$.

3 Sequences

3.1 Convergence

A sequence of real numbers is a function $s : I \to \mathbb{R}$ where the indexing set I is an infinite subset of the natural numbers.

Theorem 3.1 (Ross 7.1)**.** *The sequence s_n converges to $s \in \mathbb{R}$ if*

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in I : (n \geq N) \implies |s_n - s| < \varepsilon.$$

We write $s = \lim(s_n)_{n \in I}$ [2]

Example 3.1. $\lim(\frac{1}{n})_{n=1}^{\infty} = 0$ □

Proof. Let $\varepsilon > 0$. By the Archimedean property, $\exists N \in \mathbb{N} \setminus \{0\}$ such that $0 < \frac{1}{N} < \varepsilon$. Let $n \geq N$. Then $0 < \frac{1}{n}\frac{1}{N}$. Thus $\forall n \geq N$, $0 < \frac{1}{n} \leq \frac{1}{N} \leq \varepsilon \implies \left|\frac{1}{n} - 0\right| < \varepsilon$. ■

Theorem 3.2 (Ross 9.1)**.** *Every convergent sequence is bounded.*

A bounded sequence is one such that $\exists M \geq 0, \forall n \in I : |s_n| \leq M$.

A convergent sequence is one such that $\exists s \in \mathbb{R}, \forall \varepsilon > 0, \exists N, \forall n \geq N : |s_n - s| < \varepsilon$.

Proof. Let $\varepsilon = 1$. Then $\exists N, \forall n : (n \geq N) \implies |s_n - s| < 1$. This implies that $\forall n \geq N : |s_n| < |s| + 1$ (because $|s_n| = |s_n - s + s| \leq |s_n - s| + |s| < 1 + |s|$).

Let $M - \max(|s_0|, |s_1|, \dots |s_{N-1}|, |s + 1|)$. M is an upper bound on the sequence, since

$$|s_k| \leq \begin{cases} \max(|s_0|, \dots, |s_{N-1}|), & k < N \\ |s| + 1, & k \geq N \end{cases}$$

■

[2]Sometimes we may drop the index and just write s_n instead of $(s_n)_n$ for simplicity's sake.

Theorem 3.3 (Ross 9.2). *If $k \in \mathbb{R}$ and $(s_n)_n$ converges to s, then $(ks_n)_n$ converges to ks.*

Proof.
Case 1: If $k = 0$, then $0 = 0$.

Case 2: If $k \neq 0$, then $\forall \varepsilon > 0$, $\exists N, \forall n \geq N : |s_n - s| < \frac{\varepsilon}{|k|}$. This means that $\forall \varepsilon > 0, \exists N, \forall n \geq N : |k||s_{n-s}| < \varepsilon$ ∎

Theorem 3.4 (Ross 9.3). *If s_n converges to s and t_n converges to t, then $(s_n + t_n)_n$ converges to $s + t$.*

Proof. Let there be an $\varepsilon > 0$. We can find N' such that $\forall n \geq N' : |s_n - s| < \frac{\varepsilon}{2}$. We can also find N'' such that $\forall n \geq N'' : |t_n - t| < \frac{\varepsilon}{2}$. Take $N = \max(N', N'')$. Then $\forall n \geq N$,

$$\begin{aligned} |(s_n + t_n) - (s + t)| &= |(s_n - s) + (t_n - t)| \\ &\leq |s_n - s| + |t_n - t| \\ &< \varepsilon \end{aligned}$$

∎

Theorem 3.5 (Ross 9.4). *If s_n converges to s and t_n converges to t, then $(s_n \cdot t_n)_n$ converges to $s \cdot t$.*

Proof. Let there be $\varepsilon > 0$. By Theorem 3.2, we can find some $M > 0, M \geq |s_n|$ for all $n > N$. Choose N such that $|t_n - t| < \frac{\varepsilon}{2M} \wedge |s_n - s| < \frac{\varepsilon}{2(|t|+1)}$ [3].

$$\begin{aligned} |s_n t_n - st| &= |s_n t_n - s_n t - st + s_n t| \\ &\leq |s_n||t_n - t| + |t||s_n - s| \\ &< M \cdot \frac{\varepsilon}{2M} + |t| \cdot \frac{\varepsilon}{2(|t| + 1)} \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

∎

Theorem 3.6 (Ross 9.5). $\lim \frac{1}{s_n} = \frac{1}{\lim s_n}, s \neq 0$. *Proof omitted.*

Theorem 3.7 (Ross 9.6). $\lim \frac{s_n}{t_n} = \frac{\lim s_n}{\lim t_n}, s \neq 0$. *Proof omitted.*

3.2 Monotone and Cauchy sequences

A sequence is called *monotone* if $\forall n, s_{n+1} \leq s_n$, or $\forall n s_{n+1} \leq s_n$.

Theorem 3.8 (Ross 10.2). *If s_n is bounded and monotone, then it converges.*

Proof. Assume that s_n is increasing. There is $M \in \mathbb{R}, \forall n, s_n \leq M$. So $\{s_n \mid n \in I\}$ is not empty and is bounded above. $\exists L \in \mathbb{R}, L = \sup(\{s_n \mid n \in I\})$.

[3]The choice of $|t + 1|$ is for the case when $t = 0$.

Claim. $\lim s_n = L$.

Let $\varepsilon > 0$. Then there exists $N, L - \varepsilon < s_N \leq L$. Since s_n is monotonously increasing, $\forall n \geq N, L - \varepsilon < s_n \leq L$. Then,

$$\begin{gathered} L - \varepsilon < s_n \leq L < L + \varepsilon \\ -\varepsilon < s_n - L < \varepsilon \\ |s_n - L| < \varepsilon \end{gathered}$$

■

3.3 Limits superior and inferior

Let $(s_n)_n$ be a sequence of real numbers. Then

$$\begin{aligned} \limsup(s_n)_n &= \lim_{N\to\infty}[\sup\{s_n \mid n \geq N\}] \\ \liminf(s_n)_n &= \lim_{N\to\infty}[\inf\{s_n \mid n \geq N\}] \end{aligned}$$

For convenience just let $\sup\ldots$ be v_N and $\inf\ldots$ be u_N in the future.

Example 3.2. Consider $s_n = \frac{(-1)^n(n+1)}{n} = \left(-\frac{2}{1}, +\frac{3}{2}, -\frac{4}{3}, +\frac{5}{4}, \ldots\right)$

$$\begin{aligned} v_1 &= \sup\{s_n \mid n \geq 1\} = \frac{3}{2} \\ v_2 &= \sup\{s_n \mid n \geq 2\} = \frac{3}{2} \\ v_3 &= \sup\{s_n \mid n \geq 3\} = \frac{5}{4} \\ &\vdots \end{aligned}$$

$(v_N)_N$ is always decreasing, so it has a limit. $(u_N)_N$ is also always decreasing, so it too has a limit. We might gather that $\limsup s_n = 1$, and $\liminf s_n = -1$. □

Example 3.3. Consider $s_n = \left(\frac{(-1)^n n}{n+1}\right)_{n=0}^{\infty}$.

$\limsup s_n = \lim(1, 1, 1, \ldots) = 1$. $\liminf s_n = -1$ □

Example 3.4. Consider the sequence $\left(\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}. \frac{1}{5}, \frac{2}{5}, \ldots\right)$

$\limsup = 1$, $\liminf = 0$. □

Theorem 3.9 (Ross 10.7)**.** $\lim(s_n)_n = L \in \mathbb{R} \cup \{\pm\infty\} \iff \limsup(s_n)_n = \liminf(s_n)_n = \lim(s_n)_n$.

Proof.
($\implies$) We only handle the case where L is real. Let $\varepsilon > 0$. We know that $(s_n)_n$ converges to L. This means that $\exists N, \forall n : n \geq N \implies |s_n - L| < \varepsilon \implies L - \varepsilon < s_n < L + \varepsilon$.

It follows that $v_N = \sup\{s_n : n \geq N\} \leq L + \varepsilon$. Also, $u_N \geq L - \varepsilon$. Since v_N is decreasing and u_N is increasing, $L - \varepsilon \leq u_n \leq v_n \leq L - \varepsilon$. Taking $\varepsilon \to 0$, we get that $\limsup = \liminf$.

($\Longleftarrow$) Again, we only handle the case where L is real. Suppose $\limsup s_n = \liminf s_n = L \in \mathbb{R}$.

Then $\varepsilon > 0, \exists N', \forall N : N \geq N' \implies |v_N - L| < \varepsilon \implies \sup\{s_n : n \geq N\} < L + \varepsilon$.

This means that:

$$\begin{aligned} &\exists N', \forall N \geq N', \forall n \geq N : s_n < L + \varepsilon \\ &\implies \exists N', \forall n \geq N' : s_n < L + \varepsilon \end{aligned}$$

Doing the same for u_N we gather that

$$\exists N'', \forall n \geq N' : s_n > L - \varepsilon$$

Take $N^* = \max(N', N'')$. Then $\forall n \geq N^*, L - \varepsilon < s_n < L + \varepsilon \implies |s_n - L| < \varepsilon$. ■

Definition 3.1. A sequence $(s_n)_n$ is called *Cauchy* if

$$\forall \varepsilon > 0, \exists N, \forall n, m : (n, m \geq N) \implies |s_n - s_m| < \varepsilon$$

□

Theorem 3.10 (Ross 10.10)**.** *Cauchy sequences are bounded.*

Proof. Let $\varepsilon = 1$. Then $\exists N, \forall n, m : n, m \geq N \implies |s_n - s_m| < 1$. In particular, $\forall n : n \geq N \implies |s_n - s_N| < 1$.

Let $M = \max\{|s_0|, |s_1|, \ldots, |s_N| + 1\}$. Then M bounds s_n from above. ■

Corollary 3.10.1. *If* $(s_n)_n$ *is Cauchy then* $\limsup, \liminf \in \mathbb{R}$

Theorem 3.11 (Ross 10.11)**.** *A sequence in* $\mathbb{R}$ *is convergent iff it is Cauchy.*

Proof.
($\implies$) Suppose we have a convergent sequence with $\lim(s_n)_n = s \in \mathbb{R}$. Let $\varepsilon > 0$. $\exists N, \forall m, n \geq N : |s_n - s| < \frac{\varepsilon}{2}, |s_m - s| < \frac{\varepsilon}{2}$.

$$\begin{aligned} |s_n - s_m| &\leq |s_n - s| + |s_m - s| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

($\Longleftarrow$) Suppose $(s_n)_n$ is Cauchy. Let $\varepsilon > 0$ and choose N large enough so $\forall n, m : n, m \geq N \implies |s_n - s_m| < \varepsilon$.

This means:

$$\begin{aligned}
\forall n, m : n, m \geq N &\implies s_n < s_m + \varepsilon \\
\forall m : m \geq N &\implies \sup\{s_n : n \geq N\} \leq s_m + \varepsilon \\
&\implies v_N - \varepsilon \leq s_m \\
&\implies v_N - \varepsilon \leq \inf\{s_m : m \geq N\} \\
&\implies v_N - \varepsilon \leq u_N
\end{aligned}$$

This means that

$$\forall \varepsilon > 0, \exists N : \limsup s_n \leq v_N \leq u_N + \varepsilon \leq \liminf s_n + \varepsilon$$

Since this holds for all $\varepsilon > 0$, this means that $\limsup s_n \leq \liminf s_n$. However the opposite inequality is also true, hence we conclude that $\limsup s_n = \liminf s_n = \lim s_n$. ■

A quick look back at our journey so far: completeness axiom $\iff$ bounded monotone sequences converge $\iff$ all Cauchy sequences converge.

A metric space in which all Cauchy sequences converge is called *complete*. We shall cover more about this later.

3.4 Subsequences

A sequence is $(s_n)_{n \in I}$, where I is an infinite subset of $\mathbb{N}$. We can do this: for any infinite $J \subseteq I$, we consider a new $(s_n)_{n \in J}$.

Even if the sequence does not converge, a subsequence might be able to.

Practically, a subsequence is any sequence formed by reading some of the terms from left to right. A subsequence of $(s_n)_{n=0}^{\infty}$ is any $(s_{n_k})_{k=0}^{\infty}$ with $n_0 < n_1 < n_2 < \dots$

Example 3.5. $\left(\frac{(-1)^n n}{n+1}\right)_{n=0}^{\infty}$. The even-indexed subsequence: $\left(\frac{(-1)^{2n} 2n}{2n+1}\right)_{n=0}^{\infty}$ □

Example 3.6. $\left(\frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}. \frac{1}{5}, \frac{2}{5}, \dots\right)$. □

Remark. If $r \in [0, 1]$, then there exists a subsequence that converges to r.

If r is rational, this is possible since $(r, r, r, \dots)$ is a subsequence that converges to r.

The general case follows from the density of $\mathbb{Q}$ in $\mathbb{R}$.

Definition 3.2. Let $\lambda \in \mathbb{R} \cup \{\pm\infty\}$. λ is a *subsequential limit* of $(s_n)_n$ if there exists a subsequence $(s_{n_k})_k$ such that $\lambda = \lim_{k \to \infty}(s_{n_k})$. If $\lambda \in \mathbb{R}$, then λ is called a *limit point*. □

Definition 3.3. Let $SSL(s_n)_n = \{\text{all subsequential limits of } (s_n)_n\} (\subseteq \mathbb{R} \cup \{\pm\infty\})$.

Let $LP(s_n)_n = \{\text{all limit points}\}$. $LP = SSL \cap \mathbb{R}$. □

Example 3.7. For the example directly above, $SSL = LP = [0, 1]$ □

Theorem 3.12 (Ross 11.2i)**.** *Let $(s_n)_n$ be a sequence and let $\lambda \in \mathbb{R}$. λ is a limit point of $(s_n)_n$ $\iff \forall \varepsilon > 0, \exists_\infty n \in I : |s_n - \lambda| < \varepsilon$.*

Proof.
$(\implies)$ Let $\lambda \in LP(s_n)_n$. Then $\exists (n_k)_k$ such that $\lim_{k\to\infty}(s_{n_k})_k = \lambda$.

$\forall \varepsilon > 0, \exists K, \forall k : k \geq K \implies |s_{n_k} - \lambda| < \varepsilon$. So the set of all $\{n_k \mid k \geq K\}$ is infinite.

$(\impliedby)$ Suppose $\exists \varepsilon > 0$ such that $\{n \in I \mid |s_n - \lambda| < \varepsilon\}$ is finite. Then $\exists N, \forall n \geq N : |s_n - \lambda| \geq \varepsilon$. But then $\forall n \geq N : (s_n \geq \lambda + \varepsilon) \vee (s_n \leq \lambda - \varepsilon)$. ■

Definition 3.4. Let S be a subset of $\mathbb{R}$, possibly empty. Let $p \in \mathbb{R}$. We say that p is adherent to S if $\text{dist}(p, S) = \inf\{\text{dist}(p, x) \mid x \in S\} = 0$. A set that contains all of its adherent points is called *closed.* □

For non-empty S the infimum always exists and is bounded below. If S is empty then it has no adherent points. It is clear that any $p \in S$ is always adherent to S. But for example $0 \neq (0, 1)$ but 0 is adherent.

Theorem 3.13. *Let $S \in \mathbb{R}$. The following are equivalent:*

i. For all convergent sequences with terms in S, its limit is also in S.

ii. S is closed.

Proof.
$(i \implies ii)$ Let p be adherent to S. We will build a sequence of terms in S that converges to p.

$\forall n \geq 1$, we can choose $s_n \in S$ such that $|s_n - p| < \frac{1}{n}$, by adherence. Then $p - \frac{1}{n} < s_n < p + \frac{1}{n}$. By the squeeze lemma, $\lim s_n = p$. We know $p \in S$ from point i. Therefore S is closed.

$(ii \implies i)$ Suppose S is closed. Let $(s_n)_n$ be a convergent sequence in S. Suppose $p = \lim s_n$. $\forall \varepsilon > 0, \exists N, \forall n \geq N : |s_n - p| < \varepsilon$. So $\forall \varepsilon > 0, \exists s \in S : |s - p| < \varepsilon$. Therefore $\inf\{|s - p|, s \in S\} < \varepsilon$, and since S is closed, p is also in S. ■

Theorem 3.14 (Title)**.** *$LP(s_n)_s$ is closed.*

Proof. We have to show that if $(\lambda_1, \lambda_2, \lambda_3, \ldots) \to \lambda$ where all λ_n are limit points of $(s_n)_n$, then λ must also be a limit point of $(s_n)_n$.

Without loss of generality, we may assume that $\forall n : |\lambda_n - \lambda| < \frac{1}{2n}$.

Pick $(t_{n,k})_k$ subsequences of (s_n) where $\forall k, |t_{n,k} - \lambda_n| < \frac{1}{2k}$.

$$\begin{array}{cccc} t_{11} & t_{12} & t_{13} & \ldots \to \lambda_1 \\ t_{21} & t_{22} & t_{23} & \ldots \to \lambda_2 \\ & & \vdots & \end{array}$$

Claim. $\lim(t_{n,n}) = \lambda$.

Let $\varepsilon > 0$. Choose n such that $\frac{1}{n} \leq \varepsilon$. Then,

$$|t_{n,n} - \lambda| \leq |t_{n,n} - \lambda_n| + |\lambda_n - \lambda| < \frac{1}{2n} + \frac{1}{2n} \leq \varepsilon$$

■

Theorem 3.15 (Ross 11.3)**.** $\lim(s_n)_n = \lambda \iff SSL(s_n)_n = \{\lambda\}$.

Proof. Let us say $s = \lim(s_n)_n$, and let $(s_{n_k})_k$ be a subsequence of $(s_n)_n$. By the way in which we simply read off the terms from left to right, $n_k \geq k$.

Take $\varepsilon > 0$. Then $\exists N : n > N \implies |s_n - s| < \varepsilon$. However this also means $n_k > N$, and therefore also $|s_{n_k} - s| < \varepsilon$.

The proof for the other direction is omitted. ■

Lemma 3.16. *Every sequence has a monotone subsequence (for $\mathbb{R}$).*

Proof. Without loss of generality, we assume that the sequence is indexed by $\mathbb{N}$. We will call $N \in \mathbb{N}$ *dominant* if $\forall m \geq N : s_m \leq s_N$. Then let D be the set of all dominant indices.

Case 1: If D is infinite, then we are done. $D = \{N_0 < N_1 < \dots\}$, and $(s_{n_k})_{k \in D}$ is decreasing, since $s_{n_{k+1}} \leq s_{n_k}$.

Case 2: If D is finite, then $\exists M$ such that $\forall N \geq M$, N is not dominant: $\forall N \geq M, \exists m \geq N : s_m > s_N$. Then we construct a subsequence as follows:

Take $N = N_0$. $\exists M_1 \geq M_0$, $s_{M_1} > s_{M_0}$. Then we can keep going, with $s_{M_0} < s_{M_1} < s_{M_2} < \dots$ ■

Lemma 3.17. *Every sequence has a subsequential limit.*

Proof. Let $(s_{n_k})_k$ be a monotonous subsequence. All monotonous sequences have limits. ■

Theorem 3.18 (Bolzano-Weierstrass theorem)**.** *If $(s_n)_n$ is bounded then $LP(s_n)_n$ is non-empty (i.e. every bounded sequence has a convergent subsequence).*

Proof. Let $(s_{n_k})_k$ be a monotonous subsequence. It also happens to be bounded. Hence, by Lemma 3.4, it converges. ■

3.5 Summary

A sequence $(s_n)_n$ is convergent if

$$\exists s \in \mathbb{R}, \underbrace{\forall \varepsilon > 0, \exists N, \forall n : n \geq N \implies |s_n - s| \leq \varepsilon}_{\lim(s_n)_n = s}$$

A sequence is bounded if $\exists M \geq 0, \forall n : |s_n| \leq M$.

A sequence is monotone if $(\forall n : s_{n+1} \geq s_n) \vee (\forall n : s_{n+1} \leq s_n)$.

Convergent $\implies$ bounded; bounded $\wedge$ monotone $\implies$ convergent.

Regarding sequences that do not converge:

$$\forall M, \exists N, \forall n : n \geq N \implies s_n \geq M \equiv \text{“}\lim(s_n)_n = +\infty\text{”}$$

It is similar with a limit of $-\infty$. We could also say that $\lim(-s_n)_n = +\infty$.

Corollary 3.18.1 (Ross 10.5)**.** *All monotone sequence have a limit* $\in \mathbb{R} \cup \{\pm\infty\}$

4 Topological concepts

5 Closed and open sets

We briefly touched on closed sets in Def. 3.4. Perhaps we should expand on this a little more. Previously we have defined the *distance* between two points $x, y \in \mathbb{R}$ with $\operatorname{dist}(x, y) = |x - y|$. Similarly we can also define the distance between a point $x \in \mathbb{R}$ and a non-empty set $S \subseteq \mathbb{R}$ as

$$\operatorname{dist}(x, S) = \inf\{\operatorname{dist}(x, s) \mid s \in S\}.$$

We say x is *adherent* to S if $\operatorname{dist}(x, S) = 0$. If $s \in S$ then s is adherent to S automatically, but the converse is not necessarily true. For example 0 is adherent to $(0, 1]$ even though it is not in the interval.

Then as per Def. 3.4, a set is called *closed* if it contains all of its adherent points.

Definition 5.1. Let $S \subseteq \mathbb{R}$. The *closure* of S, denoted as $\overline{S}$, is given by

$$\overline{S} = \{x \in \mathbb{R} \mid \operatorname{dist}(x, S) = 0\}.$$

That is, $\overline{S}$ contains all of the points that are adherent to S. □

Theorem 5.1. *Let* $S \subseteq \mathbb{R}$*.* $\overline{S}$ *is closed.*

Proof. We want to show that $\forall x \in \mathbb{R}, \forall s \in S, \forall \overline{s} \in \overline{S} : \operatorname{dist}(x, \overline{S} = 0) \implies \operatorname{dist}(x, S) = 0$ [4]. In other words, this states that every adherent point of $\overline{S}$ is also an adherent point of S and hence is in $\overline{S}$.

$$\begin{aligned}
\operatorname{dist}(x, \overline{S}) &= \inf\{\operatorname{dist}(x, \overline{s})\} \\
0 &= \inf\{|x - \overline{s}|\} \\
&\leq \inf\{|x - s| + |s - \overline{s}|\} \\
&= \inf\{\operatorname{dist}(x, s)\} + \inf\{\operatorname{dist}(\overline{s}, s)\} \\
&= \operatorname{dist}(x, S)
\end{aligned}$$

[4] We suppress the predicates for the rest of the proof, it is understood that we take the infimum over all $\overline{s} \in \overline{S}$.

The last line is because $\overline{S}$ contains all the points adherent to S, hence its distance from any $s \in S$ is 0. Now

$$\begin{aligned}
\operatorname{dist}(x, S) &= \inf\{\operatorname{dist}(x, s)\} \\
&= \inf\{|x - s|\} \\
&\leq \inf\{|x - \overline{s}| + |\overline{s} - s|\} \\
&= \inf\{\operatorname{dist}(x, \overline{s})\} + \inf\{\operatorname{dist}(s, \overline{s})\} \\
&= 0
\end{aligned}$$

Hence $\operatorname{dist}(x, S) = 0$, which completes the proof. ■

Theorem 5.2. $\overline{\mathbb{Q}} = \mathbb{R}$.

Proof. $\overline{\mathbb{Q}} \subseteq \mathbb{R}$ simply because it contains elements selected from $\mathbb{R}$ by definition.

Now we want to show that $\overline{Q} \supseteq \mathbb{R}$. To do this, we can show that $\forall x \in \mathbb{R} : \inf\{\operatorname{dist}(x, q) \mid q \in \mathbb{Q}\} = 0$, then which would then suggest that $x \in \overline{\mathbb{Q}}$.

Take any $x \in \mathbb{R}$, and let $D = \{\operatorname{dist}(x, q) \mid \forall q \in \mathbb{Q}\}$. Let there be an $\varepsilon > 0$, and take the smallest $n \in \mathbb{N}$ such that $10^{-n} < \varepsilon$. Then let x' be x but truncated to the n-th digit after the decimal point (if it has less decimals then we do nothing). $x' \in \mathbb{Q}$ since it has a finite number of decimal points, and is easily expressible as a fraction. Then $|x - x'| < 10^{-n} < \varepsilon$. Since $\forall \varepsilon > 0, \exists d \in D : d < 0 + \varepsilon$, we conclude that $\inf D = 0$. ■

It would be too easy if we said that a set that is not closed is open, since there are odd edge cases that we are ignoring (such as the empty set) this way.

Definition 5.2. Let $\subseteq \mathbb{R}$. A point $x \in S$ is *interior* to S if there is an $\varepsilon > 0$ such that $(x - \varepsilon, x + \varepsilon) \subseteq S$. This also means there is some open neighbourhood around x that is contained entirely within S. □

Definition 5.3. A set is called *open* if all of its elements are interior points. □

Theorem 5.3. *Let $U \subseteq \mathbb{R}$. U is open iff U^c is closed.*

Proof.

$(\implies)$ We need to check if there are any adherent points of U^c outside of U^c, i.e. inside U. Say there was some $u \in U$ that is adherent to U^c. Then $\operatorname{dist}(u, U^c) = 0$ and by the property of the infimum $\forall \varepsilon > 0, \exists u' \in U^c : |u - u'| < \varepsilon$, which also mean that $u' \in (u - \varepsilon, u + \varepsilon) \not\subseteq S$. Thus all adherent points of U^c is contained within itself.

$(\impliedby)$ Since U^c is closed, it contains all of its adherent points, so none of them exist in U. Thus for any $u \in U$, we can let $\varepsilon = \operatorname{dist}(u, U^c) > 0$. Then necessarily $(u - \varepsilon, u + \varepsilon) \subseteq U$. ■

So far we have been constantly using open intervals $(a, b) = \{x \in \mathbb{R} | a < x < b\}$ even before we have talked about the openness of a set, and it may be disturbing. However this is not circular as open intervals were not defined with any notion of the openness that we have just discussed. Nevertheless, it is easy to show that the name "open" interval is valid.

Theorem 5.4. *Every open interval $(a, b) \subset \mathbb{R}$ is open.*

Proof. For any $x \in (a, b)$, let $\varepsilon = \min(x - a, b - x)$. Then $(x - \varepsilon, x + \varepsilon) \subseteq (a, b)$. ■

$$\text{Ordered field} \underbrace{\to}_{\substack{\text{Discard}\\ \text{everything}\\ \text{but norm}}} \text{Normed space} \underbrace{\to}_{\substack{\text{Discard}\\ \text{everything but}\\ \text{metric function}}} \text{Metric space} \underbrace{\to}_{\substack{\text{Discard}\\ \text{everything except}\\ \text{closed or open}}} \text{Topological space} - \cdots \to \text{Set}$$

A topological space is a set $X(= \mathbb{R})$ together with τ, a set of subsets of X, called a topology, consisting of sets that are closed. They satisfy these properties:

i. $\emptyset, X \in \tau$: they are closed.

ii. $v_1, v_2 \in \tau \implies v_1 \cup v_2 \in \tau$

iii. $\{v_i : i \in I\} \subseteq \tau \implies \left(\bigcap_{i \in I} v_i\right) \in \tau$

Note that we only specify finite unions but intersections can potentially be infinite.

Definition 5.4. A point $p \in \mathbb{R}$ is *adherent* to $S \subseteq \mathbb{R}$ if $\forall \varepsilon > 0, \exists x \in S : |x - p| < \varepsilon$. □

Definition 5.5. A set S is *closed* if p adherent to S implies $p \in S$. □

Proof. i. $\emptyset$ has no adherent points, so it is vacuously true. Otherwise, if p is adherent to $\mathbb{R}$ then $p \in \mathbb{R}$.

ii. Suppose v_1 and v_2 are closed subsets of $\mathbb{R}$. Suppose there is a p adherent to $(v_1 \cup v_2)$. Then $\forall \varepsilon > 0, \exists x \in v_1 \cup v_2 : |x - p| < \varepsilon$.

Pick $x_n \in v_1 \cup v_2$ such that $|x_n - p| < \frac{1}{n}$. Then there exists a subsequence where all terms are in v_1 or in v_2. So p is in either v_1 or v_2.

iii. If p is adherent to $\bigcap_{i \in I} v_i$ then $\forall \varepsilon > 0, \exists x \in \bigcap_i v_i$, such that $|x - p| < \varepsilon$. Then $\forall i \in I, \exists x \in v_i : |x - p| < \varepsilon$, so $\forall i \in I : (p \text{ adh } v_i)$. It follows that $\forall i \in I : p \in v_i \implies p \in \bigcap_i v_i$.

■

Example 5.1. $\mathbb{Z}$ is closed. □

Example 5.2. In $\mathbb{R}$ there are infinite families of closed sets whose union is not closed. Consider $\bigcup_{n=2}^{\infty} \left[\frac{1}{n}, 1\right] = \left[\frac{1}{2}, 1\right] \cup \left[\frac{1}{3}, 1\right] \cup \ldots = (0, 1])$ □

Example 5.3. Let us define the set C_{n+1} to be C_n but without all the middle thirds of C_n. In other words:

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

Since C_n is an union of 2^n closed intervals, it is also closed.

Consider the *Cantor set* $C_\infty = \bigcap\limits_{n=0}^{\infty} C_n$.

C_∞ is also closed. It is non-empty. However C_∞ contains no open intervals, and it has no interior points.

C_∞ consists of all $r \in [0,1]$ such that $r = (0.d_1d_2d_3\ldots d_i)_3$, with $d_i \in \{0,1\}$. So $(0.2020\ldots)_3 = \left(\frac{3}{4}\right)_{10} \in C_\infty$. $\square$

5.1 Cardinality

Intuitively, the cardinality of a set represents how many items are in it. So $|\{1,2,4\}| = |\{2,1,1,4\}| = 3$.

Let A and B be sets.

- $|A| = |B|$ if there is a bijection $f : A \to B$.
- $|A| \leq |B|$ if there is an injection $f : A \to B$.
- $|A| < |B|$ if $|A| \leq |B|$ but $|A| \neq |B|$

Definition 5.6. S is *countable infinite* if S is finite, or if $|S| = |\mathbb{N}|$. We also define $|\mathbb{N}| = \aleph_0$ as the smallest infinite cardinal. $\square$

In other words, S is countable iff there exists a sequence of real numbers $(s_n)_{n=0}^{\infty}$ such that $S = \{s_n \mid n \in \mathbb{N}\}$

Example 5.4. $\mathbb{N}$ is countably infinite. $\mathbb{Z}$ is also countably infinite since we can enumerate them as $(1, -1, 2, -2, \ldots)$.

We claim that $\mathbb{Q}$ is also countably infinite. We first show that $\mathbb{Q}^+$ (all positive rationals) are countably infinite, the extension to all rationals follow easily afterwards. We can enumerate $\mathbb{Q}^+$ in this manner:

$\square$

Example 5.5. The set of all infinite binary sequences $2^{\mathbb{N}}$ is uncountable.

Suppose instead that there exists a surjection from $\mathbb{N} \to 2^{\mathbb{N}}$, and we can list the elements with $f(n) = (s_{n,i})_{i=0}^{\infty}$. Consider $(1 - s_{n,n})_{n=0}^{\infty}$. We have found a sequence that was not present in the original. This is precisely Cantor's diagonal argument – we take the diagonal entries and flip the bits to create a totally new sequence. $\square$

Example 5.6. C_∞ is uncountable. Recall that C_∞ is also the set that contains all $r \in [0,1]$ that can be represented in base 3 without using the digit 1. Then we can defines an injective function $f : 2^{\mathbb{N}} \to C_\infty$ as simply $(d_1d_2d_3\ldots) \to 0.(2d_1)(2d_2)(2d_3)\ldots$.

f is also injective, because $(d_1d_2\ldots) \neq (d_1'd_2'\ldots) \implies f(d_1d_2\ldots) \neq f(d_1'd_2'\ldots)$

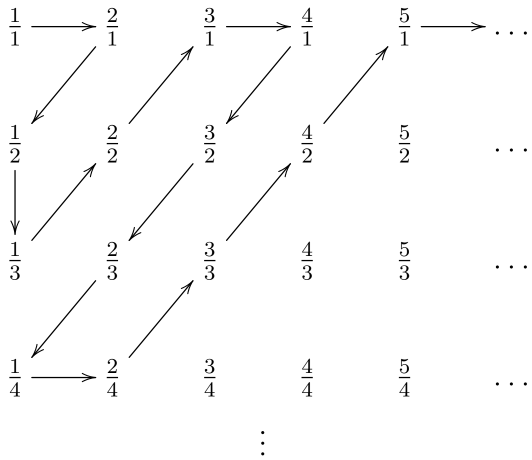

Figure 1: Enumerating the rationals.

This also means that $|C_\infty| = \left|2^{\mathbb{N}}\right|$.

Obviously, $|C_\infty| \leq |\mathbb{R}|$. But there also exists injections $\mathbb{R} \to [0,1] \to 2^{\mathbb{N}}$. Hence $|\mathbb{R}| \leq 2^{\aleph_0}$. Therefore we conclude that $|\mathbb{R}| = 2^{\aleph_0}$

Here we perhaps encounter a paradox. We know that $\mathbb{Q}$ is dense in $\mathbb{R}$. In other words, every non-empty subset of $\mathbb{R}$ meets $\mathbb{Q}$: $\mathbb{Q}$ is topologically big. Yet $|\mathbb{Q}| < |\mathbb{R}|$.

C_∞ on the other hand is "nowhere dense". If $x, y \in C_\infty$, then there exists an open set $U : x < U < y$ [5] and $U \cap C_\infty = \emptyset$. It is topologically small, and yet $|C_\infty| = |\mathbb{R}|$. □

6 Infinite series

An *infinite series* is an expression of the form $\sum_{n=N}^{\infty} a_n$ [6].

We call $(s_M)_{M=N}^{\infty} = \sum_{n=N}^{M} a_n$ a *partial sum*.

The series converges iff the sequence of partial sums is a convergent sequence.

Example 6.1.

- $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges.
- $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges to $\frac{\pi^2}{6}$
- $\sum_{n=1}^{\infty} \frac{1}{n^3}$ converges to ???

□

[5] Here we use the inequality $x < U$ as shorthand to mean x is smaller than every element in U.
[6] We occasionally drop the bounds for simplicity.

Definition 6.1 (Cauchy Criterion)**.** $\sum_{n=0}^{\infty} a_n$ converges iff the partial sums are Cauchy.

$$\forall \varepsilon > 0, \exists N, \forall m, n \geq N : \left| \sum_{k=0}^{n} a_k - \sum_{k=0}^{m} a_k \right| < \varepsilon$$

In other words,

$$\forall \varepsilon > 0, \exists N, \forall m, n \geq N : \left| \sum_{k=m}^{n} a_k \right| < \varepsilon$$

□

This means that if $\sum a_n$ converges, then also $\forall \varepsilon > 0, \exists N, \forall n \geq N : |\sum_{k=n}^{n} a_k| < \varepsilon$. Therefore $\lim a_n = 0$. The converse is clearly not true.

Example 6.2 (Geometric Series)**.**

$$\sum r^k = \begin{cases} \frac{1}{1-r}, & \text{if } |r| < 1 \\ \text{diverges otherwise} \end{cases}$$

□

Theorem 6.1 (Comparison Test)**.** *If $(a_n)_n$ is a sequence of non-negative numbers, and $\sum a_n$ converges and $|b_n| \leq a_n$, then $\sum b_n$ converges.*

Proof.

$$\sum_{k=n}^{m} b_k \leq \sum_{k=n}^{m} |b_k| \leq \sum_{k=n}^{m} a_k < \varepsilon$$

Then we just choose N such that $\forall n, m \geq N, |\sum_{k=n}^{m} a_k| < \varepsilon$. ■

Theorem 6.2 (Absolute Convergence)**.** *If $\sum |a_n|$ converges, then $\sum a_n$ does so as well.*

Theorem 6.3 (Ratio Test)**.** *Consider $\sum a_n$.*

- *If* $\limsup \left| \frac{a_{n+1}}{a_n} \right| < 1$, *the series converges absolutely,*
- *If* $\liminf \left| \frac{a_{n+1}}{a_n} \right| > 1$, *the series diverges,*
- *Otherwise no information.*

Theorem 6.4 (Root Test)**.** *Consider $\sum a_n$. Let $\alpha = \limsup |a_n|^{\frac{1}{n}}$.*

- *If $\alpha < 1$, the series converges absolutely,*
- *If $\alpha > 1$, the series diverges,*
- *If $\alpha = 1$, no information.*

Proof. Choose ε such that $\alpha + \varepsilon < 1$. Then there exists N such that $\alpha - \varepsilon < \sup\{|a_n|^{\frac{1}{n}} : n \geq N\} < \alpha + \varepsilon$

Then $\forall n \geq N$,

$$\begin{aligned} |a_n|^{\frac{1}{n}} &< \alpha + \varepsilon \\ |a_n| &< (\alpha + \varepsilon)^n \\ \sum_{n=N}^{\infty} |a_n| &\leq \sum_{n=N}^{\infty} (\alpha + \varepsilon)^n \end{aligned}$$

Hence the LHS converges. ■

7 Continuity

7.1 Continuous Functions

Intuitively, a discontinuity in a function might be caused by a jump, vertical asymptotes, or removable discontinuities (e.g. $\frac{\sin x}{x}\big|_0$).

From Calculus we might recall the fact that a function f is said to be continuous at a point α if f is defined at α and $\lim_{x\to\alpha} f(x) = f(\alpha)$.

We will be concerned with function $f : D \to \mathbb{R}$. We say f is real-valued on $D(\subseteq \mathbb{R})$.

Definition 7.1. Let $f : D \to \mathbb{R}$ be a real valued function with domain D. Let α be in D.

i. We say that f is continuous at the point α if:
   $((x_n)_n$ is a sequence in D and $\lim(x_n)_n = \alpha) \implies \lim(f(x_n))_n = f(\alpha)$.

ii. We say that if f is continuous on the set $S \subseteq D$ if:
   $\forall \alpha \in S$, f is continuous at α.

iii. We say that f is continuous if it is continuous on its domain. This is called *sequential continuity.*

□

Definition 7.2. f is called ε-δ continuous at α if:

$$\forall \varepsilon > 0, \exists \delta > 0, \forall x \in D : |x - \alpha| < \delta \implies |f(x) - f(\alpha)| < \varepsilon$$

□

Theorem 7.1 (Ross 17.2)**.** *f is sequentially continuous at α iff f is ε-δ continuous at α.*

Proof.

($\implies$) Assume f is not ε-δ continuous.

$$\exists \varepsilon > 0, \forall \delta > 0, \exists x \in D : (|x - \alpha| < \delta) \wedge (|f(x) - f(\alpha)| \geq \varepsilon)$$
$$\exists \varepsilon > 0, \forall n \geq 1, \exists x_n \in D : \left(|x_n - \alpha| < \frac{1}{n}\right) \wedge (|f(x_n) - f(\alpha)| \geq \varepsilon)$$

Notice that $\lim(x_n)_n = \alpha$. $\lim(f(x_n))_n \neq f(\alpha)$. So f is not sequentially continuous at α.

($\Longleftarrow$) Suppose that f is ε-δ continuous, and let $(x_n)_n \to \alpha$, $x_n \in D$.

$$\forall \varepsilon > 0, \exists N, \forall n : (n \geq N) \implies |x_n - \alpha| < \varepsilon$$
$$\forall \delta > 0, \exists N, \forall n : (n \geq N) \implies |x_n - \alpha| < \delta$$

Then we have

$$\exists N, \forall n : (n \geq N) \implies (|x_n - \alpha| < \delta \implies |f(x_n) - f(\alpha)| < \varepsilon)$$

from ε-δ continuity of f. So $\lim f(x_n)_n = \lim f(\alpha)$. ∎

Theorem 7.2. *Suppose that $f, g : D \to \mathbb{R}$ are continuous at $\alpha \in D$. Then:*

i. *$|f|$ is continuous at α.*

ii. *kf is continuous at α.*

iii. *$f + g$ is continuous at α.*

iv. *$f \cdot g$ is continuous at α.*

v. *f/g is continuous at α provided $g(\alpha) \neq 0$.*

Proof. We will quickly prove a few of them.

i. $|f(x_n) - f(\alpha)| \leq |f(x_n) - f(\alpha)|$

ii. Suppose $(x_n)_n \to \alpha$, $x_n \in D$.

$$\lim((kf)(x_n))_n = \lim(kf(x_n)))n = k \lim f(x_n)_n = kf(\alpha)$$

iii.

$$\lim((f + g)(x_n))_n = \lim(f(x_n) + g(x_n))_n = \ldots = (f + g)(\alpha)$$

v. Suppose $(x_n)_n \to \alpha$, $x_n \in D$. By continuity of f, g, $f(x_n) \to f(\alpha)$, $g(x_n) \to g(\alpha) \neq 0$. $\exists N', \forall n \geq N', g(x_n) > 0$.

$$\lim_{n \geq N'} \frac{f(x_n)_n}{g(x_n)_n} = \frac{\lim f(x_n)_n}{\lim g(x_n)_n} = \frac{f(\alpha)}{g(\alpha)} = \left(\frac{f}{g}\right)(\alpha)$$

∎

Theorem 7.3 (Ross 17.5)**.** *If f is continuous at α and g is continuous at $f(\alpha)$, then $g \circ f$ is continuous.*

Example 7.1.

- Constants are continuous.
- x is continuous: $\forall \varepsilon > 0, \exists \delta > 0, \forall x \in \mathbb{R} : |x - \alpha| < \delta \implies |x - \alpha| < \varepsilon$
- x^n, $n \in \mathbb{Z}^+$ is continuous.
- Polynomials are continuous.
- $\frac{1}{x}$ is continuous on its domain $\mathbb{R} \setminus \{0\}$.
- Rational functions are continuous on their domain.
- If $f : X \to Y$ is continuous and invertible, then so is its inverse.
- $\frac{x}{|x|}$ is continuous on its domain.

□

7.2 Properties of continuous functions

Theorem 7.4 (Extreme Value Theorem)**.** *Suppose D is closed and bounded (compact), and suppose $f : D \to \mathbb{R}$ is continuous.*

i. f is bounded.

ii. f has its minimum or maximum points in D. i.e. $\exists \alpha, \beta \in D, \forall x \in D : f(\alpha) \leq f(x) \leq f(\beta)$.

Proof. i. Assume f is not bounded. For simplicity we assume it is not bounded above. Then $\forall n \geq 1, \exists x_n \in D : f(x_n) \geq n$. Consider $(x_n)_n$. It is bounded as well. Then there is a convergent subsequence $(x_{n_k})_k \to \lambda \in \mathbb{R}$, by the B-W theorem (3.18). $\lambda \in D$ since D is closed. But $\lim f(x_{n_k})_k = \infty$ because $f(x_{n_k}) \geq n_k$. So f is not continuous.

ii. Let $M = \sup\{f(x) \mid x \in D\}$. M exists since the set is bounded and non-empty. We want to show that $\exists \beta \in D, f(\beta) = M$.

$$\forall n \geq 1, \exists y_n \in D : M - \frac{1}{n} < f(y_n) \leq M$$

Since y_n is bounded, there is a subsequence that converges to β. Then $\lim(y_{n_k})_k = \lim(y_n) = M$. So by continuity, $f(\beta) = M$.

■

Theorem 7.5 (Intermediate Value Theorem)**.** *Let $f : I \to \mathbb{R}$ be continuous, where I is an interval. Suppose $a, b \in I$ and $a < b$. Then $\forall y$ between $f(a)$ and $f(b)$, $\exists x^* \in [a, b] : f(x^*) = y$.*

Proof. Assume $f(a) < y < f(b)$. Let $S = \{x \in [a,b] \mid f(x) < y\}$. Since $a \in S$, $S \neq \emptyset$ and since $S \leq b$, S is bounded. Let $x^* = \sup S$. Since $[a,b]$ is closed, $x^* \in [a,b]$. We can then build a sequence $(s_n)_n$ such that $s_n \to s^*$, $x^* - \frac{1}{n} < s_n \leq x^*$. Then $\lim f(s_n) = f(x^*) \leq y$. Now let $t_n = \min\{b, x^* + \frac{1}{n}\}$. Then $(t_n)_n$ is a sequence in $([a,b] \setminus S) \subseteq I$ and $\to x^*$ so $f(x^*) \geq y$. Then it follows that $f(x^*) = y$. ■

Corollary 7.5.1. *If $f : I \to \mathbb{R}$ is continuous and I is an interval, then so is $f(I)$.*

7.3 Uniform Continuity

We say $f : D \to \mathbb{R}$ is continuous if

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

We can say $\delta = \delta(f, \varepsilon, y)$.

Definition 7.3 (Uniform Continuity)**.** We say that f is *uniformly continuous* on D if

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

□

Here δ depends only on f and ε.

Example 7.2. Consider $y = \frac{1}{x}$ on $(0, \infty)$. It is not uniformly continuous although it is continuous. If you pick some δ that works for ε and y, it will stop working as $y \to 0^+$. □

Example 7.3. Consider $f(x) = sin(\frac{1}{x})$ on $(0, \infty)$. If is also continuous but not uniformly continuous. □

Theorem 7.6 (Ross 19.2)**.** *If $f : D \to \mathbb{R}$ is continuous and D is compact, then f is uniformly continuous.*

Proof. Suppose not. Then

$$\exists \varepsilon > 0, \forall \delta > 0, \exists x, y \in D : (|x - y| < \delta) \wedge (|f(x) - f(y)| \geq \varepsilon)$$
$$\exists \varepsilon > 0, \forall n \geq 1, \exists x_n, y_n \in D : \left(|x_n - y_n| < \frac{1}{n}\right) \wedge (|f(x_n) - f(y_n)| \geq \varepsilon)$$

$(x_n)_n$ is a bounded sequence, so it has a convergent subsequence, $(x_{n_k})_k \to \alpha$. $\alpha \in D$, D is closed. Since $|x_n - y_n| \to 0$, $(y_{n_k})_k \to \alpha$

So $f(\alpha) = \lim f(x_{n_k}) = \lim f(y_{n_k})$. This is a contradiction since $|f(x_n) - f(y_n)| \geq \varepsilon$. ■

Theorem 7.7. *If $(s_n)_n$ is a Cauchy sequence in D, and $f : D \to \mathbb{R}$, is uniformly continuous, then $(f(s_n))_n$ is Cauchy.*

Theorem 7.8. *Let $f : D \to \mathbb{R}$. The following are equivalent:*

i. f is uniformly continuous.

ii. There exists and unique continuous function $\tilde{f} : \overline{D} \to \mathbb{R}$[7] *so that* $\forall x \in D, f(x) = \tilde{f}(x)$.

Definition 7.4 (Lipschitz Continuity)**.** A function $f : D \to \mathbb{R}$ is Lipschitz continuous if

$$\exists k \geq 0, \forall x, y \in D : |f(x) - f(y)| \leq k|x - y|$$

In other words, secants are never steeper than k. $\square$

All Lipschitz continuous functions are uniformly continuous (pick $\delta = \frac{\varepsilon}{k}$)

8 Integration

The definite, proper *integral* of a function f is written as such:

$$\int_a^b f(x)\,\mathrm{d}x\,,$$

where f is a function $f : [a, b] \to \mathbb{R}$. f is bounded.

If we let $\mathscr{S}$ be the set of all closed subintervals of $[a, b]$, and $\mathscr{F}$ be the set of all "integrable" functions on $[a, b]$, then $\int : \mathscr{S} \times \mathscr{F} \to \mathbb{R}$

We want $\int$ to satisfy the following properties:

- Linearity. If f, g are integrable, and $\lambda \in \mathbb{R}$, then
  - $\int_I \lambda f = \lambda \int_I f$.
  - $(f + g) = \int_I f + \int_I g$
- Additivity over almost disjoint domains.
  - $\int_{I \cup J} f = \int_I f + \int_J f$ if $|I \cap J| \leq 1$
- Monotonicity.
  - If $f \leq g$ on I, then $\int_I f \leq \int_I g$.
- It should compute areas.
  - $\forall c \in \mathbb{R}$ and all I (closed subintervals of $[a, b]$), $\int_I c = c|I|$.

[7] Here $\overline{D}$ denotes the closure of D.

8.1 Riemann-Darboux Integration

Let S be a subset of the real numbers. The indicator function of S is $\mathbb{1} : \mathbb{R} \to \{0, 1\}$:

$$\mathbb{1}_S(x) = \begin{cases} 1, & \text{if } x \in S \\ 0, & \text{otherwise} \end{cases}$$

A *step function* on $[a, b]$ is a finite linear combination of indicator functions of subintervals of $[a, b]$ (not necessarily closed):

$$\varphi(x) = \sum_{k=1}^{n} c_k \cdot \mathbb{1}_{I_k}(x)$$

A step function always has a finite range. So $\phi([a, b])$ is a finite set.

Definition 8.1 (Integration)**.**

$$\int_a^b \varphi = \sum_{y \in \varphi([a,b])} y \cdot \left|\varphi^{-1}(y)\right|$$

□

This is very cumbersome to use. We can also say that:

Theorem 8.1.

$$\int_a^b \varphi = \sum_{k=1}^{n} c_k |I_k|$$

We can check that any way of writing φ gives the same integral.

Corollary 8.1.1. *The integral is linear for step functions.*

Corollary 8.1.2. *The integral is monotonic for step functions.*

An *interval partition* of $[a, b]$ is $\{I_k\}_{k=1}^{n}$, such that

$$[a, b] = I_i \cup \ldots \cup I_n \text{ and } i \neq j \implies I_i \cap I_j = \emptyset$$

An interval partition of $[a, b]$ is *compatible* with φ if $\forall k, \varphi$ is constant on I_k.

8.1.1 Integrating Step functions

There are a few ways we can go about integrating a step function.

i. Graphical methods

ii. Definition 8.1: $\int_a^b \varphi = \sum_{y \in \varphi([a,b])} y \cdot |\varphi^{-1}(y)|$

iii. Theorem 8.1: $\int_a^b \sum_{k=1}^n c_k \mathbb{1}_{I_k} = \sum_{k=1}^n \int_a^b c_k \mathbb{1}_{I_k} = \sum_{k=1}^n c_k |I_k|$

Theorem 8.2. *$\int_a^b \varphi = \sum c_k |I_k|$ and does not depend on how we write φ*[8].

Theorem 8.3. *$\int_a^b : \mathrm{Step}([a,b]) \to \mathbb{R}$*[9] *is linear, monotonic, additive, and non-trivial.*

Proof. Let $\varphi = \sum_{i=1}^n c_i \mathbb{1}_{I_i}$, $\phi = \sum_{j=1}^m d_j \mathbb{1}_{J_j}$. Let $\lambda \in \mathbb{R}$.

- $\lambda\varphi = \sum_{i=1}^n \lambda c_i \mathbb{1}_{I_i}$. Therefore $\int_a^b \lambda\varphi = \sum_{i=1}^n \lambda c_i |I_i| = \lambda \sum_{i=1}^n c_i |I_i| = \lambda \int_a^b \varphi$.
- $\int_a^b \varphi + \psi = \int_a^b \left(\sum_{i=1}^n c_i \mathbb{1}_{I_i} + \sum_{j=1}^m d_j \mathbb{1}_{J_j} \right) = \int_a^b \varphi + \int_a^b \psi$.
- Let φ and ψ be step functions in $\mathrm{Step}([a,b])$, with $\forall x \in [a,b] : \varphi(x) \leq \psi(x)$ (or we just write $\varphi \leq \psi$). Assume that $I_1, I_2, \ldots, I_n$ are pairwise disjoint, i.e. $i \neq j \implies I_i \cap I_j = \emptyset$. Also assume that $J_1, J_2, \ldots, J_m$ are also pairwise disjoint.

  $$\psi - \varphi = \sum_{j=1}^m \sum_{i=1}^n (d_j - c_i) \mathbb{1}_{I_i \cap J_j}.$$

  If $x \in I_i \cap J_j$, then $\psi(x) - \varphi(x) = d_j - c_j \geq 0$. Thus

  $$\int_a^b (\psi - \varphi) = \sum_{j=1}^m \sum_{i=1}^n (d_j - c_i) |I_i \cap J_j| \geq 0.$$

  Therefore, by linearity, $\int_a^b \psi \leq \int_a^b \varphi$.

■

8.1.2 Integrating non-step functions

Now we try to integrate non=step functions. Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Let us define two sets

$$\mathscr{L}(f[a,b]) = \left\{ \int_a^b \varphi \mid \varphi \text{ is a step function } \leq f \right\}$$

$$\mathscr{U}(f[a,b]) = \left\{ \int_a^b \varphi \mid \varphi \text{ is a step function } \geq f \right\}$$

Definition 8.2. The lower integral of f on $[a,b]$ is defined as

$$\underline{\int_a^b} f = \sup \mathscr{L}(f[a,b])$$

The upper integral of f on $[a,b]$ is defined as

$$\overline{\int_a^b} f = \inf \mathscr{U}(f[a,b])$$

□

[8]This proof is actually non-trivial but is quite long winded and tedious so we skip it for now.
[9]Here $\mathrm{Step}([a,b])$ gives the set of all step functions on $[a,b]$

Theorem 8.4. *The integrals exist.*

Proof. Because f is bounded, $\exists M \geq 0, \forall x \in [a,b], |f(x)| \leq M$. The constant function M is a step function (similarly for $-M$). $M \cdot \mathbb{1}_{[a,b]}$ is a step function $\geq f$. Since $\int_a^b M = M(b-a) \in \mathscr{U}(f[a,b])$, so $\mathscr{U}(f[a,b])$ is non-empty, and it is bounded from below by $-M(b-a)$.

Similarly, $\mathscr{L}([a,b])$ is non-empty because it contains $-M(b-a)$, and it is bounded from above by $M(b-a)$. Therefore both the upper and lower integrals exist. ■

Definition 8.3. A function is said to be *integrable* if

$$\underline{\int}_a^b f = \overline{\int}_a^b f \qquad \text{and} \qquad \int_a^b = \underline{\int}_a^b = \overline{\int}_a^b$$

□

Claim. *If $f : [a,b] \to \mathbb{R}$ is bounded, then $\underline{\int}_a^b f \leq \overline{\int}_a^b f$.*

Proof. Suppose that $\varphi \leq f \leq \psi$. Then $\int_a^b \varphi \leq \int_a^b \psi$. So $\mathscr{L}(f[a,b]) \leq \mathscr{U}(f[a,b])$, and thus $\sup \mathscr{L} \leq \inf \mathscr{U}$ ■

Theorem 8.5. *Let $f, g : [a,b] \to \mathbb{R}$ be bounded and $\lambda \in \mathbb{R}$. The following are true:*

i. *Scaling:*
   - *If $\lambda \geq 0$, then $\underline{\int}_a^b \lambda f = \lambda \underline{\int}_a^b f$ and $\overline{\int}_a^b \lambda f = \lambda \overline{\int}_a^b f$.*
   - *If $\lambda \leq 0$, then $\underline{\int}_a^b \lambda f = \lambda \overline{\int}_a^b f$ and $\overline{\int}_a^b \lambda f = \lambda \underline{\int}_a^b f$.*

ii. *Addition:*
   - $\underline{\int}_a^b (f+g) \geq \underline{\int}_a^b f + \underline{\int}_a^b g$.
   - $\overline{\int}_a^b (f+g) \leq \overline{\int}_a^b f + \overline{\int}_a^b g$.

iii. *Monotonicity: If $f \leq g$ on $[a,b]$ then $\underline{\int}_a^b f \leq \underline{\int}_a^b g$, and $\overline{\int}_a^b f \leq \overline{\int}_a^b g$.*

iv. $\underline{\int}_a^b = \underline{\int}_a^c + \underline{\int}_c^b$ *and* $\overline{\int}_a^b = \overline{\int}_a^c + \overline{\int}_c^b$

Proof. We will prove the properties of the lower integrals since it is fairly straightforward to derive those for the upper integrals from there.

i. Let $\lambda > 0$. We want to show $\underline{\int}_a^b \lambda f = \lambda \underline{\int}_a^b f$.

   $\underline{\int}_a^b \lambda f = \sup \left\{ \int_a^b \varphi \mid \varphi \leq \lambda f \right\}$. However notice that $\{\varphi \mid \varphi \leq \lambda f\} = \{\lambda \varphi_0 \mid \varphi_0 \leq f\}$.

   Therefore $\left\{ \int_a^b \varphi \mid \varphi \leq \lambda f \right\} = \left\{ \lambda \int_a^b \varphi_0 \mid \varphi_0 \leq f \right\} = \lambda \left\{ \int_a^b \varphi_0 \mid \varphi_0 \leq f \right\}$.

   Then $\underline{\int}_a^b \lambda f = \lambda \sup(\mathscr{L}(f)) = \lambda \underline{\int}_a^b f$.

ii. We want to show that $\underline{\int_a^b}(f+g) \geq \underline{\int_a^b} f + \underline{\int_a^b} g$.

Notice $\{\varphi_f + \varphi_g \mid \varphi_f \leq f, \varphi_g \leq g\} \subseteq \{\varphi \mid \varphi \leq f+g\}$. Then

$$\begin{aligned}
\left\{\int_a^b \varphi \mid \varphi \leq f+g\right\} &\supseteq \left\{\int_a^b (\varphi_f + \varphi_g) \mid \varphi_f \leq f, \varphi_g \leq g\right\} \\
&= \left\{\int_a^b \varphi_f + \int_a^b \varphi_g \mid \varphi_f \leq f, \varphi_g \leq g\right\} \\
&= \left\{\int_a^b \varphi_f \mid \varphi_f \leq f\right\} + \left\{\int_a^b \varphi_g \mid \varphi_g \leq g\right\}
\end{aligned}$$

Taking the supremum on both sides, we can see that $\sup(\mathscr{L}(f+g)) \geq \sup(\mathscr{L}(f) + \mathscr{L}(g)) = \sup(\mathscr{L}(f)) + \sup(\mathscr{L}(g))$. So $\underline{\int_a^b}(f+g) \geq \underline{\int_a^b} f + \underline{\int_a^b} g$.

iii. We want to show that $f \leq g \implies \underline{\int_a^b} f \leq \underline{\int_a^b} g$.

$\{\varphi \leq f\} \subseteq \{\phi \leq g\}$. Taking supremum, $\underline{\int_a^b} f \leq \underline{\int_a^b} g$.

■

Theorem 8.6. *Let $f : [a,b] \to \mathbb{R}$ be a bounded function. Consider the following functions:*

$$\underline{F} : [a,b] \to \mathbb{R} : x \mapsto \underline{\int_a^x} f \qquad\qquad \overline{F} : [a,b] \to \mathbb{R} : x \mapsto \overline{\int_a^x} f$$

Claim. *$\underline{F}$ and $\overline{F}$ are Lipschitz continuous.*

Proof. Let $x, y \in [a,b]$, $x \leq y$. Let M satisfy $|f| \leq M$. Then $\underline{F}(y) - \underline{F}(x) = \underline{\int_a^y} f - \underline{\int_a^x} f = \underline{\int_x^y} f$.

By monotonicity, $-\underline{\int_x^y} M \leq \underline{\int_x^y} f \leq \underline{\int_x^y} M$. So $-M(y-x) \leq \underline{F}(y) - \underline{F}(x) \leq M(y-x)$. So $|\underline{F}(x) = \underline{F}(y)| \leq M(x-y)$. Hence $\underline{F}$ is Lipschitz continuous. The proof for $\overline{F}$ follows similarly. ■

8.1.3 Integrability

Integrability was defined in Definition 8.3.

Theorem 8.7 (Cauchy Criterion for integrability)**.** *f is integrable iff $\forall \varepsilon > 0, \exists \varphi \leq f, \exists \psi \geq f, \int_a^b (\psi - \varphi) \leq \varepsilon$.*

Proof. Suppose $\exists \varepsilon > 0, \forall \varphi \leq f, \forall \phi \geq f : \int_a^b \phi - \int_a^b \varphi \geq \varepsilon$. Then

$$\begin{aligned}
\int_a^b \varphi &\leq \int_a^b \phi + \varepsilon \\
\underline{\int_a^b} f &\leq \overline{\int_a^b} f + \varepsilon \\
\underline{\int_a^b} f &= \overline{\int_a^b} f
\end{aligned}$$

■

Theorem 8.8. *We have* $\int_a^b f = L$ *iff* $\exists \varphi_1, \varphi_2, \ldots, \psi_1, \psi_2, \ldots \in \mathrm{Step}([a,b])$*, such that* $\varphi_1 \leq \varphi_2 \leq \ldots \leq f \leq \ldots \leq \phi_2 \leq \phi_1$*, and* $\lim_{n\to\infty} \int_a^b \varphi_n = L = \lim_{n\to\infty} \int_a^b \phi_n$.

Proof. By the Cauchy Criterion. ■

Theorem 8.9. *Let* $f, g : [a,b] \to \mathbb{R}$ *be bounded and integrable, and* $\lambda \in \mathbb{R}$.

i. λf *is integrable and* $\int_a^b \lambda f = \lambda \int_a^b f$.

ii. $f + g$ *is integrable and* $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

iii. *If* $f \leq g$ *on* $[a,b]$*, then* $\int_a^b f \leq \int_a^b g$.

iv. *If* $c \in [a,b]$*, then* f *is integrable on* $[a,c]$ *and* $[c,b]$*, and* $\int_a^b f = \int_a^c f + \int_c^b f$.

Proof. Most of them follow from Theorem 8.5.

i. If $\lambda \geq 0$, then $\underline{\int}_a^b \lambda f = \lambda \underline{\int}_a^b f = \lambda \overline{\int}_a^b f = \overline{\int}_a^b \lambda f$.
   If $\lambda < 0$, then $\underline{\int}_a^b \lambda f = \lambda \overline{\int}_a^b f = \lambda \underline{\int}_a^b f = \overline{\int}_a^b \lambda f$.

ii. $\overline{\int}_a^b (f+g) \leq \overline{\int}_a^b f + \overline{\int}_a^b g = \underline{\int}_a^b f + \underline{\int}_a^b g \leq \underline{\int}_a^b (f+g)$.
   $\overline{\int}_a^b (f+g) = \underline{\int}_a^b (f+g) = \int_a^b (f+g) = \int_a^b f + \int_a^b g$.

iii. $\underline{\int}_a^b$ and $\overline{\int}_a^b$ are both monotone.

iv. $\int_a^b f = \underline{\int}_a^b f = \underline{\int}_a^c f + \underline{\int}_c^b f$, and $\int_a^b f = \overline{\int}_a^b f = \overline{\int}_a^c f + \overline{\int}_a^c f$.

$$\underline{\int}_a^c f + \underline{\int}_c^b f = \overline{\int}_a^c f + \overline{\int}_c^b f$$
$$\underbrace{\underline{\int}_a^c f - \overline{\int}_c^b f}_{\geq 0} = \underbrace{\overline{\int}_a^c f - \underline{\int}_c^b f}_{\leq 0}$$

So both sides are equal to 0. This shows integrability. The other part follows.

■

We have spoken at length about integrable functions.

Example 8.1. A non-integrable function:

$$\mathbb{1}_{\mathbb{Q}} \text{ on } [0,1]$$

□

Proof. Suppose $\psi \in \text{Step}([0,1])$ and $\psi \geq \mathbb{1}_{\mathbb{Q}}$. We can write $\psi(x) = \sum_{k=1}^{n} c_k \mathbb{1}_{I_k}(x)$, where $c_1, \dots, c_n \in \mathbb{R}$ and $I_1, \dots, I_n$ are non-overlapping subintervals of $[0,1]$, and $\bigcup_{k=1}^{n} I_k = [0,1]$.

$\psi(x) = c_k$ on I_k. We claim that $c_k \geq 1$. $\forall x \in [0,1], \psi(x) \geq f(x)$. If I_k is not degenerate, there are always rational numbers in it. Then $\exists q \in \mathbb{Q} \cap I_k$. Then $\psi(q) = c \geq \mathbb{1}_{\mathbb{Q}}(q) = 1$. So $\int_0^1 \psi = \sum_{k=1}^{n} c_k |I_k| \geq \sum_{k=1}^{n} 1 \cdot |I_k| = 1$

By a similar argument, if $\varphi \leq \mathbb{1}_{\mathbb{Q}}(x)$, then $\int_a^b \varphi \leq 0$. ■

Remark. It is possible to have the sum of two non-integrable functions to be integrable. Take $\mathbb{1}_{\mathbb{Q}}$ and $\mathbb{1}_{\mathbb{R}\setminus\mathbb{Q}}$ for example.

8.1.4 Sufficient conditions for integrability

Here we want to show:

- Piecewise monotone functions are integrable.
- Piecewise continuous functions are integrable.

Theorem 8.10. *If $f : [a,b] \to \mathbb{R}$ is bounded and monotone, then f is integrable.*

Proof. We will show this for the case where f is increasing, and $f(a) < f(b)$.

Let $\varepsilon > 0$. Choose an interval partition of $[a,b] = I_1, \cup \ldots \cup I_n$ where $I_1, \ldots, I_n$ are non-overlapping, such that $|I_k| < \frac{\varepsilon}{f(b)-f(a)}$. Let $t_0, \ldots, t_n$ be the endpoints of the intervals, i.e. $I_1 = [t_0, t_1)$, $I_2 = [t_1, t_2)$, etc.

$\inf_{x \in I_k} f(x) = f(t_{k-1})$ and $\sup_{x \in I_k} f(x) = f(t_k)$. Now let $\varphi = \sum_{k=1}^{n} f(t_{k-1}) \mathbb{1}_{I_k}$ and $\psi = \sum_{k=1}^{n} f(t_k) \mathbb{1}_{I_k}$. Then $\int_a^b (\psi - \varphi) = \sum_{k=1}^{n} [f(t_k) - f(t_{k-1})] |I_k| < \frac{\varepsilon}{f(b)-f(a)} [f(t_1) - f(t_0) + f(t_2) - f(t_1) + \ldots + f(t_n) - f(t_{n-1})] = \varepsilon$

Thus f is integrable by the Cauchy Criterion. ■

Theorem 8.11. *All continuous $f : [a,b] \to \mathbb{R}$ are integrable.*

Proof. Let $\varepsilon > 0$. We may choose $\delta > 0$ such that $\forall x, y \in [a,b], |x - y| < \delta \implies |f(x) - f(y)| < \frac{\varepsilon}{b-a}$ (assume $a < b$). Partition $[a,b]$ into intervals $I_1, I_2, \ldots I_n$ so that each $|I_k| < \delta$. Let $\overline{I}_k$ be the closure of I_k. $f : \overline{I}_k \to \mathbb{R}$ is a continuous function on a closed bounded interval. Then $\exists p_k, q_k \in \overline{I}_k, \forall x \in \overline{I}_k : f(p_k) \leq f(x) \leq f(q_k)$. But also $\forall x \in I_k : f(p_k) \leq f(x) \leq f(q_k)$.

Let $\varphi = \sum_{k=1}^{n} f(p_k) \mathbb{1}_{I_k}$, $\psi = \sum_{k=1}^{n} f(q_k) \mathbb{1}_{I_k}$. So $\varphi \leq f \leq \psi$.

$\int_a^b (\psi - \varphi) = \sum [f(q_k) - f(p_k)] |I_k|$. Since $|I_k| < \delta$, $\left|\overline{I}_k\right| < \delta$. Therefore, $|p_k - q_k| < \delta$. Then $|f(p_k) - f(q_k)| < \frac{\varepsilon}{b-a}$. $\int_a^b (\psi - \varphi) < \varepsilon$. ■

9 Derivatives

Definition 9.1. Let $f : (a, b) \to \mathbb{R}$ [10]. We say that f is *differentiable* at $y \in (a, b)$ if

$$\lim_{x \to y} \frac{f(x) - f(y)}{x - y}$$

exists and is real, and we call the limit the *derivative*. □

Example 9.1. Show the derivative of $x^3 + 2x$ is $3x^2 + 2$. □

Proof.

$$\begin{aligned}
\lim_{x \to y} \frac{f(x) - f(y)}{x - y} &= \lim_{h \to 0} \frac{f(x + h) - f(x)}{h} \\
&= \lim_{h \to 0} \frac{(x + h)^3 + 2(x + h) - x^3 - 2x}{h} \\
&= 3x^2 + 2
\end{aligned}$$

■

Definition 9.2 (Limit Definition of continuity)**.** A function $f : [a, b] \to \mathbb{R}$ is continuous at $y \in [a, b]$ if $\lim_{x \to y} f(x) = f(y)$. □

Theorem 9.1 (Ross 28.2)**.** *Let $f : [a, b] \to \mathbb{R}$ be differentiable at $y \in (a, b)$. Then f is continuous at y.*

Proof.

$$\begin{aligned}
\lim_{x \to y} \left[(x - y) \frac{f(x) - f(y)}{x - y} \right] &= \lim_{x \to y} (x - y) \lim_{x \to y} \frac{f(x) - f(y)}{x - y} \\
\lim_{x \to y} [f(x) - f(y)] &= 0 \cdot f'(y) \\
&= \lim_{x \to y} f(x) - \lim_{x \to y} f(y) \\
&= \lim_{x \to y} f(x) - f(y) \\
\lim_{x \to y} f(x) &= f(y)
\end{aligned}$$

■

Theorem 9.2 (Derivative Rules)**.** *Let $f, g : (a, b) \to \mathbb{R}$, let $\lambda \in \mathbb{R}$, let $yin(a, b)$. Suppose f, g are differentiable at y.*

i. $(\lambda f)'(y) = \lambda f'(y)$.

ii. $(f + g)'(y) = f'(y) + g'(y)$.

iii. Product rule: $(fg)'(y) = f(y)g'(y) + f'(y)g(y)$.

[10] Notice how integration deals with functions on closed intervals while differentiation deals with functions on open intervals.

iv. *Quotient rule:* $(\frac{f}{g})'(y) = \frac{g'(y)f(y)-f'(y)g(y)}{g(y)^2}$.

v. *Chain rule:* $(g \circ f)'(y) = g'(f(y))f'(y)$

Proof. We will prove a few of them.

i.
$$\lim_{x\to y} \frac{(\lambda f)(x) - (\lambda f)(y)}{x-y} = \lambda \lim_{x\to y} \frac{f(x)-f(y)}{x-y} = \lambda f'(y)$$

ii.
$$\lim_{x\to y} \frac{(f+g)(x) - (f+g)(y)}{x-y} = \lim_{x\to y} \frac{f(x)+g(x)-f(y)-g(y)}{x-y} = f'(x) + g'(x)$$

iii.
$$\begin{aligned}
\lim_{x\to y} \frac{(fg)(x) - (fg)(y)}{x-y} &= \lim_{x\to y} \frac{f(x)g(x) - f(y)g(y)}{x-y} \\
&= \lim_{x\to y} \frac{f(x)g(x) - f(y)g(y) - f(y)g(y) + f(x)g(y)}{d} \\
&= \lim_{x\to y} \frac{f(x)[g(x)+g(y)]}{x-y} - \lim_{x\to y} \frac{g(y)[f(x)-f(y)]}{x-y} \\
&= f(y)g'(y) - f'(y)g(y)
\end{aligned}$$

■

Theorem 9.3 (Ross 29.1)**.** *Suppose $f : [a,b] \to \mathbb{R}$ is continuous. Let $p, q \in [a,b]$ satisfy $\forall x \in [a,b] : f(p) \leq f(x) \leq f(q)$. Suppose also that f is differentiable in $[a,b]$. Then $f'(p) = 0$ and $f'(q) = 0$.*

Proof. Take the local maximum q. We know that the derivative at q exists.

Suppose $f'(q) > 0$. Then $\exists \delta > 0, |x-q| < \delta \implies \frac{f(x)-f(q)}{x-q} > 0$. Then if $x \in (q, q+\delta)$, $\frac{f(x)-f(q)}{x-q} > 0$ and $x - q > 0$. Then $f(x) - f(q) > 0$. Contradiction.

Suppose $f'(q) < 0$. Then $\exists \delta > 0, |x-q| < \delta \implies \frac{f(x)-f(q)}{x-q} < 0$. Then if $x \in (q-\delta, q)$, $\frac{f(x)-f(q)}{x-q} < 0$ and $x - q < 0$. Then $f(x) - f(q) > 0$. Contradiction.

Therefore $f'(q) = 0$. ■

Theorem 9.4 (Rolle's Theorem)**.** *If $f : [a,b] \to \mathbb{R}$ is continuous, and f is differentiable on (a,b), and $f(a) = f(b)$, then $\exists y \in (a,b) : f'(y) = 0$.*

Proof. By the extreme value theorem (Theorem 7.4), $\exists p, q, \forall x \in [a,b] : f(p) \leq f(x) \leq f(q)$.

Suppose p, q are at both endpoints $\{p,q\} \subseteq \{a,b\}$, then $f(a) = f(b) \implies f$ is constant.

If $p, q \in (a,b)$, then the derivative at p or q is 0. ■

Theorem 9.5 (Mean Value Theorem)**.** *Suppose $f : [a, b] \to \mathbb{R}$ is continuous, and f is differentiable on (a, b). Then*

$$\exists y \in (a, b) : f'(y) = \frac{f(b) - f(a)}{b - a}$$

Proof. Take $L(x) = \frac{f(b)-f(a)}{b-a}$. Consider $f(x) - L(x)$. This is 0 at a and b. Then by Rolle's theorem (Theorem 9.4), $\exists y, f'(y) = L'(y) = \frac{f(b)-f(a)}{b-a}$. ■

Theorem 9.6 (Ross 29.4)**.** *If $f' = 0$ on (a, b) then f is constant.*

Proof. Assume f is not constant. Then $\exists y_1, y_2 \in (a, b) : f(y_1) \neq f(y_2)$. Then by the mean value theorem (Theorem 9.5), $\exists z \in (y_1, y_2), f'(z) = \frac{f(y_2)-f(y_1)}{y_2-y_1} \neq 0$. ■

Corollary 9.6.1. *If $f' = g'$ on (a, b), then $f = g + C$.*

Proof. Apply Theorem 9.6 on $f - g$. ■

10 The Fundamental Theorem of Calculus

Definition 10.1. Suppose $\{I_k\}_{k=1}^n$ is an interval partition of $[a, b]$, i.e. I_ks do not overlap and $\bigcup I_k = [a, b]$. Let this be $\mathcal{P}$

Let $f : [a, b] \to \mathbb{R}$.

- The upper Darboux step function for $\mathcal{P}$ is $\psi = \sum_{k=1}^n \sup_{x \in \bar{I}_k} f(x) \mathbb{1}_{I_k}$
- The lower Darboux step function for $\mathcal{P}$ is $\varphi = \sum_{k=1}^n \inf_{x \in \bar{I}_k} f(x) \mathbb{1}_{I_k}$

□

It may be helpful to remember that if f is continuous then sup = max and inf = min.

We say f is *integrable* on (a, b) if every extension of f to $[a, b]$ is integrable.

The Fundamental Theorem of Calculus comes in two parts.

Theorem 10.1 (Fundamental Theorem of Calculus Part 1)**.** *Suppose $g : [a, b] \to \mathbb{R}$ is continuous and differentiable on (a, b). If $g' : (a, b) \to \mathbb{R}$ is integrable, then*

$$\int_a^b g' = g(b) - g(a)$$

Proof. Let $\varepsilon > 0$. By integrability of g', $\exists \psi, \varphi \in \text{Step}([a, b]), \varphi \leq g' \leq \psi$ and $\int_a^b (\psi - \varphi) \leq \varepsilon$.

We can replace ψ and φ with the Darboux step functions. Then $\int_a^b (\psi - \varphi) < \varepsilon$ still. $\psi = \sum_{k=1}^n \max_{x \in \bar{I}_k} g'(x) \mathbb{1}_{I_k}$ and $\varphi = \sum_{k=1}^n \min_{x \in \bar{I}_k} g'(x) \mathbb{1}_{I_k}$

Let $\overline{I}_k = [t_{k-1}, t_k]$. For each $k = 1, \dots, n$, choose $x_k \in (t_{k-1}, t_k)$ such that $g'(x_k) = \frac{g(t_k) - g(t_{k-1})}{t_k - t_{k-1}}$.

$$\begin{aligned} g(b) - g(a) &= \sum_{k=1}^{n} [g'(t_k) - g'(t_{k-1})] \\ &= \sum_{k=1}^{n} g'(x_k)(t_k t_{k-1}) \\ &= \sum_{k=1}^{n} g'(x_k)|I_k| \end{aligned}$$

Then

$$\begin{gathered} \sum_{k=1}^{n} \min g'(x_k)|I_k| \le g(b) - g(a) \le \sum_{k=1}^{n} \max g'(x_k)|I_k| \\ \int_a^b \varphi \le g(b) - g(a) \le \int_a^b \psi \end{gathered}$$

Also $\int_a^b \varphi \le \int_a^b g' \le \int_a^b \psi$. Then $\left| g(b) - g(a) - \int_a^b g' \right| < \varepsilon$. So $\int_a^b g' = g(b) - g(a)$. ■

Theorem 10.2 (Fundamental Theorem of Calculus Part 2)**.** *Let f be bounded and integrable on $[a, b]$. Then f is integrable on $[a, x]$ for any $x \in [a, b]$. Let $F = \int_a^x f(t)\,\mathrm{d}t$. $F : [a, b] \to \mathbb{R}$. Then:*

- *F is Lipschitz (and thus uniformly) continuous on $[a, b]$.*
- *If f is continuous at x, then F is differentiable at x and $F'(x) = f(x)$.*

Proof. $\overline{F}(x) = \overline{\int_a^x} f$ is Lipschitz continuous and $\underline{F}(x) = \underline{\int_a^x} f$ is also Lipschitz continuous. If f is integrable on $[a, b]$, it is integrable on $[a, x]$ for all $x \in [a, b]$. So $\int_a^x f = \overline{\int_a^x} f = \underline{\int_a^x} f$ and therefore $F(x)$ is also Lipschitz continuous.

Next we want to show that $L = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = f(x)$. First, suppose h is positive. Then consider $L - f(x)$:

$$\begin{aligned} L - f(x) &= \lim_{h \to 0} \frac{\int_x^{x+h} f(t)\,\mathrm{d}t}{h} - f(x) \\ &= \frac{1}{h} \int_x^{x+h} f(t)\,\mathrm{d}t - \frac{1}{h}(hf(x)) \\ &= \frac{1}{h} \int_x^{x+h} f(t) - f(x)\,\mathrm{d}t \end{aligned}$$

f is continuous at x. Let $\varepsilon > 0$. Then $\exists \delta > 0$, such that $|t - x| < \delta \implies |f(t) - f(x)| < \frac{\varepsilon}{2}$. If $0 < h < \delta$, $\left| \frac{1}{h} \int_x^{x+h} f(t) - f(x)\,\mathrm{d}t \right| \le \frac{1}{h} \int_x^{x+h} |f(t) - f(x)|\,\mathrm{d}t < \varepsilon$.

Next, suppose $h < 0$. Then

$$\begin{aligned} L - f(x) &= \lim_{h\to 0} \frac{\int_{x+h}^{x} f(t)\,\mathrm{d}t}{h} - f(x) \\ &= \frac{1}{|h|}\int_{x+h}^{x} f(t)\,\mathrm{d}t - \frac{1}{|h|}(|h|f(x)) \\ &= \frac{1}{|h|}\int_{x+h}^{x} f(t) - f(x)\,\mathrm{d}t \end{aligned}$$

By the same property on continuity, let $\varepsilon > 0$, then $\exists \delta > 0$, such that $|t - x| < \delta \implies |f(t) - f(x)| < \frac{\varepsilon}{2}$. $\frac{1}{|h|}\int_{x+h}^{x} f(t) - f(x)\,\mathrm{d}t \leq \frac{1}{|h|}\int_{x+h}^{x} |f(t) - f(x)|\,\mathrm{d}t < \varepsilon$.

■

11 Aside

Some interesting things that may not be directly of relevance to the rest of the material.

11.1 Transcendentals

Let us define a function $\ln(x) = \int_1^x \frac{1}{t}\,\mathrm{d}t$. Since $\frac{1}{t}$ is continuous on $(0, \infty)$, $\frac{1}{t}$ is integrable on every $[1, x]$ $(x \geq 1)$, and every $[x, 1]$ $(0 < x \leq 1)$.

Then $\ln : (0, \infty) \to \mathbb{R}$. Also we know that ln is Lipschitz continuous on every $[a, b] \subset (0, \infty)$. This implies that it is continuous on $(0, \infty)$. ln is also differentiable and $(\ln x)' = \frac{1}{x}$.

Proposition 11.1. $\forall x, y \in (0, \infty)$

i. $\ln(xy) = \ln x + \ln y$.

Proof. Let $h(x) = \ln(xy) - \ln x - \ln y$. We want to show that $h(x) = 0$.
$h'(x) = \frac{1}{xy} \cdot y - \frac{1}{x} - 0 = 0$. Thus h is a constant. But $h(1) = 0$. So $h = 0$. ■

ii. $\ln(x^q) = q \ln x$, *where* $q \in \mathbb{Q}$.

Proof.
Case 1: $q \in \mathbb{N}$. Proof by induction: $\ln(x^{q+1}) = \ln(x^1) + \ln(x)$.

Case 2: $q = -1$. $\ln \frac{1}{x} + \ln x = \ln(1) = 0$. Then $\ln \frac{1}{x} = \ln x$.

Case 3: $q = -n$, $n \in \mathbb{N}$. Then $\ln \frac{1}{x^n} + \ln x^n = 0 \ldots$

Case 4: $q = \frac{1}{n}$, $n \in \mathbb{N}, n \neq 0$. Then $\ln x = \ln\big((x^{1/n})^n\big) = n \ln x^{1/n}$. So $\ln x^{1/n} = \frac{1}{n} \ln x$.

Case 5: $q = \frac{m}{n}$, $m, n \in \mathbb{N}, n \neq 0$. It follows from the above few cases. ■

Let $\lim_{n\to\infty}(1 + \frac{1}{n})^n = e$. We claim that $\ln e = 1$.

Proof.

$$\begin{aligned}
\ln e &= \ln\left(\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n\right)\\
&= \lim_{n\to\infty}\ln\left(1+\frac{1}{n}\right)^n\\
&= \lim_{n\to\infty} n\ln\left(\frac{n+1}{n}\right)\\
&= \lim_{n\to\infty} n(\ln(n+1)-\ln n)\\
&= \lim_{n\to\infty} n\left(\int_n^{n+1}\frac{1}{t}\,\mathrm{d}t\right)
\end{aligned}$$

$\frac{n}{n+1} \le n\ln(n+1) - n\ln(n) \le 1$. As $n\to\infty$, $\ln e = 1$. ■

$\ln : (0,\infty)\to\mathbb{R}$ is strictly increasing since $\frac{1}{x}$ is a positive function. It follows that ln is injective. Then there exists a function $\exp:\mathbb{R}\to(0,\infty)$ which is the inverse of ln.

Theorem 11.1 (Inverse function theorem)**.** *Let f be an injective function on an open interval I, then let $J=f(I)$. If f is differentiable at $x\in I$, and $f'(x)\neq 0$, then f^{-1} is also differentiable at $y=f(x)$ and $(f^{-1})'(y)=\frac{1}{f'(x)}$.*

$\exp:\mathbb{R}\to(0,\infty)$ is differentiable on its domain because its inverse has derivative $\frac{1}{x}$ and $\frac{1}{x}$ is never 0 on $(0,\infty)$. Then $\exp'(y)=\frac{1}{\ln'(x)}=\frac{1}{1/x}=\exp y$.

Suppose $b>0$ and $x\in\mathbb{R}$. We want to define b^x. We define $b^x := \exp(x\ln b)$. $x\mapsto b^x$ is continuous, and is also unique.

11.2 Takagi's function

Consider the following:

$$\begin{aligned}
\sigma(x) &= \operatorname{dist}(x,\mathbb{Z})\\
\sigma_1(x) &= \operatorname{dist}(x,2^{-1}\mathbb{Z}) = \frac{1}{2}\sigma(2x)\\
\sigma_k(x) &= \operatorname{dist}(x,2^{-k}\mathbb{Z}) = 2^{-k}\sigma(2^k x)
\end{aligned}$$

Definition 11.1. $T(x)=\sum_{k=0}^{\infty}\sigma_k(x)$ is Takagi's function[11] □

Firstly, this is continuous. $|\sigma_k(x)|\le 2^{-(k+1)}$. Since $\sum_{k=0}^{\infty}2^{-(k+1)}$ converges, $T(x)$ converges uniformly on $\mathbb{R}$. Since each of the terms are continuous, by uniform continuity, T is continuous. An interesting and perhaps surprising fact is that despite it being continuous, it is not differentiable!

[11] It is also called the *Blancmange* curve due to its resemblance to the dessert.

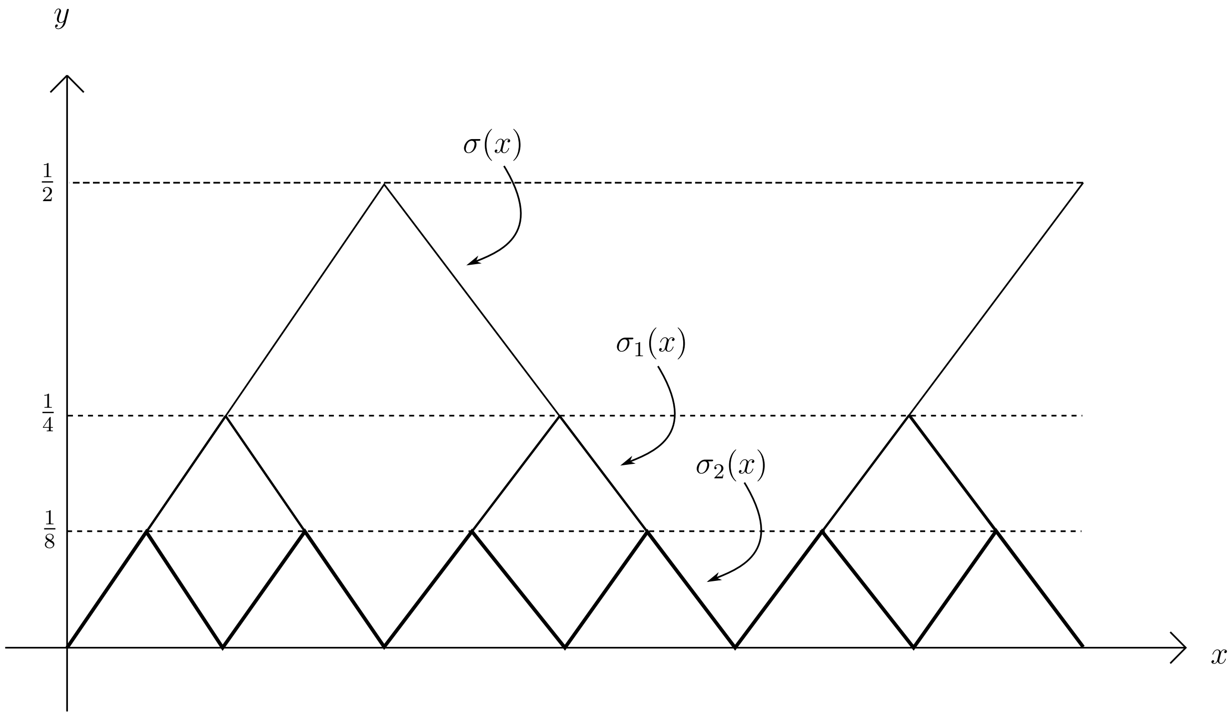

Figure 2: A sketch of a few σ's

Theorem 11.2. *$\forall x \in \mathbb{R}$, $T(x)$ is not differentiable at x.*

Proof. Take any $x \in \mathbb{R}$. For every $n \geq 0$, we can find $\frac{a}{2^n} \leq x < \frac{n+1}{2^n}$. Let $u_n = \frac{a}{2^n}$ and $v_n = \frac{a+1}{2^n}$. As $n \to \infty$, $u_n \to x \leftarrow v_n$. If T was differentiable at x, then $\frac{T(v_n)-T(u_n)}{v_n-u_n}$ will have to tend to some limit as $n \to \infty$. But

$$\frac{T(v_n) - T(u_n)}{v_n - u_n} = \sum_{k=0}^{\infty} \frac{\sigma_k(v_n) - \sigma_k(u_n)}{v_n - u_n} = \sum_{k=0}^{n-1} \frac{\sigma_k(v_n) - \sigma_k(u_n)}{v_n - u_n}$$

because $u_n, v_n \in 2^{-n}\mathbb{Z}$. But the terms are always ± 1, so the series cannot converge. ■

Intuitively it may be explained as such: though the curve is continuous, yet it is infinitely "spiky". You will never be able to find the gradient at a certain point because the spikiness throws off your secant line as you keep trying to make it shorter and shorter.

12 Sequence of functions

Fix a domain $D \subseteq \mathbb{R}$. A sequence of functions on D is $(f_0, f_1, f_2, \ldots)_{n=0}^{\infty}$ where $\forall n : f_n : D \to \mathbb{R}$. For all points $p \in D$ we can consider $(f_n(p))_{n=0}^{\infty}$.

Definition 12.1. Let $(f_n)_n$ be a sequence of functions $D \to \mathbb{R}$. Let $f : D \to \mathbb{R}$. We say that $(f_n)_n$ converges pointwise to f if $\forall p \in D : \lim_{n\to\infty} f_n(p) = f(p)$. i.e.

$$\forall p \in D, \forall \varepsilon > 0, \exists N, \forall n : (n \geq N) \implies |f_n(p) - f(p)| < \varepsilon$$

□

Example 12.1. $(x, x^2, x^3, \ldots)$ on $[0, 1]$. If $0 \leq p < 1$ then $p^n \to 0$ as $n \to \infty$. If $p = 1$ then $p^n \to 1$ as $n \to \infty$. Then the pointwise limit of this sequence of functions on $[0, 1]$ is

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

Here, every term is continuous, but the pointwise limit is not. □

Example 12.2. Let $f_n : [0, 1] \to \mathbb{R}$ be the function whose graph makes a triangle with base $[0, 2^{-n}]$ and height 2^{n+1}. See Figure 3.

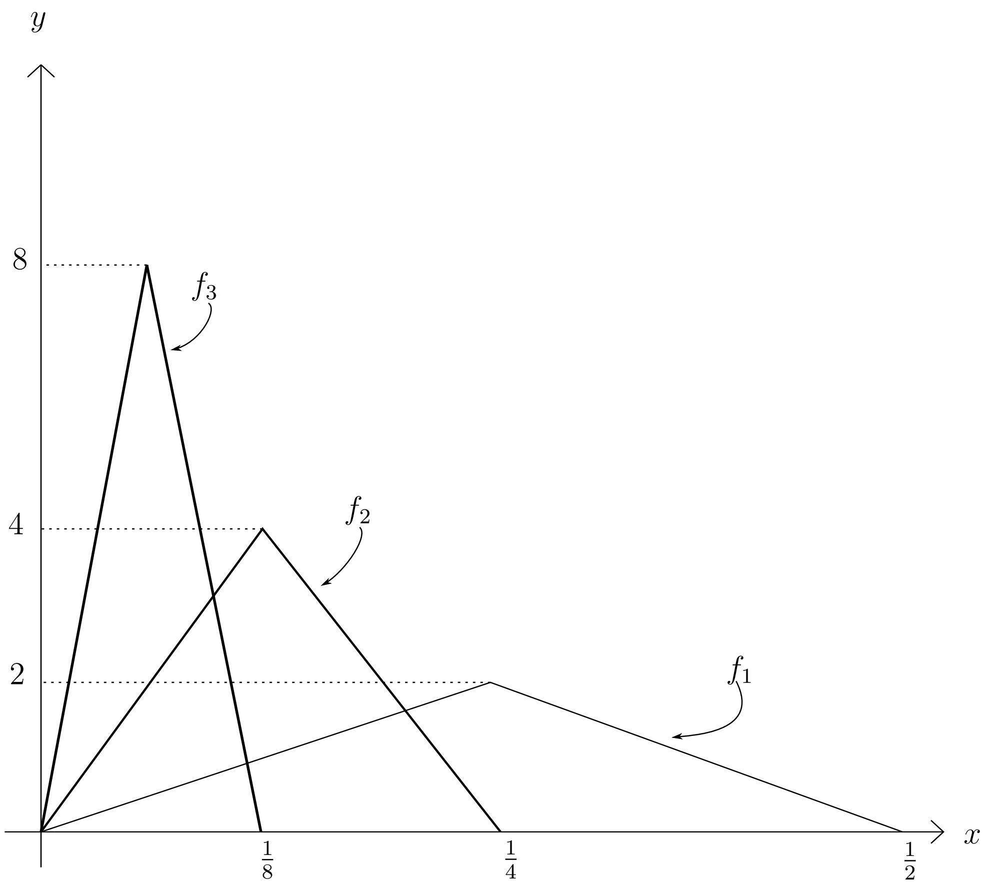

Figure 3: A sketch of a few successive elements of the series.

Then $\lim f_n(p) = 0$ for all p, because of $p > 0$ then eventually $2^{-n} < p$ so $f_n(p) = 0$ after that point. Here $\int_0^1 f_n = 1$ but $\int_0^1 f = 0$ □

Example 12.3. Let $(q_n)_{n=0}^{\infty}$ be an enumeration of $\mathbb{Q} \cap [0, 1]$: $(0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \ldots)$, similar to that shown in Figure 1.

Let $f_n(x) = \begin{cases} 1, & \text{if } x = q_k, k \leq n \\ 0, & \text{otherwise} \end{cases}$.

Here $\int_0^1 f_n = 0$, but the limit is not integrable. □

This perhaps suggests that we need a stronger form of convergence.

Definition 12.2 (Uniform convergence)**.**

$$\forall \varepsilon > 0, \exists N, \forall p \in D, \forall n : (n \geq N) \implies |f_n(p) - f(p)| < \varepsilon$$

□

Theorem 12.1. *If $(f_n) \to f$ uniformly and all f_n's are continuous, then f is also continuous. The uniform limits of continuous functions are continuous.*

Proof. First we make the observation that $|f(x) - f(p)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(p)| + |f_n(p) - f(p)|$.

Let $\varepsilon > 0$. Choose N large enough such that $\forall y \in D, n \geq N \implies |f_n(y) - f(y)| < \frac{\varepsilon}{3}$. By continuity of f_n at p, $\exists \delta > 0$ such that $|x - p| < \delta \implies |f_n(x) - f_n(p)| < \frac{\varepsilon}{3}$. Then $|f(x) - f(p)| < \varepsilon$. ■

Theorem 12.2. *If $(f_n)_n$ is a sequence of integrable functions on $[a, b]$ and $(f_n)_n \to f$ uniformly on $[a, b]$, then f is integrable and $\int_a^b f = \lim_{n\to\infty} \int_a^b f_n$.*

Proof. Let $\varepsilon > 0$. Uniform convergence means:

$$\exists N, \forall n \geq N, \forall x \in [a, b] : |f_n(x) - f(x)| \leq \frac{\varepsilon}{2(b-a)}$$
$$\forall x \in [a, b] : f_n(x) - \frac{\varepsilon}{2(b-a)} \leq f(x) \leq f_n(x) + \frac{\varepsilon}{2(b-a)}$$
$$\int_a^b f_n(x) - \frac{\varepsilon}{2} = \int_a^b f_n(x) - \frac{\varepsilon}{2(b-a)} \leq \underline{\int_a^b} f(x) \leq \overline{\int_a^b} f_n(x) \leq \int_a^b f_n + \frac{\varepsilon}{2(b-a)} = \int_a^b f_n + \frac{\varepsilon}{2}$$

Therefore, $\forall \varepsilon > 0$, $\overline{\int_a^b} f - \underline{\int_a^b} f \leq \varepsilon$. Therefore $\overline{\int_a^b} f = \underline{\int_a^b} f$.

Integrating,

$$f(x) - \frac{\varepsilon}{2(b-a)} \leq f_n(x) \leq f(x) + \frac{\varepsilon}{2(b-a)}$$
$$\int_a^b f - \frac{\varepsilon}{2} \leq \int_a^b f_n \leq \int_a^b f + \frac{\varepsilon}{2}$$

Therefore $\left|\int_a^b f_n - \int_a^b f\right| \leq \frac{\varepsilon}{2} < \varepsilon$, and $\lim_{n\to\infty} \int_a^b f_n = \int_a^b f$. ■

Theorem 12.3. *Suppose $(f_n)_n$ on (a, b) with the properties:*

i. All differentiable on (a, b)

ii. $(f_n)_n \to f$ uniformly.

iii. $(f'_n)_n \to g$ uniformly.

Then f is differentiable and $f' = g$.

13 Series of functions

Example 13.1. $\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!} = \sin(x)$ for all x because the series converges to $\sin(x)$ pointwise on $\mathbb{R}$. □

Definition 13.1. We say that $\sum_{k=0}^{\infty} t_k(x)$ converges pointwise/uniformly to $F(x)$ if

$$\left(\sum_{k=0}^{\infty} t_k(x)\right)_{n=0}^{\infty}$$

converges pointwise/uniformly to $F(x)$ on D. □

Theorem 13.1 (Weierstrass M-test)**.** *If $\exists (M_k)_{k=0}^{\infty}$ of non-negative real numbers such that $\sum_{k=0}^{\infty} M_k$ converges absolutely, and $\forall k, |t_k(x)| \leq M_k$ when $x \in D$, then $\sum_{k=0}^{\infty} t_k(x)$ converges uniformly to some function.*

Proof. An extension of Definition 6.1 follows.

Suppose $(f_n)_n$ is a sequence of functions. It converges uniformly to some function if $\forall \varepsilon > 0, \exists N, \forall x \in D, \forall m, n : (m, n \geq N) \implies |f_n(x) - f_m(x)| < \varepsilon$.

For a series of functions:

$$\forall \varepsilon > 0, \exists N, \forall x \in D, \forall m, n : (N \leq m \leq n) \implies \left|\sum_{k=m}^{n} t_k(x)\right| < \varepsilon$$

Let $\varepsilon > 0$, by the Cauchy criterion for convergence of $\sum_{k=0}^{\infty} M_k$, $\exists N : \forall m, n, (N \leq m \leq n) \implies |\sum_{k=m}^{n} M_k| < \varepsilon \implies \sum_{k=m}^{n} M_k < \varepsilon$. Let $N \leq m \leq n$. Then

$$\left|\sum_{k=m}^{n} t_k(x)\right| \leq \sum_{k=m}^{n} |t_k(x)| \leq \sum_{k=m}^{n} M_k < \varepsilon$$

■

Example 13.2. Power series $F(x) = \sum_{n=0}^{\infty} \frac{x^n}{3^n (n+1)^2}$.

We can perform the ratio test to find out where this is defined.

$$R(x) = \lim_{n \to \infty} \left|\frac{t_{n+1}(x)}{t_n}\right| = \ldots = \frac{|x|}{3}$$

The series converges absolutely if $\frac{|x|}{3} < 1$ or $|x| < 3$. □

Proposition 13.1. *Let $\sum_{c=0k}^{\infty} (x - \gamma)^k$ centred at γ. Let R be the radius of convergence ($\frac{1}{R} = \limsup_{k \to \infty} |c_k|^{\frac{1}{k}}$).*

i. The series converges absolutely in $\gamma - R, \gamma + R$.

ii. The series diverges outside the interval.

iii. *It could have any behaviour at* $\gamma \pm R$.

iv. *By the Weierstrass M-test, if* $[a, b] \subset (\gamma - R, \gamma + R)$*, the series converges uniformly on* $[a, b]$.

v. *It is differentiable and integrable within the radius.*

Example 13.3. Consider the previous example by on the interval $[-2, 2]$.

$$\left|\frac{x^n}{3^n(n+1)^2}\right| \leq \frac{x^n}{3^n(n+1)^2} < \left(\frac{2}{3}\right)^n = M_n$$

By the Weierstrass M-test, the series converges uniformly on $[-2, 2]$. □

14 Measure Theory and Lebesgue Integration

What is wrong with the Riemann integral?

- The integral is only possible defined if the domain of the function is bounded.
- If f is bounded then upper and lower integrals exist.
- f is only integrable when the upper and lower integrals match.
- $(f_n) \to f$, $\lim \int_a^b f_n = \int_a^b f$ only when convergence is uniform.

14.1 Measure Theory

A *measure space* is a set X equipped with a notion of how big subsets of X are.

A measure space (X, Σ, μ):

- X is a set
- Σ is a set of subsets of X called *measurable* sets, closed under countable unions, countable intersections, and complements. $\emptyset \in \Sigma$, $X \in \Sigma$. This is also called a σ-algebra.
- μ, the measure, $\mu : \Sigma \to [0, \infty]$, satisfying
  - $\mu(\emptyset) = 0$.
  - If $(E_1, E_2, \ldots)$, with $E_n \in \Sigma$ for all n, is a sequence of pairwise disjoint sets, then $\mu(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k)$.

Proposition 14.1. *If* $A, B \in \Sigma$*, and* $A \subseteq B$*, then* $\mu(A) = \mu(B)$.

Proof.

$$B = A \cup (B \setminus A) = A \cup (B \cap A^c)$$
$$\mu(B) = \mu(A) + \mu(B \cap A^c) \geq \mu(A)$$

■

Example 14.1 (Counting measure)**.** $(X, \Sigma, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \text{number of elements})$ [12]. $\square$

Example 14.2 (Lebesgue measure)**.** $(X, \Sigma, \mu) = (\mathbb{R}, \sigma, \lambda)$

- λ is the best possible measure that agrees with a naive notion of length $(\sup(I) - \inf(I))$. $\lambda(x + E) = \lambda(E)$ for $x \in \mathbb{R}$ and $E \in \Sigma$. This is also what is called "translation invariant".

$\square$

Example 14.3 (Dirac measure)**.** Let $p \in \mathbb{R}$, then $\delta_p(E) = \begin{cases} 0, & \text{if } p \notin E \\ 1 & \text{if } p \in E \end{cases}$ $\square$

Example 14.4. The following subsets of $\mathbb{R}$ are measurable:

- All intervals, finite or countable unions or intersections.
- The Cantor Set (Ex 5.3). It is made up of countable intersections and finite unions of intervals.
  - If $C_0 \supset C_1 \supset C_2 \supset \dots$ and $C_\infty = \bigcap_{n=0} C_n$ then $\lambda(C_\infty) = \lim_{n\to\infty} \lambda(C_n) = \lim_{n\to\infty} \left(\frac{2}{3}\right)^n = 0$.
  - In fact, if E is countable, then $\lambda(E) = 0$. This is because you could enumerate E with say $(e_k)_{k=0}^{\infty}$, and since E is a disjoint countable union of e_k's, $\lambda(E) = \sum_{k=0}^{\infty} \lambda([e_k, e_k]) = 0$.

$\square$

14.2 Integration with measures

Consider a function, with X as a measure space, $f : X \to \mathbb{R}$. f is called *measurable* if

$$\forall c \in \mathbb{R}, f^{-1}((c, \infty)) = \{x \in X \mid f(x) > c\}$$

is a measurable subset of $X \in \Sigma$.

Consider a simple function $s : X \to \mathbb{R}$ such that s is a non-negative linear combination of indicators of measurable sets:

$$s(x) = \sum_{k=1}^{n} c_k \mathbb{1}_{E_k}(x)$$

where $c_k > 0, E_k \in \Sigma$. Then, if D is measurable, we define $\int_D s \,\mathrm{d}\mu = \sum_{k=1}^{n} c_k \mu(E_k \cap D)$.

If f is measurable and $f \geq 0$ on D, we define

$$\int_D f \,\mathrm{d}\mu = \sup \left\{ \int_D s \,\mathrm{d}\mu \mid s \text{ is simple and } 0 \leq s \leq f \right\}$$

If $f : D \to \mathbb{R}$ is any measurable function, $f = f^+ - f^-$ [13], with f^+, f^- both being measurable. We attempt to define $\int_D f \,\mathrm{d}\mu = \int_D f^+ \,\mathrm{d}\mu \int_D f^- \,\mathrm{d}\mu$.

[12] Here $\mathcal{P}$ refers to the power set, or the set of all subsets.

[13] Here $f^+ = |f| + f$ and $f^- = |f| - f$, the positive and negative components of f respectively.

- If $\int f^+$ and $\int f^- \in \mathbb{R}$, then f is μ-integrable.
- If $\int f = \pm\infty$, then f is not integrable.
- If $\int f =$"$\infty - \infty$", then the integral is not defined.

Example 14.5 (Integrating the Dirichlet function)**.** Previously we could not integrate $\mathbb{1}_{\mathbb{Q}}(x)$ since $\overline{\int_0^1} \mathbb{1}_{\mathbb{Q}} \neq \underline{\int_0^1} \mathbb{1}_{\mathbb{Q}}$. However, $\int_{[0,1]} \mathbb{1}_{\mathbb{Q}} \,\mathrm{d}\lambda = \lambda(\mathbb{Q} \cap [0,1]) = 0$. □

Some remarks about Lebesgue integration.

- It satisfies the desired properties of an integral (linearity, additivity, monotonicity, triviality).
- If $f : [a,b] \to \mathbb{R}$ is Riemann integrable, then it is Lebesgue integrable and $\int_a^b f(x)\,\mathrm{d}x = \int_{[a,b]} f \,\mathrm{d}\lambda$.
- $\lim \int f_n = \int \lim f_n$.

Theorem 14.1 (Monotone/dominated convergence theorem)**.** *Suppose $(f_n)_n \to f$ pointwise, and all f_n are Lebesgue integrable. Then f is Lebesgue integrable and $\int_D f \,\mathrm{d}\lambda = \lim_{n\to\infty} \int_D f_n \,\mathrm{d}\lambda$ if*

- *Monotonic: $0 \leq f_1 \leq f_2 \leq \ldots \leq f$ on D, or*
- *Dominated: If there exists a Lebesgue integrable g and $|f_n| \leq g$.*

Example 14.6 (Vitali's construction)**.** Equip $\mathbb{R}$ with a relation $\sim$. We say $x \sim y$ if $x - y \in \mathbb{Q}$. It is reflexive, symmetric, and transitive. We claim $\mathbb{R} =$ the union of $\sim$ equivalence classes. The equivalence class of x, denoted by $[x]_\sim = \{y \in \mathbb{R} \mid x \sim y\}$. If $x \not\sim t$, then $[x]_\sim \cap [t]_\sim$ and $\mathbb{R} = \bigcup_{x\in\mathbb{R}} [x]_\sim$. Note that each $[x]_n$ is countable.

It follows that there are uncountably many $\sim$ equivalence classes. Let V_0 be a choice of representatives from each equivalence class with $0 \in V_0$ and $V_0 \in [0,1]$. $\forall x \in \mathbb{R}, \exists! y \in V_0 : x \sim y$. Also we note that V_0 can only be defined, not constructed, but it requires the Axiom of Choice.

Let $(q_k)_{k=0}^\infty$ be an enumeration of $\mathbb{Q} \cap [-1,1], q_0 = 0$. Let $V_k = q_k + V_0$.

i. $[0,1] \subseteq \bigcup_{k=0}^\infty V_k \subseteq [-1,2]$.
   If $x \in \bigcup V_k$ then $x = q_k + v$ for $q_k \in \mathbb{Q} \cap [-1,1]$ and $v \in V_0$ and $v \in [0,1]$. Let $t \in [0,1]$. Because V_0 is a set of representatives for the $\sim$ equivalence classes, $\exists! v \in V_0 : t \sim v$. Then $t - v \in \mathbb{Q} \cap [-1,1]$. Thus $t - v = q_k$ for some k. Then $t = v + q_k \in q_k + V_0 = V_k$.

ii. All V_k are disjoint. If they overlapped at x then $x = q_j + v \in V_0 = q_k + w \in V_0$, contradiction.

iii. $\lambda(\bigcup_{k=0}^\infty V_k) = \sum_{k=0}^\infty \lambda(V_k)$.

$$1 \leq \sum_{k=0}^{\infty} \lambda(q_k + V_0) \leq 3$$

$$1 \leq \sum_{k=0}^{\infty} \lambda(V_0) \leq 3$$

This shows that $\lambda(V_0)$ does not exist since the sum will have to either be 0 or infinite. Hence V_0 is not measurable.

$\square$