

18.100B Notes

1 Real Numbers

Definition 1.1. A *field* is a set F equipped with operations $+$ and \times such that

- $(F, +)$ and $(F \setminus \{0\}, \times)$ are Abelian groups.
- $x(y + z) = xy + xz$ for all $x, y, z \in F$. (Distributivity)

Properties:

1. Additive identity is unique.
2. $x \cdot 0 = 0$ ($\forall x \in F$).
3. $x, y \neq 0 \Rightarrow xy \neq 0$ (i.e. fields are *integral domains*).

Example 1.1.

1. The set $\mathbb{N} = \{0, 1, 2, \dots\}$ of natural numbers is not a field.
2. The set \mathbb{Z} of integers is an Abelian additive group but not a field.
3. The set \mathbb{Q} of rationals is a field.
4. The binary field $\mathbb{Z}/2\mathbb{Z} = \{0, 1\}$ with mod 2 operations is a field.
5. $\mathbb{Z}/4\mathbb{Z}$ is not a field, e.g. $2 \cdot 2 = 0$ but $2 \neq 0$.

Definition 1.2. A field F is *ordered* if there exists a relation $<$ on F (with $x > y$ meaning $y < x$, $x \leq y$ meaning $x < y$ or $x = y$, etc) such that for all $x, y, z \in F$,

- Exactly one of $x = y$, $x < y$, $x > y$ holds. (Trichotomy)
- $x < y$ and $y < z$ implies $x < z$. (Transitivity)
- $x < y$ implies $x + z < y + z$. (Additivity)
- $x < y$ and $z > 0$ implies $xz < yz$. (Multiplicativity)

We define $P = \{x \in F : x > 0\}$.

Properties:

1. $x > y \Rightarrow x - y \in P$.
2. $x^2 \geq 0$ for all $x \in F$.
3. $x > 0 \Rightarrow x^{-1} > 0$. (*Hint: First prove $1 > 0$.*)

Definition 1.3. Let F be an ordered field.

- $u \in F$ is an *upper bound* for a subset $S \subseteq F$ if $x \leq u$ for all $x \in S$. If an upper bound for S exists, we say S is *bounded above*.
- $\ell \in F$ is a *lower bound* for a subset $S \subseteq F$ if $x \geq \ell$ for all $x \in S$. If an upper bound for S exists, we say S is *bounded below*.
- If $S \subseteq F$ is bounded above and below, we say that it is *bounded*.
- $u \in F$ is the *maximum* of S , denoted $\max S$, if u is an upper bound and $u \in S$.
- $\ell \in F$ is the *minimum* of S , denoted $\min S$, if ℓ is a lower bound and $\ell \in S$.
- $u \in F$ is the *supremum* of S , denoted $\sup S$, if it is the least upper bound for S . More precisely, we say that S has supremum

$$\sup S = \min\{x \in F : x \text{ is an upper bound for } S\} \quad \text{if it exists.}$$

- $\ell \in F$ is the *infimum* of S , denoted $\inf S$, if it is the greatest lower bound for S . More precisely, we say that S has infimum

$$\sup S = \max\{x \in F : x \text{ is an lower bound for } S\} \quad \text{if it exists.}$$

- By convention, $\inf \emptyset = \infty$ and $\sup \emptyset = -\infty$. If S is unbounded above (below) we say $\sup S = \infty$ ($\inf S = -\infty$).
- We say that F is *complete* if it satisfies the *completeness axiom*: Every nonempty subset of F that is bounded above has a supremum.

Example 1.3.

1. $\{x \in \mathbb{Q} : x < 1\}$ has upper bounds but no maximum.
2. $\{x \in \mathbb{Q} : x \geq 1\}$ has no upper bounds but has a minimum.
3. $\{x \in \mathbb{Q} : x^2 < 2\}$ is bounded above but has no supremum.

Theorem 1.1. The set \mathbb{R} of real numbers is the unique complete ordered field.

No proof. To prove this we have to prove existence and uniqueness. Two ways for existence: via Dedekind cuts or via rational Cauchy sequences.

Example 1.4.

1. \mathbb{Q} is ordered but not complete (see previous example).
2. $[0, 1] := \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ has maximum 1.
3. $(0, 1) := \{x \in \mathbb{R} : 0 < x < 1\}$ has supremum 1 but no maximum.

Theorem 1.2. (Existence of $\sqrt{2}$) There exists $r \in \mathbb{R}$ with $r^2 = 2$.

Proof. Set $S = \{x \in \mathbb{R} : x^2 < 2\}$. We first prove a lemma:

Lemma. If $v > 0$ and $v^2 \geq 2$, then v is an upper bound for S .

Proof. Let $x \in S$ be any element. If $x < 0$ then $x < 0 < v$. If $x \geq 0$, we have that $x^2 < 2 \leq v^2$. So $0 < v^2 - x^2 = (v - x)(v + x) \implies 0 < v - x$. \square

Since $5 > 0$ and $5^2 \geq 2$, S is bounded above. Therefore there is a supremum $u = \sup S$.

- If $u^2 > 2$, set $a = \frac{u^2 - 2}{2u} > 0$. Then $u - a = \frac{u^2 + 2}{2u} > 0$ and

$$(u - a)^2 = u^2 - 2ua + a^2 = 2 + a^2 > 2$$

so $u - a$ is a lower upper bound for S than u , a contradiction.

- If $u^2 < 2$, set $a = \frac{2 - u^2}{5} > 0$. Since 2 is an upper bound for S , we have $0 < u < 2$. Also $u + a > u$ and

$$(u + a)^2 = u^2 + 2ua + a^2 < 2 + 4a + a = 2$$

so $u + a > u$ is in S , a contradiction.

Therefore, by trichotomy, $u^2 = 2$. ■

Theorem 1.3. (Archimedean Property) Let x, y be reals. Then

- A) $y > 0 \implies \exists n \in \mathbb{N}$ such that $ny > x$.
- B) $x < y \implies \exists q \in \mathbb{Q}$ such that $x < q < y$. (\mathbb{Q} is dense in \mathbb{R})

Proof.

- A) We show that $S = \{ny : n \in \mathbb{N}\}$ has no upper bound. Assume not, then $z = \sup S$ exists. Since $z - y < z$ is not an upper bound for S , there exists $z - y < ny \in S$. But then $z < (n + 1)y \in S$, contradicting the fact that z is an upper bound for S .

B) Pick $n \in \mathbb{N}^*$ such that $n(y - x) > 1$ using part A. Another useful lemma:

Lemma. For every $x \in \mathbb{R}$ there exists $n \in \mathbb{Z}$ such that $n - 1 \leq x < n$.

Proof. By part A applied to x and $-x$, we can find two integers $a, b \in \mathbb{Z}$ such that $a < x < b$. Since $\{n \in \mathbb{Z} : x < n \leq b\} \subseteq \{n \in \mathbb{Z} : a \leq n \leq b\}$ is finite, there exists a minimum $n \in \mathbb{Z}$ such that $x < n$. This gives $n - 1 \leq x < n$. \square

With $m - 1 \leq nx < m$ ($m \in \mathbb{Z}$), we get $nx < m \leq nx + 1 < ny \implies x < \frac{m}{n} < y$. \blacksquare

Theorem 1.4. (Principle of Induction) For a property $P(n)$ ($n \in \mathbb{N}$), if $P(0)$ and $P(n) \implies P(n + 1)$ ($n \in \mathbb{N}$) are true, then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. Assume for contradiction that there exists some $k \in \mathbb{N}$ such that $P(k)$ is false. Then

$$\{n \in \mathbb{N} : P(n) \text{ is false and } n \leq k\}$$

is a non-empty finite set. Hence it has a minimum element m . Then $m > 0$ ($P(0)$ is true), and thus $m - 1 \in \mathbb{N}$ and $P(m - 1)$ is true. But $P(m - 1) \implies P(m)$, a contradiction. \blacksquare

Exercise. Prove that $(1 + x)^n \geq 1 + nx$ for all $x \geq -1$ and $n \in \mathbb{N}$.

2 Sequences

Definition 2.1. The absolute value function is defined by

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

Theorem 2.1. (Triangle Inequality) $|x + y| \leq |x| + |y|$ for all $x, y \in \mathbb{R}$.

Proof.

- If $x, y \geq 0$, then $x + y \geq 0$, so $|x + y| = x + y = |x| + |y|$.
- If $x, y < 0$, then $x + y < 0$, so $|x + y| = -x - y = (-x) + (-y) = |x| + |y|$.
- If $x < 0 \leq y$, then $|x| + |y| = -x + y$. Note that $-x + y \geq -x - y$ and $-x + y \geq x + y$ are both true, so $|x| + |y| \geq |x + y|$ regardless. The case $y < 0 \leq x$ is analogous. ■

Definition 2.2. A sequence $\{x_n\}_{n \in \mathbb{N}} = \{x_0, x_1, \dots\}$ is an ordered list of real numbers. Explicitly, we have a function $x : \mathbb{N} \rightarrow \mathbb{R}$ and we denoted $x_n = x(n)$.

Example 2.1. The following are sequences:

1. $x_n = n^2$ for $n \in \mathbb{N}$.
2. $x_n = 1/n$ for $n \in \mathbb{N}^*$ (here the sequence is $\{x_n\}_{n \geq 1}$).
3. (Arithmetic progression) $\{x_n\}_{n \in \mathbb{N}}$ satisfying $\begin{cases} x_n = x_{n-1} + a, & n \geq 1 \\ x_0 = b \end{cases}$

Definition 2.3. Let $\{x_n\}_{n \in \mathbb{N}}$ is said to *converge* to $\ell \in \mathbb{R}$ if

$$(\forall \varepsilon > 0) (\exists N \in \mathbb{N}) (\forall n \geq N) (|x_n - \ell| < \varepsilon)$$

If this is true, we write $\lim_{n \rightarrow \infty} x_n = \ell$.

Example 2.2. $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$.

Proof. Let $\varepsilon > 0$ be arbitrary. Pick $N > 1/\varepsilon$ (Archimedean Property). For all $n \geq N$,

$$\left| \frac{1}{n} - 0 \right| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon. \quad \blacksquare$$

Theorem 2.2. If a sequence $\{x_n\}_{n \in \mathbb{N}}$ converges to both ℓ and ℓ' , then $\ell = \ell'$.

Proof. Given $\varepsilon > 0$,

- $\exists N_1 \in \mathbb{N}$ such that $|x_n - \ell| < \varepsilon/2$ for all $n \geq N_1$.
- $\exists N_2 \in \mathbb{N}$ such that $|x_n - \ell'| < \varepsilon/2$ for all $n \geq N_2$.

Then for any $n \geq \max\{N_1, N_2\}$,

$$|\ell - \ell'| \leq |\ell - x_n| + |x_n - \ell'| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \blacksquare$$

Exercise. If $\lim_{n \rightarrow \infty} x_n = \ell$ and $c \in \mathbb{R}$, then $\lim_{n \rightarrow \infty} cx_n = c\ell$ and $\lim_{n \rightarrow \infty} (x_n + c) = \ell + c$.

Definition 2.4. $\{x_n\}_{n \in \mathbb{N}}$ is *bounded* if $\exists M \in \mathbb{R}$ such that $|x_n| < M$ for all $n \in \mathbb{N}$.

Exercise. A converging sequence is bounded.

Theorem 2.3. If $\lim_{n \rightarrow \infty} x_n = \ell$ and $\lim_{n \rightarrow \infty} y_n = \ell'$, then

- $\lim_{n \rightarrow \infty} (x_n + y_n) = \ell + \ell'$
- $\lim_{n \rightarrow \infty} (x_n y_n) = \ell \ell'$
- if $\ell \neq 0$ and $x_n \neq 0$ for all $n \in \mathbb{N}$, $\lim_{n \rightarrow \infty} (x_n + y_n) = 1/\ell$

Proof of Second Point. Let $L = \max(|\ell|, |\ell'|)$. Given $\varepsilon > 0$, there exists N such that

$$|x_n - \ell| < \min\left(\frac{\varepsilon}{3L}, L\right) \quad \text{and} \quad |y_n - \ell'| < \min\left(\frac{\varepsilon}{3L}, L\right)$$

for all $n \geq N$. Then, for all $n \geq N$,

$$\begin{aligned} |x_n y_n - \ell \ell'| &= |(x_n - \ell)(y_n - \ell') + \ell(y_n - \ell') + \ell'(x_n - \ell)| \\ &\leq |(x_n - \ell)(y_n - \ell')| + |\ell| |y_n - \ell'| + |\ell'| |x_n - \ell| \\ &< \frac{\varepsilon}{3L} \cdot L + L \cdot \frac{\varepsilon}{3L} + L \cdot \frac{\varepsilon}{3L} = \varepsilon. \quad \blacksquare \end{aligned}$$

Definition 2.5. $\{x_n\}_{n \in \mathbb{N}}$ is said to *diverge to ∞* , written as $x_n \rightarrow \infty$, if for all $M \in \mathbb{R}$ there exists $N \in \mathbb{N}$ such that $x_n \geq M$ for all $n \geq N$. The case $x_n \rightarrow -\infty$ is analogous.

Exercise. If $x_n \leq y_n$ for all $n \in \mathbb{N}$, then $\lim_{n \rightarrow \infty} x_n \leq \lim_{n \rightarrow \infty} y_n$ if both limits exist.

Theorem 2.4. (Squeeze Theorem) If $\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = \ell$ and $x_n \leq z_n \leq y_n$ for all $n \in \mathbb{N}$, then $\lim_{n \rightarrow \infty} z_n = \ell$.

Proof. Since $\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = \ell$, for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that

$$|x_n - \ell| < \varepsilon \quad \text{and} \quad |y_n - \ell| < \varepsilon$$

for all $n \geq N$. Therefore for $n \geq N$,

$$\ell - \varepsilon < x_n \leq z_n \leq y_n < \ell + \varepsilon \implies |z_n - \ell| < \varepsilon. \quad \blacksquare$$

Exercise. $\lim_{n \rightarrow \infty} \frac{\sin n}{n} = 0$.

Definition 2.6. $\{x_n\}_{n \in \mathbb{N}}$ is *monotone* if it is either nonincreasing ($x_n \geq x_{n+1}$ for all $n \in \mathbb{N}$) or nondecreasing ($x_n \leq x_{n+1}$ for all $n \in \mathbb{N}$).

Theorem 2.5. (Monotone Convergence Theorem) If $\{x_n\}_{n \in \mathbb{N}}$ is nondecreasing and bounded above, then it converges. Similarly, if it is nonincreasing and bounded below, then it converges.

Proof. Since $\{x_n\}_{n \in \mathbb{N}}$ is bounded above, $\{x_n : n \in \mathbb{N}\}$ has an upper bound, with a supremum ℓ . Then for any $\varepsilon > 0$, there exists some $x_N > \ell - \varepsilon$, which means, for all $n \geq N$,

$$\ell - \varepsilon < x_N \leq x_n \leq \ell \implies |x_n - \ell| < \varepsilon. \quad \blacksquare$$

Worked Example. The sequence defined by $\begin{cases} x_0 = \sqrt{2} \\ x_{n+1} = \sqrt{2 + x_n} \end{cases} \quad n \geq 0$ converges.

Proof. We first prove by induction that $x_n \leq x_{n+1} \leq 2$ for all $n \in \mathbb{N}$. For $n = 0$,

$$x_0 = \sqrt{2} \leq \sqrt{2 + \sqrt{2}} = x_1 \leq \sqrt{2 + \sqrt{4}} = 2.$$

If $x_{n-1} \leq x_n \leq 2$, then

$$x_n = \sqrt{2 + x_{n-1}} \leq \sqrt{2 + x_n} = x_{n+1} \leq \sqrt{2 + 2} = 2.$$

Therefore $\{x_n\}_{n \in \mathbb{N}}$ is non-decreasing and bounded above by 2. Hence it converges to some $\ell \in \mathbb{R}$. Extra: *How to find ℓ ?* If we apply the limit on both sides of $x_{n+1} = \sqrt{2 + x_n}$,

$$\begin{aligned}\lim_{n \rightarrow \infty} x_{n+1} &= \lim_{n \rightarrow \infty} \sqrt{2 + x_n} \\ \ell &= \sqrt{2 + \ell} \\ \ell &= -1 \text{ or } 2\end{aligned}$$

Since all $x_n \geq 0$, we must have $\ell = 2$. ■

Definition 2.7. A *subsequence* of $\{x_n\}_{n \in \mathbb{N}}$ is any ordered infinite subset. Precisely, it is some $\{x_{n_j}\}_{j \in \mathbb{N}}$ where $n_0 < n_1 < n_2 < \dots$ are natural numbers.

Exercise. If $\lim_{n \rightarrow \infty} x_n = \ell$ then every subsequence converges to ℓ .

Theorem 2.6. Every sequence $\{x_n\}_{n \in \mathbb{N}}$ admits a monotone subsequence.

Proof. For each $m \in \mathbb{N}$ say x_m is a *tail-major* if $x_m \geq x_n$ for all $n \geq m$. If $\{x_n\}_{n \in \mathbb{N}}$ has infinitely many tail-majors, the subsequence of tail-majors is a non-increasing subsequence. Otherwise, there are finitely many tail-majors, so eventually for each x_n there always exists some $n' > n$ such that $x_n < x_{n'}$; this recursively defines an increasing subsequence. ■

Theorem 2.7. (Bolzano-Weierstrass) Every bounded sequence has a convergent subsequence.

Proof. Immediate from Theorem 2.7.

Definition 2.8. A sequence $\{x_n\}_{n \in \mathbb{N}}$ is *Cauchy* if

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall m, n \geq N)(|x_n - x_m| < \varepsilon)$$

Theorem 2.8. In \mathbb{R} , a sequence converges if and only if it is Cauchy.

Proof. (\Rightarrow) Let $\varepsilon > 0$. Then there exists N such that $|x_n - \ell| < \varepsilon/2$ for all $n \geq N$. Then for all $m, n \geq N$,

$$|x_n - x_m| \leq |x_n - \ell| + |x_m - \ell| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(\Leftarrow) We perform three steps:

- $\{x_n\}_{n \in \mathbb{N}}$ is bounded: $|x_n - x_N| \leq 1$ for all $n \geq N$ for some N , so

$$|x_n| \leq \max(1 + |x_N|, |x_1|, \dots, |x_{N-1}|).$$

- By Bolzano-Weierstrass, let $\{x_{n_j}\}_{j \in \mathbb{N}}$ be a subsequence of $\{x_n\}_{n \in \mathbb{N}}$ converging to ℓ .
- We prove that $\{x_n\}_{n \in \mathbb{N}}$ converges to ℓ too. For any $\varepsilon > 0$, there exists some N such that $|x_m - x_n| < \varepsilon/2$ and $|x_{n_j} - \ell| < \varepsilon/2$ for all $m, n, n_j \geq N$. Hence for all $n \geq N$,

$$|x_n - \ell| \leq |x_n - x_{n_j}| + |x_{n_j} - \ell| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \blacksquare$$

Definition 2.9. The *limit superior* and *limit inferior* of $\{x_n\}_{n \in \mathbb{N}}$ are defined by

$$\limsup x_n = \lim_{n \rightarrow \infty} \left(\sup_{k \geq n} x_k \right), \quad \liminf x_n = \lim_{n \rightarrow \infty} \left(\inf_{k \geq n} x_k \right)$$

Note that limsup and liminf exists for *any sequence* (allowing $\pm\infty$) because $\sup_{n \geq N} x_n$ and $\inf_{n \geq N} x_n$ are both monotone sequences.

Theorem 2.9. $\{x_n\}_{n \in \mathbb{N}}$ converges if and only if $\limsup x_n = \liminf x_n \in \mathbb{R}$.

Proof. (\Rightarrow) Assume $x_n \rightarrow \ell$. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $|x_n - \ell| < \varepsilon$ for all $n \geq N$. Then $\inf_{k \geq n} x_k \geq \ell - \varepsilon$ and $\sup_{k \geq n} x_k \leq \ell + \varepsilon$ for all $n \geq N$, giving

$$\ell - \varepsilon \leq \liminf x_n \leq \limsup x_n \leq \ell + \varepsilon$$

for any $\varepsilon > 0$ and hence $\liminf x_n = \limsup x_n$. (\Leftarrow) Assume $\limsup x_n = \liminf x_n = \ell \in \mathbb{R}$. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that

$$\left| \inf_{k \geq n} x_k - \ell \right| < \varepsilon, \quad \left| \sup_{k \geq n} x_k - \ell \right| < \varepsilon$$

for all $n \geq N$. Then for all $n \geq N$,

$$\ell - \varepsilon < \inf_{k \geq N} x_k \leq x_n \leq \sup_{k \geq N} x_k < \ell + \varepsilon$$

and hence $x_n \rightarrow \ell$. \blacksquare

3 Series

Definition 3.1. Given a sequence $\{x_n\}_{n \in \mathbb{N}}$, we define the series

$$\sum_{k=0}^n x_k = x_0 + x_1 + \cdots + x_n \quad \text{and} \quad \sum_{k=0}^{\infty} x_k = \lim_{n \rightarrow \infty} \sum_{k=0}^n x_k \text{ if it converges.}$$

Properties:

1. Linearity: $\sum_{k=0}^n cx_k = c \sum_{k=0}^n x_k$ and $\sum_{k=0}^n (x_k + y_k) = \sum_{k=0}^n x_k + \sum_{k=0}^n y_k$.
2. Distributivity: $\sum_{k=0}^n x_k \sum_{k=0}^n y_k = \sum_{k=0}^n x_k \sum_{j=0}^n y_j = \sum_{k=0}^n \sum_{j=0}^n x_k y_j$.

Cauchy Revisited. $\sum_{k=0}^n x_k$ is Cauchy if and only if

$$(\forall \varepsilon > 0)(\exists N \in \mathbb{N})(\forall n > m \geq N) \left(\left| \sum_{k=m+1}^n x_k \right| < \varepsilon \right).$$

Example 3.1.

1. **Geometric Series.** $x_k = r^k$ where $r > 0$. Then $\sum_{k=0}^n r^k = \frac{1-r^{n+1}}{1-r}$ for $r \neq 1$, so
 - If $r > 1$, then $r^{n+1} \rightarrow \infty$ and hence $\sum_{k=0}^n r^k$ diverges.
 - If $r = 1$, then $\sum_{k=0}^n r^k = n + 1$ also diverges.
 - If $0 < r < 1$, then $r^{n+1} \rightarrow 0$ and hence $\sum_{k=0}^n r^k$ converges to $\frac{1}{1-r}$.

Exercise. If all $x_k \geq 0$, then $\sum_{k=0}^{\infty} a_k$ converges if and only if the partial sums $\sum_{k=0}^n a_k$ are bounded for all n . As a corollary, if $0 \leq a_k \leq b_k$ for all $k \geq N_0$ and $\sum_{k=0}^n a_k$ diverges, then $\sum_{k=0}^n b_k$ diverges too.

Theorem 3.1. (Comparison Test)

If $|a_k| \leq b_k$ for all $k \geq N_0$ and $\sum_{k=0}^{\infty} b_k$ converges, then $\sum_{k=0}^{\infty} a_k$ converges.

Proof. We prove that $\sum_{k=0}^n a_k$ is Cauchy. Let $\varepsilon > 0$. Then there exists $N \in \mathbb{N}$ such that

$$\left| \sum_{k=m+1}^n b_k \right| < \varepsilon \quad \text{for all } n > m \geq N.$$

Hence for all $n > m \geq N$,

$$\left| \sum_{k=m+1}^n a_k \right| \leq \sum_{k=m+1}^n |a_k| \leq \sum_{k=m+1}^n b_k < \varepsilon. \quad \blacksquare$$

Definition 3.2. The series $\sum_{k=0}^{\infty} a_k$ *converges absolutely* if $\sum_{k=0}^{\infty} |a_k|$ converges.

By the Comparison Test, if a series converges absolutely then it converges too.

Theorem 3.2. (Alternating Series Test)

If $x_k \geq 0$ is non-increasing and $x_k \rightarrow 0$, then $\sum_{k=0}^{\infty} (-1)^k x_k$ converges.

Proof. Let $S_n = \sum_{k=0}^n (-1)^k x_k$. Observe

- $S_{2n+2} = S_{2n} - x_{2n+1} + x_{2n+2} \leq S_{2n}$
- $S_{2n+1} = S_{2n-1} + x_{2n} - x_{2n+1} \geq S_{2n-1}$
- $S_{2n+1} = S_{2n} - x_{2n+1} \leq S_{2n}$
- $|S_{2n+1} - S_{2n}| = |x_{2n+1}| \rightarrow 0$

Therefore $S_1 \leq S_3 \leq S_5 \leq \dots \leq S_4 \leq S_2 \leq S_0$. By the Monotone Convergence Theorem, $\{S_{2n}\}_{n \in \mathbb{N}}$ and $\{S_{2n+1}\}_{n \in \mathbb{N}}$ both converge. By the fourth bullet point, they must converge to the same value ℓ . Hence $S_n \rightarrow \ell$. \blacksquare

Example 3.2. $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln 2.$ (not obvious!)

Theorem 3.3. (Ratio Test)

If all $x_k \neq 0$ and $\lim_{n \rightarrow \infty} \left| \frac{x_{k+1}}{x_k} \right| < 1$, then $\sum_{k=0}^{\infty} x_k$ converges.

Proof. Say the limit is $0 \leq \ell < 1$. Then there exists $\ell < \beta < 1$ and $N \in \mathbb{N}$ such that $\left| \frac{x_{k+1}}{x_k} \right| \leq \beta$ for all $k \geq N$. This recursively gives $|x_{k+N}| \leq \beta^k |x_N|$ for all $k \geq 0$. By the Comparison Test, $\sum_{k=0}^{\infty} |x_{k+N}|$ converges by comparing it to the geometric series $\sum_{k=0}^{\infty} \beta^k |x_N|$ which converges. Therefore $\sum_{k=0}^{\infty} x_k$ is absolutely convergent and thus convergent. ■

More General Form of the Ratio Test.

- If $\limsup \left| \frac{x_{k+1}}{x_k} \right| < 1$, then $\sum_{k=0}^{\infty} x_k$ converges.
- If $\liminf \left| \frac{x_{k+1}}{x_k} \right| > 1$, then $\sum_{k=0}^{\infty} x_k$ diverges.

Note that we cannot conclude convergence nor divergence when the limit is exactly 1.

Example 3.3.

$$1. \lim_{n \rightarrow \infty} \left| \frac{1/(n+1)}{1/n} \right| = 1 \text{ but } \sum_{k=1}^{\infty} \frac{1}{k} \text{ diverges while } \sum_{k=1}^{\infty} \frac{(-1)^k}{k} \text{ converges.}$$

Definition 3.3. The *exponential function* is defined as

$$\exp(x) = e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Exercise. Prove that $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ converges for all $x \in \mathbb{R}$. *Hint: Ratio Test*

Theorem 3.4. $e := \exp(1)$ is irrational.

Proof. Assume not, then $\frac{m}{n} = \sum_{k=0}^{\infty} \frac{1}{k!}$ for some integers $m, n > 0$. Then

$$\begin{aligned} \left| m(n-1)! - \sum_{k=0}^n \frac{n!}{k!} \right| &= n! \left| e - \sum_{k=0}^n \frac{1}{k!} \right| \\ &= n! \left(\frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots \right) \\ &< \frac{1}{n+1} + \frac{1}{(n+1)^2} + \frac{1}{(n+1)^3} + \cdots \\ &= \frac{1}{n} \leq 1 \end{aligned}$$

is an integer strictly between 0 and 1, a contradiction! ■

Theorem 3.5. $e^x = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$ for all $x \in \mathbb{R}$.

Proof. We first prove a Lemma:

Lemma. $\left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k}{n}\right) \geq 1 - \frac{k(k+1)}{2n}$ for any positive integers $k \leq n$.

Proof. Induct on k . For $k = 1$, equality holds. Assume it is true for some $k < n$, then

$$\begin{aligned} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k+1}{n}\right) &\geq \left(1 - \frac{k(k+1)}{2n}\right) \left(1 - \frac{k+1}{n}\right) \\ &= 1 - \frac{(k+1)(k+2)}{2n} + \frac{k(k+1)^2}{2n^2} \\ &\geq 1 - \frac{(k+1)(k+2)}{2n}. \end{aligned} \quad \square$$

Let $\varepsilon > 0$. Then pick $N_1 > \frac{|x|^2 e^{|x|}}{\varepsilon}$. For all $n \geq \max(2, N_1)$,

$$\begin{aligned} \left| \sum_{k=0}^n \frac{x^k}{k!} - \left(1 + \frac{x}{n}\right)^n \right| &= \left| \sum_{k=2}^n \frac{x^k}{k!} \left[1 - \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)\right] \right| \\ &\leq \sum_{k=2}^n \frac{|x|^k}{k!} \cdot \frac{(k-1)k}{2n} \\ &\leq \frac{|x|^2}{2n} \cdot \sum_{k=0}^{n-2} \frac{|x|^k}{k!} \leq \frac{|x|^2 e^{|x|}}{2n} < \frac{\varepsilon}{2}. \end{aligned}$$

Also there exists an $N_2 > 2$ such that for all $n \geq N_2$,

$$\left| e^x - \sum_{k=0}^n \frac{x^k}{k!} \right| < \frac{\varepsilon}{2}$$

Hence, for all $n \geq \max(N_1, N_2)$,

$$\left| e^x - \left(1 + \frac{x}{n}\right)^n \right| \leq \left| e^x - \sum_{k=0}^n \frac{x^k}{k!} \right| + \left| \sum_{k=0}^n \frac{x^k}{k!} - \left(1 + \frac{x}{n}\right)^n \right| < \varepsilon. \quad \blacksquare$$

Theorem 3.6. (Products of Series)

If $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ converge absolutely, then $\sum_{k=0}^{\infty} \left(\sum_{\ell=0}^k a_{\ell} b_{k-\ell} \right) = \sum_{k=0}^{\infty} a_k \sum_{k=0}^{\infty} b_k$.

Proof. $\sum_{k=0}^n \left| \sum_{\ell=0}^k a_{\ell} b_{k-\ell} \right| \leq \sum_{k=0}^n \sum_{\ell=0}^k |a_{\ell}| |b_{k-\ell}| \leq \sum_{k=0}^n |a_k| \sum_{k=0}^n |b_k| \leq \sum_{k=0}^{\infty} |a_k| \sum_{k=0}^{\infty} |b_k|$ converges monotonically, so $\sum_{k=0}^{\infty} \left(\sum_{\ell=0}^k a_{\ell} b_{k-\ell} \right)$ converges absolutely. Taking $n \rightarrow \infty$ in

$$\begin{aligned} \left| \sum_{k=0}^n \left(\sum_{\ell=0}^k a_{\ell} b_{k-\ell} \right) - \sum_{k=0}^n a_k \sum_{k=0}^n b_k \right| &= \left| \sum_{k=n+1}^{2n} \left(\sum_{\ell=k-n}^n a_{\ell} b_{k-\ell} \right) \right| \\ &\leq \sum_{k=n+1}^{2n} \left(\sum_{\ell=k-n}^n |a_{\ell}| |b_{k-\ell}| \right) \\ &\leq \sum_{k=n+1}^{2n} \left(\sum_{\ell=0}^k |a_{\ell}| |b_{k-\ell}| \right) \end{aligned}$$

gives the desired result since $\sum_{k=0}^n \left(\sum_{\ell=0}^k |a_{\ell}| |b_{k-\ell}| \right)$ is Cauchy. \blacksquare

Example 3.4. The assumption of absolute convergence is necessary. Consider

$$a_k = b_k = \frac{(-1)^k}{\sqrt{k+1}}.$$

Then $\sum a_k$ and $\sum b_k$ converge by the alternating series test, but

$$\sum_{\ell=0}^k a_\ell b_{k-\ell} = (-1)^k \sum_{\ell=0}^k \frac{1}{\sqrt{(\ell+1)(k-\ell+1)}}$$

does not make a convergent series since

$$\sum_{\ell=0}^k \frac{1}{\sqrt{(\ell+1)(k-\ell+1)}} \geq \sum_{\ell=0}^k \frac{1}{k+1} = 1$$

so the series ‘oscillates with amplitude ≥ 1 ’.

Definition 3.4. A series $\sum_{k=0}^{\infty} x_k$ is *unconditionally convergent* if any reordering of the x_k gives a series converging to the same number.

The two theorems below show that absolute convergence and unconditional convergence are equivalent.

Theorem 3.7. (Dirichlet)

If $\sum_{k=0}^{\infty} x_k$ is absolutely convergent, it is unconditionally convergent.

Proof. We first treat the case where all $x_k \geq 0$. Let $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ be any bijection. Then the partial sums of $\sum_{k=0}^{\infty} x_{\sigma(k)}$ are bounded above by $\sum_{k=0}^{\infty} x_k$, so by the Monotone Convergence Theorem, it converges. Now we treat the general case.

Define $x_k^+ = \max\{0, x_k\}$ and $x_k^- = \max\{0, -x_k\}$, then $x_k = x_k^+ - x_k^-$ and $|x_k| = x_k^+ + x_k^-$. From the previous case, $\sum_{k=0}^{\infty} x_{\sigma(k)}^+$ and $\sum_{k=0}^{\infty} x_{\sigma(k)}^-$ are unconditionally convergent, so any rearranged sum can be written as

$$\sum_{k=0}^{\infty} x_{\sigma(k)} = \sum_{k=0}^{\infty} (x_{\sigma(k)}^+ - x_{\sigma(k)}^-) = \sum_{k=0}^{\infty} x_{\sigma(k)}^+ - \sum_{k=0}^{\infty} x_{\sigma(k)}^-.$$

■

Theorem 3.7. (Riemann)

If $\sum_{k=0}^{\infty} x_k$ converges but not absolutely, then for any $\ell \in \mathbb{R}$ or $\ell = \pm\infty$ there exists some rearrangement σ such that $\sum_{k=0}^{\infty} x_{\sigma(k)} = \ell$.

Proof. Again define $x_k^+ = \max\{0, x_k\}$ and $x_k^- = \max\{0, -x_k\}$. Now partition \mathbb{N} into

$$P = \{k \in \mathbb{N} : x_k \geq 0\} = \{k \in \mathbb{N} : x_k^+ \geq 0, x_k^- = 0\}$$

$$N = \{k \in \mathbb{N} : x_k < 0\} = \{k \in \mathbb{N} : x_k^+ = 0, x_k^- > 0\}$$

Since $\sum_{k=0}^{\infty} x_k$ converges but not absolutely, we have

$$\sum_{k=0}^{\infty} |x_k| = \infty, \quad \sum_{k=0}^{\infty} x_k^+ = \infty, \quad \sum_{k=0}^{\infty} x_k^- = -\infty, \quad \lim_{k \rightarrow \infty} x_k^+ = \lim_{k \rightarrow \infty} x_k^- = 0.$$

So the idea is

- If $\ell \in \mathbb{R}$, we keep choosing indices from P (or N) until we accumulate to a number close to ℓ , and then we alternate between P and N to get arbitrarily close to ℓ .
- If $\ell = \infty$, we keep choosing indices from P , but occasionally adding a term from N so that the series always grows more than it drops, and that we eventually can include everything from N .
- If $\ell = -\infty$, swap the roles of P and N .

We leave the formalities as an exercise. ■

4 Topology of \mathbb{R}

Definition 4.1.

- An *open interval* of \mathbb{R} is $(a, b) = \{x \in \mathbb{R} : a < x < b\}$ for some $a, b \in \mathbb{R} \cup \{\pm\infty\}$.
- A *closed interval* of \mathbb{R} is $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ for some $a, b \in \mathbb{R} \cup \{\pm\infty\}$.

For a given set $E \subseteq \mathbb{R}$, we say that $p \in E$ is

- an *interior point* of E if there exists $a < p < b$ such that $(a, b) \subseteq E$.
- an *isolated point* of E if there exists $a < p < b$ such that $(a, b) \subseteq E = \{p\}$.
- a *boundary point* if for all $a < p < b$, (a, b) intersects both E and E^c .
- a *limit point* (or accumulation point) if for all $a < p < b$, $(a, b) \cap E$ is infinite.

and we say E is

- *open* if every $p \in E$ is an interior point of E .
- *closed* if E contains all limit points of E .

Example 4.1.

1. p is a limit point if for all $a < p < b$, $(a, b) \cap E \neq \{p\}$ (this definition works for \mathbb{R} but not other topological spaces)
2. An interior point of E must be a limit point of E .
3. For $E = [0, 1]$ or $(0, 1)$, the point 0 is a limit point and boundary point, but not an interior nor isolated point. The point 0.5 is an interior point of E .
4. Open intervals are open. Closed intervals are closed.

Definition 4.2.

- The *interior* of E , denoted $\overset{\circ}{E}$ or $\text{int}(E)$, is the set of its interior points.
- The *closure* of E , denoted \overline{E} , is the union of E and its limit points.

Properties:

1. (Pset) $\overset{\circ}{E}$ is the largest open set $\subseteq E$ and \overline{E} is the smallest closed set $\supseteq E$.
2. E is open if and only if E^c is closed.

3. Finite intersections or arbitrary unions of open sets are open.
4. Arbitrary intersections or finite unions of closed sets are closed.

Definition 4.2.

- The *interior* of E , denoted $\overset{\circ}{E}$ or $\text{int}(E)$, is the set of its interior points.
- The *closure* of E , denoted \overline{E} , is the union of E and its limit points.

Tangent: Countability

Definition 4.3. A set S is *countable* if there exists a surjection $f : \mathbb{N} \rightarrow S$.

Example 4.2.

1. Finite sets and \mathbb{N} are countable.
2. If X, Y are countable, $X \times Y$ is countable. Hence \mathbb{Q} is countable.
3. A countable union of countable sets is countable.

Theorem 4.1. \mathbb{R} is not countable (*uncountable*).

Proof. We use a trick called Cantor's diagonalization. Suppose that there exists a surjective $f : \mathbb{N} \rightarrow (0, 1)$. Every number in $(0, 1)$ has a unique decimal expansion. We write

$$\begin{aligned} f(0) &= 0.a_{00}a_{01}a_{02}a_{03}\cdots \\ f(1) &= 0.a_{10}a_{11}a_{12}a_{13}\cdots \\ f(2) &= 0.a_{20}a_{21}a_{22}a_{23}\cdots \\ &\vdots \end{aligned}$$

and construct a number r that is different from $f(n)$ at the $(n+1)$ -th decimal place for all $n \in \mathbb{N}$. We can construct this by letting the $(n+1)$ -th decimal place of r be $a_{nn} + 1$ if $a_{nn} < 9$ or 0 if $a_{nn} = 9$. Then r does not have a preimage, contradicting surjectivity. ■

Back to Topology

Theorem 4.2. Every open set of \mathbb{R} is a countable union of disjoint open intervals.

Proof. Let E be an open set. For every $x \in E$, define

$$\begin{aligned} a_x &= \inf \{y \in E : (y, x] \subseteq E\} \\ b_x &= \sup \{z \in E : [x, z) \subseteq E\} \\ I_x &= (a_x, b_x) \end{aligned}$$

Since E is open, $a_x < x < b_x$ for all $x \in E$.

Claim 1. $a_x, b_x \notin E$.

Proof. If $a_x \in E$, then since E is open there exists $y < a_x$ such that $(y, a_x] \in E$, but then y is a smaller lower bound of $\{y \in E : (y, x] \subseteq E\}$, a contradiction. Similar for b_x . \square

Claim 2. $I_x = (a_x, b_x) \subseteq E$.

Proof. Let $y \in (a_x, b_x)$. Then there exists $z \in (a_x, y)$ such that $(z, x] \in E$ since $a_x < z$ is the infimum. Then $y \in (z, x] \subseteq E$. Since y was arbitrary, $(a_x, b_x) \subseteq E$. \square

Claim 3. If $I_x \cap I_y = \emptyset$, then $I_x = I_y$.

Proof. WLOG $a_x \leq a_y$. Since I_x, I_y overlap, we have $a_x \leq a_y < b_x$. Now if $a_x < a_y$, then $a_y \in I_x \subseteq E$ but Claim 1 says $a_y \notin E$, a contradiction. \square

Therefore $\{I_x : x \in E\}$ is a set of disjoint intervals whose union is E . To prove that it is countable, simply pick a rational in each I_x . Since the I_x are disjoint, each I_x maps to a different rational, hence embedding $\{I_x : x \in E\}$ into a subset of \mathbb{Q} which is countable. \blacksquare

Definition 4.4.

- An *open cover* U of $E \subseteq \mathbb{R}$ is a collection of open sets $\{O_\alpha\}_{\alpha \in I}$ such that $E \subseteq \bigcup_{\alpha \in I} O_\alpha$.
- $K \subseteq \mathbb{R}$ is (covering) *compact* if every open cover of K admits a finite subcover.
- $K \subseteq \mathbb{R}$ is *sequentially compact* if every sequence in K admits a converging subsequence in K .

Theorem 4.3. Let $K \subseteq \mathbb{R}$. The following are equivalent:

1. K is compact.
2. K is sequentially compact.
3. K is closed and bounded.

Example 4.3. $[0, 1]$ and finite sets are compact. $(0, 1)$ and \mathbb{R} are not compact.

Theorem 4.4. (Cantor's Intersection Theorem)

Let $\{K_n\}_{n \in \mathbb{N}}$ be a sequence of nonempty compact sets in \mathbb{R} such that $K_0 \supseteq K_1 \supseteq K_2 \supseteq \dots$. Then $K = \bigcap_{n \in \mathbb{N}} K_n$ is compact and nonempty.

Proof. $K \subseteq K_0$ is bounded. K is also closed because an arbitrary intersection of closed sets is closed. To prove K is non-empty, pick a $x_n \in K_n$ for each n and use the Bolzano-Weierstrass theorem. ■

Example 4.4. The *Cantor set* K is defined recursively as follows:

- $K_0 = [0, 1]$.
- Remove the middle third of each interval in K_n to get K_{n+1} , so

$$K_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1], \quad K_2 = [0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1], \quad \dots$$

- $K = \bigcup_{n \in \mathbb{N}} K_n$.

Each K_n is made up of 2^n closed intervals of length $1/3^n$, so the ‘total length’ of K_n is $(2/3)^n$, which goes to 0 as $n \rightarrow \infty$! Exercise:

- K is uncountable. (*Hint:* The points in C are exactly the reals in $[0, 1]$ that can be written with digits 0 and 2 in base 3, but be careful of things like $0.022\dots = 0.1$.)
- K is *perfect* (closed without isolated points) with an empty interior.

5 Metric Spaces

Definition 5.1. A *metric space* (X, d) is a set X equipped with a *metric* d , which is a function $d : X \times X \rightarrow \mathbb{R}_{\geq 0}$ such that for all $x, y, z \in X$,

- $d(x, y) = 0 \Leftrightarrow x = y$
- $d(x, y) = d(y, x)$ (Symmetry)
- $d(x, z) \leq d(x, y) + d(y, z)$ (Triangle Inequality)

Example 3.1.

1. \mathbb{R} with $d(x, y) = |x - y|$. (We have been working in this metric space so far)
2. \mathbb{R}^n with $d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$. (Euclidean metric)
3. \mathbb{R}^n with $d(\mathbf{x}, \mathbf{y}) = \sup_{1 \leq i \leq n} |x_i - y_i|$. (Uniform metric)
4. Any set X with $d(x, y) = \mathbf{1}(x \neq y)$. (Discrete metric)
5. $\mathcal{L}_2 = \{\{x_n\}_{n \in \mathbb{N}} : \sum_{n \in \mathbb{N}} x_n^2 < \infty\}$ with $d(\mathbf{x}, \mathbf{y}) = d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{\infty} (x_i - y_i)^2}$.
6. $\mathcal{L}_1 = \{\{x_n\}_{n \in \mathbb{N}} : \sum_{n \in \mathbb{N}} |x_n| < \infty\}$ with $d(\mathbf{x}, \mathbf{y}) = d_1(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{\infty} |x_i - y_i|$.
7. $\mathcal{L}_{\infty} = \{\{x_n\}_{n \in \mathbb{N}} : \text{bounded}\}$ with $d(\mathbf{x}, \mathbf{y}) = d_{\infty}(\mathbf{x}, \mathbf{y}) = \sup_{1 \leq i \leq n} |x_i - y_i|$.

We can generalize many definitions from the real topology to general metric spaces:

Definition 5.2.

- Convergence: $(\forall \varepsilon > 0) (\exists N \in \mathbb{N}) (\forall n \geq N) (d(x_n, \ell) < \varepsilon)$.
- Cauchy sequence: $(\forall \varepsilon > 0) (\exists N \in \mathbb{N}) (\forall m, n \geq N) (d(x_n, x_m) < \varepsilon)$.
- Open/Closed balls: $\mathcal{B}(x, r) = \{y : d(x, y) < r\}$, $\overline{\mathcal{B}}(x, r) = \{y : d(x, y) \leq r\}$.
- Open set: $(\forall x \in E) (\exists r > 0) (\mathcal{B}(x, r) \subseteq E)$. Closed set: E^c is open.
- Neighborhood of $x \in X$: Any open set containing x .
- Diameter of E : $\text{diam}(E) = \sup \{d(x, y) : x, y \in E\}$. Bounded set: $\text{diam}(E) < \infty$.
- Limit point of E : Any neighborhood of it intersects E infinitely much.
- Isolated point of E : Exists some neighbourhood that intersects E at only itself.

Definition 5.2 contd.

- Closure of E : $\overline{E} = E \cup \{\text{limit points of } E\}$.
- Interior of E : $\overset{\circ}{E} = \{x \in E : \text{exists neighborhood of } x \text{ contained in } E\}$.
- E is *dense* in F if $F \subseteq \overline{E}$. (Equivalently, all neighborhoods of all points in F must intersect E .)
- $K \subseteq X$ is *compact* if every open cover of K admits a finite subcover.
- $K \subseteq X$ is *totally bounded* if $(\forall \varepsilon > 0) (\exists x_1, \dots, x_n) (K \subseteq \mathcal{B}(x_1, \varepsilon) \cup \dots \cup \mathcal{B}(x_n, \varepsilon))$.
- $K \subseteq X$ is *complete* if every Cauchy sequence converges.
- $K \subseteq X$ is *separable* if it has a countable dense subset.

Example 5.2.

1. $(\mathbb{R}, |\bullet - \bullet|)$ is complete and separable ($\overline{\mathbb{Q}} = \mathbb{R}$).
2. $(\mathcal{L}_\infty, d_\infty)$ is not separable.

Proof. Consider $A = \{\text{sequences of 0s and 1s}\}$ which is uncountable. Then $d_\infty(x, y) = 1$ for all $x \neq y$, so $\{\mathcal{B}(x, 0.5) : x \in A\}$ is an uncountable collection of disjoint open neighborhoods. Any dense subset has to intersect each ball, hence must be uncountable.

3. Totally bounded \Rightarrow bounded. The converse is not true; check discrete metric.

Exercise. In $(\mathbb{R}, |\bullet - \bullet|)$,

- Totally bounded \Leftrightarrow Bounded.
- Complete \Leftrightarrow Closed.

Exercise. In (X, d) , a sequence $\{x_n\}_{n \in \mathbb{N}}$ converges to ℓ if and only if

$$\ell \in \bigcap_{n \in \mathbb{N}} \overline{\{x_n, x_{n+1}, \dots\}}$$

Theorem 5.1. Let $K \subseteq \mathbb{R}$. The following are equivalent:

1. K is compact.
2. K is sequentially compact.
3. K is complete and totally bounded.

Proof. (1) \Rightarrow (3): Assume K is compact. Fix $\varepsilon > 0$. Then $K \subseteq X = \bigcup_{x \in X} \mathcal{B}(x, \varepsilon)$, so there exists a finite subcover $\{\mathcal{B}(x_i, \varepsilon) : 1 \leq i \leq n\} \supseteq K$. Hence K is totally bounded. Consider a Cauchy sequence $\{x_n\}_{n \in \mathbb{N}}$ in X . Assume it does not converge. For each $n \in \mathbb{N}$ define $U_n = X \setminus \overline{\{x_n, x_{n+1}, \dots\}}$. By the exercise above, $\{U_n : n \in \mathbb{N}\}$ is an open cover of K , so it admits a finite subcover, whose union is $X \setminus \overline{\{x_N, x_{N+1}, \dots\}}$ for some $N \in \mathbb{N}$. Hence

$$\{x_0, x_1, x_2, \dots\} \subseteq K \subseteq X \setminus \overline{\{x_N, x_{N+1}, \dots\}}$$

which is a contradiction.

(3) \Rightarrow (2): Let $\{x_n\}_{n \in \mathbb{N}}$ be a sequence in K . Since K is totally bounded, write $K \subseteq \mathcal{B}(x_1, 1) \cup \dots \cup \mathcal{B}(x_N, 1)$. Then there exists a $\mathcal{B}(x_i, 1)$ containing infinitely many elements of $\{x_n\}_{n \in \mathbb{N}}$, corresponding to a subsequence $\{x_n^{(0)}\}_{n \in \mathbb{N}}$ of $\{x_n\}_{n \in \mathbb{N}}$. Repeat the same argument with balls of radius $1/2$, giving a subsequence $\{x_n^{(1)}\}_{n \in \mathbb{N}}$ of $\{x_n^{(0)}\}_{n \in \mathbb{N}}$, and so on for radii $1/2^n$ for $n \in \mathbb{N}$. Then the diagonal sequence $\{x_n^{(n)}\}_{n \in \mathbb{N}}$ is Cauchy: For all n , they will eventually be contained in some ball of radius $1/2^n$. Therefore it is a converging subsequence.

(2) \Rightarrow (1): Let K be sequentially compact.

Lemma. K is totally bounded.

Proof. Pick $\varepsilon > 0$. Assume the union of any finite collection of open ε -balls does not contain K . We generate a sequence that does not converge in K , namely a sequence $\{x_n\}_{n \in \mathbb{N}}$ such that $d(x_i, x_j) \geq \varepsilon$ for all $i \neq j$:

- Assume x_0, \dots, x_k are chosen such that $d(x_i, x_j) \geq \varepsilon$ for all $i \neq j$. Then K has an element that is not in $\mathcal{B}(x_0, \varepsilon) \cup \dots \cup \mathcal{B}(x_k, \varepsilon)$, and we pick that as x_{k+1} . \square

Let $\mathcal{I} = \{I_0, I_1, I_2, \dots\}$ be the union, over all $n \in \mathbb{N}^*$, of finite sets of open $(1/n)$ -balls that cover K . Let U be an open cover of K . We first show that there is a countable subcover U' : For each $k \in \mathbb{N}$ we choose, if it exists, some $O_k \in U$ such that $I_k \subseteq O_k$. Write $U' = \{O_k : k \in \mathbb{N}\}$.

Lemma. U' covers K .

Proof. Pick any $x \in K$, then there exists some $O \in U$ that contains x . Since O is open, there exists some neighborhood $\mathcal{B}(x, \varepsilon) \subseteq O$. Pick some $I_k \in \mathcal{I}$ which is an open ball of radius $< \varepsilon/2$ and contains x . Then $x \in I_k \subseteq \mathcal{B}(x, \varepsilon) \subseteq O$, so there exists some O_k (e.g. O) to be chosen when building U' . Since x was arbitrary, U' covers K . \square

Hence $U' = \{O_0, O_1, \dots\}$ is a countable subcover of K . It suffices to prove that U' admits a finite subcover. Assume not, then for all $k \in \mathbb{N}$,

$$\bigcup_{i=0}^k O_i \not\supseteq K \implies \exists x_k \in K \setminus \bigcup_{i=0}^k O_i$$

Now, a subsequence of $\{x_k\}_{k \in \mathbb{N}}$ converges to x . So for all $k \in \mathbb{N}$, there exists a sequence in $K \setminus \bigcup_{i=0}^k O_i \subseteq \left(\bigcup_{i=0}^k O_i \right)^c$ that converges to $x \in K$ (a sufficiently far tail of the subsequence).

But $\left(\bigcup_{i=0}^k O_i \right)^c$ is closed, so it contains the limit point x . Since $k \in \mathbb{N}$ was arbitrary,

$$x \in \left(\bigcup_{i=0}^{\infty} O_i \right)^c = K^c$$

which is a contradiction since x was in K . \blacksquare

Theorem 5.2. (Baire) Let (X, d) be a complete metric space and O_n is open and dense in X for all $n \in \mathbb{N}$. Then $O = \bigcup_{n \in \mathbb{N}} O_n$ is dense in X .

Example 5.3. Enumerate the rational numbers $\mathbb{Q} = \{q_0, q_1, \dots\}$ and let $O_n = \mathbb{R} \setminus \{q_n\}$. Then $\bigcup O_n = \mathbb{R} \setminus \mathbb{Q}$ is dense in \mathbb{R} (not open!).

Proof. Let U be any open subset. We want to prove that U intersects O .

- Since O_1 is dense, there exists $x_1 \in O_1 \cap U$. Since O_1 is open, there exists a neighborhood whose closure $\mathcal{B}(x_1, r_1) \subseteq O_1 \cap U$.
- Recursively, pick $x_n \in O_n \cap \mathcal{B}(x_{n-1}, r_{n-1})$, and pick some $\overline{\mathcal{B}}(x_n, r_n) \subseteq O_n \cap \mathcal{B}(x_{n-1}, r_{n-1})$ such that $0 < r_n < r_{n-1}/2$.

So we have x_1, x_2, \dots and $r_1 > 2r_2 > 4r_3 > \dots$, so $\{x_n\}_{n \in \mathbb{N}}$ is Cauchy, so it converges to some $x \in \bigcap_{n \in \mathbb{N}} \overline{\mathcal{B}}(x_n, r_n)$ which is contained in both U and O . \blacksquare

6 Continuous Functions

Definition 6.1.

- Let $(X, d_X), (Y, d_Y)$ be metric spaces. We say $f : X \rightarrow Y$ is *continuous at* $x \in X$ if for every $x_n \rightarrow x$ we have $f(x_n) \rightarrow f(x)$.
- $f : X \rightarrow Y$ is *continuous* if it is continuous at every $x \in X$.

Theorem 6.1. $f : X \rightarrow Y$ is continuous at x if and only if

$$(\forall \varepsilon > 0) (\exists \delta > 0) (\forall d_X(x, y) < \delta) (d_Y(f(x), f(y)) < \varepsilon).$$

Proof. (\Rightarrow) Let $x_n \rightarrow x$ and $\varepsilon > 0$, pick the associated δ . Then there exists some $N \in \mathbb{N}$ such that $d_X(x, x_n) < \delta$ for all $n \geq N$, so $d_Y(f(x), f(x_n)) < \varepsilon \implies f(x_n) \rightarrow f(x)$.

(\Leftarrow) Assume there is an $\varepsilon > 0$ such that there is no such δ . Then for each $n \in \mathbb{N}$ we pick x_n such that $d_X(x, x_n) < 1/n$ and $d_Y(f(x), f(x_n)) \geq \varepsilon$. Then $x_n \rightarrow x$ but $f(x_n) \not\rightarrow f(x)$. ■

Theorem 6.2.

$f : X \rightarrow Y$ is continuous if and only if for all open sets U in Y , $f^{-1}(U)$ is open in X .

Proof. Say f is continuous. Take U open in Y , and any $x \in f^{-1}(U)$. Since U is open, there exists $\mathcal{B}(f(x), \varepsilon) \subseteq U$. Since $f(x_n) \rightarrow f(x)$ for all $x_n \rightarrow x$, we can find a $\delta > 0$ such that

$$f(\mathcal{B}(x, \delta)) \subseteq \mathcal{B}(f(x), \varepsilon)$$

so $\mathcal{B}(x, \delta) \subseteq f^{-1}(U)$ and hence $f^{-1}(U)$ is open. Conversely, fix $\varepsilon > 0$. Then $\mathcal{B}(f(x), \varepsilon)$ is open in Y and hence $f^{-1}(\mathcal{B}(f(x), \varepsilon))$ is open in X , so there exists a neighborhood $\mathcal{B}(x, \delta) \subseteq f^{-1}(\mathcal{B}(f(x), \varepsilon))$ and thus $f(\mathcal{B}(x, \delta)) \subseteq \mathcal{B}(f(x), \varepsilon)$. ■

Example 6.1. Continuous functions:

1. Isometries: $f : X \rightarrow Y$ such that $d_Y(f(x), f(y)) = d_X(x, y)$ for all $x, y \in X$.
2. L -Lipschitz functions: $f : X \rightarrow Y$ such that there exists $L > 0$ such that $d_Y(f(x), f(y)) \leq L d_X(x, y)$ for all $x, y \in X$.
3. α -Hölder functions: $f : X \rightarrow Y$ such that there exists $L > 0, 0 < \alpha < 1$ such that $d_Y(f(x), f(y)) \leq L d_X(x, y)^\alpha$ for all $x, y \in X$.
4. $f(x) = |x|$ on \mathbb{R} is 1-Lipschitz. $f(x) = \sqrt{x}$ on $\mathbb{R}_{\geq 0}$ is 0.5-Hölder.

Theorem 6.3. (Banach Fixed Point Theorem)

Let (X, d) be complete and $f : X \rightarrow X$ be α -Lipschitz for some $0 < \alpha < 1$ (such functions are called *contractions*). Then f has a unique fixed point: $f(a) = a$.

Proof. Uniqueness is easy: If $f(x) = x$ and $f(y) = y$ then

$$d(x, y) = d(f(x), f(y)) \leq \alpha d(x, y) \implies x = y.$$

We prove existence by starting at any $x_0 \in X$ and considering the sequence $x_n = f(x_{n-1})$.

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x_{n-1})) \leq \alpha d(x_n, x_{n-1}) \\ \therefore d(x_{n+1}, x_n) &\leq \alpha^n d(x_1, x_0). \end{aligned}$$

With this, $\{x_n\}_{n \in \mathbb{N}}$ is Cauchy (exercise), so it converges to some x . Taking $n \rightarrow \infty$ on both sides of $x_{n+1} = f(x_n)$ (allowed because Lipschitz is continuous!) gives $f(x) = x$. ■

Definition 6.2. $f : X \rightarrow Y$ is *uniformly continuous* if

$$(\forall \varepsilon > 0) (\exists \delta > 0) (\forall d_X(x, y) < \delta) (d_Y(f(x), f(y)) < \varepsilon).$$

Remark: Here δ does not depend on x !

Example 6.2.

1. Hölder functions are uniformly continuous.
2. $f(x) = x^2$ on \mathbb{R} is not uniformly continuous:

Proof. Say $\varepsilon = 1$. For any chosen δ , we see that for $x > 1/\delta$,

$$\left| f\left(x + \frac{\delta}{2}\right) - f(x) \right| = \delta x + \frac{\delta^2}{4} > 1 = \varepsilon.$$

Theorem 6.4. If X is compact and $f : X \rightarrow Y$ is continuous, then $f(X)$ is compact.

Proof. Let $\{U_\alpha\}_{\alpha \in I}$ be an open cover of $f(X) \subseteq Y$. Since f is continuous, $\{f^{-1}(U_\alpha)\}_{\alpha \in I}$ is an open cover of X and hence there exists some finite subcover $\{f^{-1}(U_1), \dots, f^{-1}(U_k)\}$ of X . Then $\{U_1, \dots, U_k\}$ is a finite subcover of $f(X)$. ■

Theorem 6.5. (Heine-Cantor)

If X is compact and $f : X \rightarrow Y$ is continuous, then f is uniformly continuous.

Proof. Fix $\varepsilon > 0$. Since f is continuous, for every x there exists $\delta_x > 0$ such that $d_y(f(y), f(x)) < \varepsilon/2$ whenever $d_X(y, x) < \delta_x$. Consider the finite subcover of $\{\mathcal{B}(x, \delta_x/2) : x \in X\}$

that covers X , say $\{\mathcal{B}(x_i, \delta_{x_i}/2) : 1 \leq i \leq n\}$. We then define $\delta = \min_{1 \leq i \leq n} \delta_{x_i}/2$.

Now take any $x, y \in X$ with $d_X(x, y) < \delta$. Then there exists x_i such that $x \in \mathcal{B}(x_i, \delta_{x_i}/2)$. That means $y \in \mathcal{B}(x_i, \delta_{x_i})$, so

$$\begin{aligned} d_Y(f(x), f(y)) &\leq d_Y(f(x), f(x_i)) + d_Y(f(x_i), f(y)) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned} \quad \blacksquare$$

Let's now focus on functions whose image is in \mathbb{R} .

Exercise.

1. If $f, g : X \rightarrow \mathbb{R}$ are continuous, then $f + g, fg, f \circ g$ are continuous.
2. Intervals are *connected*: For any two disjoint open sets O_1, O_2 whose union is the interval, the interval is completely contained in one of O_1, O_2 . (Pset 7)

Theorem 6.6.

If X is compact, $f : X \rightarrow \mathbb{R}$ is continuous, then $f(X)$ has a maximum and minimum.

Proof. By Theorem 6.5, $f(X)$ is compact, so it is closed and bounded (Theorem 4.3). Since it is bounded, $f(X)$ has a supremum and an infimum. Since it is closed, the supremum and infimum are in $f(X)$. ■

Theorem 6.7. (Intermediate Value Theorem)

If $f : [a, b] \rightarrow \mathbb{R}$ is continuous and $f(a) < \mu < f(b)$, there exists $c \in [a, b]$ with $f(c) = \mu$.

Proof. Assume $\mu \notin f([a, b])$. Then $f^{-1}((-\infty, \mu)) \cup f^{-1}((\mu, \infty)) = f^{-1}((-\infty, \mu) \cup (\mu, \infty))$ are two disjoint open sets whose union is $[a, b]$, contradicting connectedness. ■

Definition 6.3.

If X is compact, we define the *uniform metric* on $\mathcal{C}(X) = \{f : X \rightarrow \mathbb{R} \text{ continuous}\}$:

$$d(f, g) = \sup \{|f(x) - g(x)| : x \in X\}$$

Exercise. Check that $(\mathcal{C}(X), d)$ is a metric space.

Definition 6.4. Let $\{f_n : X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ be a sequence of continuous functions.

- We say f_n *converges pointwise* to f if $f_n(x) \rightarrow f(x)$ for all $x \in X$.
- We say f_n *converges uniformly* to f if $\sup_{x \in X} |f_n(x) - f(x)| \rightarrow 0$ as $n \rightarrow \infty$.

This is equivalent to f_n converging in $(\mathcal{C}(X), d)$, so we can write $f_n \xrightarrow{d} f$.

Example 6.3. Set $X = [0, 1]$.

1. $f_n(x) = 1/n$ converges uniformly to 0.
2. $f_n(x) = x^n$ converges pointwise to $\mathbf{1}(x = 1)$ but does not converge uniformly.

Theorem 6.8. $(\mathcal{C}(X), d)$ is complete.

Proof. Let $\{f_n : X \rightarrow \mathbb{R}\}_{n \in \mathbb{N}}$ be Cauchy. Then for all $x \in X$, $\{f_n(x)\}_{n \in \mathbb{N}}$ is Cauchy and hence f_n converges pointwise, say to f . We now have to prove f is continuous and $f_n \xrightarrow{d} f$.

Let $\varepsilon > 0$. Then there exists N such that $|f_n(x) - f_m(x)| < \varepsilon$ for all $m, n \geq N$ and $x \in X$. Taking $m \rightarrow \infty$ gives $|f_n(x) - f(x)| < \varepsilon$, which is the criteria of uniform convergence. To check that f is continuous, let $\varepsilon > 0$ again and fix x .

- There exists $N \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon/3$ for all $n \geq N$.
- Since f_N is continuous, $\exists \delta > 0$ such that $|f_N(x) - f_N(y)| < \varepsilon/3$ for all $d_X(x, y) < \delta$.

Therefore,

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f(y) - f_N(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

■

Definition 6.5.

- A set $K \subseteq \mathcal{C}(X)$ is *uniformly bounded* if there exists an $M \in \mathbb{R}$ such that $f(x) \leq M$ for all $f \in K$ and $x \in X$.
- A set $K \subseteq \mathcal{C}(X)$ is *(uniformly) equicontinuous* if

$$(\forall \varepsilon > 0) (\exists \delta > 0) (\forall f \in K, d_X(x, y) < \delta) (d_Y(f(x), f(y)) < \varepsilon).$$

Theorem 6.9. (Arzelà-Ascoli)

Let X be compact. $K \subseteq \mathcal{C}(X)$ is *relatively compact* (i.e. \overline{K} is compact) if and only if it is uniformly bounded and uniformly equicontinuous.

Proof. We make three observations when X is compact:

1. A continuous $f : X \rightarrow \mathbb{R}$ must be uniformly continuous and bounded.

2. X is separable.

Proof. Since X is totally bounded, for all $n \in \mathbb{N}$ there exists a finite set S_n of points whose $(1/n)$ -ball neighborhoods cover X . Then $S = \bigcup_n S_n$ is a countable dense set: Given any open $U \subseteq X$, there is some $\mathcal{B}(u, 1/N) \subseteq U$ and some $s \in S_{2N}$ with $u \in \mathcal{B}(s, 1/2N)$. This means $s \in \mathcal{B}(u, 1/N) \subseteq U$. \square

3. \overline{K} is compact if and only if \overline{K} is complete and totally bounded.

Assume \overline{K} is compact.

1. *Proving K is uniformly bounded.*

Since \overline{K} is totally bounded, there exists f_1, \dots, f_n such that $\overline{K} \subseteq \bigcup_{i=1}^n \mathcal{B}(f_i, 1)$. By Obv 1, each f_j is bounded by some M_j . Now let $M = \max_{1 \leq j \leq n} M_j + 1$. Hence, for any $f \in K \subseteq \overline{K}$, there exists some $\mathcal{B}(f_j, 1)$ that contains f , so

$$|f(x)| \leq |f_j(x)| + |f(x) - f_j(x)| < M_j + 1 \leq M$$

2. *Proving K is uniformly equicontinuous.*

Let $\varepsilon > 0$. There exists f_1, \dots, f_n such that $\overline{K} \subseteq \bigcup_{i=1}^n \mathcal{B}(f_i, \varepsilon/3)$. For each $1 \leq j \leq n$, since f_j is uniformly continuous there exists $\delta_j > 0$ such that $|f_j(x) - f_j(y)| < \varepsilon/3$ whenever $d_X(x, y) < \delta_j$. Let $\delta = \min_{1 \leq j \leq n} \delta_j$. Now take any $f \in K \subseteq \overline{K}$. Then $f \in \mathcal{B}(f_j, \varepsilon/3)$ for some j . For any $d_X(x, y) < \delta$,

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_j(x)| + |f_j(x) - f_j(y)| + |f_j(y) - f(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Now we prove the other direction. We will prove \overline{K} is sequentially compact, but we first note that we just need to prove for sequences strictly in K :

Lemma. To prove that \overline{K} is sequentially compact, we just need to prove that all sequences in K (instead of in \overline{K}) has a convergent subsequence (in \overline{K} automatically).

Proof. Let $\{y_n\}_{n \in \mathbb{N}}$ be any sequence in \overline{K} . Then every y_n is the limit of some sequence $\{y_{nj}\}_{j \in \mathbb{N}}$ in K , so for every n there exists N_n such that $d(y_{nj}, y_n) < 1/(n+1)$ for all $j \geq N_n$. Then the diagonal sequence $\{y_{nN_n}\}_{n \in \mathbb{N}}$ is in K , so it admits a subsequence $\{y_{n_k N_{n_k}}\}_{k \in \mathbb{N}}$ that converges to say y . We claim that $\{y_{n_k}\}_{k \in \mathbb{N}}$ converges to y too: Let $\varepsilon > 0$. Choose N such that $1/(N+1) < \varepsilon/2$ and $d(y_{n_k N_{n_k}}, y) < \varepsilon/2$ for $k \geq N$. Then

$$d(y_{n_k}, y) \leq d(y_{n_k}, y_{n_k N_{n_k}}) + d(y_{n_k N_{n_k}}, y) < \frac{1}{n_k + 1} + \frac{\varepsilon}{2} < \varepsilon. \quad \square$$

Hence, let $\{f_n\}_{n \in \mathbb{N}} \subseteq K$. Since X is separable, say $\{x_k\}_{k \in \mathbb{N}}$ is dense in X .

- Given x_0 , we can extract a subsequence $\{f_{0j}\}_{j \in \mathbb{N}}$ such that $f_{0j}(x_0) \xrightarrow{j \rightarrow \infty} g(x_0) \in \mathbb{R}$ by the Bolzano-Weierstrass Theorem.
- We then extract a subsequence $\{f_{1j}\}_{j \in \mathbb{N}}$ of $\{f_{0j}\}_{j \in \mathbb{N}}$ such that $f_{1j}(x_1) \xrightarrow{j \rightarrow \infty} g(x_1) \in \mathbb{R}$ by the Bolzano-Weierstrass Theorem. Note that $f_{1j}(x_0) \xrightarrow{j \rightarrow \infty} g(x_0)$ still.
- Repeat for x_2, x_3, \dots .

We then consider the diagonal sequence $\{f_{jj}\}_{j \in \mathbb{N}}$. Then $f_{jj}(x_k) \xrightarrow{j \rightarrow \infty} g(x_k)$ for all k . Rename the initial sequence $\{f_n\}$ to be the subsequence $\{f_{jj}\}_{j \in \mathbb{N}}$. We want to show that $\{f_n\}$ is Cauchy.

- Since K is equicontinuous, $\exists \delta > 0$ such that $|f(x) - f(y)| < \varepsilon/3$ for all $d_X(x, y) < \delta$.
- Let $\{\mathcal{B}(x_j, \delta) : 0 \leq j \leq J\}$ be a finite subcover of X . Then there exists N such that $|f_n(x_j) - f_m(x_j)| < \varepsilon/3$ for all $m, n \geq N$ and $0 \leq j \leq J$.

Therefore for all $m, n \geq N$ and $x \in X$, we have some $x \in \mathcal{B}(x_j, \delta)$ and so

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x) - f_n(x_j)| + |f_n(x_j) - f_m(x_j)| + |f_m(x_j) - f_m(x)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \\ \therefore d(f_n, f_m) &= \sup_{x \in X} |f_n(x) - f_m(x)| \leq \varepsilon. \end{aligned}$$

and hence $\{f_n\}_{n \in \mathbb{N}}$ is Cauchy, so it converges to some $g \in \overline{K}$. ■

7 Derivatives

Definition 7.1.

- Let $f : I \rightarrow \mathbb{R}$ where $I \subseteq \mathbb{R}$. Then we say $\lim_{x \rightarrow x_0} f(x) = \ell$ if for all $\varepsilon > 0$, there exists $\delta > 0$ such that $|f(x) - \ell| < \varepsilon$ for all $x \in I$ with $0 < |x - x_0| < \delta$.
- Let I be an open interval. We say that $f : I \rightarrow \mathbb{R}$ is *differentiable at x_0* if

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\delta \rightarrow 0} \frac{f(x_0 + \delta) - f(x_0)}{\delta} \in \mathbb{R}$$

exists, in which case we denote the limit by $f'(x_0)$, called the *derivative* at x_0 . We say f is *differentiable* if f is differentiable at all points in I .

- $\frac{f(x) - f(x_0)}{x - x_0}$ is called the *difference quotient* and represents the slope.

Exercise. $\lim_{x \rightarrow x_0} f(x) = \ell$ if and only if $\lim_{n \rightarrow \infty} f(x_n) = \ell$ for any sequence $x_n \rightarrow x_0$.

Example 7.1. $f(x) = x^2$ is differentiable: $\lim_{\delta \rightarrow 0} \frac{(x + \delta)^2 - x^2}{\delta} = \lim_{\delta \rightarrow 0} 2x_0 + \delta = 2x$.

Theorem 7.1. If f is differentiable at x_0 , then it is continuous at x_0 .

Proof. Assume we have a sequence $x_n \rightarrow x$. Then

$$\begin{aligned} \lim_{n \rightarrow \infty} |f(x_n) - f(x)| &\leq \lim_{n \rightarrow \infty} \frac{f(x_n) - f(x)}{x_n - x} \cdot (x_n - x) \\ &= \lim_{n \rightarrow \infty} \frac{f(x_n) - f(x)}{x_n - x} \cdot \lim_{n \rightarrow \infty} (x_n - x) \\ &= f'(x) \cdot 0 = 0. \end{aligned}$$

and hence $f(x_n) \rightarrow f(x)$. ■

Properties:

1. \mathbb{R} -linearity: $(cf)' = cf'$ for all $c \in \mathbb{R}$.
2. Leibniz (Product) Rule: $(fg)' = f'g + fg'$.
3. Quotient Rule: If $g'(x_0) \neq 0$, $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$.

Proof of (2).

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \frac{f(x_0 + \delta)g(x_0 + \delta) - f(x_0)g(x_0)}{\delta} \\ &= \lim_{\delta \rightarrow 0} \frac{f(x_0 + \delta) - f(x_0)}{\delta} g(x_0 + \delta) + f(x_0) \frac{g(x_0 + \delta) - g(x_0)}{\delta} \\ &= f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

Exercise. $f(x) = x^n$ ($n \in \mathbb{N}$) $\implies f'(x) = nx^{n-1}$.

Theorem 7.2. (Chain Rule)

If f, g are differentiable at x_0 , then $f \circ g$ is differentiable at x_0 , with

$$(f \circ g)'(x_0) = f'(g(x_0))g'(x_0).$$

Proof. Take $x_n \rightarrow x$ with $x_n \neq x$ for all n . Then

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{f(g(x_n)) - f(g(x))}{x_n - x_0} &= \lim_{n \rightarrow \infty} \frac{f(g(x_n)) - f(g(x))}{g(x_n) - g(x_0)} \cdot \frac{g(x_n) - g(x_0)}{x_n - x_0} \\ &= f'(g(x_0))g'(x_0) \end{aligned}$$

if $g(x_n) \neq g(x)$ eventually. If $g(x_n) = g(x)$ eventually, then it evaluates to 0 anyway and $g'(x_0) = 0$ too. ■

Example 7.2.

1. With $f = g^{-1}$, we get $f'(g(x)) = 1/g'(x)$.
2. Say $f(x) = \sqrt{x}$ and $g(x) = x^2$, then $f'(x^2) = \frac{1}{2x}$.
3. $f(x) = |x|$ is not differentiable at 0: $\lim_{\delta \rightarrow 0^-} \frac{|\delta|}{\delta} = -1 \neq 1 = \lim_{\delta \rightarrow 0^+} \frac{|\delta|}{\delta}$.

Definition 7.2.

$f : \Omega \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ is said to have *directional derivative* at $x_0 \in \Omega$ in direction $v \in \mathbb{R}^n$ if

$$Df(x_0)[v] := \lim_{\delta \rightarrow 0} \frac{f(x_0 + \delta v) - f(x_0)}{\delta}$$

exists. We say f is *differentiable* at x_0 if $Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear map.

Theorem 7.3. If $f : [a, b] \rightarrow \mathbb{R}$ is differentiable, then the maximum of f occurs at either a, b or a point x_0 with $f'(x_0) = 0$. *Note:* Maximum exists since $[a, b]$ is compact.

Proof. If it does not occur at a nor b , then it occurs at an interior point $x_0 \in (a, b)$. Then

$$f'(x_0) = \lim_{\delta \rightarrow 0^+} \frac{f(x_0 + \delta) - f(x_0)}{\delta} \leq 0$$

and

$$f'(x_0) = \lim_{\delta \rightarrow 0^-} \frac{f(x_0 + \delta) - f(x_0)}{\delta} \geq 0$$

and hence $f'(x_0) = 0$. ■

Theorem 7.4. (Rolle's)

If $f : [a, b] \rightarrow \mathbb{R}$ is continuous, f is differentiable on (a, b) , and $f(a) = f(b)$, then there exists $c \in (a, b)$ with $f'(c) = 0$.

Proof. If f is constant then the result is trivial. Otherwise, a maximum or minimum occurs at the interior, and we can use Theorem 7.3. ■

Theorem 7.5. (Mean Value Theorem)

If $f : [a, b] \rightarrow \mathbb{R}$ is continuous, f is differentiable on (a, b) , then there exists $c \in (a, b)$ with $f'(c) = \frac{f(b) - f(a)}{b - a}$.

Proof. Define $F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$ and apply Rolle's Theorem. ■

Exercise.

1. If $f' = 0$ then f is constant.
2. If $|f'| \leq L$ then f is L -Lipschitz. *Hint:* Use the Mean Value Theorem.

Theorem 7.6. (L'Hôpital's Rule) Let f, g be differentiable on I , and let $x_0 \in I$ such that $f(x_0) = g(x_0) = 0$, and $g'(x) \neq 0$ on some $\mathcal{B}(x_0, \varepsilon)$, and $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$ exists.

$$\text{Then} \quad \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Proof. Take some $x_1 \in \mathcal{B}(x_0, \varepsilon)$. Consider $\Phi(x) = f(x_1)g(x) - g(x_1)f(x)$. By Rolle's Theorem, there exists some c between x_0 and x_1 such that $\Phi'(c) = f(x_1)g'(c) - g(x_1)f'(c) = 0$. Hence for all $x \in \mathcal{B}(x_0, \varepsilon)$, there exists some c_x between x and x_0 such that $\frac{f(x)}{g(x)} = \frac{f'(c_x)}{g'(c_x)}$. Taking the limit $x \rightarrow x_0$ (and hence $c_x \rightarrow x_0$) gives the result. ■

Definition 7.3.

- A function $f : I \rightarrow \mathbb{R}$ is *convex* if for all $x_1 < x_2$ in I and any $0 < \alpha < 1$,

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

We say that f is *strictly convex* if the inequality is always strict.

- A function $f : I \rightarrow \mathbb{R}$ is *concave* if for all $x_1 < x_2$ in I and any $0 < \alpha < 1$,

$$f(\alpha x_1 + (1 - \alpha)x_2) \geq \alpha f(x_1) + (1 - \alpha)f(x_2).$$

We say that f is *strictly concave* if the inequality is always strict.

- Define the *right and left derivative*

$$f'_+(x_0) = \lim_{\delta \rightarrow 0^+} \frac{f(x_0 + \delta) - f(x_0)}{\delta}, \quad f'_-(x_0) = \lim_{\delta \rightarrow 0^-} \frac{f(x_0 + \delta) - f(x_0)}{\delta}$$

Exercise. If f is convex on I , then $x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$ is nondecreasing on I .

Theorem 7.7.

Say f is convex on I . Then $f'_-(x) \leq f'_+(x) \leq f'_-(y) \leq f'_+(y)$ for all $x < y$ in I .

Proof. Firstly, $\frac{f(x_0 + \delta) - f(x_0)}{\delta}$ is nondecreasing as $\delta \rightarrow 0^-$ and hence $f'_-(x_0)$ (and similarly $f'_+(x_0)$) exists for any $x_0 \in I$. The result then follows from the Exercise above. ■

Corollary.

1. When f is differentiable, then f is convex if and only if f' is nondecreasing.
2. When f, f' are both differentiable, then f is convex if and only if $f'' \geq 0$.

Theorem 7.8. If f is convex, f' exists except at countably many points.

Proof. Whenever f is not differentiable at $x \in I$, we have $f'_-(x) < f'_+(x)$, so we can pick a $q_x \in \mathbb{Q}$. We then have an injective map $x \mapsto q_x$ on nondifferentiable points. ■

Example 7.3.

1. $f(x) = |x|$ is convex, with $f'(x) = \text{sgn}(x)$ when $x \neq 0$ and $f'_-(0) = -1, f'_+(0) = 1$.
2. $f(x) = e^x$ is convex because $f''(x) = e^x > 0$.

Definition 7.4.

- A function $f : I \rightarrow \mathbb{R}$ is in C^1 if it is differentiable and f' is continuous.
- If $f'(x_0) = 0$, we say x_0 is a *critical point* and $f(x_0)$ is a *critical value*.
- We say $y \in \mathbb{R}$ is a *regular value* if it is not a critical value.
- A set $S \subseteq \mathbb{R}$ has *measure zero* if for all $\varepsilon > 0$ there exists countably many intervals that (i) covers S and (ii) have total combined length $< \varepsilon$.

Exercise.

1. A subset of a measure zero set has measure zero.
2. Every finite or countable subset has measure zero.
3. The Cantor set (uncountable!) has measure zero.
4. A countable union of measure zero sets has measure zero.
Hint: Take a covering of total length $< \varepsilon/2^{n+1}$ for the n -th set.

Theorem 7.9. (Sard's Theorem)

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be in C^1 . Then $\{\text{critical values of } f\} \subseteq \mathbb{R}$ has measure zero.

Proof. It suffices to prove that the set of critical values of f on a closed interval $[a, b]$ has measure zero, because to get the full set of critical values we just apply to $[-n, n]$ for all $n \in \mathbb{N}$ and take the countable union. WLOG we will prove for $[0, 1]$. Let $\varepsilon > 0$

Since $f' : [0, 1] \rightarrow \mathbb{R}$ is continuous, it is uniformly continuous and hence there exists $N \in \mathbb{N}$ such that $|f'(x) - f'(y)| < \varepsilon/2$ for all $|x - y| < 1/N$. Partition $[0, 1]$ into $I_k = [\frac{k}{N}, \frac{k+1}{N}]$ for $k = 0, \dots, N-1$. For every k where I_k has a critical point x_k , for all $x, y \in I_k$ we have

$$|f(x) - f(y)| \stackrel{\text{MVT}}{=} |f'(c)| |x - y| = |f'(c) - f'(x_k)| |x - y| < \frac{\varepsilon}{2} \cdot \frac{1}{N}$$

and hence the length of $f(I_k)$ is $< \varepsilon/2N$. Taking all $0 \leq k \leq N-1$ for which I_k has a critical point, we get a covering with total length $< \varepsilon$. ■

Example 7.4.

1. The constant function has critical points everywhere, with exactly one critical value, and hence $\{\text{critical value}\}$ has measure zero.

Theorem 7.10. Any regular value of $f : [a, b] \rightarrow \mathbb{R}$ in C^1 has a finite pre-image.

Proof. Let y_0 be a regular value ($f'(y) \neq 0$ for any $f(y) = y_0$). Then $f^{-1}(\{y_0\}) \subseteq [a, b]$ is closed and hence compact. If $f^{-1}(\{y_0\})$ were infinite, then it admits a sequence $\{x_n\}_{n \in \mathbb{N}}$ converging to some \bar{x} . But then $f(x_n) = f(\bar{x}) = y_0$ and hence by Rolle's there always exists a x'_n between x_n and \bar{x} such that $f'(x'_n) = 0$. Then $0 = f'(x'_n) \rightarrow f'(\bar{x}) \neq 0$ which is a contradiction since $f(\bar{x}) = y_0$. ■

8 Riemann Integral

Definition 8.1.

- A *partition* of $[a, b]$ is a finite set of points $\sigma = \{a = x_0 < \cdots < x_N = b\}$.
- The *size* $|\sigma|$ of σ is $\max_{1 \leq i \leq N} |x_i - x_{i-1}|$.
- A partition σ' is a *refinement* of σ if $\sigma' \supseteq \sigma$.
- Given a bounded $f : [a, b] \rightarrow \mathbb{R}$ and a partition σ of $[a, b]$,
 - The *upper (Riemann) sum* is $S(f, \sigma) = \sum_{i=1}^N (x_i - x_{i-1}) \sup_{x \in [x_{i-1}, x_i]} f(x)$.
 - The *lower (Riemann) sum* is $s(f, \sigma) = \sum_{i=1}^N (x_i - x_{i-1}) \inf_{x \in [x_{i-1}, x_i]} f(x)$.
- Given a bounded $f : [a, b] \rightarrow \mathbb{R}$,
 - The *upper (Riemann) integral* is $\mathcal{I}^+(f) = \inf_{\forall \sigma} S(f, \sigma)$.
 - The *lower (Riemann) integral* is $\mathcal{I}^-(f) = \sup_{\forall \sigma} s(f, \sigma)$.
- A bounded $f : [a, b] \rightarrow \mathbb{R}$ is *Riemann integrable* if $\mathcal{I}^-(f) = \mathcal{I}^+(f) := \int_a^b f(x) \, dx$.
Denote by $\mathcal{R}(a, b)$ the set of all Riemann integrable functions on $[a, b]$.
- Given $f : [a, b] \rightarrow \mathbb{R}$ and $I \subseteq [a, b]$ an interval, define $\text{osc}_I f = \sup_I f - \inf_I f$.

Remark.

1. Given two partitions σ_1, σ_2 of $[a, b]$, there is always a partition that is refinement of both, e.g. $\sigma_3 = \sigma_1 \cup \sigma_2$.
2. $s(f, \sigma) \leq \sum_{i=1}^N (x_i - x_{i-1}) f(\xi_i) \leq S(f, \sigma)$ for any choice $\xi_i \in [x_{i-1}, x_i]$ for all i .
3. **Exercise.** If σ_3 is a refinement of both σ_1, σ_2 , then

$$s(f, \sigma_1) \leq s(f, \sigma_3) \leq S(f, \sigma_3) \leq S(f, \sigma_2) \implies s(f, \sigma_1) \leq S(f, \sigma_2) \quad \forall \sigma_1, \sigma_2$$

which implies $\mathcal{I}^-(f) \leq \mathcal{I}^+(f)$ for any (bounded) f .

Theorem 8.1. The following are equivalent:

1. $f \in \mathcal{R}(a, b)$.
2. $(\forall \varepsilon > 0) (\exists \sigma) (S(f, \sigma) - s(f, \sigma) < \varepsilon)$.
3. $(\forall \varepsilon > 0) (\exists \delta > 0) (\forall |\sigma| < \delta) (S(f, \sigma) - s(f, \sigma) < \varepsilon)$.
4. $(\forall \varepsilon > 0) (\exists N > 0) (\forall n \geq N) (S(f, \sigma_n) - s(f, \sigma_n) < \varepsilon)$ where

$$\sigma_n = \left\{ a + \frac{k}{n}(b-a) : 0 \leq k \leq n \right\} \quad (\text{equipartition})$$

5. $(\exists \mathcal{I} \in \mathbb{R}) (\forall \varepsilon > 0) (\exists \delta > 0) (\forall |\sigma| < \delta) (\forall \xi_i \in [x_{i-1}, x_i]):$

$$\left| \sum_{i=1}^N (x_i - x_{i-1}) f(\xi_i) - \mathcal{I} \right| < \varepsilon.$$

Proof. We just prove (2) \Leftrightarrow (3). (\Leftarrow) is trivial. (\Rightarrow)

Assume (2) is true. Let $\varepsilon > 0$. Then there exists some $\sigma = \{x_0 < \dots < x_N\}$ with $S(f, \sigma) - s(f, \sigma) < \varepsilon$. Since f is bounded, let $|f(x)| \leq M$ for all x .

We pick $\delta = \varepsilon/(2MN)$. Then let $\sigma' = \{y_0 < \dots < y_{N'}\}$ be any partition of size $< \delta$. Note that any interval $Y_i = [y_{i-1}, y_i]$ is either

(A) entirely contained within some $X_{f(i)} = [x_{f(i)-1}, x_{f(i)}]$, or

(B) contains some x_j for some j . (There are at most N such intervals)

$$\begin{aligned} \therefore S(f, \sigma') - s(f, \sigma') &= \sum_{i=1}^{N'} (y_i - y_{i-1}) \operatorname{osc}_{Y_i} f \\ &\leq \sum_{i:(A)} (y_i - y_{i-1}) \operatorname{osc}_{X_{f(i)}} f + \sum_{i:(B)} (y_i - y_{i-1}) \operatorname{osc}_{Y_i} f \\ &= \sum_{j=1}^N \sum_{i:(A), Y_i \subseteq X_j} (y_i - y_{i-1}) \operatorname{osc}_{X_j} f + \sum_{i:(B)} (y_i - y_{i-1}) \operatorname{osc}_{Y_i} f \\ &\leq \sum_{j=1}^N (x_j - x_{j-1}) \operatorname{osc}_{X_j} f + N(\delta)(2M) \\ &= S(f, \sigma) - s(f, \sigma) + 2MN\delta \leq 2\varepsilon. \end{aligned}$$

■

Theorem 8.2. Continuous functions are Riemann integrable.

Proof. Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous and $\varepsilon > 0$. Then f is uniformly continuous and hence exists an $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for all $|x - y| < \delta$. Then for any partition σ of size $< \delta$,

$$S(f, \sigma) - s(f, \sigma) = \sum_{i=1}^N (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f \leq \sum_{i=1}^N (x_i - x_{i-1}) \varepsilon = \varepsilon(b - a). \quad \blacksquare$$

Example 8.1. The *Dirichlet function* $\varphi(x) = \mathbf{1}(x \in \mathbb{Q})$ on $[0, 1]$ is not Riemann integrable, with $\mathcal{I}^+(\varphi) = 1$ and $\mathcal{I}^-(\varphi) = 0$.

Properties: If $f, g \in \mathcal{R}(a, b)$, then

1. $f + g, \lambda f$ ($\lambda \in \mathbb{R}$) are in $\mathcal{R}(a, b)$ and $\int_a^b f + g = \int_a^b f + \int_a^b g$, $\int_a^b \lambda f = \lambda \int_a^b f$.
2. $fg, \max(f, g), \min(f, g) \in \mathcal{R}(a, b)$.
3. $f/g \in \mathcal{R}(a, b)$ if $\inf g > 0$.
4. $f \leq g \implies \int_a^b f \leq \int_a^b g$.
5. $|f| \in \mathcal{R}(a, b)$ with $\int_a^b |f| \geq \left| \int_a^b f \right|$. (Triangle inequality)
6. If $c \in (a, b)$, then $f \in \mathcal{R}(a, c) \cap \mathcal{R}(c, b)$ with $\int_a^b f = \int_a^c f + \int_c^b f$.

Remark: We denote $\int_a^a f = 0$ and $\int_b^a f = -\int_a^b f$ if $a < b$.

Theorem 8.3. (Fundamental Theorem of Calculus / FTC)

If $f : [a, b] \rightarrow \mathbb{R}$ is continuous, then $F(x) = \int_a^x f$ is differentiable with $F' = f$.

Proof. For any $x \in (a, b)$ and $h > 0$,

$$\begin{aligned} \left| \frac{F(x+h) - F(x)}{h} - f(x) \right| &= \frac{1}{h} \left| \int_x^{x+h} f - hf(x) \right| \\ &= \frac{1}{h} \left| \int_x^{x+h} (f - f(x)) \right| \\ &\leq \frac{1}{h} \int_x^{x+h} |f - f(x)| \\ &\leq \frac{1}{h} \sup_{[x, x+h]} |f - f(x)| \xrightarrow{h \rightarrow 0^+} 0 \end{aligned}$$

The case where $h < 0$ is similar. ■

Theorem 8.4. (Integral Form of FTC)

If $F : [a, b] \rightarrow \mathbb{R}$ is in C^1 , then $\int_a^b F' = F(b) - F(a)$.

Proof. Apply FTC to $f = F'$, giving $G(x) = \int_a^x F'$ with $G' = F'$. But $(G - F)' = 0 \implies G(b) - F(b) = G(a) - F(a)$ and thus $\int_a^b F' = G(b) - G(a) = F(b) - F(a)$. ■

Theorem 8.5. (Integration by Parts)

If $f, g : [a, b] \rightarrow \mathbb{R}$ are in C^1 , then $\int_a^b f'g = f(b)g(b) - f(a)g(a) - \int_a^b fg'$.

Proof. Apply Theorem 8.4 to $F = fg$ (with $F' = f'g + gf'$). ■

Theorem 8.6. (Characterization of Riemann Integrability)

$f \in \mathcal{R}(a, b)$ if and only if

- f is bounded, and
- The set of points of discontinuity of f has measure zero.

Example 8.2.

1. $f : [0, 1] \rightarrow \mathbb{R}$ with $f(x) = \mathbf{1}(x = 1/2)$ is in $\mathcal{R}(a, b)$ (discontinuous only at $1/2$).
2. The Dirichlet function is discontinuous everywhere, so it is not in $\mathcal{R}(a, b)$.

Definition 8.2. The *oscillation of f at point x* is $\text{osc}(f, x) = \lim_{\delta \rightarrow 0^+} \text{osc}_{[x-\delta, x+\delta]} f \geq 0$

Exercise. $\text{osc}(f, x) = 0$ if and only if f is continuous at x .

Proof of Theorem 8.6. (\Leftarrow) Let $|f(x)| \leq M$ for all x . Denote E as the set of discontinuity points, so E has measure zero. Let $\varepsilon > 0$. Then

$$E_\varepsilon = \left\{ x \in [a, b] : \text{osc}(f, x) \geq \frac{\varepsilon}{2(b-a)} \right\} \subseteq E$$

has measure zero too. Also, E_ε is closed (Exercise! If $x \notin E_\varepsilon$, choose δ small enough so that the oscillation is still within $\varepsilon/2(b-a)$, so E_ε^c is open). Therefore, E_ε is compact and hence can be covered by finitely many disjoint closed intervals of total length $< \varepsilon/(4M)$.

We then consider a partition σ of $[a, b]$ that contains all the closed intervals chosen above

(type A), and the rest where the oscillations are $< \varepsilon/2(b-a)$ (type B). Then

$$\begin{aligned}
 S(f, \sigma) - s(f, \sigma) &= \sum_{i=1}^N (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f \\
 &= \sum_{i:(A)} (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f + \sum_{i:(B)} (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f \\
 &\leq \sum_{i:(A)} (x_i - x_{i-1}) (2M) + \sum_{i:(B)} (x_i - x_{i-1}) \frac{\varepsilon}{2(b-a)} \\
 &< \frac{\varepsilon}{4M} (2M) + (b-a) \frac{\varepsilon}{2(b-a)} = \varepsilon.
 \end{aligned}$$

(\Rightarrow) f is bounded by definition, so we just need to prove that the set of discontinuity points has measure zero. We will prove that

$$E_\delta = \{x \in [a, b] : \operatorname{osc}(f, x) \geq \delta\}$$

has measure zero. The result then follows from a union of $\delta = 1/n$ for $n \in \mathbb{N}^*$. Take any partition σ of $[a, b]$, then

$$\begin{aligned}
 S(f, \sigma) - s(f, \sigma) &= \sum_{i=1}^N (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f \\
 &= \sum_{[x_{i-1}, x_i] \cap E_\delta = \emptyset} (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f + \sum_{[x_{i-1}, x_i] \cap E_\delta \neq \emptyset} (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f \\
 &\geq \sum_{[x_{i-1}, x_i] \cap E_\delta \neq \emptyset} (x_i - x_{i-1}) \operatorname{osc}_{[x_{i-1}, x_i]} f \\
 &\geq \delta \sum_{[x_{i-1}, x_i] \cap E_\delta \neq \emptyset} (x_i - x_{i-1}) \geq \delta |E_\delta|
 \end{aligned}$$

But for any $\varepsilon > 0$ we can force $S(f, \sigma) - s(f, \sigma) < \varepsilon$ for some σ , so $|E_\delta| = 0$. ■

Definition 8.3. An *ordinary differential equation* (ODE) is a problem in the form

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0$$

where $y(x)$ is a differentiable function from $\mathbb{R} \rightarrow \mathbb{R}^n$ to be solved.

Example 8.3.

1. Newton's Law of Cooling: $\theta'(t) = \kappa \cdot (T - \theta(t))$.
2. Newton's 2nd Law: $mx''(t) = F(x(t)) \Leftrightarrow \frac{d}{dt} \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ F(x(t))/m \end{pmatrix}$

We show that most ODEs have a unique solution.

Theorem 8.7. (Picard-Lindelöf/Cauchy-Lipschitz)

Let $D \subseteq \mathbb{R}^2$ be open and $(x_0, y_0) \in D$. Let $f : D \rightarrow \mathbb{R}$ be L -Lipschitz in the second variable (namely $|f(x, y_1) - f(x, y_2)| \leq L |y_1 - y_2|$). Then for some $\varepsilon > 0$ there exists a unique solution $y : (x_0 - \varepsilon, x_0 + \varepsilon) \rightarrow \mathbb{R}$ to the ODE

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0.$$

Remark.

1. The theorem is true for higher dimensions too, i.e. $f : D \subseteq \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$.
2. While the proof is for *local* existence and uniqueness, it can be extended *globally*: There always exists a *maximal* interval I_{\max} containing x_0 where $y(x)$ exists and is unique. *Sketch*: Keep expanding $(x_0 - \varepsilon, x_0 + \varepsilon)$ by repeatedly applying this theorem to the boundaries when possible.
3. The L -Lipschitz condition of f is necessary. Consider the ODE $y'(x) = 3y^{2/3}, y(0) = 0$. Then $y = x^3$ and $y = 0$ are solutions. (In fact, there are infinitely many solutions! Can you find them? Note that scaling doesn't work.)

Proof. By FTC, the differential equation is equivalent to the *integral equation*

$$y(x) = y_0 + \int_0^x f(t, y(t)) dt$$

so for any $I = (x_0 - \varepsilon, x_0 + \varepsilon)$ let's define the functional $\mathcal{L} : \mathcal{C}(I) \rightarrow \mathcal{C}(I)$ with

$$[\mathcal{L}(y)](x) = y_0 + \int_0^x f(t, y(t)) dt$$

and use the Banach Fixed Point Theorem.

$$\begin{aligned} d(\mathcal{L}(y_1), \mathcal{L}(y_2)) &= \sup_{x \in I} \left| \int_{x_0}^x (f(t, y_1(t)) - f(t, y_2(t))) dt \right| \\ &\leq \sup_{x \in I} \left| \int_{x_0}^x |f(t, y_1(t)) - f(t, y_2(t))| dt \right| \\ &\leq L \sup_{x \in I} \left| \int_{x_0}^x |y_1(t) - y_2(t)| dt \right| \quad (\text{Lipschitz}) \\ &\leq L \sup_{x \in I} \left| \int_{x_0}^x \sup_{z \in I} |y_1(z) - y_2(z)| dt \right| \\ &\leq L|I| \sup_{z \in I} |y_1(z) - y_2(z)| = L|I|d(y_1, y_2) \end{aligned}$$

so \mathcal{L} is a contraction if we choose $|I| < 1/L$, i.e. $\varepsilon < 1/(2L)$. ■