Contents

Chapter 1

The Real Line and Euclidean Space

1.1 Ordered Fields and the Number Systems

Definition 1: Ordered Field

An ordered field is a field equipped with a subset $R \subset F \times F$ s.t. $x \leq y$ if $(x, y) \in \mathbb{R}$.

R must satisfy

1. $x \leq x$
2. $x \leq y, y \leq x \implies x = y$
3. $x \leq y, y \leq z \implies x \leq z$
4. $x, y \in F \implies x \leq y$ or $y \leq x$
5. $x \leq y \implies x + z \leq y + z$
6. $0 \leq x, 0 \leq y \implies 0 \leq xy$

For ordered fields, $x^2 \geq 0$.

Proposition 1. *Let F be an ordered field, then for $x, y \in F$*

1. $|x + y| \leq |x| + |y|$
2. $\big||x| - |y|\big| \leq |x - y|$

3. $|xy| \leq |x||y|$

Definition 2: Principle of Mathematical Induction

If $S \in \mathbb{N}$ s.t. $1 \in S$ and $k \in S \implies k+1 \in S$, then $S = \mathbb{N}$.

Definition 3: Well Ordering Principle

If $S \in \mathbb{N}$ s.t. $S \neq \{0\}$, then there exists $\sigma \in S$ s.t. $\sigma \leq n$ for all $n \in S$ (a smallest element exists).

Example 1

$\mathbb{N}$ is well-ordered, but $\mathbb{Z}$ is not.

Definition 4: Countable

A set S is **countable** if there is an injection $f : S \hookrightarrow \mathbb{N}$

Proposition 2. *$\mathbb{N} \times \mathbb{N}$ is countable.*

Proof. Set $f(m,n) = 2^m(2n-1)$. This is injective. □

Corollary 1. *If X and Y are countable, $X \times Y$ is countable.*

Proof. We know there exist $\phi : X \hookrightarrow \mathbb{N}$ and $\psi : Y \hookrightarrow \mathbb{N}$. Form the map $(\phi, \psi) : X \times Y \hookrightarrow \mathbb{N} \times \mathbb{N}$, then $f \circ (\phi, \psi) \hookrightarrow \mathbb{N}$ □

Corollary 2. *$\mathbb{Z}$ is countable.*

Proof. Let $g : \mathbb{Z} \to \mathbb{N}_0 \times \mathbb{N}_0$ be defined by $g(m) = (m, 0)$ if $m \geq 0$ and $g(m) = (0, -m)$ if $m < 0$. Since $\mathbb{N}_0$ is countable ($m \mapsto m+1$), we know $\mathbb{N}_0 \times \mathbb{N}_0$ is countable, and thus $\mathbb{Z}$ is also countable. □

Corollary 3. *$\mathbb{Q}$ is countable.*

Proof. Let $p/q \in \mathbb{Q}$ with $q > 0$ and p, q relatively prime. Define $g : \mathbb{Q} \hookrightarrow \mathbb{Z} \times \mathbb{N}$ by $g(p/q) = (p, q)$. Since $\mathbb{Z} \times \mathbb{N}$ is countable, so is $\mathbb{Q}$. □

Proposition 3. *$\mathbb{Q}$ is dense in itself: If $x, y \in \mathbb{Q}$, then there is a $z \in \mathbb{Q}$ such that $x < z < y$.*

Proof. Let $z = (x+y)/2$. This works. $\square$

Definition 5: Archimedean Property

If $x \in \mathbb{Q}$, then there exists $n \in \mathbb{N}$ such that $x < n$.

Proof. If $x < 0$, take $n = 1$. If $x = p/q$ with $p, q > 0$, take $n = p + 1$. This works because

$$\frac{p}{q} \leq \frac{p}{q} + (q-1)\frac{p}{q} = p < p + 1$$

$\square$

Definition 6: Field Isomorphism

Let S_1 and S_2 be two fields. A bijective function $\phi : S_1 \leftrightarrow S_2$ is a field isomorphism if $\phi(x+y) = \phi(x) + \phi(y)$ and $\phi(xy) = \phi(x)\phi(y)$. If such a ϕ exists, we say that S_1 and S_2 are isomorphic. Additionally, if S_1 and S_2 are ordered fields and $x \leq y \implies \phi(x) \leq \phi(y)$, then ϕ is called an ordered field isomorphism.

Proposition 4. *If F is an ordered field, then there exists $\mathbb{Z}_F \subset F$ such that $\mathbb{Z}_F$ (with inherited arithmetic) is ring isomorphic to $\mathbb{Z}$.*

Proof. $0_F, 1_F, -1_F \in F$. By the properties of fields,

$$\begin{aligned} 0_F &< 1_F \\ 1_F &< 1_F + 1_F \\ &\vdots \end{aligned}$$

Define $\phi : \mathbb{N} \leftrightarrow F$ by $\phi(n) = \underbrace{1_F + \cdots + 1_F}_{n \text{ times}}$ and $\phi(-n) = -\phi(n)$. $\square$

Proposition 5. *Similarly, there exists $\mathbb{Q}_F \in F$ such that $\mathbb{Q}_F$ is an ordered field $\simeq \mathbb{Q}$.*

Definition 7: Archimedean Ordered Field

An ordered field F is Archimedean if for all $x \in F$, there exists $n \in \mathbb{N}_F$ such that $x < n$.

Proposition 6. *The following are equivalent to the original definition of an Archimedean ordered field.*

1. $0 < y < x \implies$ *there exists* $k \in \mathbb{N}$ *such that* $x < ky$.
2. $x > 0 \implies$ *there exists* $n \in \mathbb{N}$ *such that* $0 < \frac{1}{n} < x$.

Proof.

1. Assume F is Archimedean, then $xy^{-1} < k$ for some $k \in \mathbb{N}$, which implies $x < ky$. Conversely, if (1) holds, then we have two cases. If $x \leq 1$, then it is clearly bounded above by, say, 2. If $1 < x$, then take $y = 1$, then there exists k such that $x < k$.
2. The proof for the second formulation is similar.

□

1.2 Completeness and the Real Number System

$\mathbb{Q}$ satisfies all the properties of an ordered field, so we need something more to characterize $\mathbb{R}$. This will be "completeness", where the limits of rationals are in the system too. This requires the use of sequences and limits.

Definition 8: Sequence

A function $f : \mathbb{N} \to S$, where S is a set, is a sequence in S. We often denote $f(n)$ by a_n and f by $\{a_n\}_{n=1}^{\infty}$.

Definition 9: Subsequence

Let $f : \mathbb{N} \to S$ be a sequence. Let $\sigma : \mathbb{N} \hookrightarrow \mathbb{N}$ be injective and increasing, i.e. $\sigma(n+1) > \sigma(n)$. Then $f \circ \sigma$ is called a **subsequence** of f.

Useful notation for subsequences is $\left\{x_{\sigma(n)}\right\}_{n=1}^{\infty}$ instead of $\{x_n\}_{n=1}^{\infty}$.

Definition 10: Convergence

A sequence $\{a_n\}_{n=1}^{\infty}$ in an ordered field F converges to a limit L if for all $\varepsilon > 0$ in F, there exists $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ if $n > N$.

Intuitively, this means that all but finitely many a_n lie outside of the interval $(L - \varepsilon, L + \varepsilon)$.

Proposition 7. *Squeeze Lemma If $x_n \to L$ and $z_n \to L$ and $x_n \leq y_n \leq z_n$ for any $n > n_0$, then $y_n \to L$.*

Proof. Since x_n and z_n converge to L, for any $\varepsilon > 0$ there exist N_x such that $|x_n - L| < \varepsilon$ when $n > N_x$ and N_z such that $|z_n - L| < \varepsilon$ when $n > N_z$.

Let $N = \max\{n_0, n_x, n_z\}$, then we have

$$-\varepsilon < x_n - L \leq y_n - L \leq z_n - L < \varepsilon.$$

Thus $y_n \to L$. □

Similarly, if $a \leq x_n \leq b$ and $x_n \to L$, then $a \leq L \leq b$.

Proposition 8. *In an Archimedean field, if $x_n \to x$ and $x_n \to y$, then $x = y$, i.e. limits are unique.*

Proof. By the triangle inequality, we have

$$|x - y| = |x - x_n + x_n - y| \leq |x - x_n| + |x_n - y|.$$

Suppose $|x-y| > 0$, then set $\varepsilon = |x-y|/2$. We can then choose N large enough that

$$|x - y| \leq |x - x_n| + |x_n - y| < \frac{|x-y|}{2} + \frac{|x-y|}{2} = |x - y|.$$

This is a contradiction, so $|x - y| = 0$. □

1.2.1 Bounded Sequences

Definition 11: Bounded

Let F be an Archimedean field, then $S \subset F$ is bounded if there exists $M \in F$ such that $|\sigma| \leq M$ for all $\sigma \in S$.

A function can be also just bounded below or above, which is a weaker condition.

Proposition 9. *A convergent sequence in an Archimedean field is bounded.*

Proof. Suppose $x_n \to L$, then there exists N such that $|x_n - L| < 1$ if $n > N$ (the choice of 1 here is arbitrary). Then for $n > N$,

$$|x_n| = |x_n - L + L| \leq |x_n - L| + |L| < 1 + |L|$$

Set the bound $M = \max\{|x_1|, |x_2|, \dots, |x_n|, 1 + |L|\}$. □

1.2.2 Limit Arithmetic

Proposition 10. *Some of the usual arithmetic operations also hold with limits. Let $x_n \to x$ and $y_n \to y$, then*

1. $x_n + y_n \to x + y$
2. $\lambda x_n \to \lambda x$
3. $x_n y_n \to xy$
4. $y_n \neq 0, y \neq 0 \implies x_n y_n^{-1} \to xy^{-1}$

Proof. Each of these proofs is based on simple applications of the triangle inequality.

1. Since $x_n \to x$ and $y_n \to y$, for all $\varepsilon > 0$, there exist N_x such that $|x - x_n| < \varepsilon/2$ when $n > N_x$ and N_y such that $|y - y_n| < \varepsilon/2$ when $n > N_y$. Let $N = \max N_x, N_y$, then for all $\varepsilon > 0$,
$$|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| < \varepsilon.$$
2. This is a special case of (3), so it suffices to prove (3).
3. We have
$$|x_n y_n - xy| = |x_n y_n - x_n y + x_n y - xy| \leq |x_n||y_n - y| + |y||x_n - x|.$$
Since $\{x_n\}$ converges, it is bounded, i.e. $|x_n| \leq M$ for all n. Let $\tilde{M} = \max\{M, |y|\}$. Then our inequality becomes
$$|x_n y_n - xy| \leq \tilde{M}|y_n - y| + \tilde{M}|x_n - x|.$$
For $\varepsilon > 0$, choose N such that $|x_n - x|, |y_n - y| < \varepsilon/(2\tilde{M})$ for all $n > N$, then $|x_n y_n - xy| < \varepsilon$.
4. Since $y_n, y \neq_0$, we have
$$\begin{aligned} |x_n y_n^{-1} - xy^{-1}| &= \left| \frac{x_n y - x y_n}{y_n y} \right| \\ &= \left| \frac{x_n y - xy + xy - x y_n}{y_n y} \right| \\ &\leq \frac{|y||x_n - x|}{|y_n y|} + \frac{|x||y - y_n|}{|y_n y|} \\ &\leq |y|^2 |y_n||x_n - x| + |x||y||y_n||y - y_n| \end{aligned}$$
Since $\{y_n\}$ converges, it is bounded, i.e. $|y_n| \leq M_y$ for all n. Let $M = \max\{|x|, |y|\} \cdot M_y$. Then for any $\varepsilon > 0$, we can find N large enough such that $|x_n - x|, |y_n - y| < \varepsilon/(2M)$. This yields
$$|x_n y_n^{-1} - xy^{-1}| \leq M|x_n - x| + M|y_n - y| < \varepsilon.$$

□

1.2.3 Completeness

The terms "nondecreasing", "nonincreasing", "strictly decreasing", and "strictly increasing" all describe monotone sequences.

Definition 12: Monotone Sequence Property

Let F be an ordered field. We say that F has the Monotone Sequence Property if every monotone increasing sequence bounded above converges.

Definition 13: Completeness

An ordered field is complete if it obeys the Monotone Sequence Property.

Complete ordered fields are Archimedean.

1.2.4 The Real Numbers

Theorem 1: The Real Number System

There's a unique (up to isomorphism) complete ordered field called the real number system. It is constructed as follows: Let S be defined

$$S = \{(x_1, x_2, \dots) | x_n \in \mathbb{Q}, \text{ sequence is increasing and bounded above}\}$$

and let two members of S be equivalent if their upper bounds are the same. Then $\mathbb{R}$ is the set of all equivalence classes in S. We do not include $\pm\infty$ in $\mathbb{R}$.

Proposition 11. *$\mathbb{Q}$ is dense in $\mathbb{R}$. This can be stated two ways:*

1. *$x, y \in \mathbb{R}$ with $x < y$, then there exists $r \in \mathbb{Q}$ with $x < r < y$*
2. *If $x, \varepsilon \in \mathbb{R}$ with $\varepsilon > 0$, then there exists $r \in \mathbb{Q}$ with $|x - r| < \varepsilon$.*

Proof. Suppose $x < y$, so $y - x > 0$. Then since $\mathbb{R}$ is Archimedean, there exists an integer n such that $0 < 1/n < y - x$, as well as another integer k such that $k/n > x$. By the well-ordering property, there is a smallest such k. Using this smallest k, we have $(k-1)/n \leq x < k/n$. Let $r = k/n$ (which is clearly rational), then

$$x < r = \frac{k-1}{n} + \frac{1}{n} \leq x + \frac{1}{n} < x + (y - x) = y.$$

So $x < r < y$, as required. $\square$

Although $\mathbb{Q}$ is dense in $\mathbb{R}$, there are actually many more irrationals than rationals.

Theorem 2

The interval $(0, 1)$ in $\mathbb{R}$ is uncountable.

Since the function $f(x) = a + (b - a)x$ maps $]0, 1[\mapsto]a, b[$, any interval in $\mathbb{R}$ is uncountable. Since $\mathbb{R}$ is uncountable but $\mathbb{Q}$ is countable, it must be the case that $\mathbb{C}$ is uncountable.

1.3 Infimums and Supremums

Definition 14: Least Upper Bound/Supremum

Let $S \subset \mathbb{R}$, then b is an **upper bound** for S if for all $x \in S$, $x \leq b$. Additionally, b is a **least upper bound** of B if

1. b is an upper bound of S
2. $b \leq u$ for every other upper bound m of S

If $S \subset \mathbb{R}$ is not bounded above or is empty, then $\sup S = \infty$.

Least upper bounds are unique. If b is an upper bound of S and $b \in S$, then b is the least upper bound.

Proposition 12. *Let $S \subset \mathbb{R}$ be nonempty. Then $b \in \mathbb{R}$ is the least upper bound of B if and only if b is an upper bound of S and for every $\varepsilon > 0$ there is an $x \in S$ such that $x > b - \varepsilon$.*

Proof. **Forward:** By definition, a least upper bound is an upper bound. Let $\varepsilon > 0$, then $b - \varepsilon/2 < b$ is not an upper bound. Then there exists some $x \in S$ such that if $b - \varepsilon/2 \leq x$, then $x > b - \varepsilon$.

Backward: Let $y < b$ and set $\varepsilon = b - y$. Since $y = b - \varepsilon$, there exists $x \in S$ such that $y < x$. Then y is not a least upper bound. This implies that b is the least upper bound. $\square$

Definition 15: Greatest Lower Bound/Infimum

You can guess the definition of this from how the supremum was defined. If the set S is unbounded of empty, $\inf S = -\infty$.

Proposition 13. *Let $A \subset B \subset \mathbb{R}$, then $\inf B \leq \inf A \leq \sup A \leq \sup B$.*

Proof. $\inf B \leq b$ for any $b \in B$, but any element $a \in A$ is also in B, so $\inf B \leq a$ for any $a \in A$. Thus $\inf B \leq \inf A$. Similarly, $\sup B \geq b$ for all $b \in B$, but since $a \in A \implies a \in B$, $\sup B \geq a$ for all $a \in A$. The middle inequality is trivial. □

Theorem 3

In $\mathbb{R}$ the following hold

1. Least upper bound property: Let $S \subset \mathbb{R}$ be non-empty and have an upper bound, then S also has a least upper bound.
2. Greatest lower bound property: Let $S \subset \mathbb{R}$ be non-empty and have a lower bound, then S also has a greatest lower bound.

Proof. We only prove the first half of this theorem. The second half holds by symmetry.

We can construct a decreasing sequence bounded below by the value that ends up being our least upper bound. Let M be an upper bound of S, and let n be a positive integer. Consider the sequence $M - 1/2^n, M - 2/2^n, M - 3/2^n, dots$ that steps down by $1/2$ at every point. Choose the first integer k such that $M - k/2^n$ fails to be an upper bound, and denote this by k_n (such a k_n is guaranteed to exist since S is nonempty and $\mathbb{R}$ is Archimedean).

Let $b_n = M - k_n/2^n$, so that b_n is *not* an upper bound but $b_n + 1/2^n$ *is* an upper bound. Since a larger k results in smaller steps, it is clear that $\{b_i\}$ is an increasing sequence bounded above by M. Then by the completeness of $\mathbb{R}$, $b_n \to b$ for some b, which we claim to be the least upper bound. Note that $b_n \leq b$ for all n.

To see this, assume $x \in S$ is greater than b. Then $x - b = \varepsilon$ for some $\varepsilon > 0$. Choose n large enough that $1/2^n < \varepsilon$, then

$$x = b + \varepsilon \geq b_n + \varepsilon > b_n + 1/2^n.$$

By construction, this is impossible. So by contradiction, there can be no elements of S that are larger than b. Thus b is the least upper bound of S. □

This theorem is equivalent to the completeness axiom for an ordered field.

1.4 Cauchy Sequences

Definition 16: Cauchy Sequence

A sequence x_n of real numbers is a Cauchy sequence if for all $\varepsilon > 0$, there is an integer N such that $|x_n - x_m| < \varepsilon$ whenever $n > N$ and $m > N$.

Proposition 14. *Every convergent sequence is a Cauchy sequence.*

Proof. $x_n \to x$, so choose N such that $|x_n - x| < \varepsilon/2$ if $n > N$. If $n, m > N$, then

$$\begin{aligned} |x_n - x_m| &= |x_n - x + x - x_m| \\ &\leq |x_n - x| + |x - x_m| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

□

Proposition 15. *Every Cauchy sequence is bounded.*

Proof. Let $\{x_n\}$ be Cauchy. Let $\varepsilon = 1$, then there exists N such that if $n, m > N$, then $|x_n - x_m| < 1$. If $n > N$, then

$$\begin{aligned} |x_n| &= |x_n - x_{N+1} + x_{N+1}| \\ &\leq |x_n - x_{N+1}| + |x_{N+1}| \\ &\leq 1 + |x_{N+1}|. \end{aligned}$$

Let $M = \max\{|x_0|, |x_1|, \ldots, |x_N|, 1 + |x_{N+1}|\}$, then $|x_n| \leq M$ for all n. □

Theorem 4: Bolzano-Weierstrass Property

Every bounded sequence in $\mathbb{R}$ has a subsequence that converges to some point in $\mathbb{R}$.

Proof. Proof in book. □

This means that for $a, b \in \mathbb{R}$ with $a < b$, every sequence of points in $[a, b]$ has a subsequence that converges to a point in $[a, b]$.

Proposition 16. *If a subsequence of a Cauchy sequence converges to x, then the sequence itself converges to x.*

Proof. Let $\{x_n\}$ be Cauchy, then we know it is bounded. By the Bolzano-Weierstrass property, it has a convergent subsequence $\{x_{\sigma(n)}\}$, i.e. $x_{\sigma(n)} \to L$. We claim $x_n \to L$ as well.

Given $\varepsilon > 0$, there exists N_σ such that if $j > N_\sigma$, then $|x_{\sigma(j)} - L| < \frac{\varepsilon}{2}$. Since $\{x_n\}$ is Cauchy, there exists N_c such that if $n, m > N_c$, then $|x_n - x_m| < \frac{\varepsilon}{2}$.

For $n > N_c$, we have

$$\begin{aligned} |x_n - L| &= |x_n - x_{\sigma(j)} + x_{\sigma(j)} - L| \\ &\leq |x_n - x_{\sigma(j)}| + |x_{\sigma(j)} - L| \end{aligned}$$

for every j. Now choose j sufficiently large such that $j > N_\sigma$ and $\sigma(j) > N_c$. Then this becomes

$$\begin{aligned} &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon \end{aligned}$$

Thus $x_n \to L$. □

Theorem 5: Cauchy Completeness

Every Cauchy sequence in $\mathbb{R}$ converges to an element of $\mathbb{R}$.

Proof. Every Cauchy sequence is bounded, so by the Bolzano-Weierstrass property, every Cauchy sequence has a subsequence that converges to some point in $\mathbb{R}$. But we have also shown that if a subsequence of a Cauchy sequence converges to a point, then the sequence itself converges to that point. Thus every Cauchy sequence converges to a point in $\mathbb{R}$. □

1.5 Limit Inferiors and Limit Superiors

Definition 17: Limit Superior/Limit Inferior

Let $\{x_n\}_{n=1}^{\infty} \subset \mathbb{R}$ be bounded above, and let $y_j \doteq \sup \{x_n\}_{n=j}^{\infty}$. Then $\{y_j\}_{j=1}^{\infty}$ is a non-increasing sequence. If $\lim_{j\to\infty} y_j = L$ exists, we define the **limit superior** to be

$$L = \overline{\lim}\, x_j = \limsup_{j\to\infty} x_j.$$

L clearly exists if y_j is bounded below.
Similarly, define the **limit inferior** to be

$$\underline{\lim}\, x_j = \liminf_{j\to\infty} x_j.$$

The limit inferior need not be the infimum, and the limit supremum need not be the supremum. The limit inferior is the limit of the infimums if we keep removing elements from the beginning of the sequence, and the limit superior is the limit of the supremums.

Proposition 17. *Let $\{x_n\}$ be a sequence in $\mathbb{R}$.*

1. *If $\{x_n\}$ is bounded below, a number a is equal to the limit inferior if and only if*
   (a) *For all $\varepsilon > 0$, there exists N such that $a - \varepsilon < x_n$ when $n > N$, and*
   (b) *For all $\varepsilon > 0$ and for all M, there exists $n > M$ with $x_n < a + \varepsilon$.*
2. *If $\{x_n\}$ is bounded above, a number b is equal to the limit superior if and only if*
   (a) *For all $\varepsilon > 0$, there exists N such that $x_n < b + \varepsilon$ when $n > N$, and*
   (b) *For all $\varepsilon > 0$ and for all M, there exists $n > M$ with $b - \varepsilon < x_n$.*

Proof. Do this. □

1.6 Euclidean Space

Definition 18: Euclidean Space

$\mathbb{R}^n$ constsis of all ordered n-tuples of real numbers

$$\mathbb{R}^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R}\}$$

Thus $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ is the Cartesian product of $\mathbb{R}$ with itself n times. $\mathbb{R}^n$ is a normed vector space.

Definition 19: Real Vector Space

A real vector space $\mathcal{V}$ is a set of vectors with operations of vector addition $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ and scalar multiplication $\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V}$ such that

1. v+w = w+v
2. (v+u)+w = v+(u+w)
3. there exists 0 such that $v + 0 = v$
4. for all v, there exists $-v$ such that $v + (-v) = 0$
5. For $\lambda \in \mathbb{R}$, $\lambda \cdot (v + w) = \lambda v + \lambda w$
6. For $\lambda, \gamma \in \mathbb{R}$, $\lambda(\gamma v) = \lambda\gamma \cdot v$
7. $(\lambda + \gamma) \cdot v = \lambda v + \gamma v$
8. $1 \cdot v = v$

We can make this a vector space over a general field F by substituting F for $\mathbb{R}$ throughout this definition.

A subset of $\mathcal{V}$ is a subspace if it's a vector space itself with the same operations. If $\mathcal{W} \subset \mathcal{V}$ over F, then $\mathcal{W}$ is a vector subspace over F if and only if $\lambda v + \gamma w \in \mathcal{W}$ whenever $\lambda, \gamma \in F$ and $v, w \in \mathcal{W}$.

An $(n-1)$-dimensional linear subspace of $\mathbb{R}^n$ is called a hyperplane. An affine hyperplane is a set $x + H = \{x + y \mid y \in H\}$, where H is a hyperplane and $x \in \mathbb{R}^n$.

Theorem 6

Euclidean n-space with addition and scalar multiplication is a vector space of dimension n.

Definition 20: Norm

The norm of $x \in \mathbb{R}^n$ is defined

$$|x| = \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}.$$

The distance between x and y is defined

$$d(x, y) = |x - y| = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2}.$$

The inner product is defined $\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$. Note that $|x|^2 = \langle x, x \rangle$. If $x, y \in \mathbb{R}^n$ are orthogonal, then $\langle x, y \rangle = 0 \iff \lambda \langle x, y \rangle = 0$. Two subspaces S and T are orthogonal if $\langle s, t \rangle = 0$ for all $s \in S, t \in T$. If they additionally span $\mathbb{R}^n$, then they are called orthogonal complements.

Proposition 18. *Let $v, w \in \mathbb{R}^n$, and let*

$$\rho(v, w) = \max \{ |v_1 - w_1|, |v_2 - w_2|, \ldots, |v_n - w_n| \}.$$

Then

$$\rho(v, w) \leq \|v - w\| \leq \sqrt{n}\, \rho(v, w)$$

Proof. Clearly

$$|v_i - w_i| = \sqrt{|v_i - w_i|^2} \leq \sqrt{\sum_{j=1}^{n} |v_j - w_j|^2} = \|v - w\|.$$

Additionally,

$$\|v - w\| = \sqrt{\sum_{j=1}^{n} |v_j - w_j|^2} \leq \sqrt{\sum_{j=1}^{n} \rho(v, w)^2} = \sqrt{n} \rho(v, w)$$

□

1.7 Norms, Inner Products, and Metrics

Definition 21: Metric Space

A **metric space** (M, d) is a set M and a function $d : M \times M \to \mathbb{R}$ such that

1. $d(x, y) \geq 0$
2. $d(x, y) = 0 \iff x = y$
3. $d(x, y) = d(y, x)$
4. $d(x, y) \leq d(x, z) + d(z, y)$

Example 2

1. $(\mathbb{R}^n, \|\cdot\|)$ is a metric space with $d(x, y) = \|x - y\|$.
2. The **discrete metric** is defined $d(x, y) = 1$ if $x \neq y$, $d(x, y) = 0$ if $x = y$. For any set S, (S, d) is a metric space.
3. In $\mathbb{R}^n$ with fixed basis, let $d(x, y) = \max_i |x_i - y_i|$. Then $(\mathbb{R}^n, d)$ is a metric space.

Definition 22: Normed Vector Space

A **normed vector space** $(\mathcal{V}, \|\cdot\|)$ is a vector space $\mathcal{V}$ and a function $\|\cdot\| : \mathcal{V} \to \mathbb{R}$ such that

1. $\|v\| \geq 0$
2. $\|v\| = 0 \iff v = 0$
3. $\|\lambda v\| = |\lambda| \|v\|$
4. $\|v + w\| \leq \|v\| + \|w\|$

Example 3

1. $(\mathbb{R}^n, \|x\| = \sqrt{x \cdot x})$
2. $(\mathbb{R}^n, \|x\| = \max_i |x_i|)$
3. Let l_∞ denote the vector space of bounded sequences in $\mathbb{R}$, then define $\|x\| = \sup_i \{|x_i| \mid i \in \mathbb{N}\}$.

Norms always produce metrics, since we can define a metric $d(v, w) = \|v - w\|$ on any normed vector space; however, not all metrics (e.g. discrete or bounded metrics) can be produced from norms.

Definition 23: Inner Product Space

A real vector space $\mathcal{V}$ with a function $\langle \cdot, \cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is an **inner product space** if

1. $\langle v, v \rangle \geq 0$
2. $\langle v, v \rangle = 0 \iff v = 0$
3. $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$ for all $\lambda \in \mathbb{R}$
4. $\langle v, w + u \rangle = \langle v, w \rangle + \langle v, u \rangle$
5. $\langle v, w \rangle = \langle v, w \rangle$

Example 4

1. In $\mathbb{R}$, $v \cdot w$ is an inner product.
2. Let $\ell_2 \subset \ell_\infty$ be bounded sequences $\{x_n\} \subset \mathbb{R}$ such that $\sum_{n=1}^{\infty} x_n^2$ is bounded. We can define an inner product by

$$\langle x, y \rangle = \sum_{n=1}^{\infty} x_n y_n.$$

Inner products always produce norms, since on any inner product space we can define a norm

$$\|v\| = \sqrt{\langle v, v \rangle}.$$

Two useful identities that aren't hard to prove:

1. $\langle \lambda v + \mu w, u \rangle = \lambda \langle v, u \rangle + \mu \langle w, u \rangle$
2. $\langle 0, w \rangle = \langle w, 0 \rangle = 0$

Proposition 19 (Cauchy-Schwarz Inequality). *If $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ is an inner product space, then we have $|\langle v, w\rangle| \leq \sqrt{\langle v, v\rangle}\sqrt{\langle w, w\rangle}$ for all $v, w \in \mathcal{V}$.*

Proof. We'll use the two expansions

$$0 \leq \langle x+y, x+y\rangle = \langle x, x\rangle + 2\langle x, y\rangle + \langle y, y\rangle$$
$$0 \leq \langle x-y, x-y\rangle = \langle x, x\rangle - 2\langle x, y\rangle + \langle y, y\rangle .$$

These yield the two inequalities

$$-2\langle x, y\rangle \leq \langle x, x\rangle + \langle y, y\rangle$$
$$2\langle x, y\rangle \leq \langle x, x\rangle + \langle y, y\rangle .$$

Note that this is the definition of absolute value, so

$$2|\langle x, y\rangle| \leq \langle x, x\rangle + \langle y, y\rangle$$

Now we're going to rescale this inequality. For any real number $t > 0$, we can write

$$\left|\left\langle tx, \frac{y}{t}\right\rangle\right| \leq \frac{t^2\langle x, x\rangle + \langle y, y\rangle/t^2}{2}$$

If either x or y is 0, then this inequality is trivial, so assume both are nonzero. This inequality holds for all t, so choose

$$t^2 = \frac{\sqrt{\langle y, y\rangle}}{\sqrt{\langle x, x\rangle}}$$

Plugging this in gives

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}$$

□

Now we can prove the triangle inequality for $\|\cdot\|$.

Proposition 20 (Triangle Inequality). $\|x+y\|^2 \leq \|x\|^2 + \|y\|^2$.

Proof.

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= \|x\|^2 + \|y\|^2\end{aligned}$$

□

1.8 The Complex Numbers

Fill this in from textbook.

Chapter 2

The Topology of Euclidean Space

2.1 Open Sets

Definition 24: ε-ball

Let (M, d) be a metric space. For each fixed $x \in M$ and $\varepsilon > 0$, the set

$$D(x, \varepsilon) = \{y \in M \mid d(x, y) < \varepsilon\}$$

is called the **ε-ball** about x.

Definition 25: Open

A set $A \subset M$ is **open** if for every $x \in A$, there exists an $\varepsilon > 0$ such that $D(x, \varepsilon) \subset A$.

Definition 26: Neighborhood

A **neighborhood** of x in M is an open set containing x.

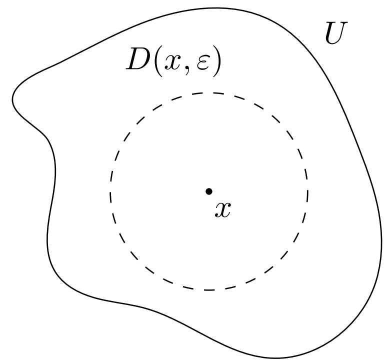

Figure 2.1: A neighborhood U of x

> **Proposition 21.** *In a metric space, every ε-ball $D(x, \varepsilon)$ is open for $\varepsilon > 0$.*

Proof. Let $y \in D(x, \varepsilon)$, then $d(x, y) < \varepsilon$. Set

$$r = \frac{\varepsilon - d(x,y)}{2}.$$

Now let $z \in D(y, r)$, then

$$\begin{aligned} d(x,z) &\leq d(x,y) + d(y,z) \\ &< d(x,y) + \frac{\varepsilon - d(x,y)}{2} \\ &= \frac{\varepsilon + d(x,y)}{2} \\ &< \varepsilon. \end{aligned}$$

Thus there exists $D(y, r) \subset D(x, \varepsilon)$ for any $y \in D(x, \varepsilon)$, so $D(x, \varepsilon)$ is open. □

> **Proposition 22.** *In (M, d) with open sets U_i,*
>
> 1. *$\bigcap_{i=1}^N U_i$ is open*
> 2. *$\bigcup_{\alpha \in A} U_\alpha$ is open*
> 3. *$\varnothing$ and M are open*

Proof.
1. Let $U_1, \dots, U_N$ be open, and let $y \in \bigcap_{i=1}^N U_i$ (if we can't find such a y_i, then the intersection is the empty set, which is open). Then $y_i \in U_i$ for all i. This implies that for each i, there exists r_i such that $D(y, r_i) \subset U_i$. Set $r = \min\{r_1, \dots, r_N\}$, then $D(y, r) \subset D(y, r_i) \subset U_i$ for all i. This implies $D(y, r) \subset \bigcap_{i=1}^N U_i$, so the intersection is open.

2. Let $x \in \bigcup_{\alpha \in A} U_\alpha$, where U_α is open for each α. This implies $x \in U_\beta$ for some $\beta \in A$, so there exists r such that $D(x, r) \subset U_\beta \subset \bigcup_{\alpha \in A} U_\alpha$. This implies that the union is open.

3. This is clear.

□

Example 5

Let $U_n = \left(-\frac{1}{n}, \frac{1}{n}\right)$, then $\bigcap_{n=1}^{\infty} U_n = \{0\}$, which is not open. Thus statement (1) does not hold for arbitrary collections of open sets.

2.2 Interior of a Set

Definition 27: Interior

If $A \subset (M, d)$, then let $A^o \doteq \{x \in A \mid \text{there exists } D(x, \varepsilon) \subset A\}$

Example 6

$[0, 1]^o = (0, 1)$.

Definition 28: Interior Point

A point $a \in A$ is an **interior point** if there's an open set U such that $a \in U \subset A$. The interior of a set is all that set's interior points.

The interior of A is the union of all open subsets of A. Since A^o is open, A^o is the largest open subset of A. Thus if A has no open subsets, then $A^o = \varnothing$. Furthermore, if A is open, then $A^o = A$.

2.3 Closed Sets

Definition 29: Closed

A set B in a metric space M is said to be **closed** if its complement $B^c = M \backslash B$ is open.

It's possible for a set to be neither open nor closed (consider $(0, 1] \in \mathbb{R}$).

Proposition 23. *In (M,d) with open sets C_i,*

1. $\bigcup_{i=1}^{N} U_\alpha$ *is closed*
2. $\bigcap_{\alpha \in A} C_\alpha$ *is closed*
3. $\varnothing$ *and* M *are closed*

Proof. The proofs of the first two statements follow almost immediately from DeMorgan's laws, so we can prove those first. Let $\{F_\alpha\}$ be a collection of closed sets. Write $F_\alpha = U_\alpha^c$ for open U_α. Then $\cup_\alpha F_\alpha = \cup_\alpha U_\alpha^c$.

Let $x \in \cup_\alpha U_\alpha^c$, then $x \in U_\beta^c$ for some β and, subsequently, $x \notin U_\beta$. This implies $x \notin \cap_\alpha U_\alpha$, so $x \in (\cap_\alpha U_\alpha)^c$.

Now let $y \in (\cap_\alpha U_\alpha)^c$, then $y \notin \bigcap_\alpha U_\alpha$. Then $y \notin U_\beta$ for some β, so $y \in U_\beta^c$. This means $y \in \cup_\alpha U_\alpha^c$.

These two implications combine to give $\cup_\alpha U_\alpha^c = (\cap_\alpha U_\alpha)^c$. Similarly, we can derive $\cap_\alpha U_\alpha^c = (\cup_\alpha U_\alpha)^c$. With DeMorgan's laws proven, we can move on to the proofs of the statements we actually care about.

1. The intersection of finite open sets is open, so $\cup_{j=1}^{N} F_j = \left(\cap_{j=1}^{N} U_j\right)^c$ must be closed by definition.
2. The union of an arbitrary collection of open sets is open, so $\cap_\alpha F_\alpha = (\cup_\alpha U_\alpha)^c$ must be closed by definition.
3. $\varnothing$ and M are each other's complements, and they are both open. Thus by definition they are both also closed.

$\square$

Example 7

Any finite set in $\mathbb{R}^n$ is closed since it is the union of finitely many single points, which themselves are closed sets.

Example 8

Let

$$F_n = \left[\frac{1}{n}, 1 - \frac{1}{n}\right].$$

The union $\bigcup_{j=1}^{\infty} F_j = (0,1)$, so the union of an arbitrary collection of closed sets is not necessarily closed.

2.4 Accumulation Points

Definition 30: Accumulation Point

A point x in metric space M is an **accumulation point** of $A \subset M$ if every open set U containing x contains some point of A other than x. Equivalently, for all $\varepsilon > 0$, $D(x, \varepsilon)$ contains $y \in A$ such that $y \neq x$. This can also be written $(D(x, \varepsilon) \setminus \{x\}) \cap A \neq \varnothing$.
The set of accumulation points of A is denoted by $\operatorname{acc}(A)$.

Other points of A get arbitrarily close to x if x is an accumulation point. This means there are infinitely many points of A that are close to x.

An accumulation point of a set doesn't need to be in the set itself. A set also doesn't need to have any accumulation points in the first place.

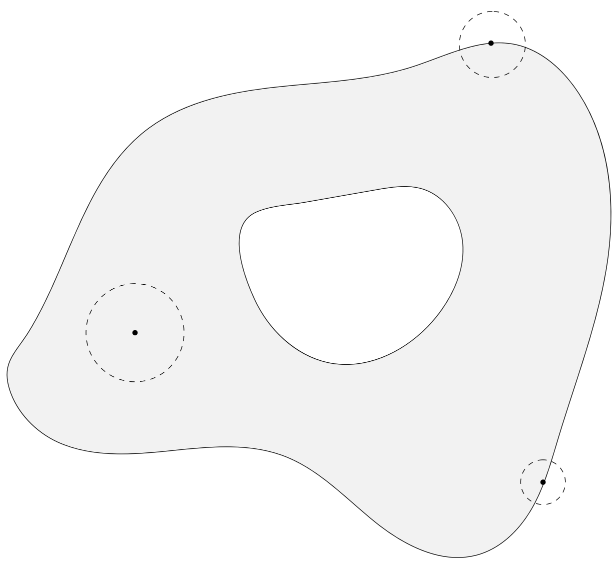

Figure 2.2: Accumulation points of a set

Example 9

1. Let $S \subset \mathbb{R}$ be bounded, and let x be a least upper bound of S. Let $D(x, \varepsilon)$ be the ε-ball around x, then if $D(x, \varepsilon) \cap S = \varnothing$, then $x - \varepsilon/2$ is an upper bound for S. This is a contradiction, so $D(x, \varepsilon) \cap S \neq \varnothing$, meaning x is an accumulation point of S.

2. Discrete metric spaces have no accumulation points.

Proposition 24. *Every point in A^o is an accumulation point of $A \subset \mathbb{R}^n$.*

Proof. Let $x \in A^o$, then there exists $\varepsilon > 0$ such that $D(x, \varepsilon) \subset A^o \subset A$. Then $(D(x, \varepsilon) \backslash \{x\}) \cap A$ is nonempty. $\square$

Proposition 25. *$A \subset (M, d)$ is closed if and only if the accumulation points of A belong to A.*

Proof. **Forward:** Let $F \subset (M, d)$ be closed, and let $x \in F^c$. Since F^c is open, there exists $\varepsilon > 0$ such that $D(x, \varepsilon) \subset F^c$. This implies $D(x, \varepsilon) \cap F = \varnothing$, so x is *not* an accumulation point of F. Thus F contains all its accumulation points.

Backward: Let $F \subset (M, d)$ be a set which contains its accumulation points. Let $p \in F^c$, then it is *not* an accumulation point of F. Then there exists an open neighborhood U of p such that $U \cap F = \varnothing$. Then there exists $\varepsilon > 0$ such that $D(p, \varepsilon) \subset U \subset F^c$, so F^c is open, so F is closed. $\square$

If a set has no accumulation points, then it satisfies this condition and is thus closed.

2.5 Closure of a Set

Definition 31: Closure

Let $A \subset (M, d)$, then $\overline{A}$ is the intersection of all closed sets containing A. It is the smallest closed set containing A.

Since the intersection of any family of closed sets is closed, $\overline{A}$ is closed.

Proposition 26. *For $A \subset M$, $\bar{A} = A \cup acc(A)$.*

Proof. Let $B = A \cup \text{acc}(A)$. By Proposition 25, any closed set containing A also contains B. If B is closed, then it will be therefore be the smallest closed set containing A, i.e. $B = \overline{A}$. To show that B is closed, we use Proposition 25 again.

Let y be an accumulation point of B. If $\varepsilon > 0$, then $D(y, \varepsilon)$ contains other points of B. Let $z \in B$ be one of them, then either $z \in A$ or $z \in \text{acc}(A)$. In the latter case, $D(z, \varepsilon - d(z, y))$ is an open set containing z, and so by definition it must contain other points of A that are distinct from y. Thus y is also an accumulation point of A, so $y \in B$. This implies that B is closed and, subsequently, equal to the closure of A. $\square$

2.6 Boundary of a Set

Definition 32: Boundary

The **boundary** of $A \subset (M, d)$ is defined $\partial A = \overline{A} \cap \overline{A^c}$.

The union of 2 closed sets is closed, so ∂A is closed. Note $\partial A = \partial A^c$.

Proposition 27. *Let $A \subset M$, then $x \in \partial A$ if and only if for all $\varepsilon > 0$, $D(x, \varepsilon)$ contains points of A and A^c (these points might include x itself).*

Proof. **Forward:** Let $x \in \partial A$. If $x \in A$, then it must be an accumulation point of A^c. If $x \in A^c$, then it must be an accumulation point of A. Either way, the conclusion follows.

Backward: The same cases apply here, and the argument is similar. □

$x \in \partial A$ need not be an accumulation point.

Example 10

1. Let $A = (0, 1)$, then $\overline{A} = [0, 1]$ and $\overline{A^c} = (-\infty, 0] \cup [1, \infty)$. Then $\partial A = \{0, 1\}$.
2. Let $A = \mathbb{Q}$, then $\overline{A} = \mathbb{R}$. $A^c = \mathbb{R} \backslash \mathbb{Q}$, so $\overline{A_c} = \mathbb{R}$. Thus $\partial \mathbb{Q} = \mathbb{R}$.

2.7 Sequences in Metric Spaces

Definition 33: Convergence

A sequence $\{x_n\}_{n=1}^{\infty} \subset (M, d)$ **converges** to $L \in M$ if for all $\varepsilon > 0$, there exists some N such that if $n > N$, then $d(x_n, L) < \varepsilon$.
Equivalently, $x_n \to L$ if for every open neighborhood U of L, there exists N such that if $n > N$, then $x_n \in U$.

Like a good mathematician, you gotta prove that these two definitions are actually equivalent.

Proof. **First implies second:** Let U be an open neighborhood of L, then there exists $\varepsilon > 0$ such that $D(L, \varepsilon) \subset U$. We know that there exists N such that if $n > N$, then $d(x_n, L) < \varepsilon$, which implies $x_n \in D(L, \varepsilon)$, which itself implies $x \in U$.

Second implies first: Fix $\varepsilon > 0$, then $D(L, \varepsilon)$ is an open neighborhood of L. Then there exists N such that for $n > N$, $x_n \in D(L, \varepsilon)$. This implies $d(x_n, L) < \varepsilon$. □

Example 11

For the discrete metric, a sequence $\{x_n\}$ converges if and only if it is constant for big enough n.

Some of the properties of limit arithmetic also carry over into the more general case, as well.

Proposition 28. *If $(\mathcal{V}, \|\cdot\|)$ is a normed vector space and $\{v_k\}, \{w_k\} \subset \mathcal{V}$ such that $v_k \to v$ and $w_k \to w$, and if $\{\lambda_k\} \subset \mathbb{R}$ such that $\lambda_k \to \lambda$, then*

1. $v_k + w_k \to v + w$
2. $\lambda_k v_k \to \lambda v$

The proofs of these don't really change when we go from $\mathbb{R}$ to $\mathcal{V}$, so hopefully you put the proofs for the earlier limit arithmetic properties in these notes somewhere.

A corollary of this is $w_k \to w \iff w_k - w \to 0$ for all sequences in normed vector spaces.

Proposition 29. *$F \subset (M, d)$ is closed if an only if for all sequences in F that converge to a point in M, that point is also in F.*

Proof. **Forward:** Let F be closed, and let $\{x_n\} \subset F$ such that $x_n \to L \in M$. Then every neighborhood of L contains elements of F. This implies that L is an accumulation point of F. Since F is closed, it contains its accumulation points. Thus $L \in F$.

Backward: Suppose that every sequence $\{x_n\} \subset F$ satisfies $x_n \to L \in M \implies L \in F$. Let p be an accumulation point of F, then for all n, we have $D(p, 1/n) \cap F \neq \varnothing$. This means we can find a point $x_n \in D(p, 1/n) \cap F$. By construction, $x_n \to p$. By assumption, this implies $p \in F$. Then F contains its accumulation points, so F is closed. □

Proposition 30. *For a set $A \subset (M, d)$, $x \in \overline{A}$ if and only if there is a sequence $x_k \in A$ with $x_k \to x$.*

Proof. This argument is similar to the previous proof. □

Example 12

Consider the open interval $(0, 1)$ with the usual metric. The sequence $\{1/n\}$ does *not* converge in this metric space since $0 \notin M$.

Proposition 31. *$v_k \to v$ in $\mathbb{R}^n$ if and only if each sequence of coordinates converges to the corresponding coordinate of v as a sequence in $\mathbb{R}$. That is, $\lim v_k = v$ in $\mathbb{R}^n$ if and only if $\lim v_k^i = v^i$ in $\mathbb{R}$ for each $i = 1, 2, \ldots, n$.*

Proof. **Forward:** If $\delta(v, v_k) = \max\left\{|v^1 - v_k^1|, \ldots, |v^n - v_k^n|\right\}$, then $\delta(v, v_k) \leq \|v - v_k\| \leq \sqrt{n}\delta(v, v_k)$ (see Proposition 18). Suppose $v_k \to v$ in $\mathbb{R}^n$ and $1 \leq j \leq n$. Let $\varepsilon > 0$, then there is an integer K such that $\|v - v_k\| < \varepsilon$ whenever $k \geq K$. Therefore, for such k, we have $|v^j - v_k^j| \leq \delta(v, v_k) \leq \|v - v_k\| < \varepsilon$, and so $v_k^j \to v^j$ in $\mathbb{R}$.

Backward: Suppose that for each $j = 1, \ldots, n$, we have $\lim v_k^j = v^j$. Let $\varepsilon > 0$, then for any $\varepsilon_0 > 0$, there are integers $K_1, K_2, \ldots, K_n$ such that

$$\begin{gathered}
|v^1 - v_k^1| < \varepsilon_0 \text{ whenever } k \geq K_1 \\
|v^2 - v_k^2| < \varepsilon_0 \text{ whenever } k \geq K_2 \\
\vdots \\
|v^n - v_k^n| < \varepsilon_0 \text{ whenever } k \geq K_n
\end{gathered}$$

There are only finitely many K_i, so one of them is the largest. Let $K = \max(K_1, \ldots, K_n)$. For $k \geq K$, we know $\|v - v_k\| \leq \sqrt{n}\delta(v, v_k) < \sqrt{n}\varepsilon_0$. If this is done with $\varepsilon_0 = \varepsilon/\sqrt{n}$, we obtain $\|v - v_k\| < \varepsilon$ whenever $k \geq K$, and so $v_k \to v$ in $\mathbb{R}^n$. □

2.8 Complete Metric Spaces

Definition 34: Cauchy Sequence

The sequence $\{x_n\} \subset (M, d)$ is a **Cauchy sequence** if for all $\varepsilon > 0$, there exists N such that if $n, m > N$, then $d(x_n, x_m) < \varepsilon$.

Unlike convergence, Cauchy sequences are metric-dependent. There's no real way to define them in terms of open neighborhoods. This is what the prof said, but that doesn't seem right?

Definition 35: Completeness

A metric space (M, d) is **complete** if every Cauchy sequence in M converges.

Example 13

1. $\mathbb{R}^n$ is complete.
2. Any discrete metric space is complete.

Definition 36: Bounded

A set $A \subset (M, d)$ is **bounded** if there exists some $p \in M$ and $R > 0$ such that $A \subset D(p, R)$.
A sequence is bounded if and only if its image is bounded (remember that we defined sequences to be functions, so their image is the set of all points in the sequence).

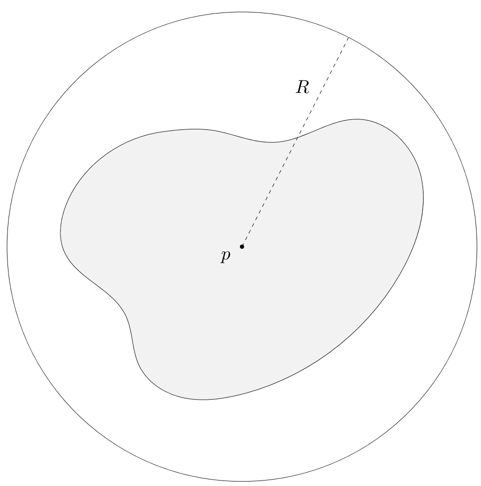

Figure 2.3: A bounded set

Proposition 32. *A convergent sequence in a normed vector space or metric space is bounded.*

Proof. If $x_n \to x$, then there exists N such that $d(x_n, x) < 1$ whenever $n > N$. Thus for all $n > N$, $x_n \in D(x, 1)$. Let $M = \max\{d(x_1, x), \ldots, d(x_N, x), 1\}$, then $d(x_i, x) \leq M$ for all i. $\square$

Proposition 33. *Just as in $\mathbb{R}$, we can develop some properties of Cauchy sequences.*

1. *Every convergent sequence in a metric space is a Cauchy sequence.*
2. *A Cauchy sequence in a metric space is bounded.*
3. *If a subsequence of a Cauchy sequence converges to x, then the sequence converges to x.*

Proof. The proofs of these are essentially the same as the proofs for $\mathbb{R}$, except absolute values are replaced with distances. $\square$

Example 14

Consider $(0, \infty)$ with metric $d(x, y) = |\ln(x/y)|$. This space is complete (although it's *not* complete under the usual metric). Now define $f : (0, \infty) \leftrightarrow \mathbb{R}$ by $x \mapsto \ln x$, then $|f(x) - f(y)| = |\ln x - \ln y| = |\ln(x/y)|$. Since $\mathbb{R}$ is complete and f is an isometry (the points in $\mathbb{R}$ that get mapped to are the same distance away as the original points in our interval), then the space $(0, \infty)$ with metric $|\ln(x/y)|$ is complete.

Theorem 7

A sequence $\{x_k\} \subset \mathbb{R}^n$ converges to a point in $\mathbb{R}^n$ if and only if it is a Cauchy sequence.

Proof. **Forward:** If $x_k \to x$, then for $\varepsilon > 0$, choose N so that $k \geq N$ implies $\|x_k - x\| < \varepsilon/2$. Then for $k, l \geq N$, $\|x_k - x_l\| = \|(x_k - x) + (x - x_l)\| \leq \|x_k - x\| + \|x - x_l\| < \varepsilon/2 + \varepsilon/2 = \varepsilon$. Thus $\{x_k\}$ is a Cauchy sequence.

Backward: Suppose $\{x_k\}$ is Cauchy. Since $|x_k^i - x_l^i| \leq \|x_k - x_l\|$, the components are also Cauchy sequences in $\mathbb{R}$. Since $\mathbb{R}$ is complete and every Cauchy sequence in $\mathbb{R}$ converges to a point in $\mathbb{R}$, $x_k^i \to x^i$ for each i. Since we have pointwise convergence, we also have convergence of the full sequence $\{x_k\}$. $\square$

Definition 37: Cluster Point

A point x in a metric space is called a **cluster point** of $\{x_k\}$ if for every $\varepsilon > 0$, there are infinitely many values of k with $d(x_k, x) < \varepsilon$.

Proposition 34. *If $\{x_k\} \subset (M, d)$ and $x \in M$, then*

1. *x is a cluster point if and only if for every $\varepsilon > 0$ and for each integer N, there is a $k > N$ with $d(x_k, x) < \varepsilon$.*
2. *x is a cluster point if and only if there is a subsequence convergent to x.*
3. *$x_k \to x$ if and only if every subsequence converges to x.*
4. *$x_k \to x$ if and only if every sebsequence of $\{x_k\}$ has a further subsequence that converges to x.*

Proof. These are also analagous to the proofs for $\mathbb{R}$. I'm sensing a pattern in this section... □

2.9 Series in Normed Vector Spaces

Let $(\mathcal{V}, |\cdot|)$ be a normed vector space and let $\{x_i\}_{i=1}^{\infty} \subset \mathcal{V}$. Set $S_n \doteq \sum_{i=1}^{n} x_i$. If $S_n \to L$, we say $\sum_{i=1}^{\infty} x_i$ is convergent and $\sum_{i=1}^{\infty} x_i = L$. If $\{S_n\}$ does *not* converge, we say $\sum_{i=1}^{\infty} x_i$ does not converge.

If $\mathcal{V} = \mathbb{R}$, we say $S_i \to \infty$ if for all M, there exists N such that if $n > N$, then $S_n > M$. If $S_i \to \pm\infty$, we say $\sum_{i=1}^{\infty} x_i = \pm\infty$ (respectively).

Definition 38: Banach Space

A **Banach space** is a complete normed vector space.

Definition 39: Hilbert Space

A **Hilbert space** is a complete inner product space.

Theorem 8

Let $\mathcal{V}$ be a complete normed vector space. A series $\sum x_k$ converges if and only if for every $\varepsilon > 0$, there is an N such that $k > N$ implies

$$\|x_k + x_{k+1} + \cdots + x_{k+p}\| < \varepsilon$$

for all integers $p = 0, 1, 2, \ldots$

Proof. Let $s_k = \sum_{i=1}^{k} x_k$. Since $\mathcal{V}$ is complete, a $\{s_k\}$ converges if and only if it is a Cauchy sequence. This is true if and only if there is an N such that $l > N$ implies $\|s_l - s_{l+q}\| < \varepsilon$ for all $q = 1, 2, \ldots$. But $\|s_l - s_{l+q}\| = \|x_{l+1} + \cdots + x_{l+q}\|$, and so the result follows with $k = l + 1$ and $p = q - 1$. □

Theorem 9

In a complete normed vector space, if $\sum x_k$ converges absolutely, then $\sum x_k$ converges.

Proof. This follows from Theorem 8 and the triangle inequality

$$\|x_k + \cdots + x_{k+p}\| \leq \|x_k\| + \cdots + \|x_{k+p}\|.$$

□

Finish this section.

Chapter 3

Compact and Connected Sets

3.1 Compactness

By "compact", we want to convey that a set is somehow closed and bounded. The standard definition of compactness, which is equivalent to the intuitive definition when we are in $\mathbb{R}^n$, is formulated in terms of open covers.

Definition 40: Cover

A **cover** of $A \subset (M, d)$ is a collection $\{U_i\}$ of sets whose union contains A. It is an **open cover** if each U_i is open (in which case the union is always also open). A **subcover** of a given cover is a subcollection of $\{U_i\}$ whose union also contains A (or "covers" A). A **finite subcover** is a subcover composed of only a finite number of sets.

Example 15

Let $A = [0, 1]$, and let $U_x = (x - 1/2, x + 1/2)$. The set $\{U_x\}_{x \in A}$ is clearly an open cover of A. The set $\left\{U_0, U_{1/2}, U_1\right\}$ clearly a finite subcover of A.

Definition 41: Sequentially Compact

$A \subset (M, d)$ is **sequentially compact** if every sequence in A has a subsequence that converges to a point in A.

Definition 42: Compact

$A \subset (M, d)$ is **compact** if every open cover of A has a finite subcover.

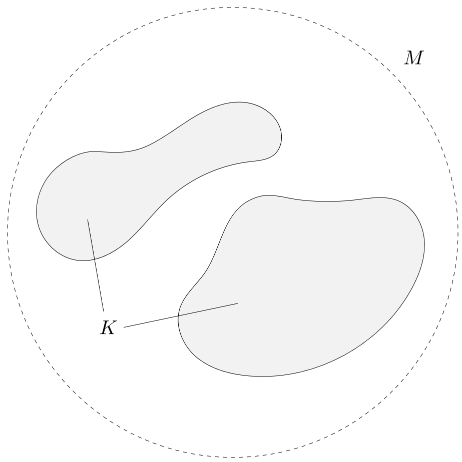

Figure 3.1: An intuitive depiction of what a compact set $K \subset (M, d)$ could look like

Theorem 10: Bolzano-Weierstrass Theorem

$A \subset (M, d)$ is compact if and only if A is sequentially compact.

Proof. Proof on page 167 of textbook. □

Definition 43: Totally Bounded

$A \subset (M, d)$ is **totally bounded** if for every $\varepsilon > 0$, there is a finite set $\{x_1, \ldots, x_{N(\varepsilon)}\} \subset M$ such that

$$A \subset \bigcup_{i=1}^{N(\varepsilon)} D(x_i, \varepsilon).$$

Proposition 35. *A compact set is closed.*

Proof. Let K be compact, let $x \in K^c$, and let $U_n = \{y \mid d(y, x) > 1/n\}$. Then $\{U_n\}_{n=1}^{\infty}$ covers $\{x\}^c$, so it also covers K. Since $U_n \subset U_{n+1}$, there is some

U_N which covers K by itself. Then $K \subset \{y \mid d(x,y) > 1/N\}$, which implies $D(x, 1/N) \subset K^c$. Thus K_c is open and, subsequently, K is closed. $\square$

Proposition 36. *A closed subset of a compact set is compact.*

Proof. Let K be compact and $A \subset K$ be closed, and let $\mathcal{U}$ be a collection of open sets that acts as an open cover of A. Since A is closed, A^c is open, so $\mathcal{U} \cup \{A^c\}$ is an open cover of K. Then there exists a finite subcover $\mathcal{U}' \cup \{A^c\}$ of K. Then $\mathcal{U}'$ covers A. Since $\mathcal{U}'$ is finite and was derived from an arbitrary open cover, A is compact. $\square$

Proposition 37. *A sequentially compact set is totally bounded.*

Proof. We can prove this by contradiction. Let A be sequentially compact, and suppose it is not totally bounded. Then there is some $\varepsilon > 0$ such that A is not covered by a finite number of ε-balls. Now let

$$\begin{aligned} y_1 &\in A \\ y_2 &\in A \backslash D(y_1, \varepsilon) \\ &\vdots \\ y_{n+1} &\in A \backslash \bigcup_{j=1}^{n} D(y_j, \varepsilon) \end{aligned}$$

We can always find a y_j in each of the above sets since A is assumed to not be totally bounded.

Since A is sequentially compact, there exists a subsequence $y_{\sigma(j)}$ such that $y_{\sigma(j)} \to y \in A$. This means the subsequence is Cauchy, i.e. there exists N such that $d(y_{\sigma(j)}, y_{\sigma(k)}) < \varepsilon/2$ for any $\varepsilon > 0$ when $\sigma(j), \sigma(k) > N$. But we constructed each y_j to be at least ε away from all other y_j'. Thus by contradiction, A is totally bounded. $\square$

Proposition 38. *A compact set is bounded.*

Proof. Let $A \subset (M,d)$ be compact, then $A \subset \bigcup_{j=1}^{\infty} D(0,j)$. Since A is compact, we must be able to find a finite subcover $\bigcup_{k=1}^{N} D(0, j_k)$ of A. Since $D(0,k) \subset D(0,l)$ when $k < l$, this implies that $D \subset D(0, j')$ for $j' = \min\{j_1, \ldots, j_N\}$. $\square$

Theorem 11

(M,d) is compact if and only if M is complete and bounded. Similarly, $A \subset (M,d)$ is compact if and only if A is closed and bounded.

Proof. Get proof from book. □

Definition 44: Finite Intersection Property

A collection of closed sets $\{K_\alpha\}$ in a metric space M has the **finite intersection property** for A if the intersection of any finite number of the K_α with A is nonempty.

Proposition 39. *$A \subset (M, d)$ is compact if and only if every collection of closed sets with the finite intersection property for A has nonempty intersection with A.*

Proof. **Forward:** Assume A is compact. Let $\{F_i\}$ be a collection of closed setes and let $U_i = F_i^c$, so that U_i is open. Suppose $A \cap (\cap_{i=1}^\infty F_i) = \varnothing$. Taking complements, this means that $\{U_i\}$ covers A. Since the cover is open, there is a finite subcover, say, $A \subset U_1 \cup \cdots \cup U_N$. Then $A \cap (F_1 \cap \cdots \cap F_N) = \varnothing$. This implies that $\{F_i\}$ doesn't have the finite intersection property. Thus $A \cap \{F_i\} \neq \varnothing$ for any finite collection $\{F_i\}$ satisfying the finite intersection property.

Backward: Let $\{U_i\}$ be an open cover of A and let $F_i = U_i^c$. Then $A \cap (\cap_{i=1}^\infty F_i) = \varnothing$, and so, by assumption, $\{F_i\}$ cannot have the finite intersection property for A. Thus $A \cap (F_1 \cap \cdots \cap F_N) = \varnothing$ for some members $F_1, \ldots, F_N$ of the collection. Then $U_1, \ldots, U_N$ is the required finite subcover, so A is compact. □

3.2 The Heine-Borel Theorem

Theorem 12: Heine-Borel Theorem

$A \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.

Proof. We already know that compact sets are closed and bounded (see lemmas in the proof of the Bolzano-Weierstrass theorem, Theorem 10, so we only need to show the backward implication.

If we show that a closed and bounded set A is sequentially compact, then by Bolzano-Weierstrass, it is also compact. We will use the fact that every bounded subset of $\mathbb{R}$ has a convergent subsequence (see the Bolzano-Weierstrass property, Theorem 4. Let $\{x_k\} = \left\{x_k^1, x_k^2, \ldots, x_k^n\right\} \subset A$ be a sequence (those are superscripts, not exponents). Since A is bounded, $\left\{x_k^1\right\}$ has a convergent subsequence, say, $\{x_{f_1(k)}^1\}$. Then $\{x_{f_1(k)}^2\}$ has a convergent subsequence, say, $\{x_{f_2(k)}^2\}$. Continuing, we get $\{x_{f_n(k)}\} = \{x_{f_n(k)}^1, \ldots, x_{f_n(k)}^n\}$, all of whose components converge. Pointwise convergence in $\mathbb{R}^n$ implies overall convergence, so

$\{x_{f_n(k)}\}$ converges to some point in $\mathbb{R}^n$. Since A is closed, the limit must lie in A. Thus A is sequentially compact, and so is compact. $\square$

3.3 Nested Set Property

Theorem 13: Nested Set Property

Let $\{F_k\}$ be a sequence of compact nonempty sets in a metric space M such that $F_{k+1} \subset F_k$, then their intersection is nonempty, i.e.

$$\bigcap_{k=1}^{\infty} F_k \neq \varnothing.$$

Proof. Here are two proofs of this theorem.

Proof 1: In the compact set $A = F_1$, the compact nonempty sets $F_1, F_2, \ldots$ have the finite intersection property, since $F_1 \cap \cdots \cap F_k = F_k$ and $F_k \cap A \neq \varnothing$ for any k. We've already proven that A is compact if and only if every collection of closed sets with the finite intersection property for A has nonempty intersection with A. Thus for any $\{F_k\}$, we have $A \cap \{F_k\} \neq \varnothing$.

Proof 2: Pick $x_k \in F_k$ for each k. The set F_1 is compact and the sequence $\{x_k\}$ lies in F_1, so by Bolzano-Weierstrass, $\{x_k\}$ has a convergent subsequence. Since each F_k is closed, the limit point must lie in all of them. $\square$

This can be inverted in a sense. Let $A_k = F_k^c$, then each U_k is open and $U_{k+1} \supset U_k$. Then $\cup_{k=1}^{\infty} U_k \neq M$. Thus if M is a metric space and the open sets U_k are increasing and have compact complements, then the union of all the U_k's is not all of M.

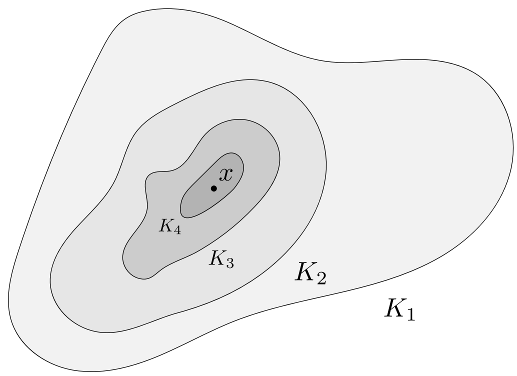

Figure 3.2: As long as each K_i is compact, x is guaranteed to exist

3.4 Path-Connected Sets

Definition 45: Continuous

For metric space (M, d), a map $\phi : [a, b] \to M$ is **continuous** if $t_k \to t$ implies $\phi(t_k) \to \phi(t)$ for every sequence $\{t_k\} \subset [a, b]$ converging to some $t \in [a, b]$.

Definition 46: Continuous Path

A **continuous path** joining two points x and y in M is a mapping $\phi : [a, b] \to M$ such that $\phi(a) = x, \phi(b) = y$, and ϕ is continuous (here x may or may not equal y, and $b \geq a$). A path ϕ **lies in a set** A if $\phi(t) \in A$ for all $t \in [a, b]$.

Definition 47: Path-Connectedness

A set is **path-connected** if every two pionts in the set can be joined by a continuous path lying in the set.

A path-connected set need not be open or closed. Consider $[0, 1], (0, 1)$, and $[0, 1)$, which are all connected.

3.5 Connected Sets

Definition 48: Connected

Let $A \subset (M, d)$, then two open sets U, V **separate** A if

1. $U \cap V \cap A = \varnothing$,
2. $A \cap U \neq \varnothing$,
3. $A \cap V \neq \varnothing$, and
4. $A \subset U \cup V$.

A is **disconnected** if such sets exist, and it is **connected** if no such sets exist.

Get rid of the A in the definition maybe? Would make it seem simpler and that's what prof did.

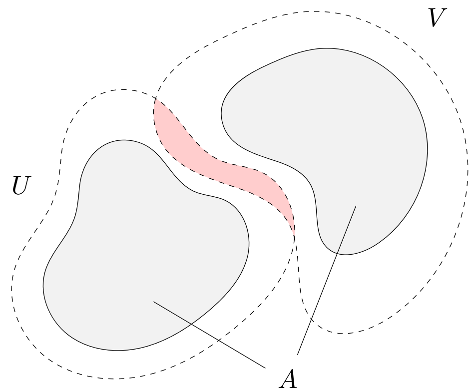

Figure 3.3: A disconnected set.

Proposition 40. *$[a, b]$ is connected.*

Proof. Suppose there exists open sets U_1, U_2 such that $[a, b] \subset U_1 \cup U_2$, $U_1 \cap U_2 = \varnothing$, $U_1 \neq \varnothing$, and $U_2 \neq \varnothing$. Without loss of generality, assume $a \in U_1$. Since U_1 is open, $[a, s) \subset U_1$ for some $s > a$. We can generalize this into a set.

Let $S \doteq \{x \in [a, b] \mid [a, x] \subset U_1\}$. We know S is nonempty and bounded above, so it has a supremum, which we denote by c. If $c = b$, then $U_2 = \varnothing$, so by contradiction, $c < b$. There are now two possibilities.

1. Suppose $c \in U_1$. Since U is open, there exists $\varepsilon > 0$ such that $(c-\varepsilon, c+\varepsilon) \subset U_1$. This contradicts c being an upper bound.
2. Suppose $c \notin U_1 \implies c \in U_2$. Since U_2 is open, there exists $\varepsilon > 0$ such that $(c-\varepsilon, c+\varepsilon) \subset U_2$. Thus $c-\varepsilon$ is an upper bound of U_1, which contradicts c being the least upper bound.

Since all possibilities led to contradiction, the assumptions were false. Thus $[a, b]$ is connected. □

Theorem 14

Path-connected sets are connected.

Proof. Look this up in book. □

Definition 49: Component

A **component** of a set A is a connected subset $A_0 \subset A$ such that there is no connected set in A containing A_0 (other than A_0 itself). Thus a component of A is a maximal connected subset of A.

Definition 50: Path Component

This is similar to components, except A_0 is the maximal path-connected subset of A.

Chapter 4

Continuous Mappings

4.1 Continuity

Definition 51: Limit of a Function

Let $f : A \subset M_1 \to M_2$. Suppose that x_0 is an accumulation point of A, then $b \in M_2$ is the **limit of f at x_0**

$$\lim_{x \to x_0} f(x) = b$$

if given any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in A$ satisfying $x \neq x_0$ and $d_1(x, x_0) < \delta$, we have $d_2(f(x), b)) < \varepsilon$.
Written in terms of open sets, this says that for all $\varepsilon > 0$, there exists $\delta > 0$ such that $f(D(x_0, \delta) \setminus \{x_0\}) \subset D(L, \varepsilon)$.

The limit of a function at any given point need not exist, but when it does exist, it is unique.

Proposition 41. *A function $f : M_1 \to M_2$ is continuous if and only if for all $p \in M_1$, $\lim_{x \to p} f(x) = f(p)$.*

Proof. **Forward:** Let $\varepsilon > 0$, then $D(f(p), \varepsilon)$ is open in M_2. Since f is continuous, $f^{-1}(D(f(p), \varepsilon))$ is open in M_1. It also contains p, so there exists some $\delta > 0$ such that $D(p, \delta) \subset f^{-1}(D(f(p), \varepsilon))$. Thus $f(D(p, \delta)) \subset D(f(p), \varepsilon)$.

Backward: Suppose $\lim_{x \to p} f(x) = f(p)$ for every $p \in M_1$. Let $U \subset M_2$, then we must show that $f^{-1}(U)$ is open. Let $p \in f^{-1}(U)$, then $f(p) \in U$. Then there exists $\varepsilon > 0$ such that $D(f(p), \varepsilon) \subset U$.

Since $f(x) \to f(p)$, apply f^{-1} to both sides of the definition of the limit of a function to show that there exists $\delta > 0$ such that

$$D(p,\delta)\setminus\{p\} \subset f^{-1}(D(f(p),\varepsilon)) \subset f^{-1}(U).$$

Since $p \in f^{-1}(U)$, this implies that $D(p,\delta) \subset f^{-1}(U)$. Thus $f^{-1}(U)$ is open, which makes f continuous. □

Definition 52: Continuity at a Point

Let $A \subset M_1$, $f : A \to M_2$, and $x_0 \in A$. We say that f is **continuous at** x_0 if either x_0 is not an accumulation point of A or $\lim_{x\to x_0} f(x) = f(x_0)$.

We can now formulate several equivalent notions of continuity over larger sets.

Theorem 15: Continuity

Let $f : A \subset M_1 \to M_2$, then the following are equivalent notions of continuity of f.

1. f is continuous at every point of A.
2. For every convergent sequence $x_k \to x$ in A, we have $f(x_k) \to f(x)$.
3. For every open set $U \subset M_2$, $f^{-1}(U)$ is open in M_1.
4. For every closed set $F \subset M_2$, $f^{-1}(F)$ is closed in M_1.

Proof. This is not a proof, but we will show that if f is continuous by definition 3, then definition 4 automatically holds. Let $f : M_1 \to M_2$ be continuous, and let $F \subset M_2$ be closed. Then F^c is open, so $f^{-1}(F)$ is open, so $f^{-1}(F)^c = f^{-1}(F^c)$ is closed.

Check the book for other stuff. □

Redo this section with the open set defn of continuity being the main one.

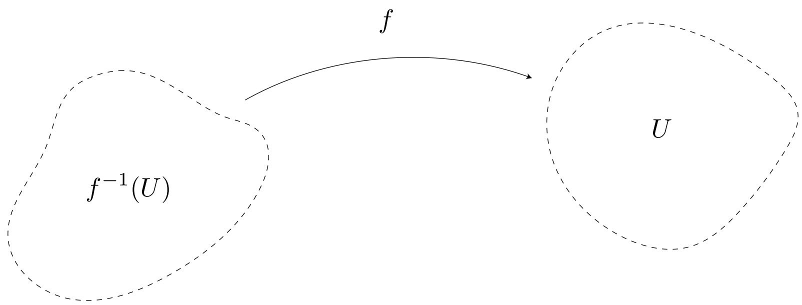

Figure 4.1: A continuous function f

4.2 Images of Compact and Connected Sets

Theorem 16

Let $f : M_1 \to M_2$ be continuous, and let $K \subset M_1$ be connected. Then $f(K)$ is also connected. Similarly, if K is path-connected, then $f(K)$ is also path-connected.

Proof. We first show the result for connectedness. Assume $f(K)$ is disconnected, then there exist open U, V such that

1. $f(K) \cap U \cap V = \varnothing$,
2. $f(K) \cap U \neq \varnothing$ and $f(K) \cap V \neq \varnothing$, and
3. $f(K) \subset U \cup V$.

We will now show that $f(K)$ being disconnected results in K being disconnected, a contradiction. Since f is continuous, $U' \doteq f^{-1}(U)$ and $V' \doteq f^{-1}(V)$ are both open. Additionally, they cover K, i.e. $K \subset U' \cup V'$. It is also clear that $f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V)$, so $K \cap U' \cap V' = \varnothing$. Finally, it is clear that $K \cap U'$ and $K \cap V'$ are both nonempty. Thus K is disconnected, which is a contradiction, so $f(K)$ must be connected.

We now show the path-connected result. Let $f(x) = \tilde{x}$ and $f(y) = \tilde{y}$, and let $\varphi : [a, b] \to K$ be a continuous map between x and y, then we claim $f \circ \varphi : [a, b] \to f(K)$ is a continuous map between $\tilde{x}$ and $\tilde{y}$. It lies in $f(K)$ since φ lies in K. Since $\varphi(a) = x$ and $\varphi(b) = y$, we have $f(\varphi(a)) = \tilde{x}$ and $f(\varphi(b)) = \tilde{y}$. Finally, $f \circ \varphi$ is continuous by Proposition 42, which we prove in the next section. Thus $f \circ \varphi$ is the desired map, and $f(K)$ is subsequently path-connected. $\square$

Theorem 17

Let $f : M_1 \to M_2$ be continuous, and let $K \subset M_1$ be compact, then $f(K)$ is compact.

Proof. Let $\{U_\alpha\}_{\alpha \in A}$ be an open cover of $f(K)$, then $\left\{f^{-1}(U_\alpha)\right\}_{\alpha \in A}$ is an open cover of K. Then there is a finite subcover $f^{-1}(U_{\alpha_1}) \cup \cdots \cup f^{-1}(U_{\alpha_N})$ of K, and since f is continuous, this subcover is open. Thus $f(K)$ is covered by the finite subcover $U_{\alpha_1} \cup \cdots \cup U_{\alpha_N}$. □

Example 16

Consider $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = 1/x$. The preimage of the compact set $[-1, 1]$ is $f^{-1}([-1, 1]) = [1, \infty)$, which is not compact. So the preimages of compact sets aren't necessarily compact under continuous maps, only the images.

4.3 Operations on Continuous Mappings

If $f : m_1 \to M_2$ is continuous and M_2 is a normed vector space, then the usual limit arithmetic theorems apply.

Proposition 42. *Let $f_1 : M_1 \to M_2$ and $f_2 : M_2 \to M_3$ be continuous, then their composition $f_2 \circ f_1 : M_1 \to M_3$ is also continuous.*

Proof. Let $U \subset M_3$ be open, then $(f_2 \circ f_1)^{-1}(U) = f_1^{-1}(f_2^{-1}(U))$. Since f_2 is continuous, $f_2^{-1}(U)$ is open. Since f_1 is continuous, $f_1^{-1}(f_2^{-1}(U))$ is open. □

Stuff on page 185 of textbook.

Product metric.

Proposition 43. *Let $\mathcal{V}$ be a normed vector space and equip $\mathcal{V} \times \mathcal{V}$ with the product metric, then*

1. $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ *is continuous*
2. $\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V}$ *is continuous*

Proof.
1. Let $U \subset \mathcal{V}$ be open. Let $(v, w) \in +^{-1}(U)$, i.e. $v + w \in U$, then there exists $\varepsilon > 0$ such that $D(v + w, \varepsilon) \subset U$. We want to find $\delta > 0$ such that $+(D((v, w), \delta)) \subset D(v + w, \varepsilon)$. If we can find such a δ, then

we will have a δ-ball (an open set) within the preimage of U. Since this open ball is being constructed for an arbitrary member of the preimage, the preimage must then be open.

Let $(x, y) \in D((v, w), \delta)$, then $d((x, y), (v, w)) < \delta$. This is equivalent to $\|x - v\| + \|y - w\| < \delta$ since $\mathcal{V} \times \mathcal{V}$ is equipped with the product metric. By the triangle inequality,

$$d(x + y, v + w) = \|x + y - (v + w)\| \leq \|x - v\| + \|y - w\| < \delta.$$

Let $\delta = \varepsilon$, then $+(D((v, w), \delta)) \subset D(v + w, \varepsilon) \subset U$. Thus $D((v, w), \delta) \subset +^{-1}(U)$ and, subsequently, $+^{-1}(U)$ is open.

2. Similar.

□

Corollary 4. *If (M, d) is a metric space, $\mathcal{V}$ is a normed vector space, and $f : M \to \mathcal{V}$, $g : M \to \mathcal{V}$, and $h : M \to \mathbb{R}$ are continuous, then*

1. *$f + g$ is continuous, and*
2. *hf is continuous.*

Proof. Both proofs use the strategy of decomposing the operations into compositions of functions that we know to be continuous.

1. Since f and g are continuous, then $(f, g) : M \times M \to \mathcal{V} \times \mathcal{V}$ defined by $(f, g)(x, y) = (f(x), g(y))$ is continuous. Addition is continuous, so $+ \circ (f, g) : M \times M \to \mathcal{V}$ is continuous. Define $\Delta : M \to M \times M$ by $\Delta(x) = (x, x)$. This is clearly continuous, so $+ \circ (f, g) \circ \Delta = f + g$ is continuous.
2. The composition $\cdot \circ (h, f) \circ \Delta$ is continuous.

□

4.4 The Boundedness of Continuous Functions on Compact Sets

Theorem 18: Maximum-Minimum Theorem

Let (M, d) be a metric space, let $A \subset M$, let $f : A \to \mathbb{R}$ be continuous, and let $K \subset A$ be compact. Then f is bounded on K, i.e. $B = \{f(x) \mid x \in K\}$ is a bounded set. Furthermore, there exists $x, y \in K$ such that $f(x) = \inf(B)$ and $f(y) = \sup(B)$.

Proof. Since K is compact and f is continuous, $f(K)$ is compact. Since $f(K) \in \mathbb{R}$, by Heine-Borel it must be closed and bounded. Since it's closed, it contains its accumulation points. Its infimum and supremem are either in the set or accumulation points, so they must lie in $f(K)$. $\square$

4.5 The Intermediate Value Theorem

Theorem 19: Intermediate Value Theorem

Let $A \subset (M, d)$, and let $f : A \to \mathbb{R}$ be continuous. Suppose that $K \subset A$ is connected and $x, y \in K$. Then for every real number $c \in \mathbb{R}$ such that $f(x) < c < f(y)$, there exists a point $z \in K$ such that $f(z) = c$.

Proof. Suppose no such z exists, then $f(A) \subset (-\infty, c) \cup (c, \infty)$. Since $y_1 < c$ and $y > c$, we know both sets in this union are nonempty. Since $f(A)$ is then clearly covered by two disjoint nonempty sets, it is disconnected. Since A was taken to be path-connected (and thus also connected), this is a contradiction, so such a z actually does exist. $\square$

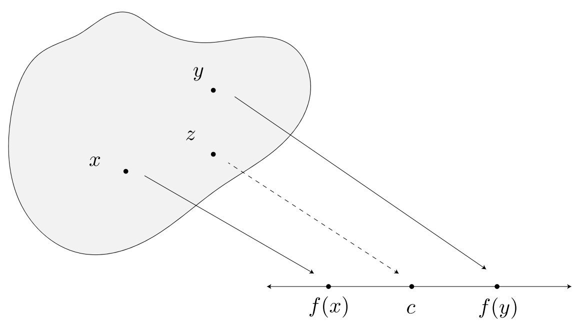

Figure 4.2: The Intermediate Value Theorem

4.6 Uniform Continuity

Definition 53: Uniform Continuity

Let (M_1, d_1) and (M_2, d_2) be metric spaces, and let $f : M_1 \to M_2$ be any function. We say that f is **uniformly continuous** if for all $\varepsilon > 0$ and for all $x, y \in M_1$, there exists $\delta > 0$ such that if $d_1(x, y) < \delta$, then $d_2(f(x), f(y)) < \varepsilon$.

It should be clear how to restrict the uniform continuity of f to certain sets. If $x, y \in A \subset M_1$ and the above condition holds, then f is uniformly continuous on A.

Note that unlike usual continuity, we have to find a δ that works for *all* x and y, so it must be independent of the inputs to the function.

Example 17

Let $f : \mathbb{R} \to \mathbb{R}$ be defined by $f(x) = x^2$. Consider

$$|f(x) - f(y)| = |x^2 - y^2| = |(x+y)(x-y)| < \varepsilon,$$

then $|x - y| < \varepsilon/|x + y|$. Since the difference between points needed to satisfy the ε constraint depends on the inputs themselves, f is *not* uniformly continuous.

Observe that if we had been working in a compact set instead of all of $\mathbb{R}$, $|x + y|$ would have been bounded and then f *would* have been uniformly continuous. This observation is formalized in the next theorem.

Theorem 20: Uniform Continuity Theorem

Let $f : M_1 \to M_2$ be continuous and let $K \subset M_1$ be compact, then f is uniformly continuous on K.

Proof. Given $\varepsilon > 0$ and $x \in K$, choose δ_x such that if $d_1(x, y) < \delta_x$, then $d_2(f(x), f(y)) < \varepsilon/2$ (we can find such δ_x because f is continuous). Then $\{D(x, \delta_x/2)\}$ is an open cover of K. Since K is compact, there is a finite subcover $\{D(x_1, \delta_{x_1}/2), \ldots, D(x_N, \delta_{x_N}/2)\}$. Let $\delta = \min\{\delta_{x_1}/2, \ldots, \delta_{x_N}/2\}$.

If $d_1(x, y) < \delta$, then there is an x_i such that $d_1(x, x_i) < \delta_{x_i}/2$ (since the disks cover K). Then by the triangle inequality,

$$d_1(x_i, y) \le d_1(x, x_i) + d_1(x, y) < \delta_{x_i}.$$

Thus for any $x, y \in K$ such that $d_1(x, y) < \delta$,

$$d_2(f(x), f(y)) \le d_2(f(x), f(x_i)) + d_2(f(x_i), f(y)) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

$\square$

4.7 Differentiation of Functions of One Variable

Definition 54: Derivative

The **derivative** of a function f at point x is defined

$$f'(x) \doteq \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}.$$

Definition 55: Differentiable

Let f be defined on some open interval containing $x \in \mathbb{R}$. The function f is **differentiable** at x if $f'(x)$ exists.
Equivalently,

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

A function is differentiable on a set A if it is differentiable at every point A.

We can rewrite the definition for differentiability that avoids the issue of division by a term that approaches 0: for any $\varepsilon > 0$, there is a $\delta > 0$ such that if $|\Delta x| < \delta$, then

$$|f(x + \Delta x) - f(x) - f'(x)\Delta x| \leq \varepsilon |\Delta x|.$$

Definition 56: Big-O Notation

Let $\phi, g : (0, a) \to \mathbb{R}$. We say ϕ is $\mathcal{O}(g)$ if

$$\left| \frac{\phi(x)}{g(x)} \right|$$

is bounded in some "deleted" neighborhood of 0, i.e. it lies in $D(0, r) \setminus \{0\}$ for some $r > 0$.
Additionally, we say ϕ is $o(g)$ if

$$\lim_{x \to 0} \frac{\phi(x)}{g(x)} = 0.$$

Based on these definitions, we can see that f is differentiable at x if there exists some $L \in \mathbb{R}$ such that $f(y) - f(x) - L(y - x)$ is $o(|y - x|)$.

Definition 57: Lipschitz

A function $f : M_1 \to M_2$ is **Lipschitz** if there exists some $L \geq 0$ such that $d_2(f(x), f(y)) \leq L d_1(x, y)$ for all $x, y \in M_1$. The function f is **locally Lipschitz** if for every compact set $K \subset M_1$, f restricted to K is Lipschitz.

Note that any Lipschitz function is also *uniformly* continuous. If we want $d_2(f(x), f(y))$ be to less than some $\varepsilon > 0$, then take x and y such that $d_1(x, y) < \varepsilon/L$.

Proposition 44. *If f is differentiable at x, then f is continuous at x.*

Proof. **Version 1:** Recall that a function f is continuous at x if and only if $\lim_{x_k \to x} f(x_k) = f(x)$. This is the case because

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

So since f is differentiable at x, it is also continuous at x.

Version 2: This version uses the second definition of differentiability. Given $\varepsilon > 0$, there is a $\delta > 0$ such that

$$|f(x + \Delta x) - f(x) - f'(x)\Delta x| \leq \varepsilon |\Delta x|$$

if $|\Delta x| < \delta$. Then by the triangle inequality,

$$|f(x + \Delta x) - f(x)| \leq |f'(x)\Delta x| + \varepsilon |\Delta x|.$$

Choosing $\varepsilon = 1$,

$$|f(x + \Delta x) - f(x)| \leq (f'(x) + 1)|\Delta x|$$

if $|\Delta x| < \delta$. This shows that f has the Lipschitz property at x, from which continuity follows. $\square$

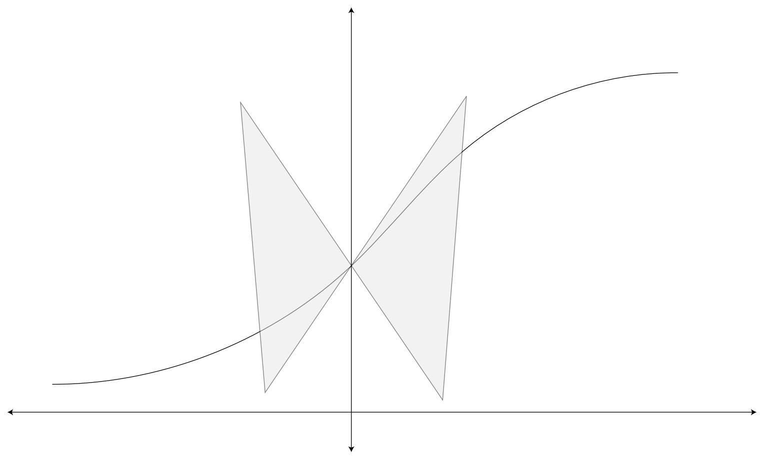

Figure 4.3: The Lipschitz property for a real-valued one variable function

Theorem 21

Suppose that f and g are differentiable at x and that $k \in \mathbb{R}$, then kf, $f+g$, and fg are differentiable at x with the expected derivatives.

1. $(kf)'(x) = kf'(x)$
2. $(f+g)'(x) = f'(x) + g'(x)$
3. $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$

Proof.
1. Do this.
2. Do this.
3. Do this.

□

Theorem 22: Chain Rule

If f is differentiable at x and g is differentiable at $f(x)$, then $g \circ f$ is differentiable at x and

$$(g \circ f)'(x) = g'(f(x))f'(x)$$

Proof. find in textbook. □

Definition 58: Differentiability Class

If f is differentiable on (a, b) and f' is continuous, we say that f is of **class** C^1. If f' is differentiable and f'' is continuous, we say that f is of class C^2.
More generally, if $f^{(n-1)}$ is differentiable and $f^{(n)}$ continuous, f is of class C^n. The function f is of class C^∞ if it is infinitely differentiable.

Definition 59: Increasing and Decreasing Functions

A function f defined in a neighborhood of x is **increasing** at x if there is an interval (a, b) containing x such that

1. If $a < y < x$, then $f(y) \leq f(x)$, and
2. If $x < y < b$, then $f(y) \geq f(x)$.

Similarly, f is **decreasing** at x if there is an interval (a, b) containing x such that

1. If $a < y < x$, then $f(y) \geq f(x)$, and
2. If $x < y < b$, then $f(y) \leq f(x)$.

Strictly increasing and decreasing functions are defined by making the above inequalities strict.

Theorem 23

Let f be differentiable at x, then

1. If f is increasing at x, then $f'(x) \geq 0$,
2. If f is decreasing at x, then $f'(x) \leq 0$,
3. If $f'(x) > 0$, then f is strictly increasing at x, and
4. If $f'(x) < 0$, then f is strictly decreasing at x.

Proof. Prove this. □

Proposition 45. *If $f : (a, b) \to \mathbb{R}$ is differentiable at $c \in (a, b)$ and if f has a maximum (or minimum) at c, then $f'(c) = 0$.*

Proof. If $f'(c) > 0$, then f is strictly increasing at c, which is a contradiction. If $f'(c) < 0$, then f is strictly decreasing at c, which is also a contradiction. Thus

$f'(c) = 0$. ☐

We can combine this with the maximum-minimum theorem (Theorem 18) to yield the familiar Rolle's Theorem.

Theorem 24: Rolle's Theorem

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

Proof. If $f(x) = 0$ for all $x \in [a,b]$, we can choose any c. Therefore assume that f is not identically zero. From the maximum-minimum theorem (Theorem 18), we know that there is a point c_1 where f assumes its maximum and there is a point c_2 where f assumes its minimum. Since f is not identically zero and $f(a) = f(b) = 0$, at least one of c_1, c_2 lies in (a,b). If $c_1 \in (a,b)$, then $f'(c_1) = 0$ by Proposition 45. The same is true if $c_2 \in (a,b)$. ☐

Theorem 25: Mean Value Theorem

If $f : [a,b] \to \mathbb{R}$ is continuous and differentiable on (a,b), there is a point $c \in (a,b)$ such that $f(b) - f(a) = f'(c)(b-a)$.

Proof. Let

$$\varphi(x) = f(x) - f(a) - (x-a)\frac{f(b)-f(a)}{b-a},$$

then apply Rolle's Theorem. ☐

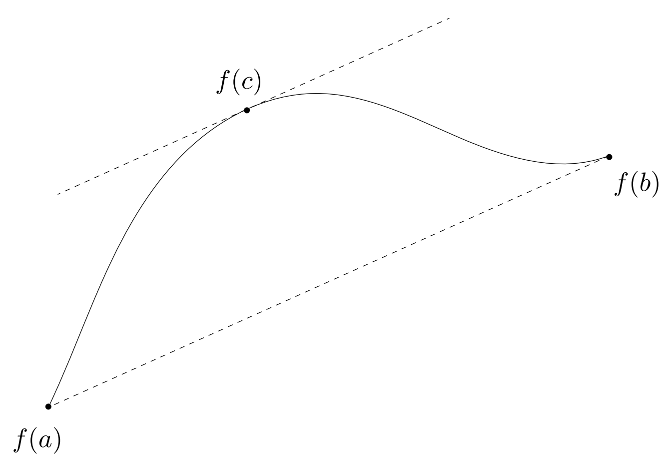

Figure 4.4: The Mean Value Theorem

Corollary 5. *If $f : [a,b] \to \mathbb{R}$ is continuous and if $f' = 0$ on (a,b), then f is constant.*

Proof. Applying the mean value theorem to f gives a point c such that $f(b) - f(a) = f'(c)(b-a) = 0$, so $f(a) = f(b)$ for all $x \in [a,b]$. Thus f is constant. □

Corollary 6. *If $f : (a,b) \to \mathbb{R}$ is differentiable with $|f'(x)| \leq M$ for every $x \in (a,b)$, then $|f(x) - f(y)| \leq M|x-y|$ for all $x, y \in (a,b)$, i.e. f is M-Lipschitz.*

Proof. For any x, y, the mean value theorem gives us a point $c \in (x,y)$ such that $|f(b) - f(a)| = |f'(c)(b-a)| \leq M|b-a|$. □

I don't know if I want to include proposition 4.7.14 on page 202 of the textbook b/c it's very similar to theorem 19. I could probably combine them to save room.

Proposition 46. *Let $f \in C([a,b])$ be differentiable on (a,b) such that $f'(x) \geq 0$ for every $x \in (a,b)$, then f is increasing on $[a,b]$. If $f'(x) \leq 0$ for every $x \in (a,b)$ instead, then f is decreasing on $[a,b]$.*

Proof. Suppose $y > x$ and $f' \geq 0$, then for some $c \in (x,y)$, $f(y) - f(x) = f'(c)(y-x) \geq 0$ since $y - x$ is positive. Thus $f(y) \geq f(x)$, so f is increasing. Similarly, $f(y) \leq f(x)$ if $f' < 0$, so f is decreasing in that situation. □

Theorem 26: Inverse Function Theorem

Suppose that $f : (a,b) \to \mathbb{R}$ is either strictly increasing or strictly decreasing over (a,b). Then f is a bijection onto its range, f^{-1} is differentiable on its domain, and $(f^{-1})'(y) = 1/f'(x)$ where $f(x) = y$.

Proof. Do this. □

Proposition 47. *Suppose that f is continuous on $[a,b]$ and twice differentiable on (a,b) and that $x \in (a,b)$, then*

1. *If $f'(x) = 0$ and $f''(x) > 0$, then x is a strict local minimum of f, and*
2. *If $f'(x) = 0$ and $f''(x) < 0$, then x is a strict local maximum of f.*

Proof. We only prove the first statement, as the second is similar. If $f''(x) > 0$, then f' is increasing at x, and so there is a $\delta > 0$ such that $f'(y) < 0$ if $x - \delta < y < x$ and $f'(y) > 0$ if $x < y < x + \delta$. Thus $f(y) > f(x)$ if $x - \delta < y < x$ and $f(y) > f(x)$ if $x < y < x + \delta$. $\square$

Somewhere in here I need to say that differentiable implies uniformly continuous.

4.8 Integration of Functions of One Variable

The integral of a function of one variable is the signed area under the curve. To define these, we'll need a notion of partitions and upper and lower sums.

Definition 60: Mesh

The **mesh** of a partition $P = \{x_1, x_2, \ldots, x_N\}$ is defined

$$|P| \doteq \sup_i |x_{i+1} - x_i|.$$

Definition 61: Refinement

Let P and Q be partitions of $[a, b]$. We say P is a **refinement** of Q if $Q \subset P$. We denote this by $Q \prec P$ or $P \succ Q$.

Consider a bounded function $f : A \subset \mathbb{R} \to \mathbb{R}$. If A is bounded, then there is some $[a, b] \supset A$. Define $f(x) = 0$ if $x \in [a, b] \backslash A$. Now partition $[a, b]$ with $P = \{x_0 = a, x_1, \ldots, x_n = b\}$ such that $x_0 < x_1 < \cdots < x_n$.

Definition 62: Upper/Lower Sums

The **upper sum** of f over P is

$$U(f, P) = \sum_{i=0}^{n-1} \sup_{x \in [x_i, x_{i+1}]} f(x)(x_{i+1} - x_i).$$

Similarly, the **lower sum** is defined

$$L(f, P) = \sum_{i=0}^{n-1} \inf_{x \in [x_i, x_{i+1}]} f(x) \cdot (x_{i+1} - x_i).$$

Note that the supremum and infimum for each subinterval exist since f is bounded. Let $-M \leq f \leq M$, then

$$-M(b-a) \leq L(f,P) \leq U(f,P) \leq M(b-a)$$

for any partition P of $[a,b]$.

Proposition 48. *If $P \succ Q$, then $L(f,Q) \leq L(f,P) \leq U(f,P) \leq U(f,Q)$.*

Proof. By induction, it suffices to consider the case that $P = Q \cup \{y\}$, say $x_j < y < x_{j+1}$. For the lower sums we have

$$\begin{aligned} L(f,P) - L(f,Q) = &\inf_{x\in[x_j,x_{j+1}]} f(x)(y-x_j) + \inf_{x\in[y,x_{j+1}]} f(x)(x_{j+1}-y). \\ &- \inf_{x\in[x_j,x_{j+1}]} f(x)(x_{j+1}-x_j) \end{aligned}$$

Since the infimum over a subset is greater than or equal to the infimum over the whole set, we have

$$\geq \inf_{x\in[x_j,x_{j+1}]} f(x)\left[(y-x_j)+(x_{j+1}-y)+(x_{j+1}-x_j)\right].$$

All the terms in brackets cancel out, leaving us with $L(f,P) - L(f,Q) \geq 0$. Thus $L(f,Q) \leq L(f,P)$. Using the fact that the supremum over a subset is less than or equal to the supremum over the whole set, we can similarly show $U(f,P) \leq U(f,Q)$. □

Let P and Q be two partitions, then neither is necessarily a subset of the other. To get around this, we note that for all P and Q, there exists a partition R which refines both, i.e. $R \succ P$ and $R \succ Q$. The set $P \cup Q$ arranged into an ordered set is one such partition.

Definition 63: Upper/Lower Integral

Given bounded function $f : A \to \mathbb{R}$ over a bounded set A, define the **upper integral** by

$$\overline{\int_A} f = \inf \{U(f,P)\}_P$$

and the **lower integral** by

$$\underline{\int_A} f = \sup \{L(f,P)\}_P .$$

Definition 64: Riemann Integral

A function f is **Riemann integrable** if $\overline{\int_A} f = \underline{\int_A} f$. The common value $\overline{\int_A} f = \underline{\int_A} f$ is denoted by $\int_A f$. If $A = [a, b]$, we write

$$\int_A f = \int_a^b f.$$

Note that this definition does *not* involve any notions of smoothness or continuity.

Theorem 27

Any increasing or decreasing function on $[a, b]$ is Riemann integrable on $[a, b]$.

Proof. Do this. □

An immediate consequence of this is that if f is continuous everywhere on $[a, b]$, then it is Riemann integrable on $[a, b]$.

Theorem 28

If $f : [a, b] \to \mathbb{R}$ is bounded and continuous at all but finitely many points of $[a, b]$, then it is Riemann integrable on $[a, b]$.

Proof. Do this. □

Proposition 49. *Let f and g be Riemann integrable on $[a, b]$, then*

1. *If $k \in \mathbb{R}$, then kf is integrable on $[a, b]$ and $\int_a^b kf = k \int_a^b f$,*
2. *$f + g$ is integrable on $[a, b]$ and $\int_a^b (f + g) = \int_a^b f + \int_a^b g$,*
3. *If $f(x) \le g(x)$ for all $x \in [a, b]$, then $\int_a^b f \le \int_a^b g$, and*
4. *If f is also integrable on $[b, c]$, then it is integrable on $[a, c]$ and $\int_a^c f = \int_a^b f + \int_b^c f$.*

Proof. Do this. □

Go through these and replace with A instead of $[a,b]$ if possible.

Corollary 7. *The absolute value of a definite integral of f is a lower bound of the definite integral of the absolute value of f, i.e.*

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

Proof. $-|f| \leq f \leq |f|$, so $-\int_a^b |f| \leq \int_a^b f \leq \int_a^b |f|$, which is the desired relation. ☐

Proposition 50. *The lower definite integral of f is a lower bound of the upper definite integral, i.e.*

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

Proof. Do this. ☐

An obvious corollary of this is that if f is integrable, then this inequality is *not* strict.

Definition 65: Antiderivative

An **antiderivative** of $f : [a,b] \to \mathbb{R}$ is a continuous function $F : [a,b] \to \mathbb{R}$ such that F is differentiable on (a,b) and $F'(x) = f(x)$ for $x \in (a,b)$.

Theorem 29: The Fundamental Theorem of Calculus

Let $f : [a,b] \to \mathbb{R}$ be continuous, then f has an antiderivative F and

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

If G is any other antiderivative of f, then we also have $\int_a^b f = G(b) - G(a)$.

Proof. Do this. ☐