

Analysis and Topology *

Zhiyuan Bai

Compiled on July 11, 2021

This document serves as a set of revision materials for the Cambridge Mathematical Tripos Part IB course *Analysis and Topology* in Michaelmas 2019. However, despite its primary focus, readers should note that it is NOT a verbatim recall of the lectures, since the author might have made further amendments in the content. Therefore, there should always be provisions for errors and typos while this material is being used.

Contents

1	Uniform Convergence	1
2	Uniform Continuity	6
3	Metric Space	7
4	Completeness	14
5	Topological Spaces	17
6	Connectedness	24
7	Compactness	27
8	Differentiation	31

1 Uniform Convergence

Definition 1.1. A complex sequence x_n is said to converge to a complex number x if $\forall \epsilon > 0, \exists N \in \mathbb{N}$ such that $\forall n > N, |x - x_n| < \epsilon$.

Definition 1.2. Let S be a set and let $f_n : S \rightarrow \mathbb{C}$ be a sequence of functions. Let $f : S \rightarrow \mathbb{C}$ be a function. We say $f_n \rightarrow f$ pointwise if for any $x, f_n(x) \rightarrow f(x)$. In other words, $\forall x \in S, \forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n > N, |f(x) - f_n(x)| < \epsilon$.

Example 1.1. Let S be the closed interval $[0, 1]$ and $f_n(x) = x^n$, then $f_n \rightarrow f$ pointwise where

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

*Based on the lectures under the same name taught by Prof. A. Zsák in Michaelmas 2019.

Note that in this example, despite the fact that all of f_n are continuous, even smooth, the resulting limit f needs not be continuous. Here is another example:

Example 1.2. Let $S = \mathbb{R}_{\geq 0}$ and let $f_n(x) = x^2 e^{-nx}$, then $f_n \rightarrow 0$ pointwise, since

$$0 \leq |f_n(x)| = \frac{x^2}{e^{nx}} = \frac{x^2}{1 + nx + \frac{n^2 x^2}{2} + \frac{n^3 x^3}{6} \dots} \leq \frac{x^2}{nx} = \frac{x}{n} \rightarrow 0$$

as $n \rightarrow \infty$.

There is another form of convergence, called uniform convergence, which is defined as follows:

Definition 1.3. Let S be a set and let $f_n : S \rightarrow \mathbb{C}$ be a sequence of functions. Let $f : S \rightarrow \mathbb{C}$ be a function. We say $f_n \rightarrow f$ uniformly if $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n > N, \forall x \in S, |f(x) - f_n(x)| < \epsilon$.

Note that the only difference between pointwise and uniform convergence is that the large integer N does not depend on x if the convergence is uniform. Note also that uniform convergence implies pointwise convergence, but not the other way around. Although it does not seem to be such a great difference in definition, in practice, it makes all the difference in the world.

Proposition 1.1. *The sequence in the first example, i.e. $f_n : [0, 1] \rightarrow \mathbb{R}$ with $f_n(x) = x^n$, does not converge uniformly.*

Proof. We can take for example $\epsilon = 1/2$. Then for any $n \in \mathbb{N}$, we can take $x = \sqrt[n]{2/3}$, so that we have

$$|f_n(x) - f(x)| = |f_n(x)| = \frac{2}{3} > \frac{1}{2}$$

So the claimed N does not exist. Therefore the sequence f_n does not converge uniformly. \square

Proposition 1.2. *The sequence in the second example, i.e. $f_n = \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$ where $f_n(x) = x^2 e^{-nx}$, converges absolutely.*

Proof. Note that

$$0 \leq f_n(x) = \frac{x^2}{e^{nx}} = \frac{x^2}{1 + nx + \frac{n^2 x^2}{2} + \dots} \leq \frac{x^2}{n^2 x^2 / 2} = \frac{2}{n^2}$$

Therefore for any $\epsilon > 0$, we can take $N = \lceil \sqrt{2/\epsilon} \rceil$, so for any $x \geq 0, n > N$, we have

$$|f(x) - f_n(x)| = |f_n(x)| = f_n(x) \leq \frac{2}{n^2} < \frac{2}{N^2} \leq \frac{2}{(\sqrt{2/\epsilon})^2} = \epsilon$$

So $f_n \rightarrow 0$ uniformly. \square

In fact, although continuous functions may not converge pointwise to a continuous function, they do converge uniformly to one.

Theorem 1.3. Let $S \subset \mathbb{C}$ be open. Suppose that $f_n : S \rightarrow \mathbb{C}$ is a sequence of continuous functions. If $f_n \rightarrow f$ uniformly, then f is continuous as well.

Informal sketch. Idea: Transfer the nice property of f_n to f . Choose large enough N such that $f_n - f$ is arbitrarily small for all $n > N$. We can always choose x' close to x where $f_n(x)$ close to $f(x)$. Then just use triangle inequality. "3- ϵ proof" \square

Proof. $\forall \epsilon > 0$, we can choose large enough N such that $\sup |f_n - f| < \epsilon/3$. We can choose $\delta > 0$ such that $|x - x'| < \delta \implies |f_n(x) - f_n(x')| < \epsilon/3$.

$$|f(x) - f(x')| \leq |f(x) - f_n(x)| + |f(x') - f_n(x')| + |f_n(x) - f_n(x')| < 3 \frac{\epsilon}{3} = \epsilon$$

As desired. \square

Remark. 1. We can use this theorem to show that x^n as in the previous example does not converge uniformly.

2. It is not true that differentiability is preserved under uniform convergence.

Theorem 1.4. Let $f_n : [a, b] \rightarrow \mathbb{R}$ be all Riemann integrable. Then if it converges uniformly, its limit is also Riemann integrable. Furthermore,

$$\int_a^b \lim_{n \rightarrow \infty} f_n(x) \, dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) \, dx$$

Recall that a function is Riemann integrable if and only if the upper and lower sums of f on the interval can be arbitrarily close.

Proof. Firstly f is bounded. Since f_n are bounded, we can just choose large enough n such that $|f_n - f| < 1$ and $|f_n| < M$, then $|f| \leq |f - f_n| + |f_n| < \epsilon + M < M + 1$ so f is bounded.

For $\epsilon > 0$ choose N such that $\sup |f_n - f| < \epsilon/(3(b-a))$ for any $n > N$. Since f_n is integrable, there is some dissection D of the interval $[a, b]$ such that $U_D(f_n) - L_D(f_n) < \epsilon/3$. We have

$$|L_D(f) - L_D(f_n)| = \sum_{(x_i) \in D} \left| \inf_{x \in [x_i, x_{i+1}]} f(x) - \inf_{x \in [x_i, x_{i+1}]} f_n(x) \right| (x_{i+1} - x_i) < \epsilon/3$$

Similarly $|U_D(f) - U_D(f_n)| < \epsilon/3$. So

$$\begin{aligned} |U_D(f) - L_D(f)| &\leq |U_D(f) - U_D(f_n)| \\ &\quad + |L_D(f_n) - L_D(f)| + |U_D(f_n) - L_D(f_n)| \\ &< 3\epsilon/3 = \epsilon \end{aligned}$$

This shows that f is integrable. Finally, we have

$$\left| \int_a^b f(x) - f_n(x) \, dx \right| \leq \int_a^b \sup |f(x) - f_n(x)| \, dx < \epsilon/3 < \epsilon$$

which completes the proof. \square

Remark. 1. For uniform convergence, we can swap the integral and the limit.

2. If $f_n \rightarrow f$ uniformly and that all f_n is bounded, then f is bounded.

Corollary 1.5. *For uniform convergence, we can swap infinite sums and integral. That is, if $f_n : [a, b] \rightarrow \mathbb{R}$ is a sequence of integrable functions whose partial sum converges uniformly to some function f , then f is integrable and*

$$\int_a^b f(x) dx = \sum_{n=1}^{\infty} \int_a^b f_n(x) dx$$

Proof. Let

$$F_n(x) = \sum_{k=1}^n f_k(x)$$

so F_n are integrable and $F_n \rightarrow f$ uniformly. Then we can just apply the preceding theorem. \square

Theorem 1.6. *Let $f_n : [a, b] \rightarrow \mathbb{R}$ be continuously differentiable on $[a, b]$. Assume that the sequence of partial sums of f'_n at every point converges uniformly, and that there is an $c \in [a, b]$ such that*

$$\sum_{n=1}^{\infty} f_n(c)$$

converges, then the sequence of partial sums of f_n converges uniformly. Furthermore, the limit f is continuously differentiable and

$$f'(x) = \sum_{n=1}^{\infty} f'_n(x)$$

Sketch of proof. Let

$$F_n(x) = \sum_{k=1}^n f_k(x), g(x) = \sum_{n=1}^{\infty} f'_n(x)$$

So we want to find a particular solution to the differential equation $f' = g$, and show that F_n converges uniformly to it. So basically we want to do

$$f(x) = \int_c^x g(t) dt + \sum_{n=1}^{\infty} f_n(c) = \lim_{n \rightarrow \infty} F_n(x) = \sum_{n=1}^{\infty} f_n(x)$$

rigorously and it would be done. \square

Proof. Let

$$g(x) = \sum_{n=1}^{\infty} f'_n(x)$$

g is continuous and hence Riemann integrable on $[a, b]$. Define $f : [a, b] \rightarrow \mathbb{R}$ by

$$f(x) = \int_c^x g(t) dt + \lambda$$

where

$$\lambda = \sum_{n=1}^{\infty} f_n(c)$$

By FTC, f is differentiable and $f'(x) = g(x)$. Since g is continuous, $f \in \mathcal{C}^1([a, b])$. It remains to show that the series sum of $f_n(x)$ converges uniformly to $f(x)$. Let $F_n(x)$ be the partial sum of the series, then by estimating its difference with f and the fact that the partial sum of derivatives of f_n converges uniformly (use FTC again), we can show that $F_n \rightarrow f$ uniformly. [Write details later] \square

Definition 1.4. Let f_n be a sequence of scalar function on a set S . We say f_n is uniformly Cauchy on S if $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall x \in X, \forall n, m > N$,

$$|f_n(x) - f_m(x)| < \epsilon$$

Theorem 1.7 (General Principle of Uniform Convergence). *A sequence of uniformly Cauchy scalar functions f_n on S converges uniformly.*

Proof. Firstly, we shall find a pointwise limit f of f_n . The existence of f is immediate since $f_n(x)$ is always Cauchy (hence converges) with x fixed. Then we shall show that this convergence is uniform. Choose any $\epsilon > 0, \exists N \in \mathbb{N}, \forall x \in X, \forall n, m > N, |f_n(x) - f_m(x)| < \epsilon/2$. Now we fix $x \in S, n > N$, since $f_n \rightarrow f$ pointwise, we can choose $m > N$ with $|f_m(x) - f(x)| < \epsilon/2$, then

$$|f(x) - f_n(x)| \leq |f(x) - f_m(x)| + |f_m(x) - f_n(x)| < 2\epsilon/2 < \epsilon$$

So $f_n \rightarrow f$ uniformly. \square

So what we did is to fix the x and the n , then let that m tend to infinity, then we can use the pointwise convergence to give the result. This is how we get pass the dependence of N on x in the pointwise convergence result.

Corollary 1.8. *Let f_n be a sequence of scalar functions on S , let*

$$\sum_{n=1}^{\infty} M_n$$

be convergent with $M_n \geq 0$. If $\sup |f_n| \leq M_n$ for any n , then

$$\sum_{n=1}^{\infty} f_n$$

converges uniformly.

Proof. Let F_n be the partial sum of f_n , F_n is uniformly Cauchy due to the convergence of the series of M_n . Essentially, $\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n, m > N$,

$$\sum_{k=n+1}^m M_k < \epsilon$$

so

$$\sum_{k=n+1}^m |f_k(x)| < \epsilon$$

Therefore it is uniformly Cauchy, so it converges uniformly. \square

Now we consider the power series

$$\sum_{n=0}^{\infty} a_n(z-a)^n$$

$(a_n)_0^{\infty}$ be a sequence of complex number. Let R be the radius of convergence. Now on the disk $|z-a| < R$, we consider

$$f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n$$

The question is: is the convergence uniform?

Example 1.3. Consider

$$f(z) = \sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$$

where $R = 1$. It does not converge uniformly. Indeed, the N^{th} partial sum is bounded by $N+1$ but $1/(1-z)$ is unbounded.

Theorem 1.9. For any r with $0 < r < R$, the series converges uniformly on $D(a, r)$.

Proof. For $w \in \mathbb{C}$ such that $r < |w-a| < R$, there is an M such that $|a_n(w-a)^n| < M$ for some $M > 0$ and any n . We have $|z-a|/|w-a| < 1$ for any $z \in D(a, r)$, hence by taking $M_n = M(r/|w-a|)^n$ shows the result. \square

The derivative of a power series (we can prove that it is complex differentiable, and we can do it term-by-term) has the same radius of convergence.

Remark. If we fix $w \in D(a, R)$, we can choose r such that $|w-a| < r < R$. Fix any $\delta > 0$ such that $|w-a| + \delta < r$, then $D(w, \delta) \subset D(a, r)$, so

$$\sum_{n=0}^{\infty} a_n(z-a)^n$$

converges uniformly on $D(w, \delta)$. We say it is locally uniformly on $D(a, R)$.

2 Uniform Continuity

Let $U \in \mathbb{C}$ and f a scalar function on U . We know what continuity means in the sense of metric space.

Definition 2.1. We say f is uniformly continuous on U if for any $\epsilon > 0, \exists \delta > 0, \forall x, y \in U, |x-y| < \delta \implies |f(x) - f(y)| < \epsilon$.

Note that the difference between uniform convergence and our initial form of convergence is that the value of δ does not depend on the point x .

Example 2.1. The standard example that a function is continuous but not uniformly continuous is that $f(x) = x^2$ on \mathbb{R} . To see why, observe that we take $\epsilon = 1$, then choose any $\delta > 0$,

$$(x + \delta/2)^2 - x^2 = \delta x + \delta^2/4$$

We can just choose $x > 1/\delta$ and the value would exceed 1.

Theorem 2.1. Let f be a scalar function defined on a closed interval $[a, b]$, then if f is continuous then it is uniformly continuous.

Proof. (Heine-Borel Theorem gives compactness of the interval which provides a direct proof.)

Assume that it is not the case, so

$$\exists \epsilon > 0, \forall \delta > 0, \exists x, y \in [a, b], |x - y| < \delta \wedge |f(x) - f(y)| > \epsilon$$

Choose such a 'bad' ϵ , consider $\delta_n = 1/n$ and choose x_n, y_n accordingly.

By Bolzano-Weierstrass, there is a subsequence x_{k_m} of x_n that converges to some x as $x \rightarrow \infty$. Note that since the interval is closed $x \in [a, b]$. Then, $|y_{k_m} - x| \leq |x_{k_m} - x| + |x_{k_m} - y_{k_m}| < \frac{1}{k_m} + \epsilon$ for any $\epsilon > 0$. so $y_{k_m} \rightarrow x$.

Since f is continuous at x , there is some δ such that for every $y \in [a, b]$, $|x - y| < \delta \implies |f(x) - f(y)| < \epsilon/2$. There is some N such that $m > N \implies |x_{k_m} - x| < \delta, |y_{k_m} - x| < \delta$. So $\epsilon < |f(x_{k_m}) - f(y_{k_m})| \leq |f(x_{k_m}) - f(x)| + |f(x) - f(y_{k_m})| < 2\epsilon/2 = \epsilon$. This is a contradiction. \square

It follows that a continuous function on a closed interval is integrable.

3 Metric Space

Definition 3.1. Let M be an arbitrary set, a metric on M is a function $d : M \times M \rightarrow \mathbb{R}_{\geq 0}$ such that the following properties hold:

1. $\forall x, y \in M, d(x, y) = 0 \iff x = y$.
2. $\forall x, y \in M, d(x, y) = d(y, x)$.
3. $\forall x, y, z \in M, d(x, y) + d(y, z) \geq d(x, z)$.

The couple (M, d) is called a metric space.

Example 3.1. 1. Let M be \mathbb{R} or \mathbb{C} , and $d(x, y) = |x - y|$ is called the usual metric on those sets.

2. Take $M = \mathbb{R}^n$ or \mathbb{C}^n , then

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt{|x_1 - y_1|^2 + \dots + |x_n - y_n|^2}$$

. This is called the ℓ^2 metric.

3. Take the same M , then we can also have the ℓ^1 metric where

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = |x_1 - y_1| + \dots + |x_n - y_n|$$

4. Also the same M , we have the ℓ^∞ metric where

$$d((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) = \max_i |x_i - y_i|$$

5. We can have the ℓ^p metric for $p \geq 1$ where

$$d((x_1, \dots, x_n), (y_1, \dots, y_n)) = \sqrt[p]{|x_1 - y_1|^p + \dots + |x_n - y_n|^p}$$

6. Let S be a set and $M = \ell_\infty S$ be the set of bounded scalar function on S . The metric we can take is the uniform metric $d(f, g) = \sup_S |f - g|$, which is well-defined since f, g are bounded.

7. Let M be any set, then define

$$d(x, y) = \begin{cases} 1, & \text{if } x = y \\ 0, & \text{otherwise} \end{cases}$$

This is called the *discrete metric* and the space (M, d) the discrete metric space.

8. Let G be a group that is generated by a symmetric set S . Define $d(x, y)$ be the least integer $n \geq 0$ such that n is the least number of generators to get from x to y .

This develops to the discipline called geometric group theory.

9. Suppose G is a connected (finite) graph, then we can define the distance between two vertices x, y to be the length of the shortest path from x to y .

10. Riemannian metric in geometry.

11. Take M to be the integers and we fix a prime p . We can define the p -adic metric $d_p(x, y)$ to be 0 if $x = y$ and $\|x - y\|_p = p^{-n}$ where n is the greatest power of p in the prime factorisation of $|x - y|$. It is obvious that it is a metric. The metric space (\mathbb{Z}, d_p) is called the p -adic integers.

12. Let M be the set of all functions from \mathbb{N} to \mathbb{R} , that is, the set of all sequences. So for $x = (x_n), y = (y_n)$, we define $d(x, y) = \sum_{n=0}^{\infty} 2^{-n} \min\{1, |x_n - y_n|\}$.

As in before, we can construct new objects from old.

Definition 3.2. Let (M, d) be a metric space and $N \subseteq M$, then $(N, d|_{N \times N})$ is a metric space and is called the metric subspace of (M, d) . Sometimes we denote $(N, d|_{N \times N})$ by (N, d) .

Example 3.2. $C[0, 1]$ be the set of real continuous functions on the unit interval having the uniform metric is a subspace of $\ell_\infty[0, 1]$.

Note that there are other metrics on $C[0, 1]$, for example

$$d(f, g) = \int_0^1 |f(x) - g(x)| dx$$

We also have the L^2 metric

$$d(f, g) = \sqrt{\int_0^1 |f(x) - g(x)|^2 dx}$$

Note that the L^∞ metric is the uniform metric.

Definition 3.3. Let $(M, d), (N, d')$ be two metric spaces, we can define the metric product space by taking the underlying set $M \times N$ and the metric

$$d_p((m_1, n_1), (m_2, n_2)) = (d(m_1, m_2)^p + d'(n_1, n_2)^p)^{1/p}$$

for some $p \geq 1$ or

$$d_\infty((m_1, n_1), (m_2, n_2)) = \max(\{d(m_1, m_2), d'(n_1, n_2)\})$$

We can generalize it to any finite product of metric space.

We need to generalize to topological spaces to have the notion of a quotient space in a metric/topological space.

We can introduce (uniform) convergence again in any metric spaces. We work in a metric space (M, d) .¹

Definition 3.4. Given a sequence x_n in M and a point $x \in M$, we say $x_n \rightarrow x$ as $n \rightarrow \infty$ if for any $\epsilon > 0, \exists N \in \mathbb{N}, \forall n > N, d(x_n, x) < \epsilon$.

Conversely, if such an x exists for a sequence x_n , we say x_n is convergent.

Lemma 3.1. Assume $x_n \rightarrow x$ and $x_n \rightarrow y$, then $x = y$.

Proof. Assume for the sake of contradiction that it is not the case. We let $\epsilon = d(x, y)$, we have a large enough $N \in \mathbb{N}$ such that $n \in \mathbb{N} \implies d(x_n, x) < \epsilon/2, d(x_n, y) < \epsilon/2$, so

$$\epsilon = d(x, y) \leq d(x, x_n) + d(y, x_n) < 2\epsilon/2 = \epsilon$$

A contradiction. □

Now we can introduce the concept of limit.

Definition 3.5. Let (x_n) be a convergent sequence which converges to x , we write

$$\lim_{n \rightarrow \infty} x_n = x$$

Example 3.3. 1. In \mathbb{R}, \mathbb{C} , this is the usual notion of convergence.

2. Take the integers under the 2-adic metric, $2^n \rightarrow 0$ as $n \rightarrow \infty$.

3. A sequence which is eventually constant converges to that constant. Obviously the converse is false in general, but true in discrete metric spaces.

4. Choose a nonempty set S , then functions that converge under the (induced) uniform metric on $\ell_\infty S$ converges uniformly as functions. This sometimes can work even on functions $\notin \ell_\infty S$, for example take $S = \mathbb{R}$, then $f_n(x) = x + 1/n$ converges uniformly to the identity function.

5. Consider the space $\mathbb{R}^{\mathbb{N}}$, the set of all sequences on \mathbb{R} with the metric

$$d((x_n), (y_n)) = \sum_{n=1}^{\infty} 2^{-n} \min\{1, |x_n - y_n|\}$$

then we can show that a sequence of sequences $(x_k^{(n)})$ (where each n gives a sequence) converges to a sequence (x_k) if and only if for each $i, x^{(n)}_i \rightarrow x_i$.

In fact, if we fix a set S , is there always a metric d on \mathbb{R}^S such that $f_n \rightarrow f$ under the metric d if $f_n \rightarrow f$ pointwise on S ? The answer is no, as we will need topological tools for it.

6. Consider $C[0, 1]$ under the uniform metric. Surely the function defined by $f_n = x^n$ do not converge, but if we equip $C[0, 1]$ with a different metric, for example the L^1 metric, that is,

$$d(f, g) = \int_0^1 |f(x) - g(x)| dx$$

In this case, the sequence f_n does converge to 0.

7. Let $(M, d), (M', d')$ be two metric spaces, we consider the metric product

¹When d is understood, we do not usually state explicitly our metric d

space $M \oplus_p M' = (M \times M', d_p)$. Then $(x_n, y_n) \rightarrow (x, y)$ if and only if $x_n \rightarrow x, y_n \rightarrow y$.

8. Consider the metric subspace $N \subset M$. If a sequence x_n converge to x in N , then $x_n \rightarrow x$ in M . The converse is not true since we can take $N = M \setminus \{x\}$ if $x_n \rightarrow x$.

Definition 3.6. Let $(M, d), (M', d')$ be two metric spaces and $f : M \rightarrow M'$ be a function. We say f is continuous at $a \in M$ if

$$\forall \epsilon > 0, \exists \delta > 0, \forall b \in M, d(a, b) < \delta \implies d'(f(a), f(b)) < \epsilon$$

If f is continuous at every $a \in N \subseteq M$, then we say f is continuous on N .

Note that if f is continuous on M , it is continuous on any $N \subseteq M$. The converse, however, is not true.

Example 3.4. We can take both metric spaces to be \mathbb{R} , then consider the function

$$f(x) = \begin{cases} 1, & \text{if } x \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

Then f is continuous on $N = \mathbb{R} \setminus \{0\}$ but not on $M = \mathbb{R}$.

Proposition 3.2. f is continuous at a if and only if for any sequence $x_n \rightarrow a$, we have $f(x_n) \rightarrow f(a)$.

Proof. If f is continuous at a , then $\forall \epsilon > 0$, we can find some $\delta > 0$ such that $d(a, b) < \delta \implies d'(f(a), f(b)) < \epsilon$. Now, find any $x_n \rightarrow a$, we can find $N \in \mathbb{N}, \forall n > N, d(a, x_n) < \delta$, but with the same $N, \forall n > N$, we have $d'(f(a), f(x_n)) < \epsilon$ by the above. So $f(x_n) \rightarrow f(a)$.

Conversely, if $x_n \rightarrow a \implies f(x_n) \rightarrow f(a)$ but f is not continuous at a , then we can find $\epsilon > 0$ such that $\forall \delta > 0$, there is some $x \in M$ such that $d(x, a) < \delta$ but $d'(f(x), f(a)) > \epsilon$. We may set $\delta_n = 1/n$ and we can obtain the corresponding x_n . Now $x_n \rightarrow a$ but $f(x_n) \not\rightarrow f(a)$. This is a contradiction. \square

The following two corollaries are then obvious.

Corollary 3.3. Let f and g be continuous scalar functions, then $f + g, f \times g$ and f/g (providing that $\forall x, g(x) \neq 0$) are all continuous.

Corollary 3.4. If $f : M \rightarrow M', g : M' \rightarrow M''$ are both continuous, then $g \circ f$ is continuous.

One can also prove them using $\epsilon - \delta$, which is not hard either.

Example 3.5. 1. Constant, identity (equipping the same metric) and inclusion (in the sense of metric subspace) functions are continuous.

2. Real and complex polynomials are continuous.

3. The metric function itself is continuous (in fact Lipschitz) with respect to the d_p metric on $M \times M$.

Definition 3.7. A function $(M, d) \rightarrow (M', d')$ is Lipschitz continuous if there is some $C \geq 0$ such that

$$\forall x, y \in M, d'(f(x), f(y)) \leq Cd(x, y)$$

we sometimes call f to be C -Lipschitz.

Proposition 3.5. *A Lipschitz function is uniformly continuous.*

Proof. Trivial. □

Definition 3.8. A map $g : (N, d) \rightarrow (N', d')$ is isometric if

$$\forall x, y \in N, d'(g(x), g(y)) = d(x, y)$$

Note that an isometric function is 1-Lipschitz. It also implies injective.

We continue with examples.

Example 3.6. 4. Let M, M' be metric spaces, fixing $y \in M'$, $f : M \rightarrow M \oplus_p M'$ by $f(x) = (x, y')$ is isometric, hence also (1-)Lipschitz.

5. Let $(M, d), (M', d')$ be metric spaces. Consider $q : M \oplus_p M' \rightarrow M, q' : M \oplus_p M' \rightarrow M'$ be the projection functions. Both of these functions are 1-Lipschitz. We can easily extend it to a finite product of metric spaces.

Now we go on to talk about the topology of metric spaces. We start with two observations. Firstly, in a product metric space $M \oplus_p M'$, convergence does not depend on the value of p . Secondly, continuity depends on the convergent sequences.

Definition 3.9. We fix a metric space (M, d) , for $x \in M$ and $r \geq 0$, the open ball $D_r(x)$ is the set $\{y \in M : d(x, y) < r\}$.

So $x_n \rightarrow x$ if and only if $\forall \epsilon > 0, \exists N \in \mathbb{N}, n > N \implies x_n \in D_\epsilon(x)$. And $f : M \rightarrow M'$ is continuous at $a \in M$ if and only if $\forall \epsilon > 0, \exists \delta > 0, \forall x \in M, x \in D_\delta(a) \implies f(x) \in D_\epsilon(f(a))$.

Definition 3.10. On (M, d) , for $x \in M$ and $r \geq 0$, the closed ball $B_r(x)$ is the set $\{y \in M : d(x, y) \leq r\}$.

Example 3.7. 1. When M is the real numbers, then an open ball is an open interval and closed ball is an closed interval.

2. In \mathbb{R}^2 , $B_1(0, 0)$ is the unit disk with boundary in d_2 , and an slanted square in d_1 , and a big square in d_∞ .

3. If M is discrete, $D_1(x) = \{x\}, B_1(x) = M$.

Note that $B_s(x) \subset D_r(x) \subset B_r(x)$ for any $s < r$.

Definition 3.11. A subset $U \subset M$ with $x \in U$ is called a neighbourhood of x (in M) if there exists some $r > 0$ with $D_r(x) \subset U$.

It does not matter if we take the closed ball instead.

Definition 3.12. Given $U \subset M$, we say U is open if $\forall x \in U, \exists r > 0, D_r(x) \subset U$.

So U is open if and only if U is a neighbourhood of x for any $x \in U$.

Lemma 3.6. *Open balls are open.*

Proof. Immediate from definition but let us write the proof anyways.

Consider $D_r(x)$, then for any $y \in D_r(x)$, since $d(x, y) < r$, if $z \in D_{r-d(x,y)}(y)$, then $d(x, z) \leq d(y, z) + d(x, y) < r \implies D_{r-d(x,y)}(y) \subset D_r(x)$. □

Proposition 3.7. *In a metric space M , the followings are equivalent:*

1. $x_n \rightarrow x$.
2. For any neighbourhood U of x , there is some $N \in \mathbb{N}, \forall n > N, x_n \in U$.
3. For any open set U containing x , there is some $N \in \mathbb{N}, \forall n > N, x_n \in U$.

Proof. 1 \implies 2: $\exists r > 0, D_r(x) \subset U$, so we can choose an $N \in \mathbb{N}, \forall n > N, d(x_n, x) < r \implies x_n \in U$.

2 \implies 3: Immediate by the preceding lemma.

3 \implies 1: Given $\epsilon > 0$, take $U = D_\epsilon(x)$, then two statement becomes identical. \square

Proposition 3.8. *Given function $f : M \rightarrow M'$, then:*

(A) *For $a \in M$, the followings are equivalent:*

1. f continuous at a .
2. For any neighbourhood V of $f(a)$, there is a neighbourhood U of a such that $f(U) \subset V$.
3. For any neighbourhood V of $f(a)$, $f^{-1}(V)$ is a neighbourhood of a .

(B) *The followings are equivalent:*

1. f is continuous.
2. The pre-image of any open set is open.

Proof. Part (A):

1 \implies 2: Given any neighbourhood V of $f(a)$, there is some r such that $D_r(f(a)) \subset V$. Since f is continuous at a , there is $\delta > 0$ with $f(D_\delta(a)) \subset D_r(f(a)) \subset V$. So $U = D_\delta(a)$ works.

2 \implies 3: Trivial since there is some neighbourhood U containing a with $f(U) \subset V \implies U \subset f^{-1}(V)$ so it is a neighbourhood of a .

3 \implies 1: Given $\epsilon > 0$, $f^{-1}(D_\epsilon(f(a)))$ contains some open ball $D_\delta(a)$ for some $\delta > 0$, so it's done.

Part (B):

1 \implies 2: Given V open in M' , for $x \in f^{-1}(V)$, we have $f(x) \in V$, so V is a neighbourhood of $f(x)$. Since f is continuous, by (A), there is an neighbourhood of x containing in it.

2 \implies 1: We shall show that it is continuous at every point. Given $\epsilon > 0, a \in M$, the ball $D_\epsilon(f(a))$ is open in M' , so $f^{-1}(D_\epsilon(f(a)))$ is open, so there is some δ with $D_\delta(a) \subset f^{-1}(D_\epsilon(f(a)))$, so we are done. \square

Definition 3.13. The topology of a metric space the collection of open subsets of it.

Proposition 3.9. *In a metric space M , we have the following:*

1. \emptyset, M are open.
2. If $\{U_i\}_{i \in I}$ are open, then $\bigcup_{i \in I} U_i$ is open.
3. If $U, V \subset M$ are open, then $U \cap V$ is open.

Proof. 1 is trivial.

For 2, given x in the union, then there is a $j \in I$ such that $x \in U_j$, but then there is some open ball $U \ni x$ such that $U \subset U_j$, then U is a subset of that union. So this union is open.

Regarding 3, given $x \in U \cap V$, then there are $\epsilon_U, \epsilon_V > 0$ such that $D_{\epsilon_U}(x) \subset U, D_{\epsilon_V}(x) \subset V$, so the ball $D_{\min\{\epsilon_U, \epsilon_V\}}(x) \subset U \cap V$. \square

Definition 3.14. A subset $A \subset M$ is closed if whenever $x_n \rightarrow x$ in M for some sequence $(x_n) \in A$, then $x \in A$.

Example 3.8. 1. Closed balls are closed.
2. So in \mathbb{R} , any closed interval is closed. Also \mathbb{R} itself is both open and closed. $[0, 1)$ is neither open nor closed.

Lemma 3.10. A subset $A \subset M$ is closed if and only if $M \setminus A$ is open.

Proof. If A is closed but $M \setminus A$ is not open, so there is some $x \in M \setminus A$ such that $D_r(x)$ is not contained in $M \setminus A$ for all $r > 0$. Hence for each $\epsilon > 0$, $\exists x_\epsilon \in M, x_\epsilon \in D_\epsilon(x)$. Taking $a_n = x_{1/n}$ gives a contradiction.

If A is not closed but $M \setminus A$ is open. So we can find $(a_n) \in A$ such that $a_n \rightarrow a \notin A$, so there is some $\epsilon > 0$ such that $D_\epsilon(a) \subset M \setminus A$, but this is a contradiction since $a_n \notin M \setminus A$ for any n but it can go ϵ -close to a . \square

Example 3.9. If (N, d) is discrete, then every subset of N is both open and closed.

Definition 3.15. Two metrics on a set are equivalent if they give the same topology.

Note that it is equivalent to say that the two metrics have the same convergence sequences since they help identify the closed sets. It also means that they have the same continuous functions, both to and from, any other spaces. Note that the two metrics induce the same topology if and only if the identity maps from both spaces are continuous.

Definition 3.16. A map $g : M \rightarrow M'$ is called a homeomorphism if it is a bijection and both g and g^{-1} are continuous.

We say g is an isometry if it is bijective and it is isometric.

We say M and M' are homeomorphic if there is a homeomorphism between them. And M, M' be isometric if there is an isometry between them.

Remark. 1. Continuous bijections may not be a homeomorphism. Take $g : \mathbb{R} \rightarrow \mathbb{R}$ where the domain is equipped with discrete metric and the codomain with the usual metric.

2. A surjective isometric function is an isometry.

Example 3.10. 1. $(0, 1), (0, \infty)$ are homeomorphic. Take $x \mapsto 1/x$.

2. \mathbb{R}^2 and \mathbb{C} are isometric.

Definition 3.17. Two metrics d, d' on M are uniformly equivalent if and only if both the identity functions $\text{id} : (M, d) \rightarrow (M, d'), \text{id} : (M, d') \rightarrow (M, d)$ are uniformly continuous.

We say d, d' are Lipschitz equivalent if and only if both the identity functions are Lipschitz.

Example 3.11. 1. On $M \times M'$, d_1, d_2, d_∞ are Lipschitz equivalent.

2. (non-example) On $C[0, 1]$ the uniform metric is not equivalent to the L^1 metric since they do not have the same convergent sequences.

4 Completeness

Definition 4.1. A metric space is called complete if every Cauchy sequence converges.

Definition 4.2. A subset $A \subset M$ where M is a metric space is called bounded if there is some $r > 0$ and $z \in M$ such that $A \subset B_r(z)$.

Lemma 4.1. *Convergent \implies Cauchy \implies Bounded.*

Proof. Suppose $x_n \rightarrow x$, then $\forall \epsilon > 0$, we can find some $N \in \mathbb{N}$ such that $\forall n > N, d(x, x_n) < \epsilon/2$, then $\forall n, m > N$,

$$d(x_m, x_n) \leq d(x_m, x) + d(x_n, x) < 2\epsilon/2 = \epsilon$$

Assume that (x_n) is Cauchy, we need that $\{x_n : n \in \mathbb{N}\}$ is contained in some ball. We know that there is some $N \in \mathbb{N}, \forall n, m > N, d(x_n, x_m) < \epsilon$. We can take $\epsilon = 1$, so $\{x_n : n \in \mathbb{N}, n > N\} \subset B_1(x_N)$. Since $(x_n)_{n < N}$ is finite, it is contained in some ball B . In particular, we can take the ball $B_{\max\{1, d(x_N, x_i) : i \leq N\}}(x_N)$ contains the sequence. \square

Remark. Bounded does not imply Cauchy and Cauchy does not imply Convergent.

Definition 4.3. A metric space M is complete if every Cauchy sequence converges.

Proposition 4.2. *If M, M' are complete, so is $M \oplus_p M'$.*

Proof. Let (a_n) be Cauchy in the product, then say that $a_i = (x_i, x'_i)$, then for all $m, n \in \mathbb{N}$, $\max\{d(x_m, x_n), d(x'_m, x'_n)\} \leq d_p(a_m, a_n)$, so $(x_n), (x'_n)$ are both Cauchy. Since both M, M' are complete, $\exists x \in M, x' \in M', x_n \rightarrow x, x'_n \rightarrow x'$. Hence $a_n \rightarrow (x, x')$. \square

Example 4.1. $\mathbb{R}^n, \mathbb{C}^n$ are complete (in Euclidean metric) for any n .

There is another very important example:

Theorem 4.3. *Let S be a non-empty set, then $\ell_\infty S$ is complete under the uniform metric.*

Proof. Theorem 1.7 shows that a uniformly Cauchy sequence of functions does converge to some scalar function on S . To see it is bounded, choose $n \in \mathbb{N}$ such that $d(f_n, f) < 1$, then since there is some $C \geq 0$ such that $\sup |f_n| \leq C$, we have $|f| \leq |f - f_n| + |f_n| < C + 1$ so it is bounded as well. \square

Proposition 4.4. *Let $N \subset M$ be a metric subspace.*

1. *If N is complete, then N is closed in M .*
2. *If N is closed and M is complete, so is N .*

So a metric subspace of a complete space is complete if and only if it is closed.

Proof. 1. If N is complete, let (x_n) be a sequence in N such that $x_n \rightarrow x$ in M , but this means that (x_n) is Cauchy by Lemma 4.1, therefore it is convergent in N due to completeness, hence $x \in N$ due to uniqueness of limit in metric space.
 2. Choose any Cauchy sequence (x_n) in N , we know that $(x_n) \rightarrow x$ for some $x \in M$ due to completeness of M , but since N is closed, $x \in N$ as well, so N is complete. \square

Theorem 4.5. *Let M be a metric space, then the space of bounded continuous scalar functions in M , $C_b(M)$ is complete in the uniform metric.*

Proof. $C_b(M)$ is a metric subspace of $\ell_\infty(M)$ which is complete. But uniform limit of continuous functions to a continuous function, so $C_b(M)$ is closed.

To spell out the proof, fix $x \in M, \epsilon > 0$, we can choose N such that $D(f_n, f) < \epsilon/3$ where D is the uniform metric. Fix any $n \geq N$, then since f_n is continuous, $\exists \delta > 0, d(x, y) < \delta \implies |f_n(x) - f_n(y)| < \epsilon/3$. Hence $d(x, y) < \delta \implies |f(x) - f(y)| \leq |f(x) - f_n(x)| + |f(y) - f_n(y)| + |f_n(x) - f_n(y)| < 3\epsilon/3 = \epsilon$. \square

Fix some $S \neq \emptyset$, a metric space (N, d') . Let $\ell_\infty(S, N)$ be the space of bounded functions $S \rightarrow N$. Then we can define the uniform metric on $\ell_\infty(S, N)$ defined by $D(f, g) = \sup_{x \in S} d'(f(x), g(x))$.

Now given a metric space (M, d) , let $C_b(M, N)$ be the set of bounded continuous functions $M \rightarrow N$, then we have

Theorem 4.6. *Let S, M, N be as above, assuming that N is complete, then $\ell_\infty(S, N)$ is complete under uniform metric, and since $C_b(M, N)$ is closed in $\ell_\infty(M, N)$ hence complete.*

Proof. Analogous to the case where $M = \mathbb{R}$ or \mathbb{C} . \square

Example 4.2. 1. For any closed and bounded interval $[a, b] \in \mathbb{R}$, then continuous functions on $[a, b]$ are the continuous and bounded functions on $[a, b]$ is complete under the uniform metric.

Definition 4.4. A map $f : M \rightarrow M'$ is a contraction mapping if f is L -Lipschitz with $L < 1$.

Theorem 4.7 (Contraction Mapping Theorem, aka Banach Fixed Point Theorem). *If f is a contraction mapping in a nonempty complete metric space, then f has an unique fixed point.*

Note that it is important for the condition listed to be satisfied.

Example 4.3. 1. If we remove the completeness criterion, $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \setminus \{0\}$ defined by $f(x) = x/2$, then f is a contraction but do not have fixed point.

2. If we remove $L < 1$, $f : \mathbb{R} \rightarrow \mathbb{R}$ by $f(x) = x + 1$ is 1-Lipschitz but do not have any fixed point.

3. $f(x) = x + 1/x, [1, \infty)$

Proof. Fix $x_0 \in M$, then define a sequence x_n by $x_{n+1} = f(x_n)$, so $x_n = f^n(x_0)$. We shall show that this sequence is Cauchy. For $n \geq 2, d(x_n, x_{n-1}) \leq Ld(x_{n-1}, x_{n-2}) \leq L^{n-1}d(x_1, x_0)$ inductively. For $m > n$,

$$\begin{aligned} d(x_m, x_n) &\leq d(x_n, x_{n+1}) + \dots + d(x_{m-1}, x_m) \\ &\leq (L^{m-1} + L^{m-2} + \dots + L^n)d(x_1, x_0) \\ &\leq \frac{L^n}{1-L}d(x_1, x_0) \end{aligned}$$

The last term, which only depends on the smaller term n , can be as small as we want when n is large enough, so the sequence is Cauchy.

Hence there is a limit x of the sequence x_n since M , but since f is continuous, $f(x_n) \rightarrow f(x)$, but $f(x_n) = x_{n+1}$, so by uniqueness of limits, $f(x) = x$.

Suppose $f(x) = x$ and $f(y) = y$, then if $x \neq y$, $|x - y| = |f(x) - f(y)| \leq L|x - y| < |x - y|$ which is a contradiction. \square

Note that $x_n \rightarrow x$ exponentially fast, so it can also be applied to numerical analysis to find an approximated solution of the fixed point.

An application of the contraction mapping theorem is to analyze the existence and uniqueness of the solution of an initial value problem.

Example 4.4. The IVP $f'(t) = f(t^2)$, $f(0) = y_0$ on $C[0, 1/2]$ is what we are interested in. Assume that f has a solution, then immediately f is continuously differentiable. By FTC,

$$f(t) = f(0) + \int_0^t f(x^2) dx$$

Let $M = C[0, 1/2]$, which is nonempty and complete, then consider the mapping $T : M \rightarrow M$ defined by $(Tg)(t) = y_0 + \int_0^t g(x^2) dx$. T is trivially well-defined since $x \in [0, 1/2] \implies x^2 \in [0, 1/4] \subset [0, 1/2]$ and that $g(x^2)$ is continuous in x . Also by FTC, $(Tg)' = g$, so Tg is continuously differentiable hence continuous. Now f solves the IVP iff f is a fixed point of T . Also we can check that T is a contraction. Indeed, take $g, h \in M$, then

$$\begin{aligned} |Tg(t) - Th(t)| &= \left| \int_0^t g(x^2) - h(x^2) dx \right| \\ &\leq \int_0^t |g(x^2) - h(x^2)| dx \leq tD(g, h) \leq D(g, h)/2 \end{aligned}$$

So it is a contraction mapping, hence by Theorem 4.7 a unique fixed point exists.

Theorem 4.8 (Lindelof-Picard Theorem). *Let $a < b, R > 0$ be real numbers and $y_0 \in \mathbb{R}^n$. Suppose $\phi : [a, b] \times B_R(y_0) \rightarrow \mathbb{R}^n$ is continuous and we have some $K > 0$ such that $\forall x, y \in B_R(y_0), \forall t \in [a, b], \|\phi(t, x) - \phi(t, y)\| \leq K\|x - y\|$. Then $\exists \epsilon > 0, \forall t_0 \in [a, b]$, the IVP*

$$f'(t) = \phi(t, f(t)), f(t_0) = y_0$$

has a unique solution in $[t_0 - \epsilon, t_0 + \epsilon] \cap [a, b]$.

Proof. Since $[a, b] \times B_R(y_0)$ is closed and bounded, $|\phi|$ is bounded above by some $C > 0$. Let $\epsilon = \min\{R/C, 1/(2K)\}$. We want to solve the IVP on $[c, d] = [t_0 - \epsilon, t_0 + \epsilon] \cap [a, b]$.

The metric space $M = C([c, d], B_R(y_0))$ is nonempty and complete, so a natural thought is to reduce the IVP to a fixed-point problem on which we can use Theorem 4.7.

Consider the mapping $T : M \rightarrow M$ given by

$$(Tg)(t) = y_0 + \int_0^t \phi(x, g(x)) dx$$

By FTC, Tg is C^1 ($(Tg)'(t) = \phi(t, g(t))$), in particular continuous, whenever g is continuous. In addition, Tg takes values in $B_R(y_0)$ since

$$\|(Tg)(t) - y_0\| = \left\| \int_{t_0}^t \phi(x, g(x)) \, dx \right\| \leq \int_{t_0}^t \|\phi(x, g(x))\| \, dx \leq \epsilon C \leq R$$

So indeed T is a well-defined function $M \rightarrow M$. It is also a contraction mapping: For $g, h \in M$,

$$\begin{aligned} \|Tg(t) - Th(t)\| &= \left\| \int_{t_0}^t \phi(s, g(s)) - \phi(s, h(s)) \, ds \right\| \\ &\leq \int_{t_0}^t \|\phi(s, g(s)) - \phi(s, h(s))\| \, ds \\ &\leq \epsilon K D(g, h) \leq D(g, h)/2 \end{aligned}$$

for any $t \in \mathbb{R}$ by the Lipschitz condition we assumed.

Note that $g \in M$ solves the IVP iff it is a fixed point of T , so we are done by Theorem 4.7. \square

Remark. 1. In general, however, you cannot extend the solution guaranteed above to a global solution. But in our previous example we can extend the solution to $[0, 1)$.

2. Also, we can apply the theorem to solve higher order equations by considering the vector of derivatives.

3. If $f : [a, b] \rightarrow \mathbb{R}^n$ can be written as (f_1, f_2, \dots, f_n) , so $f' = (f'_1, f'_2, \dots, f'_n)$ assuming each component is differentiable. Similarly, the integral of a vector valued function is the vector of the integrals of the components, given that they exist. So we can do everything by components.

4. The reason that the integral of the norm is at least the norm of the integral is Cauchy-Schwarz.

5. We can show that a continuous function on a closed bounded set in \mathbb{R}^n is bounded by Bolzano-Weierstrass.

5 Topological Spaces

Definition 5.1. Consider a set X . A topology τ is a collection of subsets of X such that the following axioms hold:

1. $\emptyset, X \in \tau$.
2. $\forall i \in I, U_i \in \tau \implies \bigcup_{i \in I} U_i \in \tau$.
3. $U, V \in \tau \implies U \cap V \in \tau$.

A topological space is a pair (X, τ) where X is a set and τ is a topology on X .

Note that the third axiom can be extended to any finite set of elements of τ .

Members of τ are called open sets of X .

Example 5.1. For any metric space, we can induce the metric topology by Proposition 3.9. For example, the Euclidean distance on \mathbb{R}^n induce the usual topology on \mathbb{R}^n .

Definition 5.2. A topological space X (or the topology of X) is called metrizable if it can be induced by some metric on X .

In the case where the topology is metrizable, any other metric that is equivalent to the previous metric gives the same topology.

Example 5.2. The indiscrete topology on a set X is $\{\emptyset, X\}$.

Definition 5.3. Give topologies τ_1, τ_2 on X , we say τ_1 is coarser than τ_2 or τ_2 is finer than τ_1 if $\tau_1 \subset \tau_2$.

We know that the indiscrete topology is coarser than any topology on X . It is then immediate that if $|X| \geq 2$, then the indiscrete topology is not metrizable. Indeed, suppose $x, y \in X$, the open ball $D_{d(x,y)}(x)$, then it contains x but not y and is open under the metric topology under d , so d cannot induce the indiscrete topology on X .

Example 5.3. The discrete topology on a set X is $\tau = 2^X$. This is metrizable. Indeed, it can be induced by the discrete metric. It is also the finest topology on X .

Example 5.4. The cofinite topology on a set X consists of all subsets of X whose complement is finite and the empty set. When X is finite, this topology is just the discrete topology. If it is infinite, it is not metrizable. Fix $x \neq y \in X$, whenever there is open sets U, V such that $x \in U, y \in V$, we know that $X \setminus (U \cap V)$ is finite, thus $U \cap V$ is not empty, but it would mean that the topological space that is not Hausdorff (which we will define below), but any metric space is (also below), so it is not metrizable.

Definition 5.4. A topological space X is called Hausdorff if any two distinct elements x, y in X , there are open sets U, V such that $x \in U, y \in V$ and $U \cap V = \emptyset$.

Proposition 5.1. *Any metric space is Hausdorff.*

Proof. Consider $U = D_{d(x,y)/2}(x), V = D_{d(x,y)/2}(y)$, they obviously contain x, y respectively and have empty intersection due to triangle inequality. \square

Definition 5.5. A subset of X is called closed if its complement is open.

This coincides with the definition of closed sets in a metric space by Lemma 3.10.

- Proposition 5.2.**
1. \emptyset, X are closed
 2. If A_i is closed for all $i \in I$, then $\bigcap_{i \in I} A_i$ is closed.
 3. If A, B are closed, then $A \cup B$ is closed.

Proof. Trivial. \square

Again the last one can be generalized to any finite indices.

Example 5.5. In cofinite topology, a subset is closed if and only if it is finite.

Definition 5.6. For a topological space X , $x \in X$ and $U \subset X$. We say U is a neighbourhood of x if $\exists V \subset X$ open such that $x \in V \subset U$.

Note again that in a metric space, this reduced to our previous definition. The proof of this is trivial.

Proposition 5.3. *Let $U \subset X$, then U is open if and only if every $x \in U$ has a neighbourhood contained in U .*

Proof. Completely trivial. □

Definition 5.7. A sequence $(x_n) \in X$ converges to some $x \in X$, or $x_n \rightarrow x$, if for any neighbourhood V of x , $\exists N \in \mathbb{N}, \forall n > N, x_n \in V$.

This again and again coincides with previous definition in metric spaces by Proposition 3.7.

Example 5.6. In a indiscrete space, any sequence converge to any element.

Theorem 5.4. *In a Hausdorff space, limits are unique.*

Proof. If $x_n \rightarrow x$ and $x_n \rightarrow y$ but $x \neq y$, then there are disjoint open sets U, V containing x, y respectively. But then there is some $N_1 \in \mathbb{N}, \forall n > N_1, x_n \in U$, and there is also some $N_2 \in \mathbb{N}, \forall n > N_2, x_n \in V$, but then for any $n > \max\{N_1, N_2\}$, $x_n \in U \cap V = \emptyset$, contradiction. □

Remark. In a metric space, A is closed if and only if whenever x_n converges in the metric space, then its limit is in A . The \implies part is true in all topological space, but not necessarily \impliedby .

Definition 5.8. Let X be a topological space and $A \subset X, x \in X$. x is called an accumulation point (aka limit point/cluster point) of A if for any neighbourhood U of x , $(A \setminus \{x\}) \cap U \neq \emptyset$.

The derived set A' of A is the set of all accumulation points of A .

Example 5.7. In \mathbb{R} , suppose $A = [0, 1) \cup \{2\}$, then $A' = [0, 1]$. Also $\mathbb{Q}' = \mathbb{R}$, and $\mathbb{Z}' = \emptyset$.

Proposition 5.5. *Let X be a topological space and $A \subset X$, then A is closed if and only if $A' \subset A$.*

Proof. If A is closed, then $U = X \setminus A$ is open, so for any $x \in X \setminus A$, U is a neighborhood of x but $U \cap A = \emptyset$, so every accumulation points of A are inside of A .

Conversely, given $x \in X \setminus A$, then $x \notin A'$, so there is a neighbourhood U of x with $U \cap A = \emptyset$, so $x \in U \subset X \setminus A$, so $X \setminus A$ is open, hence A is closed. □

Definition 5.9. Let A be a subset of a topological space X , then the interior of A , $\text{int } A$ or A° is defined by

$$\text{int } A = \bigcup_{U \subset A, U \text{ open}} U$$

The closure, $\text{cl } A$ or \bar{A} , is defined by

$$\text{cl } A = \bigcap_{A \subset F, F \text{ closed}} F$$

Note that $A^\circ \subset A \subset \bar{A}$, and $A^\circ = A$ if and only if A is open, $\bar{A} = A$ if and only if A is closed.

Proposition 5.6.

$$\begin{aligned} A^\circ &= \{x \in X : A \text{ is a neighbourhood of } x\} \\ \bar{A} &= \{x \in X : \forall U \subset X \text{ such that } U \text{ is a neighbourhood of } x, U \cap A \neq \emptyset\} \\ &= A \cup A' \end{aligned}$$

Proof. Exercise. □

Example 5.8. In \mathbb{R} , $\overline{[0, 1] \cup \{2\}} = [0, 1] \cup \{2\}$, $([0, 1] \cup \{2\})^\circ = (0, 1)$, $\bar{\mathbb{Q}} = \mathbb{R}$, $\mathbb{Q}^\circ = \emptyset = \mathbb{Z}^\circ$, $\bar{\mathbb{Z}} = \mathbb{Z}$.

Remark. Convergent sequences determine the metric topology, since $x \in \bar{A} \iff \exists (x_n) \in A, x_n \rightarrow x$. Again we have the \Leftarrow direction for all topological spaces but not necessarily for the \Rightarrow direction.

Definition 5.10. Let X be a topological space and $A \subset X$. We say A is dense if $\bar{A} = X$.

We say X is separable if there is a countable dense set in X .

Definition 5.11. 1. \mathbb{R}^n is separable since $\bar{\mathbb{Q}}^n = \mathbb{R}^n$.
2. (non-example) An uncountable set in the discrete topology is not separable.

As usual we can try to construct new spaces from old.

Definition 5.12. Let (X, τ) be a topological space and $Y \subset X$. The subspace (or relative) topology on Y is the collection $\{U \cap Y : U \in \tau\}$. This is also called the topology on Y induced by τ .

One can check that this is indeed a topology.

Example 5.9. Let $X = \mathbb{R}$, $Y = [0, 2]$, then $U = (1, 2]$ is open in Y since $U = Y \cap (1, 3)$. Note that U is not open in X .

Remark. 1. If $Z \subset Y \subset X$, then the topology on Z induced by the topology on X is the topology on Z induced by the topology on Y which is induced by the topology on X . So the subspace of a subspace is a subspace.

2. If $N \subset M$ where N, M are metric spaces, then the metric on N induced by the metric on M induces the metric topology on N induced by the metric topology on M .

Proposition 5.7. Let Y be the subspace of a topological space X .

1. $A \subset Y$ is closed in Y if and only if there is closed set $B \subset X$ such that $B \cap Y = A$.
2. $\forall A \subset Y, \bar{A}^Y = Y \cap \bar{A}^X$.

Remark. The analogy of 2 on interiors does not always work. Take $X = \mathbb{R}, Y = \{0\}$.

Proof. Trivial. □

Definition 5.13. A base for a topological space (X, τ) is a family $\mathcal{B} \subset \tau$ such that $\forall U \in \tau, \exists \mathcal{C} \subset \mathcal{B}$ such that

$$U = \bigcup_{B \in \mathcal{C}} B$$

In other words, the topology τ consists of the arbitrary unions of some family of open sets which is a subset of \mathcal{B} . So a base determines topology.

Example 5.10. 1. The set of all open intervals is a base of the usual topology on \mathbb{R} . In general, the collection of all open balls in a metric space is a base for the metric topology on it.

However, what we want to do is not to construct \mathcal{B} from τ , but the other way around.

Lemma 5.8. *let X be a set and $\mathcal{B} \subset 2^X$. Assume that*

1. $X = \bigcup_{B \in \mathcal{B}} B$.
2. $\forall B_1, B_2 \in \mathcal{B}, \forall x \in B_1 \cap B_2, \exists B \in \mathcal{B}, x \in B \subset B_1 \cap B_2$.

Then there is an unique topology on X that is generated by the base \mathcal{B} .

Proof. We must have the topology

$$\tau = \left\{ \bigcup_{B \in \mathcal{C}} B : \mathcal{C} \subset \mathcal{B} \right\}$$

It is immediate that τ is a topology on X . Indeed, $\emptyset, X \in \tau$ and it is closed under arbitrary union. For intersection, consider

$$U_1 = \bigcup_{B \in \mathcal{C}_1} B, U_2 = \bigcup_{B \in \mathcal{C}_2} B$$

Given $x \in U_1 \cap U_2$, so $\exists B_1 \in \mathcal{C}_1, B_2 \in \mathcal{C}_2$, so there is some $B_x \in \mathcal{B}$ such that $x \in B_x \subset B_1 \cap B_2 \subset U_1 \cap U_2$, thus

$$U_1 \cap U_2 = \bigcup_{x \in U_1 \cap U_2} B_x$$

By definition \mathcal{B} is a base for τ . □

Definition 5.14. A topological space is called second-countable if it has a countable base.

Example 5.11. The set of all open balls of rational radii and centres is a countable base for \mathbb{R}^n . So \mathbb{R}^n is second-countable.

Definition 5.15. A map $f : (X, \tau) \rightarrow (Y, \rho)$ is continuous if $V \in \rho \implies f^{-1}(V) \in \tau$.

This extends our previous definition of continuity in metric space by Proposition 3.8.

Proposition 5.9. *Let $f : X \rightarrow Y$ be a map between topological spaces, then*

1. *f is continuous if and only if the preimage of any closed set is closed.*
2. *If \mathcal{B} is a base for Y , then f is continuous if and only if for all $B \in \mathcal{B}$, $f^{-1}(B)$ is open in X .*
3. *Composition of continuous functions is continuous.*

Proof. Trivial. □

Example 5.12. Constant, identity and inclusion are always continuous. Hence the restriction of a continuous map is continuous.

Definition 5.16. Let $f : X \rightarrow Y$ be a map between topological spaces, then we say f is a homeomorphism if f is a bijection and both f, f^{-1} are continuous. We say X, Y are homeomorphic, or $X \cong Y$, if there is a homeomorphism between them.

Definition 5.17. f is a open map if for every U open in X , $f(U)$ is open in Y . So f is a homeomorphism if and only if f is a continuous open bijection.

Definition 5.18. A property P of topological spaces is called a topological property if it is preserved under homeomorphisms.

Definition 5.19. Let $(X, \tau), (Y, \rho)$ be topological spaces and let

$$\mathcal{B} = \{U \times V : U \in \tau, V \in \rho\}$$

Then $X \times Y \in \mathcal{B}$ and $U_1 \times V_1 \cap U_2 \times V_2 = (U_1 \cap U_2) \times (V_1 \cap V_2) \in \mathcal{B}$. Thus there is a unique topology on $X \times Y$ with base \mathcal{B} . This is called the product topology.

So a set W in the product topological space is open if and only if $\forall(x, y) \in W, \exists U \in \tau, V \in \rho, x \in U \times V \subset W$.

Example 5.13. \mathbb{R}^2 in the usual topology is homeomorphic to $\mathbb{R} \times \mathbb{R}$ in the product topology. In general, the topology induced by the (p -)product metric is the product topology of metric topologies. So products of metrizable topologies are metrizable.

Proposition 5.10. Consider $\pi_X : X \times Y \rightarrow X, \pi_Y : X \times Y \rightarrow Y$ be the projections. Then π_X, π_Y are continuous and if Z is a topological space, and $f : Z \rightarrow X \times Y$ is continuous if and only if $\pi_X \circ f, \pi_Y \circ f$ are both continuous.

Note that $f(z) = (\pi_X \circ f(z), \pi_Y \circ f(z))$.

Proof. Given an open set $U \subset X, \pi_X^{-1}(U) = U \times Y$, which is open in $X \times Y$, so π_X is continuous. Similarly, π_Y is continuous.

Given such an f , if f is continuous, then both of $\pi_X \circ f, \pi_Y \circ f$ are continuous since composition of continuous functions is continuous. Conversely, if both of $\pi_X \circ f, \pi_Y \circ f$ are continuous, then it is enough to check that any member of the base $U \times V \subset X \times Y$ has an open preimage. Indeed,

$$\begin{aligned} f^{-1}(U \times V) &= f^{-1}(U \times Y) \cap f^{-1}(X \times V) \\ &= f^{-1}(\pi_X^{-1}(U)) \cap f^{-1}(\pi_Y^{-1}(V)) \\ &= (\pi_X \circ f)^{-1}(U) \cap (\pi_Y \circ f)^{-1}(V) \end{aligned}$$

which is open by assumption. □

It is trivial to extend all the above to finite products. It is interesting to know that $(X \times Y) \times Z \cong X \times (Y \times Z)$, and $X \times Y \cong Y \times X$. Now we turn to quotient topology.

Definition 5.20. Start with a topological space (X, τ) and let R be an equivalence relation on X . We let X/R be the set of equivalence classes (the “quotient set”). Let $q : X \rightarrow X/R$ be the quotient map sending $x \mapsto [x]$ where $[x] = \{y \in X : yRx\}$ is the equivalence class containing x . The quotient topology on X/R is the family

$$\tau_R = \{V \subset X/R : q^{-1}(V) \in \tau\}$$

Proposition 5.11. *the quotient topology is indeed a topology.*

Proof. $q^{-1}(X/R) = X, q^{-1}(\emptyset) = \emptyset$, so $\emptyset, X \in \tau_R$.

$$\forall (V_i)_{i \in I} \in \tau_R, q^{-1}\left(\bigcup_{i \in I} V_i\right) = \bigcup_{i \in I} q^{-1}(V_i)$$

which is open.

$$\forall U, V \in \tau_R, q^{-1}(U \cap V) = q^{-1}(U) \cap q^{-1}(V)$$

which is also open. □

Remark. 1. Note that q is surjective and continuous under τ_R .

2. For $x \in X, t \in X/R, x \in t \iff q(x) = t$, hence

$$\forall V \subset X/R, q^{-1}(V) = \{x \in X : q(x) \in V\} = \{x \in X : \exists t \in V, q(x) = t\} = \bigcup_{t \in V} t$$

Example 5.14. $\mathbb{Q} \leq \mathbb{R}$ as (additive) groups, so \mathbb{R}/\mathbb{Q} gives a equivalence relation. So we can induce a quotient topology on \mathbb{R}/\mathbb{Q} , which immediately we can find to be the indiscrete topology which is not metrizable, which is why we do not do quotients in metric spaces.

Consider $q : X \rightarrow X/R$ the quotient map and any map $f : X \rightarrow Y$ such that $xRy \implies f(x) = f(y)$, then there is a map $\tilde{f} : X/R \rightarrow Y$ such that $\tilde{f} \circ q = f$. That is, the following diagram commutes.

$$\begin{array}{ccc} X & \xrightarrow{f} & Y \\ q \downarrow & \nearrow \tilde{f} & \\ X/R & & \end{array}$$

If f is surjective, so is \tilde{f} . Also, if $f(x) = f(y) \iff xRy$, then \tilde{f} is injective.

Proposition 5.12. *Let X, Y be topological spaces, R an equivalence relation on X , $q : X \rightarrow X/R$ the quotient map, $f : X \rightarrow Y$ some map with $xRy \implies f(x) = f(y)$, then let \tilde{f} be as above, then*

1. *If f is continuous so is \tilde{f} .*
2. *If f is an open map so is \tilde{f} .*

Proof. 1. Let V be open in Y , then $f^{-1}(V)$ is open in X , so $q^{-1}(\tilde{f}^{-1}(V)) = f^{-1}(V)$ is open, so $\tilde{f}^{-1}(V)$ is open, hence \tilde{f} is continuous.

2. Given open $V \in X, U = q^{-1}(V)$ is open in X , and $V = q(U)$, so $\tilde{f}(V) = f(U)$ which is open. □

Corollary 5.13. *If $f(x) = f(y) \iff xRy$, f is surjective, continuous and open, then \tilde{f} is a homeomorphism.*

Remark. Work “upstairs”!

Example 5.15. Take \mathbb{R}/\mathbb{Z} where the equivalence relation is as if they are additive groups. $\mathbb{R}/\mathbb{Z} \cong S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Indeed, consider the map $f(t) = e^{2\pi it}$, then the induced \tilde{f} is a homeomorphism by the preceding corollary. If f is not open, then there is some open $U \in \mathbb{R}$ such that $f(U)$ is not open. $\exists(z_n) \in S^1 \setminus f(U)$ such that $z_n \rightarrow z$, then due to surjectivity we know that there is some $x \in U$ such that $f(x) = z$ and $(x_n) \in [x - 1/2, x + 1/2]$ such that $f(x_n) = z_n$ and we know that $x_n \notin U$, but x_n has a convergent subsequence $(x_{k_n}) \rightarrow y \in \mathbb{R} \setminus U$ which is closed, so due to continuity we must have $f(x) = f(y) \implies x - y \in \mathbb{Z} \implies x = y \notin U$, which is a contradiction.

6 Connectedness

An interval I in \mathbb{R} has the defining property that $\forall x, y, z, x < y < z$, then $x, z \in I \implies y \in I$. We know that a real continuous function maps intervals to intervals due to the intermediate value theorem. But it may not work if the (restricted) domain is not an interval.

Definition 6.1. A topological space X is disconnected if there are open $U, V \subset X$ such that $U \neq \emptyset$ and $V \neq \emptyset$ partitions X , that is $U \cap V = \emptyset$ and $U \cup V = X$. In this case, we say U, V disconnect x .

A topological space X is connected if it is not disconnected.

Lemma 6.1. *The image of continuous function on connected space is connected.*

Proof. Suppose $f : X \rightarrow Y$ is continuous. Note that if we consider f as $f : X \rightarrow \text{Im } f$ then it is still continuous. Then if U, V disconnect $\text{Im } f$, then $f^{-1}(U), f^{-1}(V)$ disconnect X . \square

Theorem 6.2. *For a topological space X , the followings are equivalent:*

1. X is connected.
2. If $f : X \rightarrow \mathbb{R}$ is continuous, then $f(X)$ is an interval.
3. Every continuous function $f : X \rightarrow D$, where D is discrete and $|D| \geq 2$, is constant. ²

Proof. 1 \implies 2: Obvious due to the preceding lemma and the trivial fact that an open set in \mathbb{R} is connected if and only if it is an interval.

2 \implies 3: Immediate, also from the preceding lemma.

3 \implies 1: We shall prove the contrapositive. Suppose that U, V disconnects X , then choose $d, e \in D$ with $d \neq e$, then the function f defined by

$$f(x) = \begin{cases} d, & \text{if } x \in U \\ e, & \text{otherwise, that is if } x \in V \end{cases}$$

is continuous but is not constant, contradiction. \square

²Most of the time we take $D = \mathbb{Z}$

- Example 6.1.** 1. \emptyset and singletons are connected.
 2. Any indiscrete topological space is connected.
 3. The cofinite topology on an infinite set is connected.
 4. The discrete topology is disconnected if it is not a singleton.

Lemma 6.3. *A subspace $Y \subset X$ is disconnected if and only if there are open sets $U, V \in X$ such that $U \cap Y \neq \emptyset, V \cap Y \neq \emptyset, U \cap V \cap Y = \emptyset, Y \subset U \cup V$.*

Proof. Trivial. □

Proposition 6.4. *Let Y be a connected subspace of X , then \bar{Y} is connected.*

Proof. Assume not, then by the preceding lemma, there exists open sets U, V in X such that $U \cap \bar{Y} \neq \emptyset, V \cap \bar{Y} \neq \emptyset, U \cap V \cap \bar{Y} = \emptyset, \bar{Y} \subset U \cup V$. It follows that $U \cap V \cap Y = \emptyset, Y \subset U \cup V$, so we must have, WLOG, $U \cap Y = \emptyset$, then $Y \subset X \setminus U \implies \bar{Y} \subset X \setminus U \implies \bar{Y} \cap U = \emptyset$, contradiction. □

Remark. 1. Alternatively, we can use the third part of Theorem 6.2. 2. In fact, for any Z with $Y \subset Z \subset \bar{Y}$ is connected since the closure of Z is \bar{Y} .

Alternative proof of Lemma 6.1. Let $f : X \rightarrow Y$ be continuous, for convenience we can just assume f is surjective using the same argument as the original proof, then consider any continuous $g : Y \rightarrow \mathbb{Z}$, then $g \circ f$ is continuous hence constant since f is connected, but f is surjective, so g is constant, then it is done by Theorem 6.2. □

Remark. 1. Connectedness is a topological property.
 2. If $f : X \rightarrow Y$ is continuous and $A \subset X$ and A is connected, then $f(A)$ is connected.

Corollary 6.5. *If X is connected and R an equivalence relation on X , then X/R is connected.*

Proof. The quotient map is continuous and surjective. □

Example 6.2. let $Y = \{(x, \sin(1/x)) : x > 0\} \subset \mathbb{R}^2$ is connected since it is the image of $f(x) = (x, \sin(1/x))$, which is continuous since its components are connected, over $\mathbb{R}_{>0}$.

By Proposition 6.4, $\bar{Y} = Y \cup (\{0\} \times [-1, 1])$ is also connected. This is called the Topologist's Sine Wave.

Lemma 6.6. *Let \mathcal{A} be a family of connected subset of a topological space X such that $\forall A, B \in \mathcal{A}, A \cap B \neq \emptyset$, then $\bigcup_{A \in \mathcal{A}} A$ is connected.*

Proof. Suppose $f : \bigcup_{A \in \mathcal{A}} A \rightarrow \mathbb{Z}$ is connected, then $f|_A$ is continuous for any $A \in \mathcal{A}$, thus it is constant, say it is n_A , then $\forall A, B \in \mathcal{A}$, then $n_A = n_B$ since $A \cap B \neq \emptyset$. Thus f is constant, hence $\bigcup_{A \in \mathcal{A}} A$ is connected. □

Proposition 6.7. *If X, Y are connected, so is $X \times Y$.*

Proof. Observe that $\forall x \in X, \{x\} \times Y \cong Y$ is connected and $\forall y \in Y, X \times \{y\} \cong X$ is connected as well, so since $(x, y) \in (\{x\} \times Y) \cap (X \times \{y\}) \neq \emptyset$, by the preceding lemma $A_{x,y} = (\{x\} \times Y) \cup (X \times \{y\})$ is connected. Now obviously $(x, y') \in A_{x,y} \cap A_{x',y'} \neq \emptyset$, so $X \times Y = \bigcup_{x \in X, y \in Y} A_{x,y}$ is connected by the preceding lemma. □

Definition 6.2. Let X be a topological space, we define an equivalence relation R by xRy if and only if there is a connected $U \subset X$ such that $x, y \in U$. One can check that this is an equivalence relation by Lemma 6.6, and the partition of X by R is called the connected components of X .

Let C_x be the equivalence class containing x .

Proposition 6.8. *Connected components are nonempty and are maximal (wrt inclusion) connected subset of X , also they are closed.*

Proof. Let C be a connected component, so it is the equivalence class of some x , so $C = C_x$, so C is nonempty since it contains x . So given $y \in C$, $\exists A_y \ni x, y$ such that A_y is connected. $A_y \in C$ by definition of the relation. Now $\forall y, z \in C$, $x \in A_y \cap A_z \neq \emptyset$, therefore by Lemma 6.6, hence $C = \bigcup_{y \in C} A_y$ is connected. If $C \subset D$ and D is connected, then $\forall y \in D$, $x, y \in D$, thus since D is connected $y \in C$, so $D \subset C \implies C = D$.

Hence since \bar{C} is connected and contains C , by maximality $C = \bar{C}$, therefore C is closed. \square

Definition 6.3. A topological space X is called path-connected if $\forall x, y \in X$, $\exists \gamma : [0, 1] \rightarrow X$ continuous, $\gamma(0) = x, \gamma(1) = y$.

Theorem 6.9. *Any path-connected space is connected.*

Proof. Suppose not, then X is path-connected but not connected, so there are open U, V disconnects X . Then fixing $x \in U, y \in V$, there exists a continuous $\gamma : [0, 1] \rightarrow X$ such that $\gamma(0) = x, \gamma(1) = y$. Thus $\gamma^{-1}(U), \gamma^{-1}(V)$ are nonempty, open, and partitions $[0, 1]$, thus $[0, 1]$ is disconnected by them, which is a contradiction. \square

The converse, however, is not true.

Example 6.3. Take the Topologist's Sine Wave, $X = \{(x, \sin(1/x)) : x > 0\} \cup (\{0\} \times [-1, 1])$. We have already shown it is connected. But it is not path-connected. Indeed, pick points $(0, 0), (1, \sin(1)) \in X$. Assume that $\gamma : [0, 1] \rightarrow X$ is continuous and $\gamma(0) = (0, 0) = x, \gamma(1) = (1, \sin(1)) = y$. Let γ_1, γ_2 be the components of γ , which are continuous. For $\gamma_1(t) > 0$, then $[0, \gamma_1(t)] \subset \gamma_1([0, t])$ by IVT, so $\exists n \in \mathbb{N}, (2\pi n)^{-1}, (2\pi n + \pi/2)^{-1} \in (0, \gamma_1(t)) \subset \gamma_1([0, t])$. So there is some a, b with $\gamma_1(a) = (2\pi n)^{-1}, \gamma_1(b) = (2\pi n + \pi/2)^{-1}$, hence $\gamma_2(a) = 0, \gamma_2(b) = 1$, so we can thus find a sequence $1 > t_1 > t_2 > \dots > 0$ with

$$\gamma_2(t_n) = \begin{cases} 1, & \text{if } n \text{ is even} \\ 0, & \text{otherwise} \end{cases}$$

So t_n converges but $\gamma_2(t_n)$ does not. This is a contradiction.

Lemma 6.10 (Gluing Lemma). *Let $f : X \rightarrow Y$ be a function between topological spaces. If $X = A \cup B$ where A, B are closed and $f|_A, f|_B$ are continuous, then f is continuous.*

Proof. Given closed V in Y ,

$$f^{-1}(V) = (f^{-1}(V) \cap A) \cup (f^{-1}(V) \cap B) = (f|_A)^{-1}(V) \cup (f|_B)^{-1}(V)$$

which is closed since A, B are closed. Hence f is continuous. \square

Corollary 6.11. *Let X be a topological space. Define the relation R by xRy if and only if there is a continuous $\gamma : [0, 1] \rightarrow X$ such that $\gamma(0) = x, \gamma(1) = y$. Then this is an equivalence relation.*

Proof. Trivial. □

Theorem 6.12. *Let $U \subset \mathbb{R}^n$ be open, then U is connected if and only if U is path-connected.*

Proof. It suffice to show every open connected subset of \mathbb{R}^n is path-connected. WLOG $U \neq \emptyset$, fix $x_0 \in U$, let V be the path-connected component containing x_0 . We shall show that $V, U \setminus V$ are both open, so by assumption $V = U$, thus the proof will be done.

V open: Since U is open, for any $x \in U$, there is $r > 0$ such that $D_r(x) \in U$. But any ball is path connected, so $\forall x \in V, \exists r_x > 0, D_{r_x}(x) \in V$, so V is open.

$U \setminus V$ open: Fix by the same proof as above, any path-connected components in V is open, so since $U \setminus V$ is the union of some of them (the ones except V), it is open. □

Example 6.4. For $n \geq 2$, \mathbb{R}^n is not homeomorphic to \mathbb{R} . Assume $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a homeomorphism. Fix $x \in \mathbb{R}^n$, and let $y = f(x)$, then $f|_{\mathbb{R}^n \setminus \{x\}}$ is still a homeomorphism to $\mathbb{R} \setminus \{y\}$. But then $\mathbb{R}^n \setminus \{x\}$ is connected by the preceding theorem, but $\mathbb{R} \setminus \{y\}$ is not, contradiction.

7 Compactness

Recall that a continuous, real-valued function on a closed bounded interval is bounded and attains its bound. The question is, for which topological space X is it true that every continuous real functions is bounded.

Example 7.1. 1. For finite X , every function $X \rightarrow \mathbb{R}$ is bounded.
2. If for all continuous $f : X \rightarrow \mathbb{R}, \exists n \in \mathbb{N}, \exists A_1, A_2, \dots, A_n \subset X$ such that $X = \bigcup_i A_i$ and f is bounded on each A_i , then f is bounded on X .

Note that given continuous $f : X \rightarrow \mathbb{R}$, for $x \in X, U_x = f^{-1}((f(x) - 1, f(x) + 1))$ is open and f is bounded there. So if there is some finite subset of $\{U_x : x \in X\}$ that still covers X , then f must be bounded.

Definition 7.1. An open cover of a topological space X is a family of open sets $\mathcal{U} = \{U_i\}_{i \in I}$ in X such that $X = \bigcup_{i \in I} U_i$.

A subcover of \mathcal{U} is a subset $\mathcal{V} \subset \mathcal{U}$ that is also an open cover of X . \mathcal{V} is called a finite subcover if it is finite.

X is compact if every open cover of X has a finite subcover.

Theorem 7.1. *If $X \neq \emptyset$ is compact and $f : X \rightarrow \mathbb{R}$ is continuous, then f is bounded and attains its bound.*

Proof. By continuity of f and compactness of X , there is a finite subset of $\{f^{-1}((f(x) - 1, f(x) + 1)) : x \in X\}$ that covers X , which means that f is bounded on any set in a finite family, so f is bounded in the union of that family, which is X . To show that f attains its bound, let $m = \{f(x) : x \in X\}$ which exists since $X \neq \emptyset$ and f is bounded. Suppose that there is not an x

with $f(x) = m$, so for any $x \in X$, $f(x) > m$ so $\exists m_x$ such that $f(x) > m_x > m$. Let $U_x = f^{-1}((m_x, \infty))$ which is open and contains x , and $\inf_{U_x} f \geq m_x > m$. Note that the family of all U_x is an open cover of X , so there is a finite subcover $\{U_x\}_{x \in F}$, so $\forall y \in X$, $f(y) \geq \min_{x \in F} m_x > m$, contradiction. \square

Note that for a subspace $Y \subset X$, Y is compact iff whenever \mathcal{U} is a family of open set in X whose union contains Y , there is a finite subset $\mathcal{V} \subset \mathcal{U}$ such that the union of elements in \mathcal{V} contains Y .

Theorem 7.2. $[0, 1]$ is compact.

Proof. Let \mathcal{U} be a set of open sets in \mathbb{R} that contains $[0, 1]$, assume that there does not exist finite subcover that contains $[0, 1]$, then if $0 \leq a < b \leq 1$ and $[a, b]$ cannot be covered by any finite $\mathcal{V} \subset \mathcal{U}$, then let $C = (a+b)/2$, then one of $[a, c]$, $[c, b]$ cannot be covered by finite $\mathcal{V} \subset \mathcal{U}$.

Therefore, inductively we can find intervals $I_n = [a_n, b_n]$ such that

$$I_0 = [0, 1], I_{n+1} \subset I_n, |b_n - a_n| = 1/2^n$$

thus $a_n \rightarrow x, b_n = a_n + (b_n - a_n) \rightarrow x$ for some $x \in [0, 1]$. Now there is $U \in \mathcal{U}$ such that $x \in U$, but then there is some ϵ with $(x - \epsilon, x + \epsilon) \subset U$, therefore for sufficiently large n , $I_n \subset U$, which is a contradiction. \square

Proposition 7.3. Let X be a topological space and $Y \subset X$ a subspace, then

1. If X is compact and Y is closed in X , then Y is compact.
2. If X is Hausdorff and Y is compact, then Y is closed.

Proof. 1. Let \mathcal{U} covers Y , then $\mathcal{U} \cup \{X \setminus Y\}$ is an open cover of X . Since X is compact, there is a finite subcover $\mathcal{V} \subset \mathcal{U} \cup \{X \setminus Y\}$ that covers X , hence $\mathcal{V} \setminus \{X \setminus Y\} \subset \mathcal{U}$ is finite and covers Y .

2. We want to show that its complement is open. Indeed, for any $x \in X \setminus Y$, and for any $y \in Y$, there are disjoint U_y, V_y such that $x \in U_y, y \in V_y$, then $\{V_y\}_{y \in Y}$ is an open cover of Y , thus there is some finite set $F \subset Y$ such that $Y \subset \bigcup_{y \in F} V_y$, so $U = \bigcap_{y \in F} U_y$ is open, and by definition it is disjoint from Y , hence $x \in U \subset X \setminus Y$, which shows that $X \setminus Y$ is open. \square

Proposition 7.4. If X is compact and $f : X \rightarrow Y$ is continuous, then $f(X)$ is compact.

Proof. For any open $\{U_i\}_{i \in I}$ that covers $f(X)$, $\{f^{-1}(U_i)\}_{i \in I}$ is an open cover of X , therefore there is some finite $F \subset I$ such that $\{f^{-1}(U_i)\}_{i \in F}$ covers X , hence $\{U_i\}_{i \in F}$ covers $f(X)$. \square

Remark. 1. Compactness is a topological property.

2. Let $f : X \rightarrow Y$ and $A \subset X$. Suppose A is compact, then $f(A)$ is compact.

Example 7.2. For $a < b$, $[a, b] = f([0, 1])$ where $f(x) = (b-a)x + a$ which is continuous, thus every closed bounded interval is compact.

Corollary 7.5. If X is compact and R is an equivalence relation on X , then X/R is compact.

Proof. The quotient map is continuous and surjective. \square

Theorem 7.6 (Topological Inverse Function Theorem). *If $f : X \rightarrow Y$ is a continuous bijection and X is compact and Y is Hausdorff, then f is a homeomorphism.*

Proof. It suffices to check that f is an open map, which, since f is a bijection, is equivalent to say that f is a closed map.

Fix any closed $V \subset X$, then V is compact since X is compact, thus $f(V)$ is compact since f is continuous, therefore $f(V)$ is closed since Y is Hausdorff. The result follows. \square

Example 7.3. Consider $f : \mathbb{R} \rightarrow S^1$ by $f(t) = e^{2\pi it}$ induces a continuous bijection $\tilde{f} : \mathbb{R}/\mathbb{Z} \rightarrow S^1$. Now $\mathbb{R}/\mathbb{Z} = q([0, 1])$ (where q is the quotient map) is compact and S^1 is Hausdorff since it is a metric space, therefore \tilde{f} is a homeomorphism.

Theorem 7.7 (Tychonoff's Theorem on Finite Products³). *Finite products of compact spaces are compact.*

Proof. It suffices to show in the case where the product consists of 2 components. Assume X, Y are compact and let \mathcal{U} cover $X \times Y$. Fix $x \in X$, then for any $y \in Y$ there is some $W_y \in \mathcal{U}$ with $x, y \in W_y$, so there is some U_y open in X and V_y open in Y such that $(x, y) \in U_y \times V_y \subset W_y$. By compactness, there is a finite $F_Y \subset Y$ with $\bigcup_{y \in F_Y} V_y = Y$. Let $T_x = \bigcup_{y \in F_Y} U_y$ is open and contains x and note that $T_x \times Y \subset \bigcup_{y \in F_Y} W_y$. But then $\{T_x\}_{x \in X}$ covers X , so there is a finite $F_X \subset X$ such that $\{T_x\}_{x \in F_X}$ covers X , hence

$$X \times Y \subset \bigcup_{x \in F_X} T_x \times Y = \bigcup_{x \in F_X} \bigcup_{y \in F_Y} W_y$$

The last term is the required finite subcover. \square

Theorem 7.8 (Heine-Borel Theorem). *A subset $K \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof. If K is compact, note that $f : \mathbb{R}^n \rightarrow \mathbb{R}$ by $x \mapsto \|x\|$, thus it is bounded. K is also closed by Proposition 7.3.

Conversely, if K is closed and bounded, there is some $M > 0$ such that $K \subset [-M, M]^n$ which is compact by Theorem 7.2 and 7.7. Since K is closed in $[-M, M]^n$, it is compact by Proposition 7.3. \square

Definition 7.2. Given an open set $U \subset \mathbb{R}^n$, a sequence of functions $f_k : U \rightarrow \mathbb{R}$ converges locally uniformly on U to some function $f : U \rightarrow \mathbb{R}$ if $\forall x \in U, \exists r > 0, D_r(x) \subset U$ and $f_k \rightarrow f$ uniformly on $D_r(x)$.

Thus this happens if and only if $f_n \rightarrow f$ uniformly on any compact subset of U .

Definition 7.3. A topological space X is sequentially compact if and only if every sequence in X has a convergent subsequence.

Example 7.4. Any closed bounded subset of \mathbb{R}^n is sequentially compact by Bolzano-Weierstrass.

³It works for arbitrary products, but that case is much much harder

Definition 7.4. Fix a metric space (M, d) . For $\epsilon > 0$ and $F \subset M$. We say F is an ϵ -net for M if $\forall x \in M, \exists y \in F, d(x, y) \leq \epsilon$. That is,

$$M = \bigcup_{y \in F} B_\epsilon(y)$$

We say M is totally bounded if for any $\epsilon > 0$, there is a finite ϵ -net for M .

Note that any compact space is totally bounded, but the converse is not true by taking $[0, 1)$, but the only thing missing here is completeness.

Theorem 7.9. *The followings are equivalent for a metric space M :*

- (1) M is compact.
- (2) M is sequentially compact.
- (3) M is totally bounded and complete.

Proof. 1 \implies 2: Let (x_n) be a sequence in M , so for $n \in \mathbb{N}$, let $A_n = \{x_k : k > n\}$. It suffices to show that $\bigcap_{n \in \mathbb{N}} \bar{A}_n$ is nonempty. Assume not, then

$$\bigcup_{n \in \mathbb{N}} M \setminus \bar{A}_n = M$$

Each $M \setminus \bar{A}_n$ is open, so by compactness of M there is some $N \in \mathbb{N}$ such that $\bigcup_{n \leq N} M \setminus \bar{A}_n = M$. But $\{A_n\}$ is decreasing, so necessarily $M \setminus \bar{A}_N = M \implies \bar{A}_N = \emptyset$, contradiction.

2 \implies 3: M is complete since a Cauchy sequence with convergent subsequence is convergent. To see it is totally bounded, assume it is not, then there is some $\epsilon > 0$ such that every ϵ -net is infinite. Pick any $x_1 \in M$, and once we have already picked x_1, \dots, x_n , we pick $x_{n+1} \notin \bigcup_{k=1}^n B_\epsilon(x_k)$. This is valid since otherwise M would have a finite ϵ -net. But (x_n) cannot possibly have any Cauchy subsequence, so it has no converging subsequence.

3 \implies 1: Assume M is not compact, so there is an open cover \mathcal{U} without any finite subcover. We say $A \subset M$ is “bad” if there is no finite subcover of A in \mathcal{U} . So M is bad but \emptyset is not. Note if $A = \bigcup_{i=1}^n B_i$ is bad, then there is some i such that B_i is bad.

Next, we want to show that if A is bad and $\epsilon > 0$, then $\exists B \subset A$ such that B is bad and $\text{diam } B = \sup_{x, y \in B} d(x, y) < \epsilon$. Indeed, since M is bounded, we have a finite $\epsilon/2$ -net F , that is,

$$\bigcup_{x \in F} B_{\epsilon/2}(x) = M \implies \bigcup_{x \in F} (B_{\epsilon/2}(x) \cap A) = A$$

But this would mean that there is some $x \in F$ such that $B_{\epsilon/2}(x) \cap A$ is bad, and by triangle inequality its diameter is less than ϵ . Using this we can construct a sequence $M \supset A_1 \supset A_2 \supset \dots$ such that A_n is bad for any n and $\text{diam } A_n < 1/n$. Picking $x_n \in A_n$ gives a Cauchy sequence (x_n) which converges to some $x \in M$ by completeness. There is some $U \in \mathcal{U}$ that contains x , and it necessarily covers A_n when n is large enough. Contradiction. \square

Remark. 1. We have a new proof of Bolzano-Weierstrass now! We also have a new proof of Theorem 7.7 for metric spaces.

2. Sadly, the equivalence of sequentially compactness and compactness fails in both directions in general topological spaces.

8 Differentiation

Definition 8.1. Fix $m, n \in \mathbb{N}$, let $L(\mathbb{R}^m, \mathbb{R}^n)$ be the set of linear maps to \mathbb{R}^m to \mathbb{R}^n . Note that this space is isomorphic to \mathbb{R}^{mn} , both algebraically and topologically, as we have the metric

$$\forall T \in L(\mathbb{R}^m, \mathbb{R}^n), \|T\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |T_{ij}|^2} = \sqrt{\sum_{i=1}^m \|Te_i\|^2}$$

Lemma 8.1. (a) Given a linear map T , for every $x \in \mathbb{R}^m$, we have $\forall x \in \mathbb{R}^m, \|Tx\| \leq \|T\|\|x\|$. So T is Lipschitz hence continuous.

(b) For $S \in L(\mathbb{R}^m, \mathbb{R}^n), T \in L(\mathbb{R}^n, \mathbb{R}^p)$, $\|TS\| \leq \|T\|\|S\|$.

Proof. (a) If $x = \sum_i x_i e_i$, then

$$\|Tx\| = \left\| \sum_{i=1}^m x_i T e_i \right\| \leq \sum_{i=1}^m |x_i| \|T e_i\| \leq \sqrt{\sum_{i=1}^m |x_i|^2} \sqrt{\sum_{i=1}^m \|T e_i\|^2} = \|x\| \|T\|$$

(b) We have

$$\|TS\| = \sqrt{\sum_{i=1}^m \|T S e_i\|^2} \leq \sqrt{\sum_{i=1}^m \|T\|^2 \|S e_i\|^2} = \|T\| \|S\|$$

As desired. □

Recall that a function $f : \mathbb{R} \rightarrow \mathbb{R}$ is differentiable at a if $\lim_{h \rightarrow 0} (f(a+h) - f(a))/h$ exists. So let $\epsilon(h) = (f(a+h) - f(a))/h - f'(a)$, then $f(a+h) = f(a) + f'(a)h + \epsilon(h)h$ and $\epsilon \rightarrow 0$ as $h \rightarrow 0$. We can think of this as $\epsilon(0) = 0$ and ϵ is continuous at 0. So we want to use it to define differentiation in higher dimensions.

Definition 8.2. Given $m, n \in \mathbb{N}$ and an open set $U \subset \mathbb{R}^m$, a function $f : U \rightarrow \mathbb{R}^n$ and $a \in U$. We say f is differentiable at a if there is a linear map $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ and a function $\epsilon : \{h \in \mathbb{R}^m : a+h \in U\} \rightarrow \mathbb{R}^n$ such that

$$f(a+h) = f(a) + T(h) + \epsilon(h)\|h\|$$

where $\epsilon \rightarrow 0$ as $h \rightarrow 0$. (Or $\epsilon(0) = 0$ and ϵ is continuous at 0).

Remark.

$$\epsilon(h) = \begin{cases} 0, & \text{if } h = 0 \\ (f(a+h) - f(a) - T(h))/\|h\|, & \text{if } h \neq 0 \text{ and } a+h \in U \end{cases}$$

Since U is open, $\exists r > 0, D_r(a) \subset U$, so $D_r(a) \subset \text{Dom } \epsilon$. Note also that our condition on ϵ is also equivalent to say $\epsilon(h)\|h\| = o(\|h\|)$ as $h \rightarrow 0$.

Next, we observe that T (if it exists) is unique. Indeed, if both T, S satisfies our condition, then $(S(h) - T(h))/\|h\| \rightarrow 0$ as $h \rightarrow 0$, so by choosing $h = x/n$ for $n \in \mathbb{N}$ we have $S = T$.

Definition 8.3. This unique T is called the derivative of f at a , denoted by $f'(a)$ or $Df(a)$ or $Df|_a$, so

$$f(a+h) = f(a) + f'(a)(h) + o(\|h\|)$$

Definition 8.4. We say f is differentiable at U if it is differentiable at a for every $a \in U$. So the derivative of f on U is the map $f' : U \rightarrow L(\mathbb{R}^m, \mathbb{R}^n)$.

Remark. When $m = 1$, T is a linear map $\mathbb{R} \rightarrow \mathbb{R}^n$, so $\forall x \in \mathbb{R}, T(x) = xv$ where $v = T(1)$ (this in fact gives us a natural correspondence $L(\mathbb{R}, \mathbb{R}^n) \cong \mathbb{R}^n, T \mapsto T(1)$). Hence for open $U \subset \mathbb{R}, f : U \rightarrow \mathbb{R}^n, a \in U$, to say f is differentiable at a is to say there is some $v \in \mathbb{R}^n$ with $f(a+h) = f(a) + hv + o(h)$. In other words, the linear map $f'(a)$ takes the form $h \mapsto hv$.

Example 8.1. 1. Every constant function is differentiable as we can take $f'(a) \equiv 0 \in L(\mathbb{R}^m, \mathbb{R}^n)$.

2. Every linear map f is differentiable by taking $f'(a) = f \in L(\mathbb{R}^m, \mathbb{R}^n)$. for every a .

3. Any bilinear map $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^p$ is differentiable. Indeed, we have

$$f((a, b) + (h, k)) = f(a + b, h + k) = f(a, b) + f(a, k) + f(h, b) + f(h, k)$$

Note that $f(a, k) + f(h, b)$ is linear in (h, k) , therefore it remains to check $f(h, k) = o(\|h\|)$. Indeed,

$$\begin{aligned} \|f(h, k)\| &= \left\| f \left(\sum_{i=1}^m h_i e_i, \sum_{j=1}^n k_j e_j \right) \right\| \\ &\leq \sum_{i,j} |h_i| |k_j| \|f(e_i, e_j)\| \\ &\leq \|(h, k)\|^2 \sum_{i,j} \|f(e_i, e_j)\| \\ &= O(\|(h, k)\|^2) = o(\|(h, k)\|) \end{aligned}$$

4. Take $f : \mathbb{R}^n \rightarrow \mathbb{R}$ by $f(x) = \|x\|^2$, so

$$f(a+h) = \|a+h\|^2 = \|a\|^2 + 2\langle a, h \rangle + \|h\|^2 = f(a) + 2\langle a, h \rangle + o(\|h\|)$$

That is, $f'(a)(h) = 2\langle a, h \rangle$.

5. Let $M_n \cong L(\mathbb{R}^n, \mathbb{R}^n)$ be the collection of all $n \times n$ real matrices. Consider $f : M_n \rightarrow M_n, A \mapsto A^2$ which has

$$f(A+H) = (A+H)^2 = A^2 + AH + HA + H^2 = f(A) + AH + HA + o(\|H\|)$$

due to Lemma 8.1. So $f'(A)(H) = AH + HA$.

Proposition 8.2. *Differentiability implies continuity.*

Proof. Write $f(a+h) = f(a) + f'(a)(h) + \epsilon(h)\|h\|$ where $\epsilon(0) = 0$ and ϵ is continuous at 0. $f'(a)$ is continuous by Lemma 8.1 (which implies every linear map is continuous), so the RHS is continuous in h , therefore $h \mapsto f(a+h)$ is continuous at $h = 0$, hence f is continuous at a . \square

Proposition 8.3 (Chain Rule). *Consider open $U \in \mathbb{R}^m, V \in \mathbb{R}^n$ and functions $f : U \rightarrow \mathbb{R}^n, g : V \rightarrow \mathbb{R}^m$ and $f(U) \subset V$. If f is differentiable at a and g is differentiable at $f(a)$, then $g \circ f$ is differentiable at a and $(g \circ f)'(a) = g'(f(a)) \circ f'(a)$.*

Proof. Let $b = f(a)$ and $S = f'(a), T = g'(b)$, then

$$\begin{cases} f(a+h) = f(a) + S(h) + \epsilon(h)\|h\| \\ g(b+k) = g(b) + T(k) + \delta(k)\|k\| \end{cases}$$

Where $\epsilon(0) = 0, \delta(0) = 0$ and both of them are continuous at 0. So

$$(g \circ f)(a) = g(b + S(h) + \epsilon(h)\|h\|)$$

Let $k(h) = S(h) + \epsilon(h)\|h\|$, so it equals

$$\begin{aligned} g(b) + T(k) + \delta(k)\|k\| &= (g \circ f)(a) + T \circ S(h) \\ &\quad + \|h\|T(\delta(h)) + \delta(k(h))\|S(h) + \delta(h)\|h\| \end{aligned}$$

Due to continuity of ϵ, δ are continuous at 0, $\|h\|T(\delta(h)) = o(\|h\|)$ and $T(\epsilon(0)) = 0$, so this term is fine.

Also, $\delta(k(0)) = 0$ and $\delta \circ k$ is continuous at 0. In addition,

$$0 \leq \frac{\|S(h) + \delta(h)\|h\|}{\|h\|} \leq \frac{\|S(h)\| + \|\epsilon(h)\|h\|}{\|h\|} \leq \|S\| + \|h\|$$

by Lemma 8.1. So

$$\lim_{h \rightarrow 0} \frac{\delta(k(h))\|S(h) + \delta(h)\|h\|}{\|h\|} = 0 \implies \delta(k(h))\|S(h) + \delta(h)\|h\| = o(\|h\|)$$

Hence $g \circ f$ is differentiable and its derivative is $T \circ S = g'(b) \circ f'(a) = g'(f(a)) \circ f'(a)$. \square

Proposition 8.4. *$f : U \rightarrow \mathbb{R}^n$ ($U \in \mathbb{R}^m$ is open) is differentiable if and only if each components $f_j = \pi_j \circ f$ is differentiable at a . Also,*

$$f'(a)(h) = \sum_{j=1}^n f'_j(a)(h)e'_j$$

Proof. Note that $\pi_j(x) = \langle x, e'_j \rangle$ is linear hence differentiable, thus by chain rule the \implies direction is done. For \impliedby , we have for every j ,

$$f_j(a+h) = f_j(a) + f'(a)(h) + \epsilon_j(h)\|h\|$$

So

$$\begin{aligned} f(a+h) &= \sum_{j=1}^n f_j(a+h)e'_j \\ &= \sum_{j=1}^n (f_j(a) + f'(a)(h) + \epsilon_j(h)\|h\|)e'_j \\ &= f(a) + \left(\sum_{j=1}^n f'_j(a)(h)e'_j \right) + \left(\sum_{j=1}^n \epsilon_j(h)e'_j \right) \|h\| \end{aligned}$$

Since $\epsilon(h) = \sum_{j=1}^n \epsilon_j(h)e'_j$ has $\epsilon(0) = 0$ and is continuous at 0, hence the result. \square

Proposition 8.5. *Let $f, g : U \rightarrow \mathbb{R}^n$ where $U \subset \mathbb{R}^m$ is open and $\phi : U \rightarrow \mathbb{R}$ is differentiable at $a \in U$, then so are $f + g$ and $\phi f : x \mapsto \phi(x)f(x)$, and*

$$(f + g)'(a) = f'(a) + g'(a)$$

$$(\phi f)'(a)(h) = \phi'(a)(h)f(a) + \phi(a)f'(a)(h)$$

Proof. We have

$$f(a + h) = f(a) + f'(a)(h) + \epsilon(h)\|h\|$$

$$g(a + h) = g(a) + g'(a)(h) + \delta(h)\|h\|$$

$$\phi(a + h) = \phi(a) + \phi'(a)(h) + \eta(h)\|h\|$$

Hence

$$(f + g)(a + h) = (f + g)(a) + (f'(a) + g'(a))(h) + (\epsilon(h) + \delta(h))\|h\|$$

We can do the same thing for products as well which will provide a proof, but we shall give a different proof. Let $F : U \rightarrow \mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$ by $f(x) = (\phi(x), f(x))$ and $G : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by $(a, x) \mapsto ax$. F is differentiable by Proposition 8.4 and G is differentiable since it is bilinear, therefore $\phi f = G \circ F$ is differentiable and we can obtain the form of the derivative from the chain rule which is the formula as claimed. \square

Definition 8.5. Let $U \subset \mathbb{R}^m$ be open and $f : U \rightarrow \mathbb{R}^n$. Fix $a \in U$ and a direction (nonzero vector) $u \in \mathbb{R}^m \setminus \{0\}$. The limit

$$\lim_{t \rightarrow 0} \frac{f(a + tu) - f(a)}{t}$$

if exists, is called the directional derivative of f at a to direction u and is denoted by $D_u f(a)$.

Remark. 1. $f(a + tu) = f(a) + tD_u f(a) + o(t)$.

2. Let $\gamma(t) = a + tu$, then $(f \circ \gamma)'(0) = D_u f(a)$.

In the special case where $u = e_i$, we write $D_i f(a)$ to denote $D_{e_i} f(a)$ and it is called the i^{th} partial derivative of f at a .

Proposition 8.6. *If f is differentiable at a , then all $D_u f(a)$ exists and we have $D_u f(a) = f'(a)(u)$, so for $h = \sum_i h_i e_i$, we have*

$$f'(a)(h) = \sum_i h_i D_i f(a)$$

Proof. We have

$$f(a + h) = f(a) + f'(a)(h) + \epsilon(h)\|h\|$$

Then

$$\frac{f(a + tu) - f(a)}{t} = f'(a)(u) + \epsilon(tu) \frac{\|t\|}{t} \rightarrow f'(a)(u)$$

As $t \rightarrow 0$. The rest follows. \square

Remark. 1. Assume f is differentiable at a , then the matrix of $f'(a)$ is exactly represented by $(f'(a))_{ji} = D_i f_j(a) = (\partial f_j / \partial x_i)(a)$. This is called the Jacobian of f at a , denoted by $Jf(a)$.

2. If all partial derivatives exists, so does $D_u f_j(a), \forall j$, and we have $D_u f_j(a) = \pi_j(D_u f(a))$. So $D_u \pi_j = \pi_j D_u$.

3. The converse of the proposition fails in general.

Theorem 8.7. *Suppose $f : U \rightarrow \mathbb{R}^n$ where $U \subset \mathbb{R}^m$ is open. Assume $\exists r > 0, D_r(a) \in U$ and $D_i f$ exists in $D_r(a)$ and is continuous at a for all i , then f is differentiable at a .*

Proof. WLOG $n = 1$ by Proposition 8.4 and the second remark above. We shall prove the case $m = 2$. The general case is similar.

Let $a = (a_1, a_2)$ and consider $h = (h_1, h_2) \in D_r(0)$. Certainly we want the derivative to equal $h_1 D_1 f(a_1, a_2) + h_2 D_2 f(a_1, a_2)$, so we will try to prove

$$f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) - h_1 D_1 f(a_1, a_2) - h_2 D_2 f(a_1, a_2) = o(\|h\|)$$

Note that we can write it out in two parts

$$\begin{aligned} & f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) - h_2 D_2 f(a_1, a_2) \\ & + f(a_1 + h_1, a_2) - f(a_1, a_2) - h_1 D_1 f(a_1, a_2) \end{aligned}$$

We have $f(a_1 + h_1, a_2) - f(a_1, a_2) - h_1 D_1 f(a_1, a_2) = o(h_1) = o(\|h\|)$ as $h \rightarrow 0$. As for the first part, let $\phi(t) = f(a_1 + h_1, a_2 + t)$ for $t \in [-|h_2|, |h_2|]$, so we have

$$f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) - h_2 D_2 f(a_1, a_2) = \phi(h_2) - \phi(0) - h_2 D_2 f(a_1, a_2)$$

Note that ϕ is continuous and is differentiable in $(-|h_2|, |h_2|)$. Indeed we have $\phi'(t) = D_2 f(a_1 + h_1, a_2 + t)$. By MVT, there is some $\theta(h_1, h_2) \in (0, 1)$ such that $\phi(h_2) - \phi(0) = \phi'(\theta h_2) h_2$. Hence

$$\begin{aligned} \phi(h_2) - \phi(0) - h_2 D_2 f(a_1, a_2) &= h_2 (D_2 f(a_1 + h_1, a_2 + \theta h_2) - D_2 f(a_1, a_2)) \\ &= o(h_2) = o(\|h\|) \end{aligned}$$

as $h \rightarrow 0$. So the theorem is proved. \square

Theorem 8.8 (Mean Value Inequality). *Consider an open $U \subset \mathbb{R}^m$ and a function $f : U \rightarrow \mathbb{R}^n$. Assume f is differentiable on U and we are given $a, b \in U$ such that $[a, b] = \{(1-t)a + tb : t \in [0, 1]\} \subset U$. Suppose there is some $M > 0$ with $\forall z \in [a, b], \|f'(z)\| \leq M$, then*

$$\|f(b) - f(a)\| \leq M \|b - a\|$$

Proof. Let $v = f(b) - f(a)$. Consider $\phi : [0, 1] \rightarrow \mathbb{R}$ defined by $\phi(t) = \langle f((1-t)a + tb), v \rangle$. Then $\phi(1) - \phi(0) = \|f(b) - f(a)\|^2$ and ϕ is differentiable with $\phi'(t) = \langle f'((1-t)a + tb)(b-a), v \rangle$. By MVT, $\exists \theta \in (0, 1)$ with $\phi(1) - \phi(0) = \phi'(\theta)$, so we have

$$\begin{aligned} \|f(b) - f(a)\|^2 &= \langle f(b) - f(a), f(b) - f(a) \rangle = \phi(1) - \phi(0) = \phi'(\theta) \\ &= \langle f'((1-\theta)a + \theta b)(b-a), v \rangle \leq \|f'((1-\theta)a + \theta b)(b-a)\| \|v\| \\ &\leq \|f'((1-\theta)a + \theta b)\| \|b-a\| \|v\| \leq M \|b-a\| \|f(b) - f(a)\| \end{aligned}$$

The theorem follows. \square

Corollary 8.9. *Let $U \subset \mathbb{R}^m$ be open and connected, and $f : U \rightarrow \mathbb{R}^n$ be differentiable such that $f' \equiv 0$, then f is constant.*

Proof. It is locally constant by the preceding theorem. It is then globally constant by connectedness. \square

Remark. Suppose we have open $U \subset \mathbb{R}^m, V \subset \mathbb{R}^n$ and $f : U \rightarrow V$ is a bijection such that f is differentiable at $a \in U$ and f^{-1} at $f(a) \in V$. Let $S = f'(a), T = (f^{-1})'(f(a))$, then $ST = I_n$ and $TS = I_m$. $\text{rank}(I_n) = \text{rank}(ST) = \text{rank}(TS) = \text{rank}(I_m)$, hence $n = m$.

Theorem 8.10 (Inverse Function Theorem). *We have an open set $U \subset \mathbb{R}^n$, a C^1 (continuously differentiable) function $f : U \rightarrow \mathbb{R}^n$ and a point $a \in U$ such that $f'(a)$ is invertible, then there exists open set $V \subset U, W \subset \mathbb{R}^n$ open such that $a \in V, f(a) \in W, f|_V : V \rightarrow W$ is a bijection with a C^1 inverse $g : W \rightarrow V$ and $\forall y \in W, g'(y) = (f'(g(y)))^{-1}$.*

Proof. Step 1: We can actually assume WLOG that $a = f(a) = 0$ and $f'(a) = I$. This is because we can consider $h : U - a = \{x - a : x \in U\} \rightarrow \mathbb{R}^n$ by $h(x) = (f'(a))^{-1}(f(x + a) - f(a))$.

Now we can fix $r > 0$ such that $D_r(0) \subset U$ and $\forall x \in D_r(0), \|f'(x) - I\| \leq 1/2$ by continuity.

Step 2: $\forall x, y \in D_r(0), \|f(x) - f(y)\| \geq \|x - y\|/2$, so f is injective. To prove this, consider $h(x) = x - f(x)$ which has $\|h'(x)\| = \|I - f'(x)\| \leq 1/2$ for any $x \in D_r(0)$. So $\|h(x) - h(y)\| \leq \|x - y\|/2$ by Theorem 8.8, in other words,

$$\frac{\|x - y\|}{2} \geq \|h(x) - h(y)\| \geq \|x - y\| - \|f(x) - f(y)\|$$

which leads to the desired inequality.

Step 3: For $0 < s < r/2, D_s(0) \subset f(B_{2s}(0)) \subset f(D_r(0))$. Fix $y \in D_s(0)$ and consider $h : B_{2s}(0) \rightarrow \mathbb{R}^n$ by $x \mapsto y - f(x) + x$. We have $h'(x) = -f'(x) + I$, so $\forall x \in B_{2s}(0), \|h'(x)\| \leq 1/2$. h is then 1/2-Lipschitz by Theorem 8.8. For any $x \in B_{2s}(0), \|h(x)\| = \|h(x) - h(0) + y\| \leq \|x\|/2 + \|y\| \leq 2s$, so $h(B_{2s}(0)) \subset B_{2s}(0)$. By Theorem 4.7 there is some $x \in B_{2r}(0)$ such that $h(x) = x$, which means that $y = f(x)$.

Step 4: Fix $0 < s < r/2$, then let $W = D_s(0)$ and $V = f^{-1}(D_s(0)) \cap D_r(0)$, so V is open and $f(V) = W$ by step 3 and f is injective by step 2, so $f|_V : V \rightarrow W$ is a bijection. The inequality in step 2 also implies that the inverse $f^{-1} = g : W \rightarrow V$ is 1/2-Lipschitz, hence continuous.

If g is differentiable, then we have $I = (f \circ g)'(y)$ for any y , hence $g'(y) = (f'(g(y)))^{-1}$ by Chain Rule. We want to show that g has this as derivative. Indeed, fix $b \in W$ and $a = g(b), T = f'(a)$, we have $f(a + h) = f(a) + T(h) + \epsilon(h)\|h\|$. Fix $\delta > 0$ such that $D_\delta(b) \in W$ and $k \in D_\delta(0)$, by setting $h = h(k) = g(b + k) - g(b)$ we have $k = f(a + h) - f(a) = T(h) + \epsilon(h)\|h\|$, so $h = T^{-1}(k) - T^{-1}(\epsilon(h)\|h\|)$, thus

$$g(b + k) = g(b) + h = g(b) + T^{-1}(k) - T^{-1}(\epsilon(h)\|h\|) = g(b) + T^{-1}(k) + o(\|k\|)$$

Hence g is differentiable at b with derivative $T^{-1} = (f'(g(b)))^{-1}$ which is continuous. \square

Definition 8.6. Let $U \subset \mathbb{R}^m$ be open and $f : U \rightarrow \mathbb{R}^n$. Suppose $a \in U$, then f is twice differentiable at a if there is some open V with $a \in V \subset U$ such that f is differentiable in V and the derivative $f' : V \rightarrow L(\mathbb{R}^m, \mathbb{R}^n)$ is differentiable. $f''(a) = (f')'(a)$ is called the second derivative of f .

So we have $f'' \in L(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^n))$ where we have

$$f'(a+h) = f'(a) + f''(a)(h) + \epsilon(h)\|h\|$$

where $\epsilon \rightarrow 0$ as $h \rightarrow 0$. Note that $\epsilon(h) \in L(\mathbb{R}^m, \mathbb{R}^n)$. So $f'(a+h)(k) = f'(a)(k) + f''(a)(h)(k) + \epsilon(h)(k)\|h\|$ for each fixed $k \in \mathbb{R}^m$. Note also that $L(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^n)) \cong \text{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ by the natural correspondence $T \mapsto B$ with $B(h, k) = T(h)(k)$. So we can think of the second derivative at a as a bilinear map $\mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^n$.

In summary, f is twice differentiable at a if and only if there is a bilinear map $B : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ such that for any $k \in \mathbb{R}^m$ we have

$$f'(a+h)(k) = f'(a)(k) + B(h, k) + o(\|h\|)$$

Example 8.2. For $f : M_n \rightarrow M_n$ by $A \mapsto A^3$. It is differentiable and $f'(A)(H) = HA^2 + AHA + A^2H$. Then to find second derivative by

$$\begin{aligned} f'(A+H)(K) &= K(A+H)^2 + (A+H)K(A+H) + (A+H)^2K \\ &= f'(A)(K) \\ &\quad + KAH + KHA + AKH + HKA + AHK + HAK \\ &\quad + o(\|H\|^2) \end{aligned}$$

So the second derivative is the bilinear map $f''(A) = B(H, K) = KAH + KHA + AKH + HKA + AHK + HAK$.

Assume that f has second derivative at a under the usual setup, then

$$f'(a+h)(k) = f'(a)(k) + f''(a)(h, k) + o(\|h\|)$$

So fix $u, v \in \mathbb{R}^m \setminus \{0\}$, by putting $k = v$ we have

$$D_v f(x+h) = D_v f(a) + f''(a)(h, v) + o(\|h\|)$$

So $D_v f$ is differentiable, therefore we can write

$$D_u D_v f(a) = f''(a)(u, v)$$

Theorem 8.11. Let $U \subset \mathbb{R}^m$ be open and $f : U \rightarrow \mathbb{R}^n$ be second differentiable on U with f'' continuous at a for some $a \in U$, then $f''(a)$ is a symmetric form, that is, for any $0 \neq u, v \in \mathbb{R}^m$, $D_u D_v f(a) = D_v D_u f(a)$.

Proof. WLOG $n = 1$ since $(f_j)'' = (f'')_j$. Define

$$\begin{aligned} \phi(s, t) &= f(a + su + tv) - f(a + su) - (f(a + tv) - f(a)) \\ &= f(a + su + tv) - f(a + tv) - (f(a + su) - f(a)) \end{aligned}$$

Consider $\Psi(x) = f(a + xu + tv) - f(a + xu)$, so $\phi(s, t) = \Psi(s) - \Psi(0) = s\Psi'(\alpha s)$ where $\alpha = \alpha(s, t) \in (0, 1)$ by mean value theorem. Expanding gives $\phi(s, t) =$

$s(D_u f(a + \alpha su + tv) - D_u f(a + \alpha su))$. Consider $\psi(y) = D_u f(a + \alpha su + yv)$, so $\phi(s, t) = s(\psi(t) - \psi(0)) = st\psi'(\beta t)$, $\beta = \beta(s, t) \in (0, 1)$. In other words,

$$\begin{aligned}\frac{\phi(s, t)}{st} &= D_v D_u f(a + \alpha su + \beta tv) \\ &= f''(a + \alpha su + \beta tv)(v, u) \rightarrow f''(a)(v, u)\end{aligned}$$

by the continuity of f'' at a . Repeat the process in the other order to get

$$\frac{\phi(s, t)}{st} \rightarrow f''(a)(u, v)$$

So they are equal. \square

Suppose we have $U \subset \mathbb{R}^m$ open and $f : U \rightarrow \mathbb{R}$, we say f has a local maximum at $a \in U$ if $\exists r > 0, \forall b \in D_r(a), f(b) \leq f(a)$. We can similarly define local minima.

Definition 8.7. We say f has a stationary point at a if f is differentiable at a and $f'(a) = 0$.

It is immediate that f has stationary points at each local maximum and minimum.

Theorem 8.12. Let $U \subset \mathbb{R}^m$ be open and $f : U \rightarrow \mathbb{R}$ be twice differentiable in U and suppose $f'(a) = 0$ and f'' is continuous at a . If the symmetric form f'' is positive definite at a , then f has a local minimum at a ; if it is negative definite at a , then f has a local maximum at a .

Proof. It is a non-examinable fact that

$$f(a + h) = f(a) + f'(a)(h) + (1/2)f''(a)(h, h) + \epsilon(h)\|h\|^2$$

for some ϵ such that $\epsilon \rightarrow 0$ as $\|h\| \rightarrow 0$.

Recall that $f''(a)$ is orthogonally diagonalizable as a real symmetric form, i.e. there is an orthonormal basis $\{u_k\}$ for \mathbb{R}^m such that

$$f''(u_i, u_j) = \begin{cases} 0, & \text{if } i \neq j \\ \lambda_i, & \text{if } i = j \end{cases}$$

Assume that f'' is positive definite, then $\lambda_i = f''(a)(u_i, u_i) > 0$ for all i , hence $\mu = \min\{\lambda_i : 1 \leq i \leq m\} > 0$. Therefore we have, for any $h = \sum_i h_i u_i \in \mathbb{R}^m$,

$$f''(a)(h, h) = \sum_{i,j} h_i h_j f''(a)(u_i, u_j) = \sum_i h_i^2 \lambda_i \geq \mu \sum_i h_i^2 = \mu \|h\|^2$$

Hence $f(a + h) - f(a) \geq \mu \|h\|^2/4 \geq 0$ whenever $\|h\|$ is small enough so that $\|\epsilon(h)\| \leq \mu/4$. This means that f has a local minimum at a . The negative definite case is analogous. \square