Analysis and Topology

Cambridge University Mathematical Tripos: Part IB

17th May 2024

Contents

1 Uniform convergence

1.1 Definition

Recall that $x_n \to x$ as $n \to \infty$ (for $x \in \mathbb{R}$ or $\mathbb{C}$) if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x_n - x| < \varepsilon$$

This is essentially considering the ε-neighbourhood of x. We aim to define the same notion of convergence for functions, by defining an analogous concept of an ε-neighbourhood. In particular, each value on the domain should converge in its own ε-neighbourhood.

Definition. Let S be a set, and $f, f_n : S \to \mathbb{R}$, be functions. We say that (f_n) converges to f uniformly on S if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \forall x \in S, |f_n(x) - f(x)| < \varepsilon$$

Note. N depends only on ε, *not* on any x. Each x converges therefore at a 'similar speed', hence the name 'uniform convergence'.

Equivalently, we can write

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \sup_{x \in S} |f_n(x) - f(x)| < \varepsilon$$

The supremum condition is equivalent overall because the inequality on the right is weakened to a possible equality, but we can always decrease ε to retain the inequality. Alternatively, we could write

$$\lim_{n \to \infty} \sup_{x \in S} |f_n - f| = 0$$

For each $x \in S$, $(f_n(x))_{n=1}^{\infty} \to f(x)$. Hence, f is unique given (f_n), since limits are unique. We call f the *uniform limit* of (f_n) on S.

1.2 Pointwise convergence

Definition. (f_n) converges *pointwise* to f on S if $(f_n(x))_{n=1}^{\infty}$ converges to $f(x)$ for every $x \in S$. In other words,

$$\underbrace{\forall x \in S}_{\text{order rearranged}}, \forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |f_n(x) - f(x)| < \varepsilon$$

Now, N depends both on ε and on x. Note that the pointwise limit of (f_n) on S is also unique since limits are unique.

Remark. Uniform convergence implies pointwise convergence, and the uniform limit is the pointwise limit.

Example. Let $f_n(x) = x^2 e^{-nx}$ on $[0, \infty)$, $n \in \mathbb{N}$. Does (f_n) converge uniformly on the domain? First let us check pointwise convergence. We have $x^2 e^{-nx} \to 0$ hence pointwise convergence to $f(x) = 0$ is satisfied. Now, we need only check uniform convergence to the function $f(x) = 0$.

$$\sup_{x \in [0,\infty)} |f_n(x) - 0| = \sup_{x \in [0,\infty)} f_n(x)$$

We could differentiate f_n and find the maximum if it exists, but we might not find the maximum if it is (for example) on the endpoints. A much better method is to find an upper bound on $|f_n(x) - f(x)|$ (which, in this example, is $f_n(x)$) that does not depend on x. In this case, we can expand e^{nx} on the denominator and isolate a single term to get

$$x^2 e^{-nx} = \frac{x^2}{e^{nx}} \leq \frac{2}{n^2}; \quad \forall x$$

Hence,

$$\sup_{x \in [0,\infty)} |f_n(x) - 0| \to 0$$

and uniform convergence is satisfied.

Example. Consider $f_n(x) = x^n$ on $[0,1], n \in \mathbb{N}$. A pointwise limit is reached by

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

Consider $\sup |f_n(x) - f(x)|$ excluding 1 (since at 1 the supremum is zero). Note $f_n(x) \to 1$ as $x \to 1$ from below, for all n. Hence the supremum is always 1 by choosing an x sufficiently close to 1. So $f_n \not\to f$ uniformly on $[0,1]$, hence (f_n) does not converge at all uniformly on this domain. Or,

$$\sup f_n(x) \geq f_n\left(\left(\frac{1}{2}\right)^{1/n}\right) = \frac{1}{2}$$

Remark. If $f_n \not\to f$ uniformly on S,

$$\exists \varepsilon > 0, \forall N \in \mathbb{N}, \exists n \geq N, \exists x \in S, |f_n(x) - f(x)| \geq \varepsilon$$

In the above example, we proved something stronger:

$$\forall n, \exists x \in S, f_n(x) \geq \frac{1}{2}$$

We could have alternatively stated, for example, $f_n(x)$ is continuous so there exists some subset of $[0,1]$ greater than $\frac{1}{2}$ always.

Theorem. Let $S \subseteq \mathbb{R}, \mathbb{C}$. Let $(f_n), f : S \to \mathbb{R}(\text{or } \mathbb{C})$, where f_n is continuous and $(f_n) \to f$ uniformly on S. Then f is continuous.

Informally, the uniform limit of continuous functions is continuous.

Proof. Fix some point $a \in S, \varepsilon > 0$. We seek $\delta > 0$ such that $\forall x \in S, |x - a| < \delta \implies |f(x) - f(a)| < \varepsilon$. We fix an $n \in \mathbb{N}$ such that $\forall x \in S, |f_n(x) - f(x)| < \varepsilon$. Since f_n is continuous, there exists $\delta > 0$ such that $\forall x \in S, |x - a| < \delta \implies |f_n(x) - f_n(a)| < \varepsilon$. So, $\forall x \in S$,

$$|x - a| < \delta \implies |f(x) - f(a)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| < 3\varepsilon$$

□

Remark. The above proof is often called a 3ε-proof. Note, the proof is not true for pointwise convergence; if $f_n \to f$ pointwise and f_n continuous, f is not necessarily continuous. Further, it is not true for differentiability; f_n differentiable does not imply f differentiable (see example sheet). Another way to interpret the result of the above theorem is to swap limits:

$$\lim_{x \to a} \lim_{n \to \infty} f_n(x) = \lim_{x \to a} f(x) = f(a) = \lim_{n \to \infty} f_n(a) = \lim_{n \to \infty} \lim_{x \to a} f_n(x)$$

1.3 Uniform limit of bounded functions

Lemma. Let $f_n \to f$ uniformly on S. If f_n is bounded for every n, then so is f.

In other words, the uniform limit of bounded functions is bounded.

Proof. Fix some $n \in \mathbb{N}$ such that $\forall x \in S, |f_n(x) - f(x)| < 1$. Since f_n is bounded, $\exists M \in \mathbb{R}$ such that $\forall x \in S, |f_n(x)| < M$. Hence, $\forall x \in S, |f(x)| \leq |f(x) - f_n(x)| + |f_n(x)| \leq 1 + M$. So f is bounded. □

1.4 Integrability

Let $f : [a, b] \to \mathbb{R}$ be a bounded function. Recall that for a dissection $\mathcal{D}$ of $[a, b]$, we define the upper and lower sums of f with respect to $\mathcal{D}$ by

$$U_{\mathcal{D}}(f) = \sum_{k=1}^{n} (x_k - x_{k-1}) \sup_{[x_{k-1}, x_k]} f(x)$$

$$L_{\mathcal{D}}(f) = \sum_{k=1}^{n} (x_k - x_{k-1}) \inf_{[x_{k-1}, x_k]} f(x)$$

Riemann's integrability criterion states that f is integrable if and only if

$$\forall \varepsilon, \exists \mathcal{D}, U_{\mathcal{D}}(f) - L_{\mathcal{D}}(f) < \varepsilon$$

Equivalently, for any $I \subset [a, b]$, we have

$$\sup_I f - \inf_I f = \sup_{x,y \in I} (f(x) - f(y)) = \sup_{x,y \in I} |f(x) - f(y)|$$

This is called the oscillation of f on I. So an integrable function 'doesn't oscillate too much'.

Theorem. Let $f_n : [a, b] \to \mathbb{R}$ be integrable for all n. If $f_n \to f$ uniformly on $[a, b]$, then f is integrable and

$$\int_a^b f_n \to \int_a^b f$$

Proof. First, we prove f to be bounded, then we will check Riemann's criterion. We know f is bounded because each f_n is bounded, hence by the lemma above f is bounded. Now fix $\varepsilon > 0$, and choose $n \in \mathbb{N}$ such that $\forall x \in [a, b], |f_n(x) - f(x)| < \varepsilon$. Since f_n is integrable, $\exists \mathcal{D} : a = x_0 < x_1 < \dots < x_N = b$ of $[a, b]$ such that $U_{\mathcal{D}} - L_{\mathcal{D}} < \varepsilon$. Now, we fix $k \in \{1, \dots, N\}$ and then for any $x, y \in [x_{k-1}, x_k]$ we have

$$|f(x) - f(y)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| < 2\varepsilon + |f_n(x) - f_n(y)|$$

Taking the supremum,

$$\sup_{x,y \in [x_{k-1}, x_k]} (f(x) - f(y)) \leq \sup_{x,y \in [x_{k-1}, x_k]} |f_n(x) - f_n(y)| + 2\varepsilon$$

Multiplying by $(x_k - x_{k-1})$ and taking the sum over all k,

$$U(f) - L(f) \leq U(f_n) - L(f_n) + 2\varepsilon(b-a) \leq \varepsilon(2(b-a)+1)$$

Hence f is integrable. We can now show that

$$\left|\int_a^b f_n - \int_a^b f\right| \leq \int_a^b |f_n - f| \leq (b-a)\sup_{[a,b]}|f_n - f| \to 0$$

□

Remark. We can interpret this as

$$\int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \lim_{n\to\infty}\int_a^b f_n(x)\,\mathrm{d}x$$

This is another 'allowed' way to swap limits.

Corollary. Let $f_n\colon [a,b] \to \mathbb{R}$ be integrable for all n. If $\sum_{n=1}^\infty f_n(x)$ converges uniformly on $[a,b]$, then

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

is integrable, and

$$\int_a^b \sum_{n=1}^\infty f_n(x)\,\mathrm{d}x = \sum_{n=1}^\infty \int_a^b f_n(x)\,\mathrm{d}x$$

Proof. Let $F_n(x) = \sum_{k=1}^n f_k(x)$. By assumption, $F_n \to F$ uniformly on $[a,b]$. F_n is integrable where the integral of F_n is the sum of the integrals:

$$\int_a^b F_n = \sum_{k=1}^n \int_a^b f_k$$

Then the result follows from the theorem above. □

1.5 Differentiability

Theorem. Let $f_n\colon [a,b] \to \mathbb{R}$ be continuously differentiable for all n. Suppose $\sum_{k=1}^\infty f_k'(x)$ converges uniformly on $[a,b]$, and that $\forall c \in [a,b], \sum_{n-1}^\infty f_n(c)$ converges. Then, $\sum_{k=1}^\infty f_k(x)$ converges uniformly on $[a,b]$ to a continuously differentiable function f, and

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{k=1}^\infty f_k\right) = \sum_{k=1}^\infty \frac{\mathrm{d}}{\mathrm{d}x} f_k(x)$$

Proof. Let $g(x) = \sum_{k=1}^{\infty} f_k'(x)$, for $x \in [a,b]$. The general idea is that we want to solve the differential equation $f' = g$ subject to the initial condition $f(c) = \sum_{n=1}^{\infty} f_n(c)$. Let $\lambda = \sum_{n=1}^{\infty} f_n(c)$ and define $f\colon [a,b] \to \mathbb{R}$ by

$$f(x) = \lambda + \int_c^x g(t)\,\mathrm{d}t$$

Note that g is integrable; $\sum_{k=1}^{\infty} f_k'(x) \to g$ uniformly implies that g is continuous and hence integrable. By the fundamental theorem of calculus, $f' = g$ and $f(c) = \lambda$. So we have found such an f that satisfies the conditions set out. All that remains is to prove uniform convergence of $\sum_{k=1}^{\infty} f_k \to f$. Also by the fundamental theorem, $f_k(x) = f_k(c) + \int_c^x f_k'(t)\,\mathrm{d}t$. Let $\varepsilon > 0$. There exists $N \in \mathbb{N}$ such that $\left|\lambda - \sum_{k=1}^{N} f_k(c)\right| < \varepsilon$ and $\left|g(t) - \sum_{k=1}^{N} f_k'(t)\right| < \varepsilon$. Now, for $n \geq N$ we have

$$\begin{aligned}
\left|f(x) - \sum_{k=1}^{n} f_k(x)\right| &= \left|\lambda + \int_c^x g(t)\,\mathrm{d}t - \sum_{k=1}^{n}\left(f_k(c) + \int_c^x f_k'(t)\,\mathrm{d}t\right)\right| \\
&\leq \left|\lambda - \sum_{k=1}^{n} f_k(c)\right| + \left|\int_c^x \left(g(t) - \sum_{k-1}^{n} f_k'(t)\right)\mathrm{d}t\right| \\
&\leq \varepsilon + |x-c|\varepsilon \\
&\leq \varepsilon(b-a+1)
\end{aligned}$$

□

1.6 Conditions for uniform convergence

Recall that a scalar sequence x_n is Cauchy if

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

and that the general principle of convergence shows that any Cauchy sequence converges.

1.7 General principle of uniform convergence

Definition. A sequence (f_n) of scalar functions on a set S is called *uniformly Cauchy* if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m,n \geq N, \forall x \in S, |f_m(x) - f_n(x)| < \varepsilon$$

Theorem. A uniformly Cauchy sequence of functions is uniformly convergent.

Proof. Let $x \in S$ and we will show that $(f_n(x))_{n=1}^{\infty}$ converges. Given $\varepsilon > 0, \exists N \in \mathbb{N}, \forall m,n \geq N, \forall t \in S, |f_m(t) - f_n(t)| < \varepsilon$. In particular, $\forall m,n \geq N, |f_m(x) - f_n(x)| < \varepsilon$. So certainly $(f_n(x))_{n=1}^{\infty}$ is Cauchy and hence convergent by the general principle of convergence. Therefore f_n converges pointwise. Now, let $f(x)$ be the limit $f(x) = \lim_{n\to\infty} f_n(x)$. Then $f_n \to f$ pointwise on S. Now we must extend this to show $f_n \to f$ uniformly on S. Given $\varepsilon > 0$, we know that $\exists N \in \mathbb{N}, \forall m,n \geq N, \forall x \in S, |f_m(x) - f_n(x)| < \varepsilon$. Now, we must show $\forall n \geq N, \forall x \in S, |f_n(x) - f(x)| < 2\varepsilon$, then we are done. We will fix $x \in S, n \geq N$. Since $f_n(x) \to f(x)$, we can choose $m \in \mathbb{N}$ such that $|f_m(x) - f(x)| < \varepsilon$, and $m \geq N$. Note however that m depends on x in this statement, but this doesn't matter—we have shown that

$$|f_n(x) - f(x)| \leq |f_n(x) - f_m(x)| + |f_m(x) - f(x)| \leq \varepsilon + \varepsilon = 2\varepsilon$$

which is a result that, in itself, does *not* depend on x. □

Note. Alternatively, we could end the proof as the following. Fix $x \in S, n \geq N$. Then

$$\forall m \geq N, |f_n(x) - f_m(x)| < \varepsilon$$

Then let $m \to \infty$, and

$$|f_n(x) - f(x)| \leq \varepsilon$$

1.8 Weierstrass M-test

Theorem. Let (f_n) be a sequence of scalar functions on S. Assume that $\forall n \in \mathbb{N}, \exists M_n \in \mathbb{R}^+, \forall x \in S, |f_n(x)| \leq M_n$. In other words, (f_n) is a sequence of bounded scalar functions. Then,

$$\sum_{n=1}^{\infty} M_n < \infty \implies \sum_{n=1}^{\infty} f_n(x) \text{ is uniformly convergent on } S$$

Proof. Let $F_n(x) = \sum_{k=1}^{n} f_k(x)$ for $x \in S, n \in \mathbb{N}$. Then

$$|F_n(x) - F_m(x)| \leq \sum_{k=m+1}^{n} |f_k(x)| \leq \sum_{k=m+1}^{n} M_k$$

Hence, given $\varepsilon > 0$, we can choose $N \in \mathbb{N}$ such that $\sum_{k=N+1}^{n} M_k < \varepsilon$. Thus, $\forall x \in S, \forall n \geq m \geq N$, we have

$$|F_n(x) - F_m(x)| \leq \sum_{k=m+1}^{n} M_k < \varepsilon$$

We have shown (F_n) is uniformly Cauchy on S and hence uniformly convergent on S. □

1.9 Power series

Consider the power series

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

where $c_n \in \mathbb{C}, a \in \mathbb{C}$ are constants, and $z \in \mathbb{C}$. Let $R \in [0, \infty]$ be the radius of convergence. Recall that

$$|z-a| < R \implies \sum_{n=0}^{\infty} c_n (z-a)^n \text{ converges absolutely;}$$

$$|z-a| > R \implies \sum_{n=0}^{\infty} c_n (z-a)^n \text{ diverges}$$

Let $D(a,R) \coloneqq \{z \in \mathbb{C} \mid |z-a| < R\}$ be the open disc centred on a with radius R. Then we can create $f : D(a,\mathbb{R}) \to \mathbb{C}$ to be defined by the power series, which is well-defined. f is the pointwise limit of the power series on D. In general, the convergence of the power series is not uniformly convergent.

Example. $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$ has $R = 1$. Let $f_n : D(0,1) \to \mathcal{C}$ be defined by $f_n(z) = \frac{z^n}{n^2}$. Then for every $z \in D(0,1), |z| \leq \frac{1}{n^2}$. Since $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < \infty$, by the Weierstrass M-test, the power series converges uniformly on the disc.

Example. Consider $\sum_{n=0}^{\infty} z^n = \frac{1}{1-z}$ with $R = 1$. Now,

$$\forall z \in D(0,1), \left|\sum_{n=0}^{\infty} z^n\right| \leq N + 1$$

Therefore, the series does not converge uniformly on the disc since $\frac{1}{1-z}$ is unbounded on the disc. Alternatively, consider

$$\sup_{|z|<1}\left|\frac{1}{1-z} - \sum_{k=0}^{n} z_k\right| = \sup_{|z|<1}\left|\frac{z^{n+1}}{1-z}\right| = \infty$$

In some sense, the problem with uniform convergence here is that we are allowed to go too close too the boundary.

Theorem. Suppose the power series $\sum_{n=0}^{\infty} c_n(z-a)^n$ has radius of convergence R. Then for all $0 < r < R$, the power series converges uniformly on $D(a,r)$.

Proof. Let $w \in \mathbb{C}$ such that $r < |w-a| < R$, for instance $w = a + \frac{r+R}{2}$. Now, let $\rho = \frac{r}{|w-a|} \in (0,1)$. Since $\sum_{n=0}^{\infty} c_n(w-a)^n$ converges, we have that $c_n(w-a)^n \to 0$ as $n \to \infty$. Therefore, $\exists M \in \mathbb{R}+$ such that $|c_n(w-a)^n| \leq M$ for all $n \in \mathbb{N}$, since convergence implies boundedness. Now, for $z \in D(a,r), n \in \mathbb{N}$ we have

$$|c_n(z-a)^n| = |c_n(w-a)^n|\left(\frac{|z-a|}{|w-a|}\right)^n \leq M\left(\frac{r}{|w-a|}\right)^n = M\rho^n$$

Since the sum $\sum_{n=0}^{\infty} M\rho^n$ converges, the Weierstrass M-test shows us that $\sum_{n=0}^{\infty} c_n(z-a)^n$ converges uniformly on $D(a,r)$. □

Remark. $f : D(a,R) \to \mathbb{C}$ defined by $f(z) = \sum_{n=0}^{\infty} c_n(z-a)^n$ is the uniform limit on $D(a,r)$ of polynomials for any r such that $0 < r < R$. Hence f is continuous on $D(a,r)$. Since $D(a,R) = \bigcup_{0<r<R} D(a,r)$, it follows that f is continuous everywhere inside the radius of convergence.

Recall that the termwise derivative $\sum_{n=1}^{\infty} c_n n(z-a)^{n-1}$ has the same radius of convergence. This sequence therefore also converges uniformly on $D(a,r)$ if $0 < r < R$. Analogously to the previous result about interchanging derivatives and sums, we can show that $\sum c_n(z-a)^n$ is complex differentiable on $D(a,R)$ with derivative $\sum_{n=1}^{\infty} c_n n(z-a)^{n-1}$. This is seen in the IB Complex Analysis course.

Now, fix $w \in D(a,R)$. Then fix r such that $|w-a| < r < R$, and fix $\delta > 0$ such that $|w-a| + \delta < r$. If $|z-w| < \delta$, then $|z-a| \leq |z-w| + |w-a| < \delta + |w-a| < r$. Therefore, geometrically, $D(w,\delta) \subset D(a,r)$. Hence, $\sum_{n=0}^{\infty} c_n(z-a)^n$ converges uniformly on $D(w,\delta)$. This is known as local uniform convergence.

Definition. $U \subset \mathbb{C}$ is called open if $\forall w \in U, \exists \delta > 0, D(w, \delta) \subset U$.

Definition. Let U be an open subset of $\mathbb{C}$, and f_n be a sequence of scalar functions on U. Then f_n converges locally uniformly on U if

$$\forall w \in U, \exists \delta > 0, f_n \text{ converges uniformly on } D(w, \delta) \subset U$$

Remark. Above, we showed that power series always converge locally uniformly inside the radius of convergence, or equivalently inside the disc $D(a, R)$. We will return to this point about local uniform convergence when discussing compactness.

2 Uniform continuity

2.1 Definition

Let $U \subset \mathbb{R}, \mathbb{C}$. Let f be a scalar function on U. Then for $x \in U$, we say f is continuous at x if

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

We say f is continuous on U if f is continuous at x for all $x \in U$:

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

Note here that δ depends on ε and x.

Definition. Let U, f be as in the previous definition. We say f is *uniformly continuous* if

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

Now, δ works for all $x \in U$ simultaneously; δ depends on ε only. Certainly, uniform continuity implies continuity.

Example. Let $f : \mathbb{R} \to \mathbb{R}$ such that $f(x) = 2x + 17$. Then f is uniformly continuous; given $\varepsilon > 0$, we can find $\delta = \frac{1}{2}\varepsilon$. Then $\forall x, y \in \mathbb{R}, |y - x| < \delta \implies |f(y) - f(x)| = |2y - 2x| = 2y - x < 2\delta = \varepsilon$.

Example. Let $f : \mathbb{R} \to \mathbb{R}$, defined by $f(x) = x^2$. This is not uniformly continuous, since no δ works for all x given some 'bad' ε. Let us take $\varepsilon = 1$, and we wish to show that no δ exists. Suppose some δ does exist. Then, let $x > 0$ and $y = x + \frac{\delta}{2}$. We should have $|f(y) - f(x)| < 1$.

$$\left(x + \frac{\delta}{2}\right)^2 - x^2 = \delta x + \frac{\delta^2}{4}$$

So for $x = \frac{1}{\delta}$, this condition $|f(y) - f(x)| < 1$ is not satisfied. Hence f is not uniformly continuous.

Note. For U, f as in the above definition, f is not uniformly continuous on U if

$$\exists \varepsilon > 0, \forall \delta > 0, \exists x, y \in U, |y - x| < \delta, |f(y) - f(x)| \geq \varepsilon$$

So there are points arbitrarily close together whose difference of function values exceed some fixed ε.

2.2 Properties of continuous functions

Theorem. Let f be a scalar function on a closed bounded interval $[a, b]$. If f is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.

Proof. Suppose there exists $\varepsilon > 0$ such that $\forall \delta > 0, \exists x, y \in [a, b], |y - x| < \delta, |f(y) - f(x)| \geq \varepsilon$. In particular, we can construct a sequence (δ_n) defined by $\delta_n = \frac{1}{n}$, and we can construct sequences $x_n, y_n \in [a, b]$ such that $|y_n - x_n| < \frac{1}{n}$ but $|f(y_n) - f(x_n)| \geq \varepsilon$. By the Bolzano–Weierstrass theorem, there exists a subsequence (x_{k_n}) that converges. Now, let x be the limit of the subsequence, $\lim_{n\to\infty} x_{k_n}$. Then $x \in [a, b]$ since the interval is closed. Then, $|y_{k_n} - x| \leq |y_{k_n} - x_{k_n}| + |x_{k_n} - x| < \frac{1}{n} + |x_{k_n} - x| \to 0$. Hence $y_{k_n} \to x$. Now, since f is continuous $f(x_{k_n}), f(y_{k_n}) \to f(x)$. Now, $\varepsilon \leq |f(x_{k_n}) - f(y_{k_n})| \to |f(x) - f(x)| = 0$, which is a contradiction. □

Corollary. A continuous function $f : [a, b] \to \mathbb{R}$ is Riemann integrable.

Proof. Since a continuous function on a closed bounded interval is bounded, we have that f is bounded. Now, fix $\varepsilon > 0$, and we want to find a dissection $\mathcal{D}$ such that the difference between upper and lower sums is less than ε. By the above theorem, f is uniformly continuous. Hence,

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

So we must simply choose a dissection such that all intervals have size smaller than δ. For instance, choose some $n \in \mathbb{N}$ such that $\frac{b-a}{N} < \delta$, and then divide the interval equally into n subintervals. If I is an interval in this dissection, then $\forall x, y \in I$ we have $|y - x| < \delta$ and hence $|f(y) - f(x)| < \varepsilon$. Hence,

$$\sup_{x,y\in I} |f(y) - f(x)| \leq \varepsilon$$

Multiplying by the length of I and summing over all subintervals I,

$$U_{\mathcal{D}}(f) - L_{\mathcal{D}}(f) \leq (b - a)\varepsilon$$

Hence f is Riemann integrable. □

3 Metric spaces

3.1 Definition

Definition. Let M be a set. Then a *metric* on M is a function $d : M \times M \to \mathbb{R}$ such that
(i) (positivity) $\forall x, y \in M, d(x, y) \geq 0$, and in particular, $x = y \iff d(x, y) = 0$
(ii) (symmetric) $\forall x, y \in M, d(x, y) = d(y, x)$
(iii) (triangle inequality) $\forall x, y, z \in M, d(x, z) \leq d(x, y) + d(y, z)$.
A metric space is a set M together with a metric d on M, written as the pair (M, d).

Example. Let $M = \mathbb{R}, \mathbb{C}$ and $d(x, y) = |x - y|$. This is known as the 'standard metric' on M. If a metric is not specified, the standard metric is taken as implied.

Example. Let $M = \mathbb{R}^n, \mathbb{C}^n$, and we define the Euclidean norm (or Euclidean length) to be

$$\|x\| = \|x\|_2 = \left(\sum_{k=1}^{n} |x_k|^2\right)^{\frac{1}{2}}$$

This satisfies

$$\|x + y\| \leq \|x\| + \|y\|$$

and it then follows that we can define the metric as

$$d_2(x, y) = \|x - y\|_2$$

called the Euclidean metric. We can check that this is indeed a metric easily. This is the standard metric on $\mathbb{R}^n, \mathbb{C}^n$. The metric space (M, d) in this case is called n-dimensional real (or complex) Euclidean space, sometimes denoted ℓ_2^n. The Euclidean norm is sometimes called the ℓ_2 norm, and the Euclidean metric is the ℓ_2 metric.

Example. Let $M = \mathbb{R}^n, \mathbb{C}^n$, and we define the ℓ_1 norm to be

$$|x|_1 = \sum_{k=1}^{n} |x_k|$$

which defines the ℓ_1 metric given by

$$d_1(x, y) = \|x - y\|_1$$

(M, d_1) is denoted ℓ_1^n. We can generalise and form the metric space ℓ_p^n for all $p \in [1, \infty]$.

Example. Again, let $M = \mathbb{R}^n, \mathbb{C}^n$. We can define the ℓ_∞ norm by

$$\|x\|_\infty = \max_{1 \leq k \leq n} |x_k|$$

This defines the ℓ_∞ metric:

$$d_\infty(x, y) = \|x - y\|_\infty = \max_{1 \leq k \leq n} |x_k - y_k|$$

We denote (M, d) by ℓ_∞^n.

In this course, we will only work with $p = 1, 2, \infty$, although the calculations can be made to work for other p.

Example. Let S be a set. Let $\ell_\infty(S)$ be the set of all bounded scalar functions on S. We then define the ℓ_∞ norm of $f \in \ell_\infty(S)$ by

$$\|f\| = \|f\|_\infty = \sup_{x \in S} |f(x)|$$

The supremum exists since the function is always bounded. This is also known as the 'sup norm' or the 'uniform norm'. Note that, for $f, g \in \ell_\infty(S)$, and $x \in S$,

$$\|f + g\| \leq \sup_{x \in S} |f(x) + g(x)| \leq |f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\| + \|g\|$$

Hence $d(f, g) = \|f - g\|$ defines a metric on $\ell_\infty(S)$. This is the standard metric on this space $\ell_\infty(S)$, also called the 'uniform metric'. For example, $\ell_\infty(\{1, \dots, n\}) = \mathbb{R}^n$ with the metric ℓ_∞. Also, for $\ell_\infty(\mathbb{N})$, we typically omit the $\mathbb{N}$ and instead write ℓ_∞ for the space of scalar sequences with the uniform metric.

Example. Consider $C[a,b]$, the set of all continuous functions on $[a,b]$. For $p=1,2$, we define the L_p norm of $f \in C[a,b]$ by

$$\|f\|_p = \left(\int_a^b |f(x)|^p \,\mathrm{d}x\right)^{\frac{1}{p}}$$

which induces the L_p metric on $C[a,b]$.

Example. Let M be a set. Then

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

is a metric, called the discrete metric on M. In particular, (M,d) is called a discrete metric space.

Example. Let G be a group generated by $S \subset G$. We assume $e \notin S$ and $x \in S \implies x^{-1} \in S$. Then

$$d(x,y) = \min\{n \geq 0 : \exists s_1, \dots, s_n, y = xs_1 \dots s_n\}$$

defines a metric called the word metric.

Example. Let p be prime. Then

$$d(x,y) = \begin{cases} 0 & \text{if } x=y \\ p^{-n} & \text{otherwise, where } x-y = p^n m, n \geq 0, m \in Z, p \nmid m \end{cases}$$

defines a metric on $\mathbb{Z}$. This is known as the p-adic metric.

3.2 Subspaces

Let (M,d) be a metric space, and $N \subset M$. Then naturally we can restrict d to $N \times N$, giving a metric on N. (N,d) is called a subspace of M.

Example. Consider $\mathbb{Q}$ with the metric $d(x,y) = |x-y|$. This is clearly a subspace of $\mathbb{R}$ (implicitly with the standard metric on $\mathbb{R}$).

Example. Since every continuous function on a closed bounded interval is bounded, $C[a,b]$ is a subset of $\ell_\infty[a,b]$. Hence $C[a,b]$ with the uniform metric is a subspace of $\ell_\infty[a,b]$.

3.3 Product spaces

Let $(M,d), (M',d')$ be metric spaces. Then any of the following defines a metric on the Cartesian product $M \times M'$.

(i) $d_1((x,x'),(y,y')) = d(x,y) + d(x',y')$

(ii) $d_2((x,x'),(y,y')) = \left(d(x,y)^2 + d(x',y')^2\right)^{\frac{1}{2}}$

(iii) $d_\infty((x,x'),(y,y')) = \max\{d(x,y), d(x',y')\}$

We commonly write $(M \times M', p)$ as $M \oplus_p M'$. Note that we always have

$$d_\infty \leq d_2 \leq d_1 \leq 2d_\infty$$

We can generalise for $n \in \mathbb{N}$ and metric spaces (M_k, d_k) for $k \in \{1, \dots, n\}$, by defining

$$\left(\bigoplus_{k=1}^n M_k\right)_p = M_1 \oplus_p \cdots \oplus_p M_n = \left(M_1 \times \cdots \times M_n, d_p\right)$$

Example. $\mathbb{R} \oplus_1 \mathbb{R} = \ell_1^2$. Further, $\mathbb{R} \oplus_2 \mathbb{R} \oplus_2 R = \ell_2^3$, and other analogous results hold.

Remark. $\mathbb{R} \oplus_1 \mathbb{R} \oplus_2 \mathbb{R}$ does not make sense since we have not defined the associativity of the $\oplus$ operator. The two choices yield different metric spaces.

3.4 Convergence

Let M be a metric space, and (x_n) a sequence in M. Given $x \in M$, we say that (x_n) converges to x in M if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, d(x_n, x) < \varepsilon$$

We say that (x_n) is convergent in M if $\exists x \in M$ such that $x_n \to x$. Otherwise, we say that (x_n) is divergent. Note that $x_n \to x$ in M if and only if $d(x_n, x) \to 0$ in $\mathbb{R}$.

Lemma. Suppose we have a sequence $x_n \to x$ and $x_n \to y$ in a metric space M. Then $x = y$.

Proof. Suppose $x \neq y$. Then let $\varepsilon = \frac{d(x,y)}{3} > 0$. So, by the definition of convergence,

$$\exists N_1 \in \mathbb{N}, \forall n \geq N_1, d(x_n, x) < \varepsilon;$$

$$\exists N_2 \in \mathbb{N}, \forall n \geq N_2, d(x_n, y) < \varepsilon$$

Now, fix $N \in \mathbb{N}$ such that $n \geq N_1, n \geq N_2$, for instance $N = \max\{N_1, N_2\}$. Then

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < 2\varepsilon = \frac{2}{3} d(x, y)$$

which is a contradiction. □

Definition. Given a convergent sequence (x_n) in a metric space M, we say the *limit* of (x_n) is the unique $x \in M$ such that $x_n \to x$ as $n \to \infty$. This is denoted

$$\lim_{n \to \infty} x_n$$

Example. This definition has the usual meaning when $M = \mathbb{R}, \mathbb{C}$.

Example. The constant sequence defined by $x_n = x$ converges to x. In particular, 'eventually constant' sequences converge; let (x_n) be a sequence in M such that $\exists x \in M, \exists N \in \mathbb{N}, \forall n \geq N, x_n = x$, then $x_n \to x$. It is not necessarily true that sequences only converge if they are eventually constant. However, in a discrete metric space, the converse is true, since we can choose ε smaller than all distances.

Example. Consider the 3-adic metric. Then, $3^n \to 0$ as $n \to \infty$ since $d(3^n, 0) = 3^{-n} \to 0$.

Example. Let S be a set. Then, $f_n \to f$ in $\ell_\infty(S)$ in the uniform metric if and only if $d(f_n, f) = \|f_n - f\|_\infty = \sup_S |f_n - f| \to 0$, which is precisely the condition that $f_n \to f$ uniformly on S. Note, however, that $f_n(x) = x + \frac{1}{n}$ for $x \in \mathbb{R}, n \in \mathbb{N}$ and $f(x) = x$, then certainly $f_n \to x$ uniformly on $\mathbb{R}$. However, $f_n, f \notin \ell_\infty(\mathbb{R})$, so the uniform metric is not defined on these functions. So the notion of uniform convergence visited before is slightly more general than the idea of convergence in this metric space.

Example. Consider Euclidean space $M = \mathbb{R}^n, \mathbb{C}^n$ with the ℓ_2 metric. Then, consider

$$x^{(k)} = \left(x_1^{(k)}, \dots, x_n^{(k)}\right) \in M$$

for $k \in \mathbb{N}$, and $x = (x_1, \dots, x_n) \in M$. Then,

$$\left|x_i^{(k)} - x_i\right| \leq \left\|x^{(k)} - x\right\|_2 \leq \sum_{i=1}^n \left|x_i^{(k)} - x_i\right|$$

So $x^{(k)} \to x$ if and only if all i satisfy $x_i^{(k)} \to x_i$. This can be thought of as convergence being equivalent to coordinate-wise (or pointwise) convergence.

Example. Consider $f_n(x) = x^n$ for $x \in [0,1]$, and $n \in \mathbb{N}$. Then (f_n) is a sequence in $C[0,1]$, which converges pointwise but not uniformly. So (f_n) is not convergent in the uniform metric. However, using the L_1 metric, we have

$$d_1(f_n, 0) = \|f_n\|_1 = \int_0^1 f_n = \frac{1}{n+1} \to 0$$

So, $f_n \to 0$ in $(C[0,1], L_1)$.

Example. Let N be a subspace of a metric space M, and (x_n) be a convergent sequence in N. Then (x_n) converges in M. The converse is not necessarily true; consider $M = \mathbb{R}$ and $N = (0,\infty)$ with $(x_n) = \frac{1}{n}$. This is divergent in N but convergent in M.

Example. Let $(M,d), (M',d')$ be metric spaces. Let $N = M \oplus_p M'$. Let $a_n = (x_n, y_n) \in N$ for all $n \in \mathbb{N}$, and $a = (x,y) \in N$. Then

$$a_n \to a \text{ in } N \iff x_n \to x \text{ in } M, y_n \to y \text{ in } M'$$

Indeed,

$$\max\{d(x_n,x), d'(y_n,y)\} = d_\infty(a_n,a) \leq d_p(a_n,a) \leq 2d_1(a_n,a) = 2d(x_n,x) + 2d'(y_n,y)$$

3.5 Continuity

Definition. Let $f : M \to M'$ be a function between metric spaces $(M,d), (M',d')$. Then for $a \in M$, we say f is continuous at a if

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

We say f is continuous if f is continuous at a for all $a \in M$. In other words,

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

Note that δ depends both on ε and a.

Proposition. Let $f : M \to M'$ be as above. Let $a \in M$. Then the following are equivalent:

(i) f is continuous at a;

(ii) $x_n \to a$ in M implies $f(x_n) \to f(a)$ in M

Proof. First we show (i) implies (ii). Suppose $x_n \to a$ in M. Then fix $\varepsilon > 0$, and seek $N \in \mathbb{N}$ such that $\forall n \geq N, d'(f(x_n), f(a)) < \varepsilon$. By continuity, there exists $\delta > 0$ such that $\forall x \in M, d(x,a) < \delta \implies d'(f(x_n), f(a)) < \varepsilon$ as required. So we want N such that $\forall n \geq N, d(x,a) < \delta$, which must exist since $x_n \to a$.

Now, we show (ii) implies (i). Suppose that f is not continuous at a. Then,

$$\exists \varepsilon > 0, \forall \delta > 0, \exists x \in M, d(x,a) < \delta, d'(f(x), f(a)) \geq \varepsilon$$

So fix such an ε for which no suitable δ exists. Choose the sequence $\delta_n = \frac{1}{n}$, so

$$d(x_n, a) < \frac{1}{n}; \quad d'(f(x_n), f(a)) \geq \varepsilon$$

Then $x_n \to a$ in M but $f(x_n) \not\to f(a)$ in M, which is a contradiction. □

Proposition. Let f, g be scalar functions on a metric space M. Let $a \in M$. Then if f, g are continuous at a, so are $f + g$ and $f \cdot g$. Moreover, letting $N = \{x \in M : g(x) \neq 0\}$ and assuming $a \in N$, $\frac{f}{g}$ is continuous at a. Hence if f, g are continuous, then so are $f + g, f \cdot g, \frac{f}{g}$ where they are defined.

Proof. Suppose $x_n \to a$. Then by the above proposition, $(f \cdot g)(x_n) = f(x_n) \cdot g(x_n) \to f(a) \cdot g(a) = (f \cdot g)(a)$, and similar results hold for the other operators. □

Remark. If $f : M \to M'$ is continuous everywhere,

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right)$$

by the second proposition.

Proposition. Let $f : M \to M', g : M' \to M''$ be functions between metric spaces. If f is continuous at a and g is continuous at $f(a)$, then $g \circ f$ is continuous at a. If f, g are continuous, $g \circ f$ is continuous.

Proof. Let $\varepsilon > 0$. We want to find $\delta > 0$ such that $\forall x \in M$,

$$d(x,a) < \delta \implies d''(g(f(x)), g(f(a))) < \varepsilon$$

Since g is continuous at $f(a)$, there exists $\eta > 0$ such that $\forall y \in M'$,

$$d'(y, f(a)) < \eta \implies d''(g(y), g(f(a))) < \varepsilon$$

Now, since f is continuous at a, for this η there exists δ such that for all $x \in M$,

$$d(x,a) < \delta \implies d'(f(x) - f(a)) < \eta$$

Then
$$d(x,a) < \delta \implies d''(g(f(x)), g(f(a))) < \varepsilon$$
as required. $\square$

Example. Constant functions are continuous. For instance, let $b \in M$ and let $f(x) = b$. Then this is continuous since $d'(f(x) - f(a)) = d'(b,b) = 0$ so any $\delta > 0$ will satisfy the condition.

Example. The identity function $f : M \to M$ defined by $x \mapsto x$ is continuous. Consider $d(f(x) - f(a)) = d(x - a)$. So $\delta = \varepsilon$ will suffice.

Example. All real and complex polynomials and rational functions are continuous wherever they are defined by the propositions and examples above. In fact, using uniform convergence, the uniform limits of such functions are also continuous. For example, exponential and trigonometric functions are continuous.

Example. Let (M,d) be a metric space. Then $d : M \oplus_p M \to \mathbb{R}$, which can be viewed as a function between metric spaces $M \oplus_p M$ and $\mathbb{R}$. Then, given $v = (x,x'), w = (y,y') \in M \oplus_p M$,
$$|d(v) - d(w)| = |d(x,x') - d(y,y')| \leq d(x,y) + d(x',y') = d_1(v,w) \leq 2d_p(v,w)$$
Hence $\delta = \frac{\varepsilon}{2}$ will suffice.

3.6 Isometric, Lipschitz, and uniformly continuous functions

Definition. Let $f : M \to M'$ be a function between metric spaces. Then, f is

(i) *isometric*, if
$$\forall x,y \in M, d'(f(x), f(y)) = d(x,y)$$

(ii) *Lipschitz*, or c-Lipschitz, if
$$\exists c \in \mathbb{R}^+, \forall x,y \in M, d'(f(x), f(y)) \leq c \cdot d(x,y)$$

(iii) *uniformly continuous*, if
$$\forall \varepsilon > 0, \exists \delta > 0, \forall x,y \in M, d(x,y) < \delta \implies d'(f(x), f(y)) < \varepsilon$$

Remark. Any isometric function is 1-Lipschitz. Any Lipschitz function is uniformly continuous. Any uniformly continuous function is continuous.

Remark. If a function is isometric, it is injective, since $f(x) = f(y) \implies x = y$. For example, if $N \subset M$, the inclusion map $i : N \to M$ defined by $i(x) = x$ is isometric but not surjective. An isometric and surjective map is called an *isometry*. If there exists an isometry $M \to M'$, we say that M and M' are isometric metric spaces, or M' is an isometric copy of M.

Example. Suppose $(M,d), (M',d')$ be metric spaces. Let $y \in M'$. We define $f : M \to M \oplus_p M'$ by $x \mapsto (x,y)$. Then $d_p(f(x), f(z)) = d_p((x,y),(z,y)) = d(x,z)$. So the function f is isometric. Therefore, $M \times \{y\}$ is an isometric copy of M in $M \oplus_p M'$.

Example. Consider the projections $q : M \oplus_p M' \to M$ defined by $q(x,y) = x$ and $q' : M \oplus_p M' \to M'$ defined by $q'(x,y) = y$. These projections are both 1-Lipschitz. Indeed,
$$d(q(x,y), q(x',y')) = d(x,x') \leq d_p((x,y),(x',y'))$$

In particular, polynomials in any finite number of variables are continuous since we can multiply continuous functions together.

3.7 Generalised triangle inequality

Suppose $u, x, y, z \in M$. Then, $|d(u,x) - d(y,z)| \leq d(u,y) + d(x,z)$. First,

$$d(u,x) \leq d(u,y) + d(y,x) \leq d(u,y) + d(y,z) + d(z,x)$$

Rearranging,

$$d(u,x) - d(y,z) \leq d(u,y) + d(x,z)$$

To achieve the negative, satisfying both conditions in the absolute value term,

$$d(y,z) \leq d(y,u) + d(u,x) + d(x,z)$$

which gives

$$d(y,z) - d(u,x) \leq d(u,y) + d(x,z)$$

as required.

4 Topology of metric spaces

4.1 Open balls

Definition. Let M be a metric space, $x \in M$, $r > 0$. Then the *open ball* in M of centre x and radius r is the set

$$\mathcal{D}_r(x) = \{y \in M : d(y,x) < r\}$$

The open ball notation is a convenient syntax for denoting closeness in some metric space. Note that, for example, $x_n \to n$ in M is equivalent to saying

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, x_n \in \mathcal{D}_\varepsilon(x)$$

We can also say that $f : M \to M'$ is continuous at x if

$$\forall \varepsilon > 0, \exists \delta > 0, f(\mathcal{D}_\delta(x)) \subset \mathcal{D}_\varepsilon(f(x))$$

Definition. The closed ball of centre x and radius $r \geq 0$ is the set

$$\mathcal{B}_r(x) = \{y \in M : d(y,x) \leq r\}$$

Example. In $\mathbb{R}$, $\mathcal{D}_r(x) = (x-r, x+r)$. Further, $\mathcal{B}_r(x) = [x-r, x+r]$. In the plane $(\mathbb{R}^2, d_p)$,

$$\mathcal{B}_1(0) = \left\{x \in \mathbb{R}^2 : \|x\|_p \leq 1\right\}$$

Note. $\mathcal{D}_r(x) \subset \mathcal{B}_r(x) \subset \mathcal{D}_s(x)$ for all $r < s$.

Example. Let M be a discrete metric space. Then for $x \in M$,

$$\mathcal{D}_1(x) = \{x\}; \quad \mathcal{B}_1(x) = M$$

4.2 Neighbourhoods and openness

Definition. Let M be a metric space, and $U \subset M$. Then for $x \in M$, we say that U is a *neighbourhood* of x (in M) if

$$\exists r > 0, \mathcal{D}_r(x) \subset U \iff \exists r > 0, \mathcal{B}_r(x) \subset U$$

Definition. We say $U \subset M$ is *open* in M, or that U is an *open subset* of M, if

$$\forall x \in U, \exists r > 0, \mathcal{D}_r(x) \subset U$$

So U is a neighbourhood of all points in U.

Example. $\mathcal{D}_r(x), \mathcal{B}_r(x)$ are neighbourhoods of x.

Example. Let $H = \{z \in \mathbb{C} : \operatorname{Im} z \geq 0\}$. Let $w \in H$ and $\delta = \operatorname{Im} w$. If $\delta > 0$, then $\mathcal{D}_\delta(w) \subset H$. If $\delta = 0$, then for any r, $\mathcal{D}_\delta(w) \not\subset H$. So H is not open.

Lemma. Open balls are open.

Proof. Let $\mathcal{D}_r(x)$ be an open ball in a metric space M. We need to show that

$$\forall y \in \mathcal{D}_r(x), \exists \delta > 0, \mathcal{D}_\delta(y) \subset \mathcal{D}_r(x)$$

So let $y \in \mathcal{D}_r(x)$ and set $\delta = r - d(x, y)$. Note that $d(x, y) > 0$, and by the triangle inequality,

$$d(z, x) \leq d(z, y) + d(y, x) < \delta + (r - \delta) = r$$

as required. □

Corollary. Let M be a metric space, $U \subset M$, $x \in M$. Then U is a neighbourhood of x if and only if there exists an open subset V of M such that $x \in V \subset U$.

Proof. In the forward direction, there exists $r > 0$ such that $\mathcal{D}_r(x) \subset U$, so let $V = \mathcal{D}_r(x)$. Conversely, if V is open we can construct $r > 0$ such that $\mathcal{D}_r(x) \subset V \subset U$. So U is a neighbourhood of x. □

4.3 Continuity and convergence using topology

Proposition. In a metric space M, the following are equivalent.
(i) $x_n \to x$;
(ii) for all neighbourhoods U of x in M, $\exists N \in \mathbb{N}, \forall n \geq N, x_n \in U$;
(iii) for all open neighbourhoods U of x in M, $\exists N \in \mathbb{N}, \forall n \geq N, x_n \in U$.

Proof. First, (i) implies (ii). Let U be a neighbourhood of x. Then by definition $\exists \varepsilon > 0, \mathcal{D}_\varepsilon(x) \subset U$. Since $x_n \to x$,

$$\exists N \in \mathbb{N}, \forall n \geq N, x_n \in \mathcal{D}_\varepsilon(x)$$

hence $\forall n \geq N, x_n \in U$.

Now we show (ii) implies (iii). This is clear since any open set U with $x \in U$ is a neighbourhood of x.

Finally, (iii) implies (i). Fix $\varepsilon > 0$. By the above lemma, $U = \mathcal{D}_\varepsilon(x)$ is open, and $x \in U$. Then by (iii),

$$\exists N \in \mathbb{N}, \forall n \geq n, x_n \in U$$

hence $d(x_n, x) < \varepsilon$. □

Proposition. Let $f : M \to M'$ be a function between metric spaces.

(a) The following are equivalent for all $x \in M$.
  (i) f is continuous at x;
  (ii) for all neighbourhoods V of $f(x)$ in M', there exists a neighbourhood U of x in M such that $f(U) \subset V$;
  (iii) for all neighbourhoods V of $f(x)$ in M', $f^{-1}(V)$ is a neighbourhood of x in M.

(b) The following are equivalent.
  (i) f is continuous;
  (ii) $f^{-1}(V)$ is open in M for all open subsets V of M'.

Proof. First, we show (a)(i) implies (a)(ii). Let V be a neighbourhood of $f(x)$ in M'. By definition, $\exists \varepsilon > 0$ such that $\mathcal{D}_\varepsilon(f(x)) \subset V$. Since f is continuous at x, there exists $\delta > 0$ such that $f(\mathcal{D}_\delta(x)) \subset \mathcal{D}_\varepsilon(f(x))$. Then, $U = \mathcal{D}_\delta(x)$ is a neighbourhood of x in M, and $f(U) \subset V$.

Now, (a)(ii) implies (a)(iii). Let V be a neighbourhood of $f(x)$ in M'. By (ii), there exists a neighbourhood of x in M such that $f(U) \subset V$. Then $U \subset f^{-1}(V)$ and since U is a neighbourhood of x in M, ther exists $r > 0$ such that $\mathcal{D}_r(x) \subset U \subset f^{-1}(V)$ Thus, $f^{-1}(V)$ is a neighbourhood of x in M.

Finally, (a)(iii) implies (a)(i). Given $\varepsilon > 0$, $V = \mathcal{D}_\varepsilon(f(x))$ is a neighbourhood of $f(x)$ in V. By (iii), $f^{-1}(V)$ is a neighbourhood of x in M. So $\exists \delta > 0$ such that $\mathcal{D}_\delta(x) \subset f^{-1}(V)$. Thus, $f(\mathcal{D}_\delta(x)) \subset V = \mathcal{D}_\varepsilon(f(x))$.

Now, (b)(i) implies (b)(ii). Let V be open in M'. So pick $x \in f^{-1}(V)$. Then, $f(x) \in V$. Since V is open, $\exists \varepsilon > 0, \mathcal{D}_\varepsilon(f(x)) \subset V$. Since f is continuous at x, $\exists \delta > 0, f(\mathcal{D}_\delta(x)) \subset \mathcal{D}_\varepsilon(f(x))$. Then, $\mathcal{D}_\delta(x) \subset f^{-1}(\mathcal{D}_\varepsilon(f(x))) \subset f^{-1}(V)$.

Finally, (b)(ii) implies (b)(i). Consider $x \in M$. We must show f is continuous at x. Let $\varepsilon > 0$. Consider the ball $V = \mathcal{D}_\varepsilon(f(x))$. This is open in M' by the above lemma. By (ii), $f^{-1}(V)$ is open in M. Further, $x \in f^{-1}(V)$. So by definition, $\exists \delta > 0, \mathcal{D}_\delta(x) \subset V$, which is exactly continuity as required. □

Definition. The *topology of a metric space M* is the family of all open subsets of M.

Proposition. The topology of a metric space satisfies

(i) $\varnothing$ and M are open;
(ii) if U_i are open in M for $i \in I$ (I may be countable or uncountable), then $\bigcup_{i \in I} U_i$ is open in M;
(iii) if U, V are open then $U \cap V$ is open.

Proof. (ii): Let $x \in \bigcup_{i \in I} U_i$, then $\exists i_a \in I, x \in U_{i_a}$. Then since U_{i_a} is open, $\exists \delta > 0, \mathcal{D}_r(x) \subset U_{i_a} \subset \bigcup_{i \in I} U_i$

(iii) Given $x \in U \cap V$, since U is open then $\exists r > 0$, $\mathcal{D}_r(x) \subset U$ and $\exists s > 0$, $\mathcal{D}_s(x) \subset V$. Then let $t = \min(r, s)$, and $\mathcal{D}_t(x) = \mathcal{D}_r(x) \cap \mathcal{D}_s(x) \subset U \cap V$. □

4.4 Properties of topology of metric space

Definition. A subspace A of a metric space M is *closed in M* if for every sequence $(x_n) \in A$ that is convergent in M,

$$\lim_{n \to \infty} x_n \in A$$

Lemma. Closed balls are closed.

Proof. Consider $\mathcal{B}_r(x)$ in M. Consider further $(x_n) \in \mathcal{B}_r(x)$ such that $x_n \to z$ in M.

$$d(z, x) \leq d(z, x_n) + d(x_n, x) \leq d(z, x_n) + r \to r$$

Hence $d(z, x) \leq r$, so $z \in \mathcal{B}_r(x)$. □

Example. $[0, 1] = \mathcal{B}_{1/2}(1/2)$ is closed in $\mathbb{R}$. This is not open, for instance consider $D_r(0) \not\subset [0, 1]$.

Example. $(0, 1) = \mathcal{D}_{1/2}(1/2)$ is open in $\mathbb{R}$. This is not closed, for instance the sequence $\frac{1}{n+1}$ tends to zero in $\mathbb{R}$.

Example. $\mathbb{R}$ and $\varnothing$ are open and closed in $\mathbb{R}$.

Example. $(0, 1]$ in $\mathbb{R}$ is neither open nor closed. Consider $\mathcal{D}_r(1) \not\subset (0, 1]$ and $\frac{1}{n} \to 0 \notin (0, 1]$.

Lemma. Let $A \subset M$. Then A is closed in M if and only if $M \setminus A$ is open in M.

Proof. Let A be closed. Suppose $M \setminus A$ is not open. Then $\exists x \in M \setminus A, \forall r > 0, \mathcal{D}_r(x) \not\subset M \setminus A$, so $\mathcal{D}_r(x) \cap A \neq \varnothing$. In particular, for every n we can choose a point in $\mathcal{D}_{1/n}(x) \cap A$. Then, $d(x_n, x) < \frac{1}{n} \to 0$ and $x_n \in A$ which contradicts the fact that A is closed.

Conversely, let us assume $M \setminus A$ is open, but suppose A is not closed. Then there exists a sequence $(x_n) \in A$ such that $x_n \to x$ in M but $x \notin A$. Since $x \in M \setminus A$ and $M \setminus A$ is open, there exists $\varepsilon > 0, \mathcal{D}_\varepsilon(x) \subset M \setminus A$. Since $x_n \to x$, we must have $\exists N \in \mathbb{N}, \forall n \geq N, x_n \in \mathcal{D}_\varepsilon(x)$ and hence $x_n \in M \setminus A$, which is a contradiction. □

Example. Let M be a discrete metric space. Let $A \subset M$. Then for all $x \in A$, $\mathcal{D}_1(x) = \{x\} \subset A$. Hence A is open. So in a discrete metric space, all subsets are open. Hence every subset is closed.

4.5 Homeomorphisms

Definition. A map $f : M \to M'$ between metric spaces is called a *homeomorphism* if f is a bijection and f, f^{-1} are continuous. Equivalently, f is a bijection, and for all open sets V in M', $f^{-1}(V)$ is open in M, and for all open sets U in M, $f(U)$ is open in M'. If there exists a homeomorphism between M, M', we say that M, M' are homeomorphic.

Example. Consider $(0, \infty)$ and $(0, 1)$. Consider the map $x \mapsto \frac{1}{x+1}$ with inverse $x \mapsto \frac{1}{x} - 1$. These are continuous, so the metric spaces are homeomorphic.

Remark. Every isometry is a homeomorphism, since it is bijective by definition. It is not true that every homeomorphism is an isometry.

Consider the identity on $\mathbb{R}$ with the discrete metric to $\mathbb{R}$ with the Euclidean metric. This is a continuous bijection whose inverse is not continuous. So it is not true that a continuous bijection always has a continuous inverse.

4.6 Equivalence of metrics

Definition. Let d, d' be metrics on a set M. We say that d, d' are *equivalent*, written $d \sim d'$, if they define the same topology. In particular, $U \subset M$ is open in (M, d) if and only if U is open in (M, d'). So $d \sim d'$ if and only if $\mathrm{id} : (M, d) \to (M, d')$ is a homeomorphism.

Remark. If $d \sim d'$, then (M, d) and (M, d') have the same convergent sequences and continuous maps.

Definition. Let d, d' be metrics on M. Then we say d, d' are *uniformly equivalent*, written $d \sim_u d'$ if

$$\mathrm{id} : (M, d) \to (M, d'); \quad \mathrm{id} : (M, d') \to (M, d)$$

are uniformly continuous. We say d, d' are *Lipschitz equivalent*, written $d \sim_{\mathrm{Lip}} d'$, if the identity maps above are Lipschitz. Equivalently, $d \sim_{\mathrm{Lip}} d'$ if $\exists a > 0, b > 0, ad(x, y) \leq d'(x, y) \leq bd(x, y)$. Note, $d \sim_{\mathrm{Lip}} d' \implies d \sim_u d' \implies d \sim d'$.

Example. Given a metric space (M, d), we define $d'(x, y) = \min(1, d(x, y))$. This defines a metric on M, and $d' \sim_u d$.

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

Example. Consider $C[0, 1]$. The L_1 metric and the uniform metric are not equivalent. Consider $f_n(x) = x^n$. This is convergent to zero in the L_1 metric but is not convergent in the uniform metric.

Example. The discrete metric and Euclidean metric on $\mathbb{R}$ are not equivalent. This is because in the discrete metric all sets are open, but in the Euclidean metric there are some non-open sets.

5 Completeness

5.1 Cauchy sequences

In $\mathbb{R}, \mathbb{C}$, every Cauchy sequence is convergent. We wish to generalise this notion to an arbitrary metric space. Recall that a sequence (x_n) in $\mathbb{R}$ or $\mathbb{C}$ is bounded if there exists $c \in \mathbb{R}^+$ such that $\forall n \in \mathbb{N}, |x_n| \leq c$.

Definition. A sequence (x_n) in a metric space M is said to be *Cauchy* if

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

The sequence is bounded if

$$\exists z \in M, \exists r > 0, \forall n \in \mathbb{N}, x_n \in \mathcal{B}_r(z)$$

This is equivalent to

$$\forall z \in M, \exists r > 0, \forall n \in \mathbb{N}, x_n \in \mathcal{B}_r(z)$$

by considering the triangle inequality around the given z point. In particular, if the metric arises from a norm, (x_n) is bounded if and only if $\|x_n\|$ is bounded.

Lemma. If a sequence is convergent, it is Cauchy. If a sequence is Cauchy, it is bounded.

Proof. Let (x_n) be a sequence in M. First, we assume that (x_n) is convergent in M, so let x be the limit. Given $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\forall n \geq N$, $d(x_n, x) < \varepsilon$. Then, for all $m, n \geq N$, we have $d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) \leq 2\varepsilon$ as required. So (x_n) is Cauchy.

Now conversely, we assume (x_n) is Cauchy. There exists $n \in \mathbb{N}$ such that $\forall m, n \geq N$, we have $d(x_m, x_n) < 1$. In particular, $d(x_n, x_N) < 1$ for $n \geq N$. In other words, $x_n \in \mathcal{B}_1(x_N)$. Now, let $r = \max\{d(x_1, x_N), \dots, d(x_{N-1}, x_N), 1\}$. This r is a bound for all elements of the sequence; for all $n \in \mathbb{N}, x_n \in \mathcal{B}_r(x_N)$. □

Remark. Boundedness does not imply the sequence is Cauchy. For instance, consider the sequence $0, 1, 0, 1, \dots$ in $\mathbb{R}$. If a sequence is Cauchy, it is not necessarily convergent in an arbitrary metric space (not $\mathbb{R}, \mathbb{C}$). For instance, consider $x_n = \frac{1}{n}$ in $(0, \infty)$. This is certainly not convergent, since the limit cannot be zero.

5.2 Definition of completeness

Definition. A metric space M is called *complete* if every Cauchy sequence in M converges in M.

Example. $\mathbb{R}, \mathbb{C}$ are complete.

5.3 Completeness of product spaces

Proposition. Product spaces of complete spaces are complete. More precisely, if M, M' are complete, then so is $M \oplus_p M'$.

Proof. Let (a_n) be a Cauchy sequence in the product space $M \oplus_p M'$. We will write $a_n = (x_n, x'_n)$ for all n. Then, since (a_n) is Cauchy,

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n \in N, d_p(a_m, a_n) < \varepsilon$$

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

$$d(x_m, x_n) \leq \max\{d(x_m, x_n), d(x'_m, x'_n)\} \leq d_p(a_m, a_n) < \varepsilon$$

Hence (x_n) is Cauchy in M, and similarly (x'_m) is Cauchy in M'. Since M, M' are complete, $(x_n), (x'_n)$ are convergent in M, M' to x, x'. Now, let $a = (x, x')$. Then,

$$d_p(a_n, a) \leq d_1(a_n, a) = d(x_n, x) + d(x'_n, x') \to 0$$

So the product space is complete. □

Remark. (a_n) is Cauchy in $M \oplus_p M'$ if and only if (x_n) is Cauchy in M and (x'_n) is Cauchy in M'.

Corollary. $\mathbb{R}^n, \mathbb{C}^n$ are complete in the ℓ_p metric. In particular, n-dimensional real or complex Euclidean space is complete.

5.4 Completeness of subspaces and function spaces

Theorem. Let S be any set. Then, $\ell_\infty(S)$, the set of bounded scalar functions on S, is complete in the uniform metric D.

Proof. Let (f_n) be a Cauchy sequence in $\ell_\infty(S)$. Then,

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall m, n \geq N, D(x_m, x_n) = \sup_{x \in S} |f_m(x) - f_n(x)| < \varepsilon$$

In other words, $\forall m, n \geq N, \forall x \in S, |f_m(x) - f_n(x)| < \varepsilon$. So (f_n) is uniformly Cauchy as defined previously. As shown previously, (f_n) is uniformly convergent. Hence, there is a scalar function f on S such that $f_n \to f$ uniformly on S. We have also shown previously that the uniform limit f of bounded functions (f_n) is bounded. In other words, $f \in \ell_\infty(S)$. Now it remains to show that $f_n \to f$ in the uniform metric.

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq M, \forall x \in S, |f_n(x) - f(x)| < \varepsilon$$

Hence,

$$\forall n \geq N, \sup_{x \in S} |f_n(x) - f(x)| = D(f_n, f) \leq \varepsilon$$

which is convergence in the metric as required. □

Proposition. Let N be a subspace of a metric space M. Then,
(i) If N is complete, N is closed in M.
(ii) If M is complete and N is closed in M, then N is complete.
In other words, in a complete metric space, a subspace is complete if and only if it is closed.

Proof. To prove (i), we let (x_n) be a sequence in N and assume that $x_n \to x$ in M. We want to show that $x \in N$. We know (x_n) is convergent in M, so it is Cauchy in M. So (x_n) is Cauchy in N. Since N is complete, $x_n \to y$ in N. So $x_n \to y$ in M. By uniqueness of limits, $x = y$ as required.

Now we want to prove (ii) is complete. Let (x_n) be a Cauchy sequence in N. Then (x_n) is Cauchy in M. Since M is complete, $x_n \to x$ in M for some $x \in M$. Since N is closed in M, $x \in N$. So $x_n \to x$ in N. □

Theorem. Let (M, d) be a metric space, and define $C_b(M)$ to be the set of functions f in $\ell_\infty(M)$ such that f is continuous. This is a subspace of $\ell_\infty(M)$ in the uniform metric D. $C_b(M)$ is complete in the uniform metric.

Proof. By the above proposition, it is sufficient to show that $C_b(M)$ is closed in $\ell_\infty(M)$. Let (f_n) be a sequence in $C_b(M)$ and we assume that $f_n \to f$ in $\ell_\infty(M)$. We want to show that $f_n \in C_b(M)$. It is now sufficient to show that f is continuous, or equivalently, continuous at every point in M. Let $a \in M$, and let $\varepsilon > 0$. Since $f_n \to f$ in $\ell_\infty(M)$, we can fix $n \in \mathbb{N}$ such that $F(f_n, f) < \varepsilon$. Since f_n is continuous (at a),

$$\exists \delta > 0, \forall x \in M, d(x, a) < \delta \implies |f_n(x) - f_n(a)| < \varepsilon$$

Hence, $\forall x \in M$, if $d(x, a) < \delta$ we have

$$\begin{aligned} |f(x) - f(a)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \\ &\leq 2D(f_n, f) + |f_n(x) - f_n(a)| \\ &< 3\varepsilon \end{aligned}$$

□

Corollary. Consider $C[a, b]$, the space of continuous functions on $[a, b]$. This space is complete in the uniform metric, since $C[a, b] = C_b[a, b]$.

Definition. Let S be a set, and (N, e) be a metric space. Then we generalise $\ell_\infty(S)$ to the following definition.

$$\ell_\infty(S, N) = \{f : S \to N : f \text{ is bounded}\}$$

where f is bounded if there exists $y \in N, r > 0$ such that $\forall x \in S, f(x) \in \mathcal{B}_r(y)$. If $g : S \to N$ is a bounded function, $\forall x \in S, g(x) \in \mathcal{B}_s(z)$, then

$$\forall x \in S, e(f(x), g(x)) \leq e(f(x), y) + e(y, z) + e(z, g(x)) \leq r + e(y, z) + s$$

This is a uniform bound for all x, so we may take the supremum. So $\sup_{x \in S} e(f(x), g(x))$ exists and we denote this by

$$\mathcal{D}(f, g) = \sup_{x \in S} e(f(x), g(x))$$

This can be shown to be a metric, called the uniform metric on $\ell_\infty(S,N)$. Now, let $S = M$, where (M,d) is a metric space. We define

$$C_b(M,N) = \{f : M \to N : f \text{ continuous and bounded}\}$$

Note that $C_b(M,N)$ is a subspace of $\ell_\infty(M,N)$ with the uniform metric.

Theorem. Let S be a set, let (M,d) be a metric space, and let (N,e) be a complete metric space. Then
(i) $\ell_\infty(S,N)$ is complete in the uniform metric D;
(ii) $C_b(M,N)$ is complete in the uniform metric D.

Proof. We first prove part (i). Let (f_n) be a Cauchy sequence in $\ell_\infty(S,N)$. We first show that (f_n) is pointwise Cauchy. Let $x \in S$.

$$\forall \varepsilon > 0, \exists K \in \mathbb{N}, \forall i,j \geq K, D(f_i,f_j) < \varepsilon$$

In particular, $e(f_i(x),f_j(x)) \leq D(f_i,f_j) < \varepsilon$ for $i,j \geq K$. So the sequence $(f_k(x))_k$ is Cauchy in N. Since N is complete, $(f_k(x))_k$ converges. This holds for all $x \in S$, hence we can define $f : S \to N$ by $f(x) = \lim_{k\to\infty} f_k(x)$.

Now, we must show that f is bounded, such that $f \in \ell_\infty(S,N)$. Since f_k is Cauchy in the uniform metric D, there exists $K \in \mathbb{N}$ such that $\forall i,j \geq K, D(f_i,f_j) < 1$. In particular, for all $i \geq K, D(f_i,f_K) < 1$. Since f_K is bounded, there exists $y \in N, r > 0$ such that $\forall x \in S, f_K(x) \in \mathcal{B}_r(y)$. Then, by the triangle inequality, for a fixed $x \in S, \forall i \geq K, e(f_i(x),f_K(x)) \leq D(f_i(x),f_K(x)) < 1$. Let $i \to \infty$, then $e(f_i(x),f_K(x)) \leq 1$. Hence $e(f(x),y) \leq e(f(x),f_K(x)) + e(f_K(x),y) \leq 1 + r$. But since this is true for all x, $1 + r$ is a uniform bound; $\forall x \in S, f(x) \in \mathcal{B}_{r+1}(y)$.

Now we will show that $f_k \to f$ uniformly in D. Again, we use

$$\forall \varepsilon > 0, \exists K \in \mathbb{N}, \forall i,j \geq K, D(f_i,f_j) < \varepsilon$$

So choose $i \geq K$, and $x \in S$. Then for all $j \geq K$, $e(f_i(x),f_j(x)) \leq D(f_i,f_j) < \varepsilon$. As $j \to \infty$, $e(f(x),f_i(x)) \leq \varepsilon$, because metrics are continuous. But since x was arbitrary, we have a uniform distance $D(f,f_i) < \varepsilon$. This holds for all $i \geq K$, so we have uniform convergence.

Now we prove part (ii). By part (i) and an above proposition, it is enough to show that $C_b(M,N)$ is closed in $\ell_\infty(M,N)$. Let (f_k) be a sequence in $C_b(M,N)$ and $f_k \to f$ in $\ell_\infty(M,N)$. We require $f \in C_b(M,N)$, so it is enough to show that f is continuous. This is exactly the proof that the uniform limit of continuous functions is continuous. Let $a \in M$, $\varepsilon > 0$. Then, since $f_k \to f$ in $\ell_\infty(M,N)$, we can fix $k \in \mathbb{N}$ such that $D(f_k,f) < \varepsilon$. Since f_k is continuous, $\exists \delta > 0, \forall x \in M, d(x,a) < \delta \implies e(f_k(x),f_k(a)) < \varepsilon$.

$$\begin{aligned}\forall x \in M, f(x,a) < \delta \implies e(f(x),f(a)) &\leq e(f(x),f_k(x)) + e(f_k(x),f_k(a)) + e(f_k(a),f(a)) \\ &\leq 3\varepsilon\end{aligned}$$

□

6 Contraction mapping theorem

6.1 Contraction mappings

Definition. A function $f : M \to M'$ is called a *contraction mapping* if $\exists \lambda, 0 \leq \lambda < 1$ such that

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

so f is λ-Lipschitz.

6.2 Contraction mapping theorem

This theorem is also called Banach's fixed point theorem.

Theorem. Let M be a non-empty complete metric space. Let $f : M \to M$ be a contraction mapping. Then f has a unique fixed point:

$$\exists! z \in M, f(z) = z$$

Proof. Let λ such that $0 \leq \lambda < 1$ and $\forall x, y \in M, d'(f(x), f(y)) \leq \lambda d(x, y)$. First we show uniqueness. Suppose there were two fixed points $f(z) = z, f(w) = w$. Then $d(z, w) = d(f(z), f(w)) \leq \lambda d(z, w) < d(z, w)$. Hence $d(z, w) = 0$ so $z = w$.

Now we show existence. Fix a starting point $x_0 \in M$. Let $x_n = f(x_{n-1})$ for all $n \in \mathbb{N}$, so $x_n = f^n(x_0)$. First, observe that for all $n \in \mathbb{N}$,

$$d(x_n, x_{n+1}) = d(f(x_{n-1}), f(x_n)) \leq \lambda d(x_{n-1}, x_n) \leq \dots \leq \lambda^n d(x_0, x_1)$$

For $m \geq n$, we have

$$d(x_n, x_m) \leq \sum_{k=n}^{m-1} d(x_k, x_{k+1}) \leq \sum_{k=n}^{m-1} \lambda^k d(x_0, x_1) \leq \frac{\lambda^n}{1 - \lambda} d(x_0, x_1)$$

Since $\frac{\lambda^n}{1-\lambda} d(x_0, x_1) \to 0$,

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \frac{\lambda^n}{1 - \lambda} d(x_0, x_1) < \varepsilon$$

Hence, $\forall m \geq n \geq N, d(x_n, x_m) < \varepsilon$. So the sequence (x_n) is Cauchy. Since M is complete, (x_n) is convergent to some point $z \in M$. f is continuous since it is a contraction, so $f(x_n) \to z$ so $f(z) = z$. So the fixed point exists. $\square$

Remark. Letting $m \to \infty$ in the inequality for $d(x_n, x_m)$, $d(x_n, z) \leq \frac{\lambda^n}{1-\lambda} d(x_0, x_1)$. So $x_n \to z$ exponentially fast. Consider $f : \mathbb{R} \setminus \{0\} \to \mathbb{R} \setminus \{0\}$, and $x \mapsto \frac{x}{2}$. This is a contraction, but there is no fixed point. This is because $\mathbb{R} \setminus \{0\}$ is not complete. Consider instead $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x + 1$. This has no fixed point, since f is an isometry ($\lambda = 1$) and not a contraction. Consider further $f : [1, \infty) \to [1, \infty)$ mapping $x \mapsto x + \frac{1}{x}$. Certainly $|f(x) - f(y)| < |x - y|$. $[1, \infty)$ is closed in $\mathbb{R}$ so it is complete. However this is not a contraction; even though $|f(x) - f(y)| < |x - y|$, there is no upper bound λ. There are no fixed points.

6.3 Application of contraction mapping theorem

Let $y_0 \in \mathbb{R}$. Then the initial value problem $f'(t) = f(t^2)$ and $f(0) = y_0$ has a unique solution on $\left[0, \frac{1}{2}\right]$. In other words, there exists a unique differentiable function $f : \left[0, \frac{1}{2}\right] \to \left[0, \frac{1}{2}\right]$ such that $f(0) = y_0$ and $f'(t) = f(t^2)$ for all t in the domain.

First, observe that if f is a solution then certainly it is continuous, so $f \in C\left[0, \frac{1}{2}\right]$. Further, by the fundamental theorem of calculus, it satisfies

$$f(t) = y_0 + \int_0^t f(s^2)\,\mathrm{d}s$$

Note that $f'(s) = f(s^2)$ is continuous. Conversely, if $f \in C\left[0, \frac{1}{2}\right]$ and $f(t) = y_0 + \int_0^t f(s^2)\,\mathrm{d}s$ then f is a solution to the initial value problem.

Let $M = C\left[0, \frac{1}{2}\right]$ with the uniform metric. This is non-empty and complete. Then we define the map $T : M \to M$ by

$$(Tg)(t) = y_0 + \int_0^t g(s^2)\,\mathrm{d}s$$

Note that Tg is well-defined since $g(s^2)$ is continuous. Moreover, by the fundamental theorem of calculus, Tg is differentiable and $(Tg)'(t) = g(t^2)$. Thus, f is a solution to the initial value problem if and only if $f \in M$ and $Tf = f$.

Now, if T is a contraction, we can use the contraction mapping theorem to assert that there is exactly one fixed point. For $g, h \in M$, $t \in \left[0, \frac{1}{2}\right]$, consider

$$|(Tg)(t) - (Th)(t)| = \left|\int_0^t \left[g(s^2) - h(s^2)\right]\mathrm{d}s\right| \leq t \sup_{s \in \left[0, \frac{1}{2}\right]} |g(s^2) - h(s^2)| \leq \frac{1}{2} D(g, h)$$

Taking the supremum over t gives $D(Tg, Th) \leq \frac{1}{2} D(g, h)$, and so there is exactly one fixed point.

Remark. The above shows that for any $\delta \in (0, 1)$ there is a unique solution to the initial value problem on $[0, \delta]$, called f_δ, since $\delta < 1$ is required for the map to be a contraction. For $0 < \delta < \mu < 1$, $f_\mu|_{[0,\delta]} = f_\delta$ by uniqueness. So we can combine the solutions together to yield a unique solution on $[0, 1)$.

6.4 Lindelöf–Picard theorem

Theorem. Let $n \in \mathbb{N}$, $y_0 \in \mathbb{R}^n$, and $a, b, R \in \mathbb{R}$, such that $a < b$ and $R > 0$. Let $\phi\colon [a, b] \times \mathcal{B}_R(y_0) \to \mathbb{R}^n$ be a continuous function. Given that there exists $K > 0$ such that $\forall t \in [a, b], \forall x, y \in \mathcal{B}_R(y_0)$, such that

$$\|\phi(t, x) - \phi(t, y)\| \leq K\|x - y\|$$

Then, $\exists \varepsilon > 0$ such that $\forall t, t_0 \in [a, b]$, the initial value problem

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

has a unique solution on $[c, d] = [t_0 - \varepsilon, t_0 + \varepsilon] \cap [a, b]$.

Remark. If f is a solution of the initial value problem, implicitly this includes the assumption that $f(t) \in B_r(y_0)$ for all $t \in [c,d]$. Note that if $f : [c,d] \to \mathbb{R}^n$, we let $f_k : [c,d] \to \mathbb{R}$ be the kth component of f, and $f_k = q_k \circ f$ where q_k is the kth coordinate projection. Then, $f(t) = (f_1(t), \dots, f_n(t))$ and we define f to be differentiable if and only if all of the components are differentiable, with $f'(t) = (f_1'(t), \dots, f_n'(t))$. Note further, if f is continuous, then so are f_k, hence f_k are integrable. So we define

$$\int_c^d f(t)\,\mathrm{d}t = v = \left(\int_c^d f_1(t)\,\mathrm{d}t, \dots, \int_c^d f_n(t)\,\mathrm{d}t\right)$$

Note that we can use the Cauchy–Schwarz inequality to give

$$\begin{aligned}
\|v\|^2 &= \sum_{k=1}^n v_k^2 \\
&= \sum_{k=1}^n v_k \int_c^d f_k(t)\,\mathrm{d}t \\
&= \int_c^d \sum_{k=1}^n v_k f_k(t)\,\mathrm{d}t \\
&= \int_c^d v \cdot f(t)\,\mathrm{d}t \\
&\leq \int_c^d \|v\| \cdot \|f(t)\|\,\mathrm{d}t \\
&= \|v\| \int_c^d \|f(t)\|\,\mathrm{d}t
\end{aligned}$$

Hence,

$$\left\| \int_c^d f(t)\,\mathrm{d}t \right\| \leq \int_c^d \|f(t)\|\,\mathrm{d}t \leq (d-c) \sup_{t\in[c,d]} \|f(t)\|$$

Proof. Recall that closed balls are closed, hence $\mathcal{B}_R(y_0)$ is a closed subset of $\mathbb{R}^n$. So ϕ is a continuous function on the closed and bounded set $[a,b] \times \mathcal{B}_R(y_0)$. It follows that ϕ is bounded. Now, let $c = \sup\{\|\phi(t,x)\| : t \in [a,b], x \in \mathcal{B}_R(y_0)\}$. Let $\varepsilon = \min(\frac{R}{c}, \frac{1}{2K})$. Let $t_0 \in [a,b]$ and let $[c,d] = [t_0-\varepsilon, t_0+\varepsilon] \cap [a,b]$. We need to show that there exists a unique differentiable function $f : [c,d] \to \mathbb{R}^n$ such that $f(t_0) = y_0$ and $f'(t) = \phi(t, f(t))$ for all $t \in [c,d]$. Since $\mathcal{B}_R(y_0)$ is closed in $\mathbb{R}^n$, and since $\mathbb{R}^n$ is complete, $\mathcal{B}_R(y_0)$ is complete. Then, $M = C([c,d], \mathcal{B}_R(y_0))$ is complete in the uniform metric D. This is certainly non-empty; consider the constant function yielding y_0. f is a solution to the initial value problem if $f \in M$ and $f'(t) = y_0 + \int_{t_0}^t \phi(s, f(s))\,\mathrm{d}s$, from the fundamental theorem of calculus applied coordinatewise. We define $T : M \to M$ mapping $g \mapsto Tg$ where Tg is given by

$$(Tg)(t) = y_0 + \int_{t_0}^t \phi(s, g(s))\,\mathrm{d}s$$

We must show T is well defined. First, note that the integral is well defined; $s \mapsto \phi(s, g(s))$ is continuous so integrable. By the fundamental theorem of calculus, Tg is differentiable and the derivative is $(Tg)'(t) = \phi(t, g(t))$. In particular, $Tg : [c,d] \to \mathbb{R}^n$ is continuous. Finally, for $t \in [c,d]$,

$$\|(Tg)(t) - y_0\| = \left\| \int_{t_0}^t \phi(s, g(s))\,\mathrm{d}s \right\| \leq |t - t_0| \sup_{s\in[c,d]} \|\phi(s, g(s))\| \leq \varepsilon c \leq R$$

So $Tg \in M$. Recall that f is a solution of the initial value problem if and only if $f \in M$ and $Tf = f$. Now we must show that T has a unique fixed point, so we will show that T is a contraction. Let $t \in [c,d]$ and $g, h \in M$.

$$\|(Tg)(t) - (Th)(t)\| = \left\| \int_{t_0}^{t} [\phi(s, g(s)) - \phi(s, h(s))] \, \mathrm{d}s \right\|$$

Note that $\|\phi(s, g(s)) - \phi(s, h(s))\| \leq K\|g(s) - h(s)\| \leq K \cdot D(g,h)$.

$$\|(Tg)(t) - (Th)(t)\| = |t - t_0| \cdot K \cdot K(g,h) \leq \varepsilon K D(g,h)$$

Taking the supremum over $t \in (c,d)$,

$$D(Tg, Th) \leq \varepsilon K D(g,h) \leq \frac{1}{2} D(g,h)$$

So T is a contraction. By the contraction mapping theorem, T has a unique fixed point in M. □

Remark. For any $\delta \in (0,1)$, taking $\varepsilon = \min(\frac{R}{c}, \frac{\delta}{K})$ works. But by the uniqueness of the solution, the choice does not matter for constructing the solution. So we can construct the solution for $\varepsilon = \min(\frac{R}{c}, \frac{1}{K})$, on $(t_0 - \varepsilon, t_0 + \varepsilon) \cap [a,b]$. In general, there is no solution on $[a,b]$. Finally, note that the above theorem can handle any nth order ODE for any $n \in \mathbb{N}$.

7 Topology

7.1 Definitions

Definition. Let X be a set. A *topology* on X is a family τ of subsets of X (so $\tau \subset \mathcal{P}(X)$) such that
(i) $\varnothing, X \in \tau$;
(ii) if $U_i \in \tau$ for all $i \in I$ where I is some index set, then $\bigcup_{i \in I} U_i \in \tau$; and
(iii) if $U, V \in \tau$ then $U \cap V \in \tau$.
A *topological space* is a pair (X, τ) where X is a set and τ is a topology on X. Members of τ are called *open sets* in the topology. So we say that $U \subset X$ is *open in* X, or U is τ*-open*, if $U \in \tau$.

Remark. If $U_i \in \tau$ for $i = 1, \dots, n$, then $\bigcap_{i=1}^{n} U_i \in \tau$.

Example. Let (M,d) be a metric space. Recall that $U \subset M$ is open in the metric sense if $\forall x \in U, \exists r > 0, \mathcal{B}_r(x) \subset U$. We may say that U is d*-open*. We have already proven that the family of d-open sets is a topology on M. This is a metric topology.

Definition. Let (X, τ) be a topological space. Then we say that X is *metrisable* (or sometimes we say τ is metrisable) if there exists a metric d on X such that τ is the metric topology on X induced by d. In other words, $U \subset X$ is τ-open if and only if U is d-open. If $d' \sim d$, then d' also induces the same topology τ on X.

Example. The indiscrete topology on a set X is a topology $\tau = \{\varnothing, X\}$. If $|X| \geq 2$, then this is not metrisable. Let d be a metric on X. Then let $x \neq y \in X$, let $r = d(x,y)$, and finally let $U = \mathcal{D}_r(x)$. We know that U is d-open. But since $x \in U, y \notin U$, $U \notin \tau$.

Definition. If τ_1, τ_2 are topologies on X, we say that τ_1 is *coarser* than τ_2, or that τ_2 is finer than τ_1, if $\tau_1 \subset \tau_2$. For example, the indiscrete topology on X is the coarsest topology on X.

Example. The discrete topology on a set X is $\tau = \mathcal{P}(X)$. This is the finest topology on X. This is metrisable by the discrete metric.

Definition. A topological space X is *Hausdorff* if $\forall x \neq y$ in X, there exist open sets U, V in X such that $x \in U, y \in V, U \cap V = \varnothing$. Informally, x, y are 'separated by open sets'.

Proposition. Metric spaces are Hausdorff.

Proof. Let $x \neq y$ be points in a metric space (M, d). Let $r > 0$ such that $2r < d(x, y)$. Then let $U = \mathcal{D}_r(x)$, let $V = \mathcal{D}_r(y)$. Certainly U, V are open since they are open balls, and they have no intersection by the triangle inequality, so the metric space is Hausdorff as required. □

Example. The cofinite topology on a set X is

$$\tau = \{\varnothing\} \cup \{U \in X : U \text{ is cofinite in } X\}$$

where U is cofinite in X if $X \setminus U$ is finite. When X is finite, this topology τ is simply $\mathcal{P}(X)$. When X is infinite, τ is not metrisable. Let $x \neq y$ in X, and let $x \in U, y \in V$ where U, V are open in X. Then U and V are cofinite, and hence $U \cap V \neq \varnothing$. So this topology on an infinite set is not Hausdorff and hence not metrisable.

7.2 Closed subsets

Definition. A subset A of a topological space (X, τ) is said to be *closed* in X if $X \setminus A$ is open in X.

Remark. In a metric space, this agrees with the earlier definition of a closed subset, as proven before.

Proposition. The collection of closed sets in a topological space X satisfy

(i) $\varnothing, X$ are closed;
(ii) If A_i are closed in X for i in some non-empty index set I, then $\bigcap_{i \in I} A_i$ is closed;
(iii) If A_1, A_2 are closed in X then $A_1 \cup A_2$ is closed.

Example. In a discrete topological space, every set is closed.

Example. In the cofinite topology, a subset is closed if and only if it is finite or the full set.

7.3 Neighbourhoods

Definition. Let X be a topological space, and let $U \subset X$ and $x \in X$. We say that U is a *neighbourhood* of x in X if there exists an open set V in X such that $X \in V \subset U$.

Remark. In a metric space, we defined this in terms of open balls not open sets. However, we have already proven that the definitions agree.

Proposition. Let U be a subset of a topological space X. Then U is open if and only if U is a neighbourhood of x for every $x \in U$.

Proof. If U is open, and $x \in U$, then by letting $V = U$, V is open and $x \in V \subset U$. Conversely, if $x \in U$, there exists V_x in X such that $x \in V_x \subset U$. Then, $U = \bigcup_{x \in U} x = \bigcup_{x \in U} V_x$ is open, since each V_x is open. □

7.4 Convergence

Definition. Let (x_n) be a sequence in a topological space X. Let $x \in X$. We say that (x_n) converges to x if for all neighbourhoods U of x in X, there exists $N \in \mathbb{N}$ such that $\forall n \geq N, x_n \in U$. Equivalently, for all open sets U which contain x, there exists $N \in \mathbb{N}$ such that $\forall n \geq N, x_n \in U$.

Remark. Again, the definition in a metric space agrees with this definition.

Example. Eventually constant sequences converge. If $\exists z \in X, \exists N \in \mathbb{N}, \forall n \geq N, x_n = z$, then $x_n \to z$.

Example. In an indiscrete topological space, every sequence converges to every point.

Example. In the cofinite topology on a set X, let $x_n \to X$. Suppose that $x_n \to x$ in X. Then if $y \neq x$, $X \setminus \{y\}$ is a neighbourhood of x. Then $N_y = \{n \in N : x_n = y\}$ is finite.

Conversely, suppose (x_n) is a sequence such that for some $x \in X$ and for all $y \neq x$, N_y is finite. Then $x_n \to x$.

In particular, if N_y is finite for all $y \in X$, the sequence converges to every point.

Proposition. If $x_n \to x$ and $x_n \to y$ in a Hausdorff space, then $x = y$.

Proof. Suppose $x \neq y$, then we can choose open sets U, V such that $x \in U, y \in V, U \cap V = \varnothing$. Since $x_n \to x$, there exists $N_1 \in \mathbb{N}$ such that $\forall n \geq N_1, x_n \in U$. Similarly there exists an analogous N_2. Hence $\forall n \geq \max(N_1, N_2), x_n \in U, x_n \in V$ which is a contradiction since $U \cap V = \varnothing$. □

Remark. If $x_n \to x$ in a Hausdorff space, we write $x = \lim_{n\to\infty} x_n$ since the limit is unique.

Remark. In a metric space, for a subset A, we say that A is closed if and only if $x_n \to x$ in A implies $x \in A$. In a general topological space, any closed set is closed under limits, but not every subset that is closed under limits is closed.

7.5 Interiors and closures

Definition. Let X be a topological space, and $A \subset X$. We define the *interior* of A in X, denoted A° or $\operatorname{int}(A)$, by

$$A^\circ = \bigcup \{U \subset X : U \text{ is open in } X, U \subset A\}$$

Similarly we define the *closure* of A in X, denoted $\overline{A}$ or $\operatorname{cl}(A)$, by

$$\overline{A} = \bigcap \{F \subset X : F \text{ is closed in } X, F \supset A\}$$

Remark. Note that A° is open in X, and $A^\circ \subset A$. In particular, if U is open in X and $U \subset A$, then $U \subset A^\circ$. Hence, A° is the largest open subset of A.

Similarly, $\overline{A}$ is closed in X, and $\overline{A} \supset A$. The intersection is not empty since X is closed and $X \supset A$, so it is well-defined. We have that $\overline{A}$ is the smallest closed superset of A.

Proposition. Let X be a topological space and let $A \subset X$. Then the interior is exactly those $x \in X$ for which A is a neighbourhood of x. Similarly, the closure is those $x \in X$ such that for all neighbourhoods U of x, $U \cap A \neq \varnothing$.

Proof. If A is a neighbourhood of X, then by definition there exists an open set U such that $x \in U \subset A$, which is true if and only if $x \in A^\circ$.

For the other part, suppose $x \notin \overline{A}$. Then there exists a closed set $F \supset A$ such that $x \notin F$. Let $U = X \setminus F$. Then U is open and $x \in U$. So U is a neighbourhood of x, and $U \cap A = \varnothing$.

Conversely, suppose there exists a neighbourhood U of x such that $U \cap A = \varnothing$. Then there exists an open set V such that $x \in V \subset U$. Since $V \subset U$, $V \cap A = \varnothing$. Let $F = X \setminus V$. Then F is closed, and $A \subset F$. Hence $\overline{A} \subset F$. So $x \notin \overline{A}$. □

Example. In $\mathbb{R}$, let $A = [0,1) \cup \{2\}$. Then $A^\circ = (0,1)$, and $\overline{A} = [0,1] \cup \{2\}$. Further, $\mathbb{Q}^\circ = \varnothing$ and $\overline{\mathbb{Q}} = \mathbb{R}$. Finally, $\mathbb{Z}^\circ = \varnothing$ and $\overline{\mathbb{Z}} = \mathbb{Z}$.

Remark. In a metric space, for a subset A we have that $x \in \overline{A}$ if and only if there exists a sequence (x_n) in A such that $x_n \to x$. In a general topological space, the existence of a sequence implies $x \in \overline{A}$ but the converse is not true.

7.6 Dense subsets

Definition. A subset A of a topological space X is said to be *dense* in X if $\overline{A} = X$. X is *separable* if there exists a countable subset $A \subset X$ such that A is dense in X.

Example. $\mathbb{R}$ is separable as $\mathbb{Q}$ is dense in $\mathbb{R}$. $\mathbb{R}^n$ is separable in the same way as $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$.

Example. An uncountable discrete topological space is not separable, since the closure of any set is itself.

7.7 Subspaces

Definition. Let (X, τ) be a topological space. Let $Y \subset X$. Then the *subspace topology*, or *relative topology* on Y induced by τ is the topology

$$\{V \cap Y : V \in \tau\}$$

on Y. This is the intersection of Y with all open sets in X. We can denote this $\tau|_Y$. So, for $U \subset Y$, U is open in Y if and only if there exists an open set V in X with $U = V \cap Y$.

Example. Let $X = \mathbb{R}, Y = [0, 2]$, and $U = (1, 2]$. Then certainly $U \subset Y \subset X$. U is open in Y, since $V = (1, 3)$ is open in X and $U = V \cap Y$. However, U is not open in X, since no neighbourhood (or ball) around 2 can be constructed in X that is contained within U.

Remark. On a subset of a topological space, this is considered the standard topology. Suppose that (X, τ) is a topological space, and $Z \subset Y \subset X$. There are two natural topologies on Z: $\tau|_Z$ and $\tau|_Y\,|_Z$. One can easily check that these two topologies are equal.

Let (M, d) be a metric space, and $N \subset M$. Again, there are two natural topologies on N: $\tau(d)|_N$ and $\tau\big(\,d|_N\big)$, where $\tau(e)$ is the metric topology induced by the metric e. These two constructions coincide; indeed, for any $x \in N, r > 0$,

$$\{y \in N : d(y, x) < r\} = \{y \in M : d(y, x) < r\} \cap N$$

Proposition. Let X be a topological space, and let $A \subset Y \subset X$. A is closed in Y if and only if there exists a closed subset $B \subset X$ such that $A = B \cap Y$. Further,

$$\mathrm{cl}_Y(A) = \mathrm{cl}_X(A) \cap Y$$

This is not true for the interior of a subset in general. For instance, consider $X = \mathbb{R}, A = Y = \{0\}$. In this case, $\mathrm{int}_Y(A) = A, \mathrm{int}_X(A) = \varnothing$.

Proof. The first part is true by taking complements: $Y \setminus A$ is open in Y. By definition, $Y \setminus A = V \cap Y$ for some open V in X. So $B = X \setminus V$ is closed in X and $A = B \cap Y$. If $A = B \cap Y$, B is closed in X, then $X \setminus B$ is open in X, and hence $Y \setminus A = (X \setminus B) \cap Y$ is open in Y.

For the second part, we know $\mathrm{cl}_X(A)$ is closed in X, so by the first part, $\mathrm{cl}_X(A) \cap Y$ is closed in Y. Then $A \subset \mathrm{cl}_X(A) \cap Y$. So by definition, $\mathrm{cl}_Y(A) \subset \mathrm{cl}_X(A) \cap Y$. Similarly, since $\mathrm{cl}_Y(A)$ is closed in Y, we can write $\mathrm{cl}_Y(A) = B \cap Y$ for some closed set B in X. But $A \subset B$, and B is closed in X, so $\mathrm{cl}_X(A) \subset B$ and hence $\mathrm{cl}_Y(A) = B \cap Y \supset \mathrm{cl}_X(A) \cap Y$. □

Remark. If $U \subset Y \subset X$, and Y is open in X, then U is open in Y if and only if U is open in X.

7.8 Continuity

Definition. A function $f : X \to Y$ between topological spaces is said to be continuous if for all open sets V in Y, the preimage $f^{-1}(V)$ is open in X.

Remark. We have already proven that this agrees with the definition of continuity of functions between metric spaces.

Example. Constant functions are always continuous. Consider $f : X \to Y$ defined by $f(x) = y_0$ for a fixed $y_0 \in Y$. For any $V \subset Y$, $f^{-1}(V) = \varnothing$ if $y_0 \notin V$, and $f^{-1}(V) = X$ if $y_0 \in V$. So f is continuous.

Example. The identity map is always continuous. If $f : X \to X$ is defined by $x \mapsto x$, $f^{-1}(V) = V$ so if V is open, $f^{-1}(V)$ is trivially open.

Example. Let $Y \subset X$. Let $i : Y \to X$ be the inclusion map. Then for an open set V in X, $i^{-1}(V) = V \cap Y$ which by definition is open in Y. Hence, if $g : X \to Z$ is continuous, then $g|_Y = g \circ i : X \to Y$ is continuous, as we will see below.

Proposition. Let $f : X \to Y$ be a function between topological spaces. Then,
(i) f is continuous if and only if for all closed sets B in Y, $f^{-1}(B)$ is closed in X;
(ii) if f is continuous and $g : Y \to Z$ is continuous, then $g \circ f$ is continuous.

Proof. To prove (i), note that for any subset $D \subset Y$, $f^{-1}(Y \setminus D) = X \setminus f^{-1}(D)$. We can now use the fact that $A \subset X$ is open in X if and only if $X \setminus A$ is closed in X, and vice versa for Y.

To prove (ii), note that if W is an open subset of Z, then $g^{-1}(W)$ is open in Y since g is continuous. Hence $f^{-1}g^{-1}(W)$ is open in X since f is continuous. But then $f^{-1}g^{-1} = (g \circ f)^{-1}$, so $g \circ f$ is continuous. $\square$

Remark. There exists a notion of 'continuity at a point' for topological spaces, but it is not as useful in this course as the global continuity definition.

7.9 Homeomorphisms and topological invariance

Definition. A function $f : X \to Y$ between topological spaces is a homeomorphism if f is a bijection, and both f, f^{-1} are continuous. If such an f exists, we say that X and Y are homeomorphic. This is exactly the definition from metric spaces.

Definition. A property $\mathcal{P}$ of topological spaces is said to be a *topological property* or *topological invariant* if, for all pairs X, Y of homeomorphic spaces, X satisfies $\mathcal{P}$ if and only if Y satisfies $\mathcal{P}$.

Example. Metrisability is a topological invariant. Being Hausdorff is a topological invariant. Being completely metrisable (metrisable into a complete metric space) is *not* a topological invariant. For example, consider metrics d, d' on $\mathbb{R}$ such that $d \sim d'$ but d is complete and d' is not.

Remark. If $f : X \to Y$ is a homeomorphism, for an open set U in X, $f(U) = (f^{-1})^{-1}(U)$ is open in Y since $f^{-1} : Y \to X$ is continuous.

Definition. A function $f : X \to Y$ between topological spaces is an *open map* if for all open sets U in X, $f(U)$ is open in Y.

Remark. $f : X \to Y$ is a homeomorphism if and only if f is a continuous and open bijection.

7.10 Products

Let X, Y be topological spaces. We want to define the topology on $X \times Y$. If U is open in X and V is open in Y, then we would like $U \times V$ to be open in $X \times Y$. Certainly $\varnothing = \varnothing \times \varnothing$ and $X \times Y$ should be open. Further $(U \times V) \cap (U' \times V') = (U \cap U') \times (V \cap V')$, so intersections work. $\bigcup_{i \in I} U_i \times V_i$ must be open for open sets U_i, V_i, but this is not obvious from what we have shown so far, so we must include this in our definition.

Definition. The *product topology* on $X \times Y$ is the topology such that a subset U of $X \times Y$ is open if there exists a set I and open sets U_i, V_i in X, Y for all $i \in I$ such that

$$U = \bigcup_{i \in I} U_i \times V_i$$

Remark. For $W \subset X \times Y$, we know that W is open if and only if for all $z \in W$, there exist open sets $U \subset X, V \subset Y$, such that $z \in U \times V \subset W$. So, thinking of the product as a product of real lines, we might say that W is open if for every point $z \in W$, we can construct a 'box set' (the Cartesian product of open intervals) contained in W that has z as an element. More formally, W is a neighbourhood of z if and only if there exist neighbourhoods U of x in X and V of y in Y such that $U \times V \subset W$.

7.11 Continuity in product topology

Example. Let $(M, d), (M', d')$ be metric spaces. Then, the metric d_∞ on $M \times M'$ is

$$d_\infty((x, x'), (y, y')) = \max(d(x, y), d'(x', y'))$$

This metric is chosen since all d_p metrics induce the same metric topology, but this is easier to work with. Also, M, M' are topological spaces with their metric topologies, which induce the product topology on the product space $M \times M'$. These two constructions create the same topology. For a point $z = (x, x') \in M \times M'$ and $r > 0$, the open ball $\mathcal{D}_r(z)$ is exactly

$$\begin{aligned} \mathcal{D}_r(z) &= \{(y, y') \in M \times M' : \ d_\infty((y, y'), (x, x')) < r\} \\ &= \{(y, y') \in M \times M' : \ d(x, y) < r, d(x', y') < r\} \\ &= \mathcal{D}_r(x) \times \mathcal{D}_r(x') \end{aligned}$$

Now, let $W \subset M \times M'$. Then W is open in the product topology if and only if for all $z = (x, x') \in W$, there exist open sets U in M and U' in M' such that $(x, x') \in U \times U' \subset W$. Equivalently, for all $z = (x, x') \in W$, there exists $r > 0$ such that $\mathcal{D}_r(x) \times \mathcal{D}_r(x') \subset W$. But $\mathcal{D}_r(x) \times \mathcal{D}_r(x') = \mathcal{D}_r(z)$, so W is d_∞-open, as required. For instance, the product topology on $\mathbb{R} \times \mathbb{R}$ is the Euclidean topology on $\mathbb{R}^2$.

Proposition. Let X, Y be topological spaces. Let $X \times Y$ be given the product topology. Then, the coordinate projections $q_X : X \times Y \to X$ and $q_Y : X \times Y \to Y$ satisfy

(i) q_X, q_Y are continuous;
(ii) if Z is any topological space, and $g : Z \to X \times Y$ is a function, then g is continuous if and only if $q_X \circ g, q_Y \circ g$ are continuous.

Proof. If U is open in X, then $q_X^{-1}(U) = U \times Y$, which is the product of an open set in X and an open set in Y, so is open in $X \times Y$. Hence q_X is continuous. Similarly, q_Y is continuous.

If g is continuous then certainly $q_X \circ g, q_Y \circ g$ are continuous since the composition of continuous functions are continuous. Conversely, let $h: Z \to X$ and $k: Z \to Y$ be continuous functions with $h = q_X \circ g$ and $k = q_Y \circ g$. Then $g(x) = (h(x), k(x))$ for $x \in Z$. Now, for open sets U in X and V in Y, we have

$$z \in g^{-1}(U \times V) \iff g(z) \in U \times V \iff h(z) \in U, k(z) \in V \iff z \in h^{-1}(U) \cap k^{-1}(V)$$

So $g^{-1}(U \times V) = h^{-1}(U) \cap k^{-1}(V)$ which is open in Z as h, k are continuous. Given an arbitrary open set W in $X \times Y$, we can write $W = \bigcup_{i \in I} U_i \times V_i$, where U_i are open in X and V_i are open in Y. Thus, $g^{-1}(W) = \bigcup_{i \in I} g^{-1}(U_i \times V_i)$ which is open. $\square$

Remark. The product topology may be extended to a finite product $X_1 \times \cdots \times X_n$, consisting of all unions of sets of the form $U_1 \times \cdots \times U_n$ where U_j is open in X_j. Properties of the product topology hold in this more general case. For example, if X_j is metrisable with metric e_j for all j, then the product topology is metrisable with, for instance, the d_∞ metric.

7.12 Quotients

Let X be a set and R an equivalence relation on X. So $R \subset X \times X$, but we will write $x \sim y$ to mean $(x, y) \in R$. For $x \in X$, we define $q(x) = \{y \in X : y \sim x\}$ to be the equivalence class of x, the set of which partition X. Let X/R denote the set of all equivalence classes. The map $q: X \to X/R$ is called the quotient map.

Definition. Let X be a topological space, and R an equivalence relation on X. The *quotient topology* on X/R is given by

$$\tau = \left\{V \subset X/R : q^{-1}(V) \text{ open in } X\right\}$$

This is a topology:
(i) $q^{-1}(\varnothing) = \varnothing$ which is open, and $q^{-1}(X/R) = X$ which is open.
(ii) If V_i are open, then $q^{-1}\left(\bigcup_{i \in I} V_i\right) = \bigcup_{i \in I} q^{-1}(V_i)$ which is a union of open sets which is open.
(iii) If U, V are open, then $q^{-1}(U \cap V) = q^{-1}(U) \cap q^{-1}(V)$ which is open.

Remark. The quotient map $q: X \to X/R$ is continuous. In particular, it is the largest possible topology on X such that q is continuous.

Let $x \in X, t \in X/R$. Then $x \in t$ if and only if $t = q(x)$. For $V \subset X/R$,

$$\begin{aligned} q^{-1}(V) &= \{x \in X : q(x) \in V\} \\ &= \{x \in X : \exists t \in V, t = q(x)\} \\ &= \{x \in X : \exists t \in V, x \in t\} \\ &= \bigcup_{t \in V} t \end{aligned}$$

Example. Consider $\mathbb{R}$, an abelian group under addition, and the subgroup $\mathbb{Z}$. We can form the quotient group $\mathbb{R}/\mathbb{Z}$, which is the set of equivalence classes where $x \sim y \iff x - y \in \mathbb{Z}$. For all $x \in \mathbb{R}$, there exists $y \in [0, 1]$ such that $x \sim y$, and for all $x, y \in [0, 1]$ we have $x \sim y$ if and only if $x = y$ or $\{x, y\} = \{0, 1\}$. So we can think of the quotient topology of $\mathbb{R}/\mathbb{Z}$ as a circle. We can say that $\mathbb{R}/\mathbb{Z}$ is homeomorphic to $S^1 = \left\{(x, y) \in \mathbb{R}^2 : \|(x, y)\| = 1\right\}$, which we will prove later.

Example. Consider the subgroup $\mathbb{Q}$ of $\mathbb{R}$. Let $V \subset \mathbb{R}/\mathbb{Q}$, such that $V \neq \varnothing$ and V is open. Then $q^{-1}(V)$ is open and not empty. Therefore, there exist $a < b \in \mathbb{R}$ such that $(a,b) \subset q^{-1}(V)$. Given $x \in \mathbb{R}$, we can choose a rational r in the interval $(a-x, b-x)$. Then $r + x \in (a,b) \subset q^{-1}(V)$, so $q(x) = q(r+x) \in V$. So $V = \mathbb{R}/\mathbb{Q}$. This is the indiscrete topology, which is not metrisable or Hausdorff. So we cannot (in general) take quotients of metric spaces.

Example. Let $Q = [0,1] \times [0,1] \subset \mathbb{R}^2$. We define the equivalence relation R given by

$$(x_1, x_2) \sim (y_1, y_2) \iff \begin{cases} (x_1, x_2) = (y_1, y_2) & \text{or} \\ x_1 = y_1, \{x_2, y_2\} = \{0,1\} & \text{or} \\ x_2 = y_2, \{x_1, y_1\} = \{0,1\} & \text{or} \\ x_1, x_2, y_1, y_2 \in \{0,1\} \end{cases}$$

The space Q/R is homeomorphic to $\mathbb{R}^2/\mathbb{Z}^2$. This is a square where the top and bottom edges are identified as the same, and the left and right edges are also identified as the same. This is homeomorphic to the surface of a torus with the Euclidean topology embedded in Euclidean three-dimensional space.

Proposition. Let X be a set, and let R be an equivalence relation on X. Let $q : X \to X/R$ be the quotient map. Let Y be a set, and $f : X \to Y$ be a function. Suppose that f 'respects' R; that is, $x \sim y \implies f(x) = f(y)$. Then there exists a unique map $\widetilde{f} : X/R \to Y$ such that $f = \widetilde{f} \circ q$. For $z \in X/R$, we write $z = q(x)$ for some $x \in X$, and then define $\widetilde{f}(z) = f(x)$.

Remark. Note that $\operatorname{Im} f = \operatorname{Im} \widetilde{f}$ since q is surjective. $\widetilde{f}$ is injective if for all $x, y \in X$, $\widetilde{f}(q(x)) = \widetilde{f}(q(y))$ implies $q(x) = q(y)$. In other words, for all $x, y \in X$, $f(x) = f(y) \implies x \sim y$. We say that f *fully respects* R if, for all $x, y \in X$,

$$x \sim y \iff f(x) = f(y)$$

In this case, $\widetilde{f}$ is injective.

7.13 Continuity of functions in quotient spaces

Proposition. Let X be a topological space and let R be an equivalence relation on X. Let $q : X \to X/R$ be a quotient map, where X/R has the quotient topology. Let Y be another topological space and $f : X \to Y$ be a function that respects R. Let $\widetilde{f} : X/R \to Y$ be the unique map such that $f = \widetilde{f} \circ q$. Then

(i) if f is continuous then $\widetilde{f}$ is continuous; and

(ii) if f is an open map (the image of an open set is open) then $\widetilde{f}$ is an open map.

In particular, if f is a continuous surjective map that fully respects R, then $\widetilde{f}$ is a continuous bijection. If in addition f is an open map, then $\widetilde{f}$ is a continuous bijective open map, so is a homeomorphism.

Proof. We prove part (i). Let V be an open set in Y.

$$q^{-1}\big(\widetilde{f}^{-1}(V)\big) = (\widetilde{f} \circ q)^{-1}(V) = f^{-1}(V) \text{ is open}$$

So by definition, $\widetilde{f}^{-1}(V)$ is open in X/R. Hence $\widetilde{f}$ is continuous. Now, we prove part (ii). Let V be an open set in X/R. Let $U = q^{-1}(V)$. Then U is open in X by definition of the quotient topology. Since

q is surjective, $q(U) = q\big(q^{-1}(V)\big) = V$. Hence,

$$\widetilde{f}(V) = \widetilde{f}(q(U)) = (\widetilde{f} \circ q)(U) = f(U) \text{ is open}$$

since f is an open map. □

Example. $\mathbb{R}/\mathbb{Z}$ is homeomorphic to a circle $S^1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$. We define

$$f(t) = (\cos 2\pi t, \sin 2\pi t)$$

Then, $s - t \in \mathbb{Z}$ if and only if $f(s) = f(t)$ so f fully respects the relation, and f is surjective. f is also continuous since each component is continuous. Hence, there exists $\widetilde{f} : \mathbb{R}/\mathbb{Z} \to S^1$ such that $f = \widetilde{f} \circ q$ and $\widetilde{f}$ is a continuous bijection. Now we must show f is an open map, and then $\widetilde{f}$ will be a homeomorphism. Suppose f is not an open map, so there exists an open set U in $\mathbb{R}$ such that $f(U)$ is not open in S^1. So $S^1 \setminus f(U)$ is not closed, so there exists a sequence (z_n) in this complement and $z \in f(U)$ such that $z_n \to z$. f is surjective so for all $n \in N$ we can choose $x_n \in [0,1]$ such that $f(x_n) = z_n$. This is a bounded sequence, so by the Bolzano–Weierstrass theorem, without loss of generality we can let $x_n \to x \in [0,1]$. Since f is continuous, $f(x_n) \to f(x)$, so $z_n \to z$. But since $z_n \notin f(U)$, we have $x_n \in \mathbb{R} \setminus U$. Since the complement is closed and $x_n \to x$, we have $x \in \mathbb{R} \setminus U$ so $x \notin U$. Since $z \in f(U)$, there exists $y \in U$ such that $z = f(y)$. Hence $k = y - x \in \mathbb{Z}$. Now, $f(x_n + k) = f(x_n) = z_n \to z$, but also $x_n + k \to x + k = y \in U$. Since $z_n \notin f(U)$, we have $x_n + k \notin U$. Since $\mathbb{R} \setminus U$ is closed and $x_n + k \to y$, we have $y \in \mathbb{R} \setminus U$ which is a contradiction.

Proposition. Let X be a topological space, and R an equivalence relation on X. Then,
(a) If X/R is Hausdorff, then R is closed in $X \times X$.
(b) If R is closed in $X \times X$ and the quotient map $q : X \to X/R$ is an open map, then X/R is Hausdorff.

Proof. Let $W = X \times X \setminus R$. For part (a), we want to show W is open, so is a neighbourhood of all of its points. Given $(x, y) \in W$, we have $x \nsim y$, so $q(x) \neq q(y)$. Since the quotient is Hausdorff, there exist open sets S, T in X/R such that $S \cap T = \varnothing$ and $q(x) \in S, q(y) \in T$. Let $U = q^{-1}(S), V = q^{-1}(T)$ which are open in X, and $x \in U, y \in V$. For all $(a, b) \in U \times V$, we have $q(a) \in S, q(b) \in T$ hence $a \nsim b$. So $(x, y) \in U \times V \subset W$. Hence R is closed.

For part (b), let $z \neq w$ be elements of X/R, and we want to separate these points by open sets. Let $x, y \in X$ such that $q(x) = z, q(y) = w$. Then $(x, y) \in W$ since $x \nsim w$. Since R is closed, W is open, so there exist open sets U, V in X such that $(x, y) \in U \times V \subset W$. Since q is an open map, $q(U)$ and $q(V)$ are open in X/R, and $z = q(x) \in q(U), w = q(y) \in q(V)$. Now it suffices to show $q(U) \cap q(V) = \varnothing$. For $(a, b) \in U \times V \subset W$, $(a, b) \notin R$ hence $q(a) \neq q(b)$ so $q(U) \cap q(V) = \varnothing$. □

8 Connectedness

8.1 Definition

Recall the intermediate value theorem from IA Analysis. If $f : I \to \mathbb{R}$ is continuous, where I is an interval, and $x < y$ in I and $c \in (f(x), f(y))$, then there exists $z \in (x, y)$ such that $f(z) = c$. An interval in this context is a set I such that for all $x < y < z \in \mathbb{R}, x, z \in I \implies y \in I$. So the intermediate value theorem essentially states that the continuous image of an interval is an interval.

Example. Consider $[0,1) \cup (1,2]$. Let f be a function from this space to $\mathbb{R}$, defined by

$$f(x) = \begin{cases} 0 & x \in [0,1) \\ 1 & x \in (1,2] \end{cases}$$

This is continuous, but the image of f is not an interval.

Definition. A topological space X is *disconnected* if there exist open subsets U, V of X such that $U \cap V = \varnothing, U \cup V = X$ and $U, V \neq \varnothing$. We say that U and V *disconnect* X. We say X is *connected* if X is not disconnected.

Theorem. Let X be a topological space. Then the following are equivalent.

(i) X is connected;
(ii) if $f : X \to \mathbb{R}$ is continuous, then $f(X)$ is an interval;
(iii) if $f : X \to \mathbb{Z}$ is continuous, f is constant.

Proof. First we show (i) implies (ii). Suppose X is connected, and $f : X \to \mathbb{R}$ is continuous, but $f(X)$ is not an interval. Then there exist $a < b < c \in \mathbb{R}$ such that $a, c \in f(X)$ and $b \notin f(X)$. Let $x, y \in X$ such that $f(x) = a, f(y) = c$. Let $U = f^{-1}(-\infty, b), V = f^{-1}(b, \infty)$. U, V are open since f is continuous. U, V are non-empty since $x \in U, y \in V$. Their intersection is empty since we are taking the preimage of disjoint sets. Finally, $U \cup V = f^{-1}(\mathbb{R} \setminus b) = X$ since b is not in the image. So U, V disconnect X, which is a contradiction.

Now (ii) implies (iii). This is immediate since an interval containing an integer must only contain one integer.

Finally, (iii) implies (i). Suppose U, V disconnect X. Let $f : X \to \mathbb{Z}$ by

$$f(x) = \begin{cases} 0 & x \in U \\ 1 & x \in V \end{cases}$$

For any $Y \subset \mathbb{R}$,

$$f^{-1}(Y) = \begin{cases} \varnothing & 0, 1 \notin Y \\ U & 0 \in Y, 1 \notin Y \\ V & 0 \notin Y, 1 \in Y \\ X & 0, 1 \in Y \end{cases}$$

which is open. But f is not constant, so this is a contradiction. □

Corollary. Let $X \subset \mathbb{R}$. Then X is connected if and only if X is an interval.

Proof. Suppose X is connected. Then the inclusion map $i : X \to \mathbb{R}$ is continuous. By the theorem above, $i(X) = X$ is an interval. Conversely, suppose X is an interval. Then, for all continuous $f : X \to \mathbb{R}$, $f(X)$ is an interval by the intermediate value theorem. Then X is connected. □

Proof. This is an alternative, direct proof that intervals are connected. Suppose U, V disconnect X. Then let $x \in U, y \in V$ such that $x < y$. Let $z = \sup U \cap [x, y]$. This set is non-empty since it contains

x and is bounded above by y. So $z = [x, y] \subset X$. We will show $z \in U \cap V$, which is a contradiction. For all $n \in \mathbb{N}$, we have $z - \frac{1}{n} < n$ so there exists $x_n \in U \cap [x, y]$ which satisfies $z - \frac{1}{n} < x_n \leq z$. Hence $x_n \to z$. Also, $U = X \setminus V$ is closed, so $z \in U$. In particular, $z < y$. Now, choose $N \in \mathbb{N}$ such that $z + \frac{1}{N} < y$. Then for all $n \geq N$ we have $z < z + \frac{1}{n} < y$. Hence $z + \frac{1}{n} \in V$. However, $z + \frac{1}{n} \to z$, and V is closed, so $z \in V$, which is a contradiction. □

8.2 Consequences of definition

Example. Any indiscrete topological space is connected. Any cofinite topological space on an infinite set is connected. The discrete topological space on a set of size at least two is disconnected.

Lemma. Let Y be a subspace of a topological space X. Then, Y is disconnected if and only if there exist open subsets U, V of X such that $U \cap V \cap Y = \varnothing$ and $U \cup V \supset Y$, and $U \cap Y \neq \varnothing, V \cap Y \neq \varnothing$.

Proof. Suppose Y is disconnected. Then there exist open subsets U', V' of Y that disconnect Y. Then there exist open sets U, V in X such that $U' = U \cap Y$ and $V' = V \cap Y$. Then U, V satisfy the requirements from the lemma.

Conversely, suppose U, V are as given. Then, let $U' = U \cap Y, V' = V \cap Y$. They are open in Y by the definition of the subspace topology, and they disconnect Y. □

Remark. In the above lemma, we say subsets U, V of X disconnect Y.

Proposition. Let Y be a subspace of a topological space X. If Y is connected, then so is $\overline{Y}$.

Proof. Suppose $\overline{Y}$ is disconnected. Then there exist open sets U, V in X which disconnect $\overline{Y}$. Then $U \cap V \cap Y \subset U \cap V \cap \overline{Y} = \varnothing$ by definition. Hence $U \cap V \cap Y = \varnothing$. Also, $U \cup V \supset \overline{Y} \supset Y$. So U, V disconnect Y unless $U \cap Y = \varnothing$ or $V \cap Y = \varnothing$. But Y is connected, so without loss of generality let $V \cap Y = \varnothing$. Then $Y \subset X \setminus V$ and $X \setminus V$ is closed, so $\overline{Y} \subset X \setminus V$. Hence $V \cap \overline{Y} = \varnothing$. This is a contradiction since U, V disconnect $\overline{Y}$. □

Remark. More generally, if $Y \subset Z \subset \overline{Y}$, and Y is connected, then Z is connected. This is since $\mathrm{cl}_Z(Y) = \mathrm{cl}_X(Y) \cap Z = Z$.

Theorem. Let $f : X \to Y$ be a continuous function between topological spaces. If X is connected, then so is $f(X)$.

Proof. Let U, V be open subsets of Y which disconnect $f(X)$. For $x \in X$, $f(x) \in f(X) \subset U \cup V$. Hence, $f^{-1}(U) \cup f^{-1}(V) = X$. Also, if $x \in f^{-1}(U) \cap f^{-1}(V)$ then $f(x) \in U \cap V \cap f(X) = \varnothing$. This is a contradiction, so $f^{-1}(U) \cap f^{-1}(V) = \varnothing$. Since f is continuous, $f^{-1}(U), f^{-1}(V)$ are open in X. Since $U \cap f(X) \neq \varnothing$ and $V \cap f(X) \neq \varnothing$, $f^{-1}(U) \neq \varnothing$ and $f^{-1}(V) \neq \varnothing$ So $f^{-1}(U), f^{-1}(V)$ disconnect X. □

Remark. This shows that connectedness is a topological property. If X, Y are homeomorphic spaces, then X is connected if and only if Y is connected. Further, note that if $f : X \to Y$ is continuous and

$A \subset X$ and A is connected, then $f(A)$ is connected. This can be shown by restricting f to the domain A.

Corollary. Any quotient of a connected topological space is connected.

Example. Let

$$Y = \left\{ \left(x, \sin \frac{1}{x} \right) : x > 0 \right\} \subset \mathbb{R}^2$$

This space is connected; the function $f : (0, \infty) \to \mathbb{R}^2$ defined by $f(x) = \left(x, \sin \frac{1}{x} \right)$ is continuous. So we have that $Y = \operatorname{Im} f$ is connected. Hence, $\overline{Y}$ is connected. We claim that

$$Z \equiv Y \cup \{(0, y) : y \in [-1, 1]\} = \overline{Y}$$

Indeed, given $y \in [-1, 1]$, for all $n \in \mathbb{N}$ we have that $(0, \frac{1}{n})$ is mapped to (n, ∞) by $x \to \frac{1}{x}$, so by the intermediate value theorem there exists $x_n \in \left(0, \frac{1}{n}\right)$ such that $\sin \frac{1}{x_n} = y$. Hence,

$$\left(x_n, \sin \frac{1}{x_n} \right) = (x_n, y) \to (0, y) \in \overline{Y}$$

So $Y \subset Z \subset \overline{Y}$. If we can show Z is closed, $Z = \overline{Y}$ since $\overline{Y}$ is the smallest closed superset of Y. Suppose $(x_n, y_n) \in Z$ for all $n \in \mathbb{N}$, and $(x_n, y_n) \to (x, y)$ in $\mathbb{R}^2$. Since $y_n \in [-1, 1]$ and $y_n \to y$, we have $y \in [-1, 1]$. If $x = 0$, we have $(x, y) \in Z$. If $x \neq 0$, then $x_n \to x$ implies $x_n \neq 0$ for all sufficiently large n. Hence $y_n = \sin \frac{1}{x_n}$ for all sufficiently large n. Thus

$$(x_n, y_n) \to \left(x, \sin \frac{1}{x} \right) \in Z$$

Lemma. Let X be a topological space and $\mathcal{A}$ be a family of connected subsets of X. Suppose that $A \cap B \neq \varnothing$ for all $A, B \in \mathcal{A}$. Then $\bigcup_{A \in \mathcal{A}} A$ is connected.

Proof. Let $Y = \bigcup_{A \in \mathcal{A}} A$, and let $f : Y \to \mathbb{Z}$ be a continuous function. We must show that f is constant. For all $A \in \mathcal{A}$, $f|_A : A \to \mathbb{Z}$ is continuous and hence constant, since A is connected. For all $A, B \in \mathcal{A}$, $A \cap B \neq \varnothing$ hence $f|_A$ and $f|_B$ are both constant and have the same value. So f must be constant, and hence Y is connected. □

Theorem. Let X, Y be connected topological spaces. Then $X \times Y$ is connected (in the product topology).

Proof. Without loss of generality, let $X \neq \varnothing, Y \neq \varnothing$. Let $x_0 \in X$. Consider the function $f : Y \to X \times Y$ defined by $f(y) = (x_0, y)$. The components of f are the functions $y \mapsto x_0$ which is continuous as it is constant, and $y \mapsto y$ which is continuous as it is the identity. So f is continuous. Then, the image of f, which is $\{x_0\} \times Y$, is connected. Similarly, for all $y \in Y$, $X \times \{y\}$ is connected. For $y \in Y$, $\{x_0\} \times Y \cap X \times \{y\} = \{(x_0, y)\} \neq \varnothing$. Hence, $A_y = \{x_0\} \times Y \cup X \times \{y\}$ is connected. For all $y, z \in Y$, $A_y \cap A_z \supset \{x_0\} \times Y$ hence $A_y \cap A_z \neq \varnothing$. Hence, $\bigcup_{y \in Y} A_y = X \times Y$ is connected. □

Example. $\mathbb{R}^n$ is connected for all $n \in \mathbb{N}$.

8.3 Partitioning into connected components

Definition. Let X be a topological space. We define a relation $\sim$ on X by $x \sim y$ if and only if there exists a connected subset A of X such that $x, y \in A$. For all $x \in X$, $x \sim x$ since $\{x\}$ is connected. Symmetry is clear from the definition. If $x \sim y$ and $y \sim z$ then by definition there exist connected subsets A, B in X such that $x, y \in A$ and $y, z \in B$. In particular, $A \cap B$ is not empty since $y \in A \cap B$. Hence $A \cup B$ is connected. Since $A \cup B$ contains x, z, we have $x \sim z$ as required for transitivity. Hence $\sim$ is an equivalence relation. For $x \in X$, we write C_x for the equivalence class containing x, called the *connected component* of x. The equivalence classes are called *connected components* of X.

Proposition. The connected components of a topological space X are non-empty, maximal connected subsets of X, they are closed, and they partition X.

Proof. Let C be a connected component of X. So $C = C_x$ for some $x \in X$. Then $x \in C$ hence $C \neq \varnothing$. Suppose $C \subset A \subset X$ and A is connected. Then for all $y \in A$, since $x, y \in A$ we must have $x \sim y$. So $y \in C$. Hence $A \subset C$, giving $A = C$. For all $y \in C$, we have $y \sim x$, so there exists a connected subset $A_y \subset X$ such that $x, y \in A_y$. Let $A = \bigcup_{y \in C} A_y$. A is connected since the union of pairwise intersecting connected sets are connected. Further $A \supset C$ so $A = C$ and C is connected. Since the closure of a connected set is connected, $\overline{C}$ is connected. But $\overline{C} \supset C$, so $C = \overline{C}$ is closed. □

8.4 Path-connectedness

Definition. Let X be a topological space. For points $x, y \in X$, a *path* from x to y in X is a continuous function $\gamma\colon [0,1] \to X$ such that $\gamma(0) = x, \gamma(1) = y$. We say that X is *path-connected* if for all $x, y \in X$, there exists a path from x to y in X.

Example. In $\mathbb{R}^n$, $\mathcal{D}_r(x)$ is path-connected by a straight line segment between any two points in the ball. In particular, let $\gamma(t) = (1-t)y + tz$. This is continuous and lies entirely inside $\mathcal{D}_r(x)$, since

$$\begin{aligned}
\|\gamma(t) = x\| &= \|(1-t)t + tz - x\| \\
&= \|((1-t)y + tz) - ((1-t)x + tx)\| \\
&\leq (1-t)\|y - x\| + t\|z - x\| \\
&< r
\end{aligned}$$

In a similar way, any convex subset of $\mathbb{R}^n$ is path-connected.

Theorem. If X is path-connected, X is connected.

Proof. Suppose X is not connected. Let U, V disconnect X. Let $x \in U, y \in V$, and suppose $\gamma\colon [0,1] \to X$ is continuous with $\gamma(0) = x$ and $\gamma(1) = y$. Then $\gamma^{-1}(U)$ and $\gamma^{-1}(V)$ disconnect $[0,1]$, which contradicts the connectedness of $[0,1]$. □

Example. The converse is false in general. Recall that the space

$$X = \left\{\left(x, \sin\frac{1}{x}\right) : x > 0\right\} \cup \{(0, y) : -1 \leq y \leq 1\}$$

is connected. We will show X is not path-connected. Suppose $\gamma : [0,1] \to X$ is continuous, and $\gamma(0) = (0,0)$ and $\gamma(1) = (1, \sin 1)$. Let $\gamma = (\gamma_1, \gamma_2)$, so γ_1, γ_2 are continuous functions. Suppose $t \in [0,1]$ such that $\gamma_1(t) > 0$. Then $\gamma_1((0,t)) \supset (0, \gamma_1(t))$ by the intermediate value theorem. In particular, there exists $n \in \mathbb{N}$ such that $\frac{1}{2\pi n} \in (0, \gamma_1(t))$. Hence, there exists $s < t$ such that $\gamma_1(s) = \frac{1}{2\pi n}$ so $\gamma_1(s) = 0$. Similarly, $\frac{1}{2\pi n + \frac{\pi}{2}} \in (0, \gamma_1(t))$ so there exists a different $s < t$ such that $\gamma_1(s) = \frac{1}{2\pi n + \frac{\pi}{2}}$ hence $\gamma_2(s) = 1$. In both cases, $\gamma_1(s) > 0$. We can inductively find a sequence $1 > t_1 > t_2 > \cdots > 0$ such that $\gamma_2(t_n)$ alternates between zero and one. But then $t_n \to t$ since it is a decreasing bounded-below sequence, and γ_2 is continuous, so $\gamma_2(t_n) \to \gamma_2(t)$ which is a contradiction.

8.5 Gluing lemma

Lemma. Let X be a topological space. Suppose $X = A \cup B$ where A, B are closed in X. Let $g : A \to Y$ and $h : B \to Y$ be continuous where Y is a topological space, such that for $A \cap B$, we have $g = h$. Then $f : X \to Y$ defined by

$$f(x) = \begin{cases} g(x) & x \in A \\ h(x) & x \in B \end{cases}$$

is well defined and continuous.

Proof. First, observe that if $F \subset A$ and F is closed in A, then there exists a closed set G in X such that $F = A \cap G$. Since A is closed in X, we must have F is closed in X. The same holds for $F \subset B$. Now, let V be a closed set in Y. Then the inverse image of V under f is

$$f^{-1}(V) = (f^{-1}(V) \cap A) \cup (f^{-1}(V) \cap B) = \underbrace{g^{-1}(V)}_{\text{closed in } A} \cup \underbrace{h^{-1}(V)}_{\text{closed in } B}$$

So $f^{-1}(V)$ is closed in X. To prove continuity it suffices to show that the preimage of a closed set is closed, since that implies that the preimage of an open set is open. □

Definition. Let X be a topological space. For $x, y \in X$, we write $x \sim y$ if there exists a path from x to y in X. This is an equivalence relation:

(i) The constant function shows that $x \sim x$ for all x.

(ii) If $\gamma : [0,1] \to X$ is continuous and $\gamma(0) = x, \gamma(1) = y$, we define $t \mapsto \gamma(1-t)$, which is a path from y to x.

(iii) Finally, if $x \sim y$ and $y \sim z$, we have continuous functions γ, δ such that $\gamma(0) = x, \gamma(1) = y = \delta(0), \delta(1) = z$. Then let

$$\eta(t) = \begin{cases} \gamma(2t) & t \in \left[0, \frac{1}{2}\right] \\ \delta(2t-1) & t \in \left[\frac{1}{2}, 1\right] \end{cases}$$

These intervals are closed on $[0,1]$ and their union is $[0,1]$. On the intersection, they are equal. By the gluing lemma, η is continuous, and now since $\eta(0) = x, \eta(1) = z$ we have $x \sim z$.

We call the equivalence classes *path-connected components* of X.

Theorem. Let U be an open subset of $\mathbb{R}^n$. Then U is connected if and only if U is path-connected.

Proof. The converse is trivial. Suppose U is connected. Without loss of generality, suppose $U \neq \varnothing$. Let $x_0 \in U$. Let $P = \{x \in U : x \sim x_0\}$ be the equivalence class of x_0. We want to show $P = U$. To do this, we will show that P is open and closed in U. Then, $P, U \setminus P$ will disconnect U unless $P = \varnothing$ or $P = U$. But we know $x_0 \in P$, hence $P = U$ will be the only possibility.

To show P is open, let $x \in U$. Since U is open, there exists $r > 0$ such that $\mathcal{D}_r(x) \subset U$. Recall that for all $y \in \mathcal{D}_r(x)$, we have $y \sim x$. Now, if $x \in P$, then we have $y \sim x$ and $x \sim x_0$ so $y \sim x_0$. So $\mathcal{D}_r(x) \subset P$. So P is open.

Now, if $x \in U \setminus P$ and $y \in \mathcal{D}_r(x)$ has $y \sim x_0$, then by transitivity $x \sim x_0$. But this is a contradiction since $x \notin P$. Hence $U \setminus P$ is open. So P is open and closed, so $P = U$. □

Theorem. For $n \geq 2$, $\mathbb{R}$ and $\mathbb{R}^n$ are not homeomorphic.

The generalisation $\mathbb{R}^m \not\simeq \mathbb{R}^n$ is true, but significantly harder to prove and outside the scope of this course.

Proof. Suppose $f : \mathbb{R} \to \mathbb{R}^n$ is a homeomorphism. Let $g = f^{-1}$. Then g is continuous. Then, $f|_{\mathbb{R}\setminus\{0\}}$ is a homeomorphism from $\mathbb{R} \setminus \{0\}$ to $\mathbb{R}^n \setminus \{f(0)\}$, with inverse $g|_{\mathbb{R}^n\setminus\{f(0)\}}$. But $\mathbb{R} \setminus \{0\}$ is disconnected, but $\mathbb{R}^n \setminus \{f(0)\}$ is connected since it is path-connected. This is a contradiction. □

9 Compactness

9.1 Motivation and definition

Recall from IA Analysis that a continuous function on a closed bounded interval is bounded and attains its bounds. We wish to generalise this result to more general topological spaces.

Example. (i) If X is finite, any function $X \to \mathbb{R}$ is finite.

(ii) If, for all continuous functions $f : X \to \mathbb{R}$ there exists $n \in \mathbb{N}$ and subsets $A_1, \dots, A_n$ of X such that $X = \bigcup_{j=1}^n A_j$ and f is bounded on A_j for all j, then the property holds.

(iii) Note that continuous functions are 'locally bounded'; if $f : X \to \mathbb{R}$ is continuous, then for all $x \in X$ we have $U_x = f^{-1}((f(x)-1, f(x)+1))$ is an open set containing X, and f is bounded on U_x. So each point has an open neighbourhood on which f is bounded. Further, $X = \bigcup_{x\in X} U_x$. If there exists a finite subset $F \subset X$ such that $\bigcup_{x\in F} U_x = X$, then f is bounded on X. This is exactly the definition we will use for compactness.

Definition. Let X be a topological space. An *open cover* for X is a family $\mathcal{U}$ of open subsets of X that cover X; that is, $\bigcup_{U\in\mathcal{U}} U = X$. A *subcover* of $\mathcal{U}$ is a subset $\mathcal{V} \subset \mathcal{U}$ that covers U. This is called a *finite subcover* if $\mathcal{V}$ is finite. We say that X is *compact* if every open cover has a finite subcover.

Remark. Compactness can be thought of as the next best thing to finiteness.

Theorem. Let X be a compact topological space and $f : X \to \mathbb{R}$ be continuous. Then f is bounded, and if X is not empty f attains its bounds.

Proof. For $n \in \mathbb{N}$, let $U_n = \{x \in X : |f(x)| < n\}$. U_n is open since $x \mapsto |f(x)|$ is continuous and $(-n, n)$ is open. It is clear that $X = \bigcup_{n\in\mathbb{N}} U_n$. This is an open cover of X. Hence there exists a finite subcover $F \subset \mathbb{N}$ such that $X = \bigcup_{n\in F} U_n = U_N$ where $N = \max F$. Hence, for all $x \in X$, we have $|f(x)| < N$ so f is bounded.

Let $\alpha = \inf_X f$; this exists since f is bounded. Suppose there exists no $x \in X$ such that $f(x) = \alpha$. Then, for all $x \in X$, $f(x) > \alpha$. Then there exists $n \in \mathbb{N}$ such that $f(x) > \alpha + \frac{1}{n}$. So let

$$V_n = \left\{x \in X : f(x) > \alpha + \frac{1}{n}\right\} = f^{-1}\left(\left(\alpha + \frac{1}{n}, \infty\right)\right)$$

We can see that V_n is open. Now, since $\bigcup_{n\in\mathbb{N}} V_n = X$, there exists a finite subcover $F \subset \mathbb{N}$ such that $\bigcup_{n\in F} V_n = X = V_N$ where N is the maximal F. Then for all $x \in X$, we have $f(x) > \alpha + \frac{1}{N}$. Hence $\inf_X f \geq \alpha + \frac{1}{N}$, which is a contradiction. The same argument applies for the supremum. □

Lemma. Let Y be a subspace of a topological space X. Then Y is compact if and only if whenever $\mathcal{U}$ is a family of open sets in X such that $\bigcup_{U\in\mathcal{U}} \supset Y$, there is a finite subfamily $\mathcal{V} \subset \mathcal{U}$ with $\bigcup_{U\in\mathcal{V}} U \supset Y$.

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

Proof. Let $\mathcal{U}$ be a family of open sets in $\mathbb{R}$ that cover $[0, 1]$. For a subset $A \subset [0, 1]$, we say that $\mathcal{U}$ *finitely covers* A if there exists a finite subcover $\mathcal{V} \subset \mathcal{U}$ of A. Note that if $A = B \cup C$ and $A, B, C \subset [0, 1]$ and $\mathcal{U}$ finitely covers B and C, we can take the union of the finite subcovers to find a finite subcover of A, so U finitely covers A. Suppose that $\mathcal{U}$ does not finitely cover $[0, 1]$. Then one of the intervals $\left[0, \frac{1}{2}\right]$ and $\left[\frac{1}{2}, 1\right]$ is not finitely coverable by $\mathcal{U}$. Let this interval be $[a_1, b_1]$. Let $c = \frac{1}{2}(a_1 + b_1)$. Then one of the intervals $[a_1, c], [c, b_1]$ is not finitely coverable by $\mathcal{U}$. Inductively, we obtain a nested sequence of intervals $[a_1, b_1] \supset \cdots \supset [a_n, b_n] \supset \cdots$ which are not finitely covered by $\mathcal{U}$ and $b_n - a_n = 2^{-n}$. Now, $a_n \to x$ for some $x \in [0, 1]$ and $b_n = a^n + 2^{-n} \to x$. But since $\mathcal{U}$ covers $[0, 1]$, there exists $U \in \mathcal{U}$ such that $x \in U$. U is open in $\mathbb{R}$, so for all $\varepsilon > 0$, we have $(x - \varepsilon, x + \varepsilon) \subset U$. Since $a_n, b_n \to x$, we can choose n such that $a_n, b_n \in (x - \varepsilon, x + \varepsilon)$. This is covered by one open set U in $\mathcal{U}$, so this is a finite subcover. This is a contradiction. □

Example. Other examples of compact spaces include the following.

(i) Any finite set is compact.

(ii) On any set X, the cofinite topology is compact. Suppose without loss of generality that X is not empty, and let $\mathcal{U}$ be an open cover for X. Let $U \in \mathcal{U}$ such that $U \neq \varnothing$. Then $F = X \setminus U$ is finite. For all $x \in F$, let $U_x \in \mathcal{U}$ such that $x \in U_x$. Then $\bigcup_{x \in F} U_x \cup U$ is a finite subcover.

(iii) Let $x_n \to x$ in a topological space X. Let $Y = \{x_n : n \in \mathbb{N}\} \cup \{x\}$. Then Y is compact. Indeed, let $\mathcal{U}$ be a family of open sets in X such that $\bigcup_{U \in \mathcal{U}} U \supset Y$. In particular, let $U \in \mathcal{U}$ such that $x \in U$. Since U is open and $x_n \to x$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $x_n \in U$. So we can cover the remaining finitely many elements analogously to the previous example, and this yields a finite subcover.

(iv) The indiscrete topology on any set is compact, since there are only two open sets.

Counterexamples include the following.

(i) An infinite set X in the discrete topology is not compact. Let

$$\mathcal{U} = \{\{x\} : x \in X\}$$

This has no finite subcover.

(ii) $\mathbb{R}$ is not compact. Consider the intervals $(-n, n)$ for all $n \in \mathbb{N}$. This is an open cover with no finite subcover.

9.2 Subspaces

Theorem. Let Y be a subspace of a topological space X. Then,

(i) Let X be compact and Y be closed in X. Then Y is compact.

(ii) Let X be Hausdorff and Y be compact. Then Y is closed in X.

Proof. Let $\mathcal{U}$ be a family of open sets in X such that their union covers Y. Then $\mathcal{U} \cup (X \setminus Y)$ is an open cover for X since Y is closed. This has a finite subcover $\mathcal{V} \subset \mathcal{U}$ such that $\bigcup_{U \in \mathcal{V}} U \cup (X \setminus Y) = X$. Then $\bigcup_{U \in \mathcal{V}} U \supset Y$.

For part (ii), let $x \in X \setminus Y$. For $y \in Y$, since $x \neq y$ there exist open sets U_y, V_y in X such that $x \in U_y, y \in V_y, U_y \cap V_y = \varnothing$. Now, $\{V_y : y \in Y\}$ is an open cover of Y. Hence there exists $F \subset Y$ finite such that $\bigcup_{y \in F} V_y \supset Y$. Now, $U = \bigcap_{y \in F} U_y$ is open, further $x \in U$ and

$$U \cap Y \subset \left(\bigcap_{y \in F} U_y \right) \cap \left(\bigcup_{y \in F} V_y \right) = \varnothing$$

Hence $X \setminus Y$ is a neighbourhood of all of its points, so it is open and Y is closed. $\square$

9.3 Continuous images of compact spaces

Theorem. Let $f : X \to Y$ be a continuous function between topological spaces such that X is compact. Then $f(X)$ is compact.

Proof. Let $\mathcal{U}$ be a family of open sets in Y such that $\bigcup_{U\in\mathcal{U}} U \supset f(X)$. Then $\bigcup_{U\in\mathcal{U}} f^{-1}(U) = X$ and $f^{-1}(U)$ is open in X for all $U \in \mathcal{U}$ since f is continuous. Since X is compact, we have a finite subcover $\mathcal{V} \subset \mathcal{U}$ such that $X = \bigcup_{U\in\mathcal{V}} f^{-1}(V)$. Hence $f(X) \subset \bigcup_{U\in\mathcal{V}} U$. □

Remark. Compactness is a topological property. If $f : X \to Y$ is continuous and $A \subset X$ is compact, then $f(A)$ is compact.

Corollary. Any quotient of a compact space is compact.

Example. Let $a < b \in \mathbb{R}$. Then $[a, b] \simeq [0, 1]$ so is compact.

9.4 Topological inverse function theorem

Theorem. Let $f : X \to Y$ be a continuous bijection from a compact space X to a Hausdorff space Y. Then f^{-1} is continuous, so f is an open map. Hence f is a homeomorphism.

Proof. Let U be an open subset of X. Then $K = X \setminus U$ is closed. Since X is compact, K is compact. Further, $f(K)$ is compact. Hence $f(K)$ is closed in Y. So $f(U) = Y \setminus f(K)$ is open in Y. □

Example. $\mathbb{R}/\mathbb{Z}$ is homeomorphic to $S^1 = \left\{x \in \mathbb{R}^2 : \|x\| = 1\right\}$. Indeed, let $f : \mathbb{R} \to S^1$ by $f(t) = (\cos(2\pi t), \sin(2\pi t))$. For all s, t, we have $f(s) = f(t)$ if and only if $s \sim t$ so f fully respects $\sim$. f is continuous and surjective. Let $\widetilde{f} : \mathbb{R}/\mathbb{Z} \to S^1$ be the unique map such that $\widetilde{f} \circ q = f$. So $\widetilde{f}$ is a continuous bijection. S^1 is Hausdorff, and $\mathbb{R}/\mathbb{Z}$ is the image of $[0, 1]$ under a continuous map, hence is compact. Hence $\widetilde{f}$ is a homeomorphism.

9.5 Tychonov's theorem

Theorem. Let X, Y be compact topological spaces. Then $X \times Y$ is compact in the product topology.

Proof. Let $\mathcal{U}$ be an open cover for $X\times Y$. We want to show that there exists a finite subcover. Without loss of generality, every member of $\mathcal{U}$ can be of the form $U\times V$ where U is open in X and V is open in Y. Indeed, for $z \in X\times Y$ we can choose $W_z \in \mathcal{U}$ such that $z \in W_z$. By definition of the product topology, there exist open sets U_z in X and V_z in Y such that $z \in U_z \times V_z \subset W_z$. So $\{U_z \times V_z : z \in X \times Y\}$ is an open cover for $X \times Y$. If there exists a finite subset $F \subset X \times Y$ such that $\bigcup_{z\in F} U_z \times V_z$ covers $X \times Y$, then $\{W_x : z \in F\}$ is a finite subcover of $\mathcal{U}$.

Let $x \in X$. Recall that $\{x\} \times Y$ is the continuous image of Y under the map $y \mapsto (x, y)$. Hence, $\{x\} \times Y$ is compact, since the continuous image of a compact space is compact. Since $\{x\} \times Y$ is covered by $\bigcup_{W\in\mathcal{U}} W$, $\mathcal{U}$ finitely covers $\{x\} \times Y$. So there exists $n_x \in \mathbb{N}$ such that we can find open sets $U_{x,1}, \dots, U_{x,n_x}$ in X and $V_{x,1}, \dots, V_{x,n_x}$ in Y such that $U_{x,j} \times V_{x,j} \in \mathcal{U}$ and $\{x\}\times Y \subset \bigcup_{j=1}^{n_x} U_{x,j} \times V_{x,j}$.

Without loss of generality, let $x \in U_{x,j}$ for all j, since any other $U_{x,j}$ is not needed in the cover. Now let $U_x = \bigcap_{j=1}^{n_x} U_{x,j}$. We know $x \in U_x$ and U_x is open since it is a finite intersection of open sets. In particular, $U_x \times Y \subset \bigcup_{j=1}^{n_x} U_{x,j} \times V_{x,j}$.

Now, $\{U_x : x \in X\}$ is an open cover for X. So there exists a finite subset $F \subset X$ such that $X = \bigcup_{x\in F} U_x$. Then, $X \times Y = \bigcup_{x\in F} U_x \times Y \subset \bigcup_{x\in F} \bigcup_{j=1}^{n_x} U_{x,j} \times V_{x,j}$. Hence,

$$\left\{U_{x,j} \times V_{x,j} : x \in F, 1 \leq j \leq n_x\right\}$$

is a finite subcover of $\mathcal{U}$. □

Remark. More generally, if $X_1, \dots, X_n$ are compact spaces, then so is $X_1 \times \dots \times X_n$.

9.6 Heine–Borel theorem

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

Proof. Suppose K is compact. $\mathbb{R}^n$ is a metric space and hence Hausdorff. Hence, K is closed in $\mathbb{R}^n$. The function $x \mapsto \|x\|$ is continuous. Therefore, it is bounded on K. So K is bounded.

Conversely, if K is bounded, there exists $M \geq 0$ such that for all $x \in K$ we have $\|x\| \leq M$. Hence, $K \subset [-M, M]^n$. Note that $[-M, M]$ is compact since it is homeomorphic to $[0, 1]$. By Tychonov's theorem, $[-M, M]^n$ is compact in the product topology. Since a closed subset of a compact space is compact, K is compact. □

Example. Closed balls $\mathcal{B}_r(x)$ in $\mathbb{R}^n$ are compact. The start of the proof for the Lindelöf–Picard theorem now makes more sense.

9.7 Sequential compactness

Definition. A topological space X is *sequentially compact* if every sequence in X has a convergent subsequence. Given a sequence (x_n) and an infinite set $M \subset \mathbb{N}$, we will write $(x_m)_{m\in M}$ for the subsequence $(x_{m_n})_{n=1}^{\infty}$ where $m_1 < m_2 < \dots$ are the elements of M. Note that if $L \subset M \subset \mathbb{N}$, then $(x_n)_{n\in L}$ is a subsequence of $(x_n)_{n\in M}$.

Example. Any closed and bounded subset of $\mathbb{R}$ is sequentially compact by the Bolzano–Weierstrass theorem. Similarly, any closed and bounded subset K of $\mathbb{R}^n$ is sequentially compact. Indeed, let (x_m) be a sequence in K. Then, writing $x_m = (x_{m,1}, \dots, x_{m,n})$, since K is bounded we have that $(x_{m,j})$ is bounded for all j. Applying the Bolzano–Weierstrass theorem to the first coordinate, we find $M_1 \subset \mathbb{N}$ such that $(x_{m,1})_{m\in M_1}$ converges in $\mathbb{R}$. Now, $(x_{m,2})_{m\in M_1}$ is bounded in $\mathbb{R}$, so again applying the Bolzano–Weierstrass theorem, we can find $M_2 \subset \mathbb{N}$ such that $(x_{m,2})_{m\in M_2}$ converges. Note that $(x_{m,1})_{m\in M_2}$ converges. So inductively we can find $M_1 \supset \dots \supset M_n$ such that $(x_{m,j})_{m\in M_n}$ converges for all j. Hence $(x_m)_{m\in M_n}$ converges in $\mathbb{R}^n$. The limit is contained in K since K is closed.

Remark. In $\mathbb{R}^n$, any compact space is sequentially compact. The converse is also true; any sequentially compact subspace must be closed and bounded. We aim to show that compactness and sequential compactness are identical in metric spaces.

9.8 Compactness and sequential compactness in metric spaces

Let (M, d) be a metric space.

Definition. For $\varepsilon > 0$ and $F \subset M$, we say that F is an *ε-net for M* if for all $x \in M$, there exists $y \in F$ such that $d(y, x) \leq \varepsilon$. Equivalently, $M = \bigcup_{y \in M} \mathcal{B}_\varepsilon(y)$. This is called a *finite ε-net* if F is finite. We say that M is *totally bounded* if for all $\varepsilon > 0$, there exists a finite ε-net for M.

Example. For $\varepsilon > 0$, let n such that $\frac{1}{n} < \varepsilon$. Then $\left\{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\right\}$ is an ε-net for $(0, 1)$.

Definition. For a non-empty $A \subset M$, the *diameter* of A is $\operatorname{diam} A = \sup\{d(x, y) : x, y \in A\}$. This is finite if and only if A is a bounded set.

Example. $\operatorname{diam} \mathcal{B}_r(x) \leq 2r$.

Lemma. Suppose M is totally bounded. Let A be a non-empty closed subset of M. Let $\varepsilon > 0$. Then there exists $K \in \mathbb{N}$ and non-empty closed sets $B_1, \dots, B_K$ such that $A = \bigcup_{k=1}^K B_k$ and $\operatorname{diam} B_k \leq \varepsilon$ for all k.

Proof. Let F be a finite $\frac{\varepsilon}{2}$-net for M. So $M = \bigcup_{x \in F} B_{\varepsilon/2}(x)$ and hence $A = \bigcup_{x \in F} (A \cap B_{\varepsilon/2}(x))$. Let $G = \left\{x \in F : A \cap B_{\varepsilon/2}(x) \neq 0\right\}$. Then for $x \in G$ let $B_x = A \cap B_{\varepsilon/2}(x)$. So for $x \in G$, we have $B_x \neq \varnothing$, $B_x \subset B_{\varepsilon/2}(x)$ and so $\operatorname{diam} B_x \leq \varepsilon$, and B_x is closed. Then $A = \bigcup_{x \in G} B_x$. □

Theorem. For a metric space (M, d), the following are equivalent.
(i) M is compact;
(ii) M is sequentially compact;
(iii) M is complete and totally bounded.

Proof. We first show (i) implies (ii). Let (x_n) be a sequence in M. Then for $n \in \mathbb{N}$, let $T_n = \{x_k : k > n\}$ be the tail of the sequence. Note that the limit of any convergent subsequence (if it exists) is in the intersection of $\bigcap_{n \in \mathbb{N}} \overline{T}_n$. So first, we prove that this intersection is non-empty. Suppose that it is empty. Then, $\bigcup_{n \in \mathbb{N}} \left(M \setminus \overline{T}_n\right) = M$. But the $M \setminus \overline{T}_n$ are open, and M is compact, there is a finite subcover. So $M \setminus \overline{T}_N = M$ for some N, since the $\underline{T_n}$ are a decreasing sequence of sets. This is a contradiction since $T_N \neq \varnothing$. Now, let $x \in \bigcap_{n \in \mathbb{N}} \overline{T}_n$, and we want to show the existence of a subsequence converging to x. First, $x \in \overline{T}_1$, so $\mathcal{D}_1(x) \cap T_1 \neq \varnothing$. Hence there exists $k_1 > 1$ such that $d(x_{k_1}, x) < 1$. Now since $x \in \overline{T}_{k_1}$, $\mathcal{D}_{1/2}(x) \cap T_{k_1} \neq \varnothing$. There exists $k_2 > k_1$ such that $d(x_{k_2}, x) < \frac{1}{2}$. Inductively, we can find a strictly increasing sequence $k_1 < k_2 < \dots$ such that $d(x_{k_n}, x) < \frac{1}{n}$ for all n, so $x_{k_n} \to x$.

Now, we show (ii) implies (iii). To show M is complete, let (x_n) be a Cauchy sequence in M. Let $k_1 < k_2 < \dots$ such that x_{k_n} converges in M, and let x be the limit. We show $x_n \to x$. Indeed, for $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $\forall m, n \geq N$, we have $d(x_m, x_n) < \varepsilon$. Then $\forall m \geq N$, we have $k_n \geq m \geq N$, so for a fixed $n \geq N$ and $\forall m \geq N$, we have $d(x_n, x) \leq d(x_n, x_{k_m}) + d(x_{k_m}, x) \leq \varepsilon + d(x_{k_m}, x)$. Let $m \to \infty$, so $d(x_n, x) \leq \varepsilon$. So $x_n \to x$. To show M is totally bounded, suppose it is not. There exists $\varepsilon > 0$ such that M has no finite ε-net. Let $x_1 \in M$, and suppose we can find $x_1, \dots, x_{n-1}$ in M. Then $\bigcup_{j=1}^{n-1} \mathcal{B}_\varepsilon(x_j) \neq M$. So we can pick $x_n \in M \setminus \bigcup_{j-1}^{n-1} \mathcal{B}_\varepsilon(x_j)$. Inductively we obtain (x_n) such that

$d(x_m, x_n) > \varepsilon$ for all $n, m \in \mathbb{N}$. So (x_n) has no Cauchy subsequence. There is therefore no convergent subsequence, which is a contradiction.

Finally, we show (iii) implies (i). Let $\mathcal{U}$ be an open cover for M. We must show there exists a finite subcover. Suppose that is not true, so $\mathcal{U}$ does not finitely cover M. We construct non-empty closed subsets $A_0 \supset A_1 \supset \dots$ of M such that for all $n \geq 0$, $\mathcal{U}$ does not finitely cover A_n, and for all $n \geq 1$ we have $\operatorname{diam} A_n < \frac{1}{n}$. Let $A_0 = M$. Suppose that for some $n \geq 1$ we have already found A_{n-1}. Since M is totally bounded, we can write $A_{n-1} = \bigcup_{k=1}^{K} B_k$ where $K \in \mathbb{N}$ and the B_k are non-empty, closed, and $\operatorname{diam} B_k < \frac{1}{n}$. Since $\mathcal{U}$ does not finitely cover A_{n-1}, there exists $k \leq K$ such that $\mathcal{U}$ does not finitely cover B_k. Let A_n be this B_k. Now, for all n, pick some $x_n \in A_n$. For all N, $\forall m, n \geq N$ we have $x_m, x_n \in A_N$ hence $d(x_m, x_n) \leq \operatorname{diam} A_N \leq \frac{1}{n}$ so the sequence is Cauchy. M is complete, so $x_n \to x$ for some $x \in M$. Let $U \in \mathcal{U}$ such that $x \in U$. U is open, so there exists $r > 0$ such that $\mathcal{D}_r(x) \subset U$. But $x_n \to x$ hence there exists n such that $d(x_n, x) < \frac{r}{2}$ and $\operatorname{diam} A_n < \frac{r}{2}$. For every $y \in A_n$, $d(y, x) \leq d(y, x_n) + d(x_n, x) \leq \operatorname{diam} A_n + \frac{r}{2} < r$. Hence every point in A_n is contained within $\mathcal{D}_r(x) \subset U$. But this contradicts the fact that $\mathcal{U}$ does not finitely cover A_n, but we have constructed a cover using just one open set. □

Remark. We can now deduce the one direction of the Heine–Borel theorem from the Bolzano–Weierstrass theorem; closed and bounded subsets of $\mathbb{R}^n$ are compact. Similarly, we can check that the product of sequentially compact topological spaces is sequentially compact in the product topology. This yields a new proof for Tychonov's theorem for metric spaces. In general, there exist topological spaces that are compact but not sequentially compact, and conversely there exist topological spaces which are sequentially compact but not compact.

10 Differentiation

10.1 Linear maps

Let $m, n \in \mathbb{N}$. Recall that $L(\mathbb{R}^m, \mathbb{R}^n)$ is the vector space of linear maps from $\mathbb{R}^m$ to $\mathbb{R}^n$. This is isomorphic to $M_{n,m}$, the space of $n \times m$ real matrices. There is also an isomorphism to $\mathbb{R}^{mn}$. Let $e_1, \dots, e_m$ be the standard basis of $\mathbb{R}^m$, and similarly let $e'_1, \dots, e'_n$ be the standard basis of $\mathbb{R}^n$. Then $T \in L(\mathbb{R}^m, \mathbb{R}^n)$ is identified with the $n \times m$ matrix (T_{ji}) where $1 \leq j \leq n$ and $1 \leq i \leq m$, such that $T_{ji} = \left\langle Te_i, e'_j \right\rangle$. We can therefore view $L(\mathbb{R}^m, \mathbb{R}^n)$ as the mn-dimensional vector space $\mathbb{R}^{mn}$ with the Euclidean norm. So the norm of a linear map T is given by

$$\|T\| = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} T_{ji}^2} = \sqrt{\sum_{i=1}^{m} \|Te_i\|^2}$$

where Te_i is the ith column of T. Thus, $L(\mathbb{R}^m, \mathbb{R}^n)$ becomes a metric space together with the Euclidean distance $d(S, T) = \|S - T\|$.

Lemma. For $T \in L(\mathbb{R}^m, \mathbb{R}^n)$ and $x \in \mathbb{R}^m$,

$$\|Tx\| \leq \|T\| \cdot \|x\|$$

So T is a Lipschitz map and hence continuous. Further, if $S \in L(\mathbb{R}^n, \mathbb{R}^p)$ then

$$\|ST\| \leq \|S\| \cdot \|T\|$$

Proof. We can write

$$x = \sum_{i=1}^{m} x_i e_i$$

Hence,

$$Tx = \sum_{i=1}^{m} x_i Te_i$$

Thus,

$$\|Tx\| \leq \sum_{i=1}^{m} |x_i|\|Te_i\| \leq \left(\sum_{i=1}^{m} x_i^2\right)^{1/2} \cdot \left(\sum_{i=1}^{m} \|Te_i\|^2\right)^{1/2} = \|T\| \cdot \|x\|$$

Further, for $x, y \in \mathbb{R}^m$ we have

$$d(Tx, Ty) = \|Tx - Ty\| = \|T(x - y)\| \leq \|T\| \cdot \|x - y\| = \|T\| d(x, y)$$

So T is Lipschitz, and any Lipschitz function is continuous. Now,

$$\|ST\| = \left(\sum_{i=1}^{m} \|STe_i\|^2\right)^{1/2} \leq \left(\sum_{i=1}^{m} \|S\|\|Te_i\|^2\right)^{1/2} = \|S\|\left(\sum_{i=1}^{m} \|Te_i\|^2\right)^{1/2} = \|S\| \cdot \|T\|$$

□

10.2 Differentiation

Recall from IA Analysis that a function $f : \mathbb{R} \to \mathbb{R}$ is *differentiable* at a point $a \in \mathbb{R}$ if

$$\lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

exists. The value of this limit is called the *derivative* of f at a, and denoted $f'(a)$. Note that f is differentiable at a if and only if there exists $\lambda \in \mathbb{R}$ and $\varepsilon : \mathbb{R} \to \mathbb{R}$ such that $\varepsilon(0) = 0$ and ε is continuous at 0, and

$$f(a + h) = f(a) + \lambda h + h\varepsilon(h)$$

This is because we can define

$$\varepsilon(h) = \begin{cases} 0 & h = 0 \\ \frac{f(a+h)-f(a)}{h} - \lambda & h \neq 0 \end{cases}$$

Informally, this ε definition states that f is approximated very well (the error $h\varepsilon(h)$ shrinks rapidly since $\varepsilon \to 0$) by a linear function in a small neighbourhood of a. Recall that if f is n times differentiable at a, then

$$f(a + h) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!} h^k + o(h^n)$$

Definition. Let $m, n \in \mathbb{N}$. Then $f : \mathbb{R}^m \to \mathbb{R}^n$ and $a \in \mathbb{R}^m$. We say that f is *differentiable* at a if there exists a linear map $T \in L(\mathbb{R}^m, \mathbb{R}^n)$ and a function $\varepsilon : \mathbb{R}^m \to \mathbb{R}^n$ such that $\varepsilon(0) = 0$

and ε is continuous at 0, and

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

Note that

$$\varepsilon(h) = \begin{cases} 0 & h = 0 \\ \frac{f(a+h)-f(a)-T(h)}{\|h\|} & h \neq 0 \end{cases}$$

So f is differentiable at a if and only if there exists $T \in L(\mathbb{R}^m, \mathbb{R}^n)$ such that

$$\frac{f(a+h) - f(a) - T(h)}{\|h\|} \to 0$$

as $h \to 0$. Such a T is unique. Indeed, suppose S, T satisfy the above limit. Then, by subtracting,

$$\frac{S(h) - T(h)}{\|h\|} \to 0$$

For a fixed $x \in \mathbb{R}^m$, $x \neq 0$, we have $\frac{x}{k} \to 0$ as $k \to \infty$ so

$$\frac{S\left(\frac{x}{k}\right) - T\left(\frac{x}{k}\right)}{\left\|\frac{x}{k}\right\|} \to 0 \implies \frac{S(x) - T(x)}{\|x\|} = 0$$

So $Sx = Tx$. It follows that $S = T$. We say that if a function f is differentiable at a point a, T is the unique *derivative* of f at a. This is denoted $f'(a) = Df(a) = Df|_a$. If $f : \mathbb{R}^m \to \mathbb{R}^n$ is differentiable at $a \in \mathbb{R}^m$ for every a, we say that f is *differentiable on* $\mathbb{R}^m$. The function $f' = D : \mathbb{R}^m \to L(\mathbb{R}^m, \mathbb{R}^n)$ mapping $a \mapsto f'(a)$ is the derivative of f.

Example. Constant functions are differentiable. Let $f : \mathbb{R}^m \to \mathbb{R}^n$ such that $f(x) = b$ for $b \in \mathbb{R}^n$. Then for all $a \in \mathbb{R}^m$, we have

$$f(a+h) = f(a) + 0h + 0$$

so f is differentiable at a and the derivative is zero.

Example. Linear maps are differentiable. Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be defined by $f(x) = Tx$ for a linear map $T \in L(\mathbb{R}^m, \mathbb{R}^n)$. Then

$$f(a+h) = f(a) + f(h) + 0$$

so f is differentiable at a with derivative $T = f$. So f' is a constant function.

Example. Consider

$$f(x) = \|x\|^2$$

For $a \in \mathbb{R}^m$, we can find

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

Hence, f is differentiable with derivative

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

Note that $f' : \mathbb{R}^m \to L(\mathbb{R}^m \to \mathbb{R})$ is linear.

Example. Note $M_n \simeq \mathbb{R}^{n^2}$. The function $f : M_n \to M_n$ given by $f(A) = A^2$. For a fixed $A \in M_n$,

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

It suffices to show H^2 is $o(\|H\|)$. We have $\left\|H^2\right\| \leq \|H\|^2$, hence

$$\frac{\left\|H^2\right\|}{\|H\|} \leq \|H\| \to 0$$

So f is differentiable at A and the derivative is given by

$$f'(A)(H) = AH + HA$$

Example. Suppose $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ is bilinear. Let $(a,b) \in \mathbb{R}^m \times \mathbb{R}^n$. Then,

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

The map $\mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ given by $(h,k) \mapsto f(a,k) + f(h,b)$ is linear as the sum of two linear maps. So it suffices to show $f(h,k)$ is $o(\|(h,k)\|)$.

$$h = \sum_{i=1}^m h_i e_i; \quad k = \sum_{j=1}^n k_j e'_j$$

Hence,

$$f(h,k) = \sum_{i=1}^m \sum_{j=1}^n h_i k_j f(e_i, e'_j) \implies \|f(h,k)\| \leq \sum_{i=1}^m \sum_{j=1}^n |h_i| \cdot |k_j| \cdot \left\|f(e_i, e'_j)\right\| \leq C\|(h,k)\|^2$$

for some constant C, since $|h_i| \leq \|(h,k)\|^2$ and similarly for $|k_j|$. So

$$\frac{\|f(h,k)\|}{\|(h,k)\|} \leq C\|(h,k)\| \to 0$$

Hence f is differentiable with

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

10.3 Derivatives on open subsets

We may define the derivative on a subset of $\mathbb{R}^m$. We will use the notion of open subsets since we are typically interested in neigbourhoods of points.

Definition. Let U be an open subset of $\mathbb{R}^m$. Let $f : U \to \mathbb{R}^n$ be a function, and $a \in U$. Then we say f is *differentiable* at a if there exists a linear map $T \in L(\mathbb{R}^m, \mathbb{R}^n)$ such that

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

where $\varepsilon(0) = 0$ and ε is continuous at zero. Note that ε need only be defined on the set of h such that $a + h \in U$, or more precisely the open set $U - a$. Hence there exists $r > 0$ such that

$\mathcal{D}_r(0) \subset U_a$. Then

$$\varepsilon(h) = \begin{cases} 0 & h = 0 \\ \frac{f(a+h)-f(a)-T(h)}{\|h\|} & h \neq 0, a+h \in U \end{cases}$$

So f is differentiable at a if and only if there exists a linear map $T \in L(\mathbb{R}^m, \mathbb{R}^n)$ such that

$$\frac{f(a+h)-f(a)-T(h)}{\|h\|} \to 0$$

Remark. The linear map T is unique, and is called the *derivative* of f at a, denoted $f'(a)$. In particular,

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

Remark. If $m = 1$, the space $L(\mathbb{R}, \mathbb{R}^n)$ is isomorphic to $\mathbb{R}^n$. The linear map is defined uniquely by a vector in $\mathbb{R}^n$ which multiplies by the scalar h. Hence, if $U \subset \mathbb{R}$ is open and $f : U \to \mathbb{R}$ be a function and $a \in U$, then f is differentiable at a if there exists a vector $v \in \mathbb{R}^n$ such that

$$\frac{f(a+h)-f(a)-hv}{|v|} \to 0$$

Equivalently, there exists $v \in \mathbb{R}^n$ such that

$$\frac{f(a+h)-f(a)}{h} \to v$$

10.4 Properties of derivative

Proposition. Let $U \subset \mathbb{R}^m$ be open, $f : U \to \mathbb{R}^n$ be a function, and $a \in U$. If f is differentiable at a, f is continuous at a.

Proof. We have

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

Hence,

$$f(x) = f(a) + f'(a)(x-a) + \|x-a\|\varepsilon(x-a)$$

The functions $x \mapsto f(a)$, $x \mapsto f'(a)(x-a)$ and $x \mapsto \|x-a\|\varepsilon(x-a)$ are all continuous at a. Hence their sum is continuous. □

Proposition (chain rule). Let $U \subset \mathbb{R}^m$ and $V \subset \mathbb{R}^n$ be open, $f : U \to \mathbb{R}^n$ and $g : V \to \mathbb{R}^p$ be functions, and $a \in U, b \equiv f(a) \in V$. Suppose f is differentiable at a, and g is differentiable at b. Then $g \circ f$ is differentiable at a and

$$(g \circ f)'(a) = g'(b) \circ f'(a)$$

Proof. Let $S = f'(a)$ and $T = g'(b)$. Then by assumption

$$f(a+h) = f(a) + S(h) + \|h\|\varepsilon(h); \quad g(b+k) + g(b) + T(k) + \|k\|\zeta(k)$$

for suitable ε, ζ. Then,

$$\begin{aligned}(g \circ f)(a+h) &= g(f(a) + S(h) + \|h\|\varepsilon(h)) \\ &= g\left(b + \underbrace{S(h) + \|h\|\varepsilon(h)}_{k}\right) \\ &= g(b) + T(S(h) + \|h\|\varepsilon(h)) + \|S(h) + \|h\|\varepsilon(h)\|\zeta(S(h) + \|h\|\varepsilon(h)) \\ &= (g \circ f)(a) + (T \circ S)(h) + \|h\|T(\varepsilon(h)) + \|k\|\zeta(k)\end{aligned}$$

It suffices to show that

$$\eta(h) \equiv \|h\|T(\varepsilon(h)) + \|k\|\zeta(k)$$

satisfies $\frac{\eta}{\|h\|} \to 0$. Then the result follows. First,

$$\frac{\|h\|T(\varepsilon(h))}{\|h\|} = T(\varepsilon(h)) \to 0$$

as $\|T(\varepsilon(h))\| \leq \|T\| \cdot \|\varepsilon(h)\| \to 0$. Then,

$$\frac{\|k\|}{\|h\|} = \frac{\|S(h)\| + \|h\| \cdot \|\varepsilon(h)\|}{\|h\|} \leq \|S\| + \|\varepsilon(h)\|$$

Hence, $k = S(h) + \|h\| \cdot \varepsilon(h) \to 0$ as $h \to 0$. Thus $\zeta(k) \to 0$ as $k \to 0$. So

$$\frac{\eta(h)}{\|h\|} = T(\varepsilon(h)) + \frac{\|k\|}{\|h\|}\zeta(k) \to 0$$

as required. □

Proposition. Let $U \subset \mathbb{R}^m$ be open, $f : U \to \mathbb{R}^n$ be a function, and $a \in U$. Let f_j be the jth component of f, so $f_j = \pi_j \circ f$. Then f is differentiable at a if and only if each f_j is differentiable at a. If this holds,

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

Equivalently,

$$\pi_j[f'(a)(h)] = f_j'(a)(h)$$

Proof. If f is differentiable at a, by the chain rule the composite $\pi_j \circ f$ is differentiable at a. Since the derivative of a linear map is itself, the derivative is given by

$$f_j'(a) = \pi_j'(f(a)) \circ f'(a) = \pi_j \circ f'(a)$$

Hence

$$f'(a)(h) = \sum_{j=1}^n \pi_j\left[f'(a)(h)e_j'\right] = \sum_{j=1}^n f_j'(a)(h)e_j'$$

Conversely suppose each f_j is differentiable. Then

$$f_j(a+h) = f_j(a) + f_j'(a)(h) + \|h\|\varepsilon_j(h)$$

for suitable $\varepsilon(j)$. Now,

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

Since each ε_j tends to zero as $h \to 0$, so does their sum. □

Remark. This proposition shows that we can prove things for an image $\mathbb{R}^n = \mathbb{R}$ without loss of generality.

10.5 Linearity and product rule

Proposition. Let $U \subset \mathbb{R}^m$ be open and functions $f, g\colon U \to \mathbb{R}^n$, $\phi\colon U \to \mathbb{R}$ which are differentiable at a. Then the functions $f + g$ and $\phi \cdot f$ are also differentiable and their derivatives are

$$(f+g)'(a) = f'(a) + g'(a); \quad (\phi f)'(a)(h) = \phi(a)[f'(a)(h)] + [\phi'(a)(h)]f(a)$$

For $m = n = 1$ this is the usual product rule.

Proof. We have

$$\begin{aligned} f(a+h) &= f(a) + f'(a)(h) + \|h\|\varepsilon(h) \\ g(a+h) &= g(a) + g'(a)(h) + \|h\|\zeta(h) \\ \phi(a+h) &= \phi(a) + \phi'(a)(h) + \|h\|\eta(h) \end{aligned}$$

for suitable ε, ζ, η. The sum gives

$$(f+g)(a+h) = (f+g)(a+h) + (f'(a) + g'(a))(h) + \|h\|(\varepsilon(h) + \zeta(h))$$

It follows that $f + g$ is differentiable at a and its derivative is the sum of the derivatives of its components.

$$\begin{aligned} (\phi \cdot f)(a+h) &= \phi(a+h)f(a+h) \\ &= (\phi \cdot f)(a) + [\phi(a)f'(a)(h) + \phi'(a)(h)f(a)] + f'(a)(h)\phi'(a)(h) \\ &+ \|h\| \underbrace{(f'(a)(h)\eta(h) + \phi'(a)(h)\varepsilon(h) + \eta(h)f(a) + \phi(a)\varepsilon(h) + \|h\|\eta(h)\varepsilon(h))}_{\delta(h)} \end{aligned}$$

Now,

$$\frac{\|\phi'(a)(h) \cdot f'(a)(h)\|}{\|h\|} = \frac{|\phi'(a)(h)| \cdot \|f'(a)(h)\|}{\|h\|} \leq \frac{\|\phi'(a)\| \cdot \|h\| \cdot \|f'(a)\| \cdot \|h\|}{\|h\|} \to 0$$

Clearly $\delta \to 0$ since the same is true for all of its components. □

11 Partial derivatives

11.1 Directional and partial derivatives

Definition. Let U, f, a as before. Fix a direction $u \in \mathbb{R}^m$ where $u \neq 0$. If the limit

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

exists, then the value of this limit is the *directional derivative* of f at a in direction u, denoted $D_u f(a)$.

Remark. Note that $D_u f(a) \in \mathbb{R}^n$. Further, $f(a + tu) = f(a) + tD_u f(a) + o(t)$. Define $\gamma : \mathbb{R} \to \mathbb{R}^m$ by $\gamma(t) = a + tu$. Then $f \circ \gamma$ is defined on $\gamma^{-1}(U)$ which is open as γ is continuous, and $0 \in \gamma^{-1}(U)$. Then,

$$\frac{f(a + tu) - f(a)}{t} = \frac{(f \circ \gamma)(t) - (f \circ \gamma)(0)}{t}$$

Hence $D_u f(a)$ exists if and only if $f \circ \gamma$ is differentiable at zero, and its value is the derivative of $f \circ \gamma$. When $u = e_i$ for a standard basis vector e_i, if $D_{e_i} f(a)$ exists we call it the ith *partial derivative* of f at a, denoted $D_i f(a)$.

Proposition. Let U, f, a as before. If f is differentiable at a, then all directional derivatives $D_u f(a)$ exist. Further,

$$D_u f(a) = f'(a)(u)$$

Further,

$$f'(a)(h) = \sum_{i=1}^{m} h_i D_i f(a)$$

for all $h = \sum_{i=1}^{m} h_i e_i$.

Proof. Since f is differentiable,

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

Let $h = tu$. Then,

$$f(a + tu) = f(a) + tf'(a)(u) + |t| \cdot \|u\|\varepsilon(tu)$$

Hence,

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

The error term converges to zero, hence the limit becomes $f'(a)(u)$. Moreover, for all h defined as above,

$$f'(a)(h) = \sum_{i=1}^{m} h_i f'(a)(e_i) = \sum_{i=1}^{m} h_i D_i f(a)$$

□

alternative proof. Let $\gamma(t) = a + tu$. Then $f \circ \gamma$ is defined on the open set $\gamma^{-1}(U)$. Note that γ is differentiable and $\gamma'(t) = u$ for all t. By the chain rule, $f \circ \gamma$ is differentiable at zero, and

$$D_u f(a) = (f \circ \gamma)'(0) = f'(\gamma(0))(\gamma'(0)) = f'(a)(u)$$

□

Remark. If $D_u f(a)$ exists, then so does $D_u f_j(a)$ where $f_j = \pi_j \circ f$. Indeed, by linearity and continuity of π,

$$\frac{f_j(a+tu) - f_j(a)}{t} = \pi_j\left(\frac{f(a+tu) - f(t)}{t}\right) \to \pi_j(D_u f(a))$$

The converse of the proposition is false in general.

11.2 Jacobian matrix

Definition. Suppose f is differentiable at a. Then the Jacobian matrix of f at a, denoted $J_f(a)$, is the matrix of $f'(a)$ with respect to the standard bases. For $1 \leq i \leq m$, the ith column is

$$f'(a)(e_i) = D_i f(a)$$

In particular, for the j, i entry,

$$\left(J_f(a)\right)_{ji} = \left\langle D_i f(a), e_j'\right\rangle = \pi_j(D_i f(a)) = D_i f_j(a) = \frac{\partial f_j}{\partial x_i}$$

11.3 Constructing total derivative from partial derivatives

Theorem. Suppose there exists an open neighbourhood V of a with $V \subset U$ such that $D_i f(x)$ exists for all $x \in V$ and for all $1 \leq i \leq m$, and the map $x \mapsto D_i f(x)$ from V to $\mathbb{R}^n$ is continuous at a for all i. Then f is differentiable at a.

Proof. By considering components, without loss of generality let $n = 1$. Let $m = 2$ for convenience of notation; this does not change the proof. Let $a = (p, q)$. Let

$$\psi(h,k) = f(p+h, q+k) - f(p,q) - hD_1 f(p,q) - kD_2 f(p,q)$$

We need to show $\psi(h,k) = o(\|(h,k)\|)$, then the derivative of f can be read off from the definition of ψ. Note,

$$\psi(h,k) = [f(p+h,q+k) - f(p+h,q) - kD_2 f(p,q)] + [f(p+h,q) - f(p,q) - hD_1(p,q)]$$

We will show separately that each part is small enough to be an error term. The second term is $o(h)$ and hence $o(\|(h,k)\|)$ by the definition of $D_1 f(p,q)$. For the first term, let $\phi(t) = f(p+h, q+tk)$ for a given fixed h, k. Then ϕ is differentiable and by the chain rule we have $\phi'(t) = D_2 f(p+h, q+tk) \cdot k$. By the mean value theorem, there exists a point $t(h,k) \in (0,1)$ such that $\phi(1) - \phi(0) = \phi'(t)$. Hence, the first term becomes

$$\phi(1) - \phi(0) - kD_2 f(p,q) = k[D_2 f(p+h, q+tk) - D_2 f(p,q)]$$

As $(h,k) \to (0,0)$, we have $(p+h, q+tk) \to (p,q)$. By continuity of $D_2 f$ at a, the term is $o(k)$ and hence $o(\|(h,k)\|)$. □

11.4 Mean value inequality

The mean value theorem cannot be extended verbatim to higher dimensional spaces, since there can be multiple paths between points.

Theorem. Let $U \subset \mathbb{R}^m$ be open, and $f : U \to \mathbb{R}^n$ be differentiable at every $z \in U$. Let $a, b \in U$ such that the line segment connecting a, b given by

$$[a, b] = \{(1-t)a + tb : 0 \le t \le 1\}$$

is contained inside U. Suppose there exists $M \ge 0$ such that for all $z \in [a, b]$, we have $\|f'(z)\| \le M$. Then

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

Proof. Let $u = b - a$ and $v = f(b) - f(a)$. Without loss of generality, let $u \ne 0$. Let $\gamma(t) = a + tu$, so $f \circ \gamma$ is defined on the open set $\gamma^{-1}(U)$, and is differentiable with derivative

$$(f \circ \gamma)'(t) = f'(\gamma(t))(\gamma'(t)) = f'(a + tu)(u)$$

Now,

$$\|f(b) - f(a)\|^2 = \langle f(b) - f(a), v\rangle = \langle (f \circ \gamma)(1) - (f \circ \gamma)(0), v\rangle$$

Let $\phi(t) = \langle (f \circ \gamma)(t), v\rangle$. Note that ϕ is differentiable since the inner product is linear. The derivative is

$$\phi'(t) = \langle (f \circ \gamma)'(t), v\rangle = \langle f'(a + tu)(u), v\rangle$$

By the mean value theorem, there exists $\theta \in (0, 1)$ such that $\phi(1) - \phi(0) = \phi'(\theta)$. Then, by the Cauchy–Schwarz inequality,

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

Hence,

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

as required. □

11.5 Zero derivatives

Corollary. Let U be an open, connected subset of $\mathbb{R}^m$, and $f : U \to \mathbb{R}^n$ be differentiable at every U. If $f'(a) = 0$ for all $a \in U$, then f is constant.

Proof. If $a, b \in U$ satisfy $[a, b] \subset U$, then by the mean value inequality we have

$$\|f(b) - f(a)\| \le \|b - a\| \sup_{z \in [a,b]} \|f'(z)\| = 0$$

Hence $f(a) = f(b)$. For an arbitrary $x \in U$, there exists $r > 0$ such that $\mathcal{D}_r(x) \subset U$. This open ball is convex, so for all $y \in \mathcal{D}_r(x)$ we have $f(y) = f(x)$. Hence f is locally constant; every point has a neighbourhood on which f is constant. Since U is connected, f is constant (refer to the derivation from the example sheet). □

11.6 Inverse function theorem

Remark. Let $V \subset \mathbb{R}^m$ and $W \subset \mathbb{R}^n$ be open sets. Let $f : V \to W$ be a bijection. Let $a \in V$, and let f be differentiable at a, and the inverse $f^{-1} : W \to V$ is differentiable at $f(a)$. Denoting $S = f'(a), T = \left(f^{-1}\right)'(f(a))$, we can use the chain rule to find

$$TS = \left(f^{-1} \circ f\right)'(a); \quad ST = \left(f \circ f^{-1}\right)'(f(a))$$

The identity function is linear so its derivative is the identity. Hence TS is the identity on $\mathbb{R}^m$ and ST is the identity on $\mathbb{R}^n$. Hence, $m = \operatorname{tr}(TS) = \operatorname{tr}(ST) = n$. So in order for f to be a bijection, the dimensions of the spaces must match. Hence $f'(a)$ is an invertible matrix. This proves that $\mathbb{R}^m, \mathbb{R}^n$ are not homeomorphic in such a way that the maps between them are differentiable. We aim now to prove an inverse; if f is differentiable and f' is invertible, then f is locally a bijection between neighbourhoods.

Definition. Let $U \subset \mathbb{R}^m$ be open, and $f : U \to \mathbb{R}^n$ be a function. We say that f is differentiable on U if f is differentiable at a for all $a \in U$. Then, the *derivative of f on U* is the function $f' : U \to L(\mathbb{R}^m, \mathbb{R}^n)$ mapping points to their derivatives. We say that f is a *C^1-function on U* if f is continuously differentiable on U; f is differentiable on U and $f' : U \to L(\mathbb{R}^m, \mathbb{R}^n)$ is a continuous function.

Theorem. Let $U \subset \mathbb{R}^n$ be open. Let $f : U \to \mathbb{R}^n$ be a C^1-function. Let $a \in U$, and let $f'(a)$ be an invertible linear map $f'(a) : L(\mathbb{R}^n)$. Then there exist open sets V, W such that $a \in V, f(a) \in W, V \subset U$ and $f|_V : V \to W$ is a bijection with inverse function $g : W \to V$. Further, g is a C^1-function, and

$$g'(y) = [f'(g(y))]^{-1}$$

Proof. We first show that without loss of generality we can let $a = f(a) = 0$ and $f'(a) = I$. To see this, let $T = f'(a)$ and define $h(x) = T^{-1}(f(x + a) - f(a))$. Then, h is defined on $U - a$, which is open. In particular, $U - a$ is an open neighbourhood of zero. By the chain rule, h is differentiable with $h'(x) = T^{-1} \circ f'(x + a)$. For $x, y \in U - a$, we then have

$$\|h'(x) - h'(y)\| = \left\|T^{-1} \circ (f'(a + x) - f'(a + y))\right\| \leq \left\|T^{-1}\right\| \cdot \|f'(a + x) - f'(a + y)\|$$

It then follows that h is a C^1-function, and that $h(0) = 0, h'(0) = T^{-1} \circ T = I$. We have transformed into a coordinate system where $a = f(a) = 0$ and $f'(a) = I$. If we can prove the result for this coordinate system, we can translate back using $f(x) = T(h(x - a)) + f(a)$.

Now, let $f(0) = 0$ and $f'(0) = I$. Since f' is continuous, there exists $r > 0$ such that $\mathcal{B}_r(0) \subset U$ and for all $x \in U$, we have

$$\|f'(x) - f'(0)\| = \|f'(x) - I\| \leq \frac{1}{2}$$

We intend to show that for all $x, y \in \mathcal{B}_r(0)$, we have $\|f(x) - f(y)\| \geq \frac{1}{2}\|x - y\|$. Indeed, define $p\colon U \to \mathbb{R}^n$ by $p(x) = f(x) - x$. Then $p'(x) = f'(x) - I$. Then, $\|p'(x)\| \leq \frac{1}{2}$ for all $x \in \mathcal{B}_r(0)$. By the mean value inequality, $\|p(x) - p(y)\| \leq \frac{1}{2}\|x - y\|$ for all $x, y \in \mathcal{B}_r(0)$. Hence,

$$\|f(x) - f(y)\| = \|(p(x) + x) - (p(y) + y)\| \geq \|x - y\| - \|p(x) - p(y)\| \geq \frac{1}{2}\|x - y\|$$

So we have proven the bound as claimed. Now, let $s = \frac{r}{2}$. We will show that $f(\mathcal{D}_r(0)) \subset \mathcal{D}_s(0)$. More precisely, we will show that for all $w \in \mathcal{D}_s(0)$ there exists a unique $x \in \mathcal{D}_r(0)$ such that $f(x) = w$. Let $w \in \mathcal{D}_s(0)$ be fixed. We now define, for all $x \in \mathcal{B}_r(0)$, the function $q(x) = w - f(x) + x = w - p(x)$. Note that $f(x) = w$ if and only if $q(x) = x$. We will show that q is a contraction mapping, and that there exists a fixed point. Since $p(0) = f(0) - 0 = 0$, we have for all $x \in \mathcal{B}_r(0)$ that

$$\|q(x)\| \leq \|w\| + \|p(x)\| = \|w\| + \|p(x) - p(0)\| \leq \|w\| + \frac{1}{2}\|x - 0\| = \frac{1}{2}\|x\| < s + \frac{1}{2}r$$

Hence, $q(\mathcal{B}_r(0)) \subset \mathcal{D}_r(0) \subset \mathcal{B}_r(0)$. We now show q is a contraction mapping. For $x, y \in \mathcal{B}_r(0)$, we have

$$\|q(x) - q(y)\| = \|p(x) - p(y)\| \leq \frac{1}{2}\|x - y\|$$

Hence $q\colon \mathcal{B}_r(0) \to \mathcal{B}_r(0)$ really is a contraction mapping on the non-empty, complete metric space $\mathcal{B}_r(0)$. By the contraction mapping theorem, there exists a unique $x \in \mathcal{B}_r(0)$ such that $q(x) = x$. But since $q(\mathcal{B}_r(0)) \subset \mathcal{D}_r(0)$, we must have $x \in \mathcal{D}_r(0)$. In particular, there exists a unique $x \in \mathcal{D}_r(0)$ such that $f(x) = w$.

Now, let $W = \mathcal{D}_s(0), V = \mathcal{D}_r(0) \cap f^{-1}(W)$. Then, we will now show that $f|_V : V \to W$ is a bijection with inverse $g\colon W \to V$ which is continuous. First, W is open and $f(0) = 0 \in W$. Since f is continuous, $f^{-1}(W)$ is open. Hence V is open, as the intersection of two open sets. We have $0 \in V$. By the previous paragraph, $f|_V : V \to W$ is a bijection since for every point in W there exists a unique point in V mapping to it. Finally, let $u, v \in W$. Let $x = g(u), y = g(v)$. Then,

$$\|g(u) - g(v)\| = \|x - y\| \leq 2\|f(x) - f(y)\| = 2\|u - v\|$$

Hence g is 2-Lipschitz and hence continuous. Now it suffices to show g is C^1, and for all $y \in W$ we have $g'(y) = [f'(g(y))]^{-1}$. This part of the proof is non-examinable. □

12 Second derivatives

12.1 Definition

Definition. Let $U \subset \mathbb{R}^m$ be an open set, and $f\colon U \to \mathbb{R}^n$. Let $a \in U$. Suppose that there exists an open neighbourhood V of a contained within U, and f is differentiable on V. We say that f is *twice differentiable* at a if $f'\colon V \to L(\mathbb{R}^m \to \mathbb{R}^n)$ is differentiable at a. We write $f''(a)$ for the derivative of f' at a, called the *second derivative* of f at a. Note that $f''(a) \in L(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^n))$.

Remark. We can visualise the second derivative as a bilinear map instead of a nested sequence of linear maps. Note,

$$L(\mathbb{R}^m, L(\mathbb{R}^m, \mathbb{R}^n)) \sim \mathrm{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$$

where $\mathrm{Bil}(X \times Y, Z)$ is the vector space of bilinear maps from $X \times Y$ to Z. For $h, k \in \mathbb{R}^m$, and T is the second derivative, we can say $T(h)(k) = \widetilde{T}(h,k)$ where $\widetilde{T}$ is a bilinear map. From now on, this bilinear map notation will be used, and T and $\widetilde{T}$ will be identified as the same.

Proposition. Let $U \subset \mathbb{R}^m$ be open, $f : U \to \mathbb{R}^n$ be a function, and $a \in U$. Let f be differentiable on an open neighbourhood V of A contained in U. Then f is twice differentiable at a if and only if there exists a bilinear map $T \in \mathrm{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ such that for every $k \in \mathbb{R}^m$, we have

$$f'(a+h)(k) = f'(a)(k) + T(h,k) + o(\|h\|)$$

Then $T = f''(a)$.

Proof. Suppose f is twice differentiable at a. Then f' is differentiable at a. So,

$$f'(a+h) = f'(a) + f''(a)(h) + \|h\| \cdot \varepsilon(h)$$

All terms are linear maps $L(\mathbb{R}^m, \mathbb{R}^n)$. In particular, ε is defined on $V - a \to L(\mathbb{R}^m, \mathbb{R}^n)$ such that $\varepsilon(0) = 0$ and ε is continuous at zero. If we evaluate this equation at a fixed $k \in \mathbb{R}^m$,

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

Here, $f''(a)$ is a bilinear map. Further,

$$\|\varepsilon(h)(k)\| \leq \|\varepsilon(h)\| \cdot \|k\| \to 0$$

Hence, $\|h\| \cdot \varepsilon(h)(k) = o(\|h\|)$. Conversely, suppose T is a bilinear map and

$$\frac{f'(a+h)(k) - f'(a)(k) - T(h,k)}{\|h\|} \to 0$$

for any fixed k, as $h \to 0$. We need to show that

$$\varepsilon(h) = \frac{f'(a+h) - f'(a) - T(h)}{\|h\|} \to 0$$

in the space $L(\mathbb{R}^m, \mathbb{R}^n)$. We know that for a fixed $k \in \mathbb{R}^m$, $\varepsilon(h)(k) \to 0$ in $\mathbb{R}^n$ as $h \to 0$. It then follows that

$$\|\varepsilon(h)\| = \sqrt{\sum_{i=1}^m \|\varepsilon(h)(e_i)\|^2} \to 0$$

since we are in a finite-dimensional vector space. □

Example. Let $f : \mathbb{R}^m \to \mathbb{R}^n$ be linear. Then f is differentiable on $\mathbb{R}^m$ with $f'(a) = f$ for all a. Hence $f' : \mathbb{R}^m \to L(\mathbb{R}^m, \mathbb{R}^n)$ sends a to f for all a. So this is a constant function, so has derivative $f''(a) = 0$.

Example. Let $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^p$ be bilinear. Then f is differentiable on $\mathbb{R}^m \times \mathbb{R}^n$ and for all $(a,b) \in \mathbb{R}^m \times \mathbb{R}^n$, we have

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

Note that this is linear in (a,b) for a fixed (h,k). Hence, $f' : \mathbb{R}^m \times \mathbb{R}^n \to L(\mathbb{R}^m, \mathbb{R}^n, \mathbb{R}^p)$ is linear. Hence this is differentiable, and its derivative is

$$f''(a,b) = f' \in L(\mathbb{R}^m, \mathbb{R}^n, L(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R}^p)) \simeq \mathrm{Bil}((\mathbb{R}^m \times \mathbb{R}^n) \times (\mathbb{R}^m \times \mathbb{R}^n), \mathbb{R}^p)$$

Example. Let $f : M_n \to M_n$ be defined by $f(A) = A^3$. Let A be fixed. Then,

$$\begin{aligned} f(A+H) &= (A+H)^3 = A^3 + A^2H + AHA + HA^2 + AH^2 + HAH + H^2A + H^3 \\ &= f(A) + (A^2H + AHA + HA^2) + o(\|H\|) \end{aligned}$$

Hence f is differentiable at A and

$$f'(A)(H) = A^2H + AHA + HA^2$$

Thus, if $n = 1$, we have commutativity and hence $f'(A) = 3A^2$. So f is differentiable on M_n. For a fixed A and fixed K, the second derivative is given by

$$\begin{aligned} f'(A+H)(K) &= (A+H)^2K + (A+H)K(A+H) + K(A+H)^2 \\ &= \underbrace{(A^2K + AKA + KA^2)}_{f'(A)(K)} \\ &\quad + (AHK + HAK + AKH + HKA + KAH + KHA) + (H^2K + HKH + KH^2) \end{aligned}$$

The term $T(H,K) = (AHK + HAK + AKH + HKA + KAH + KHA)$ is bilinear in H and K as required. So the second derivative is T. In one dimension, this is equivalent to saying $f''(A) = 6A$.

12.2 Second derivatives and partial derivatives

Let U be open in $\mathbb{R}^n$, let $f : U \to \mathbb{R}^n$, and let $a \in U$. Let f be twice differentiable at a, so f is differentiable on some open neighbourhood V of a contained within U, and $f' : V \to L(\mathbb{R}^m, \mathbb{R}^n)$ is differentiable at a. Recall that

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

Evaluating at a fixed k,

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

Let $u, v \in \mathbb{R}^m \setminus \{0\}$ be directions. Let $k = v$. Then,

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

Hence, the map $D_vf : V \to \mathbb{R}^n$ maps $x \mapsto D_vf(x) = f'(x)(v)$. Then this map is differentiable at a and

$$(D_vf)'(a)(h) = f''(a)(h,v)$$

Hence there exist directional derivatives.

$$D_uD_vf(a) \stackrel{\text{def}}{=} D_u(D_vf)(a) = (D_vf)'(a)(u) = f''(a)(u,v)$$

In particular, we have

$$D_iD_jf(a) = f''(a)(e_i, e_j)$$

for $1 \leq i, j \leq m$.

12.3 Symmetry of mixed directional derivatives

Theorem. Let U be open in $\mathbb{R}^n$, let $f : U \to \mathbb{R}^n$, and let $a \in U$. Let f be twice differentiable on an open set V with $a \in V \subset U$. Let $f'' : V \to \mathrm{Bil}(\mathbb{R}^m \times \mathbb{R}^m, \mathbb{R}^n)$ be continuous at a. Then, for all directions $u, v \in \mathbb{R}^m \setminus \{0\}$, we have

$$D_u D_v f(a) = D_v D_u f(a)$$

Equivalently,

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

In other words, f'' is a symmetric bilinear map.

Proof. Without loss of generality we can let $n = 1$. Indeed, we have

$$(D_u f)_j(x) = [D_u f(x)]_j = [f'(x)(u)]_j = f_j'(x)(u) = D_u f_j(x)$$

Hence, $(D_u f)_j = D_u f_j$. For v:

$$(D_v D_u f)_j = D_v (D_u f)_j = D_v D_u f_j$$

So it is sufficient to show that $D_v D_u f_j(a) = D_u D_v f_j(a)$. Now, consider

$$\phi(s, t) = f(a + su + tv) - f(a + tv) - f(a + su) + f(a)$$

for $s, t \in \mathbb{R}$. Let s, t be fixed, and consider

$$\psi(y) = f(a + yu + tv) - f(a + yu)$$

Note that $\phi(s, t)$ can be written as

$$\phi(s, t) = \psi(s) - \psi(0)$$

The term $\psi(s) - \psi(0)$ can be interpreted as $(f(a + su + tv) - f(a + tv)) - (f(a + su) - f(a))$, which is the second difference given by the function when traversing the parallelogram with sides su, tv. By the mean value theorem, there exists $\alpha(s, t) \in (0, 1)$ such that

$$\phi(s, t) = \psi(s) - \psi(0) = s\psi'(\alpha s) = s[D_u f(a + \alpha su + tv) - D_u f(a + \alpha su)]$$

Now, applying the mean value theorem to the function $y \mapsto D_u f(a + \alpha su + yv)$, we have

$$\phi(s, t) = st D_v D_u f(a + \alpha su + \beta tv)$$

for $\beta(s, t) \in (0, 1)$. Now,

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

Since f'' is continuous at a, we can let $s, t \to 0$ and find

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

Now, we can repeat the above using

$$\psi(y) = f(a + su + yv) - f(a + yv)$$

This calculates the second difference from above, but using the other path. We can find

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

as required. $\square$