Eidgenössische Technische Hochschule Zürich

Analysis II: Multiple Variables

Spring Semester 2026

Lecture Notes

Contents

References

[1] Serra, Joaquim: Skript zur Vorlesung Analysis II. ETH Zürich. 2026

Analysis II: Multiple Variables

Fynn Krebser–fkrebser@student.ethz.ch

Spring Semester 2026

Lec 1

1 Metric Spaces

In analysis II, will be working mostly in $\mathbb{R}^n$, which is a vector space defined as

$$\mathbb{R}^n \coloneqq \{x = (x_1, \dots, x_n), x_i \in \mathbb{R}\}.$$

Here, we call $n \in \mathbb{N}$ the DIMENSION. We want to add some extra structure in particular the EUCLIDEAN STRUCTURE

Definition 1.1: Euclidean structure

On $\mathbb{R}^n$ we define the EUCLIDEAN NORM as

$$\|x\| \coloneqq \sqrt{x_1^2 + \cdots + x_n^2},$$

which describes the length of the vector x. We also define the SCALAR PRODUCT as

$$x \cdot y = \langle x, y\rangle \coloneqq \sum_{i=1}^n x_i y_i,$$

furthermore the EUCLIDEAN DISTANCE is defined as

$$d(x,y) \coloneqq \|x - y\|.$$

Lemma 1.2: Triangle inequality

For all $x, y, z \in \mathbb{R}^n$ we have

$$\|z - x\| \leq \|y - x\| + \|z - y\|.$$

The lemma states that the shortest path between two points is a straight line.

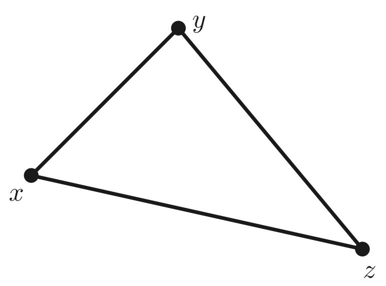

Figure 1: Triangle Inequality

Proof. Equivalently, $\|a + b\| \leq \|a\| + \|b\|$ for all $a, b \in \mathbb{R}^n$. This is because if we let $a = y - x$ and $b = z - y$, then $a + b = z - x$. Equivalently, squaring both sides, we have

$$\|a + b\|^2 \leq \|a\|^2 + \|b\|^2 + 2\|a\|\|b\|. \tag{1.1}$$

By definition of the norm, we have

$$\sum_{i=1}^n (a_i + b_i)^2 = \sum_{i=1}^n a_i^2 + b_i^2 + 2a_i b_i = \|a\|^2 + \|b\|^2 + 2a \cdot b.$$

Together with (1.1), we have that our statement is equivalent to $a \cdot b \leq \|a\|\|b\|$, which is the Cauchy-Schwarz inequality. □

Lemma 1.3: Cauchy-Schwarz inequality

For all $x, y \in \mathbb{R}^n$ we have

$$x \cdot y \leq \|x\|\|y\|.$$

Proof. If either $x = 0$ or $y = 0$, then the statement becomes $0 \leq 0$, which is true.

Let now $x, y \in \mathbb{R}^n$ be nonzero. Now for every $\lambda > 0$, we have

$$2x \cdot y = 2\sum_{i=1}^n \lambda x_i \frac{y_i}{\lambda} \leq \sum_{i=1}^n \lambda^2 x_i^2 + \frac{y_i^2}{\lambda^2} = \lambda^2\|x\|^2 + \frac{1}{\lambda^2}\|y\|^2,$$

Since $2ab \leq a^2 + b^2$. Since x, y are nonzero we can take $\lambda^2 = \frac{\|y\|}{\|x\|}$, which gives us

$$2x \cdot y \leq 2\|x\|\|y\|.$$

□

In order to not always define convergence and continuity in other subjects, we will now introduce the notion of a METRIC SPACE, which is a set with a distance function defined on it.

Definition 1.4: Metric space

A metric space is a pair (X, d), where X is a set and $d : X \times X \to [0, \infty)$ is a function such that for all $x, y, z \in X$ we have

- $d(x, y) = 0$ if and only if $x = y$,
- $d(x, y) = d(y, x)$, and
- $d(x, z) \leq d(x, y) + d(y, z)$.

Example 1.5:

1. $(\mathbb{R}^n, d_{\text{Euclidean}})$ is a metric space.
2. $(\mathbb{R}^2, d_{\text{NY}}(x, y))$ is a metric space, where the MANHATTAN METRIC is defined as

$$d_{\text{NY}}(x, y) = |x_1 - y_1| + |x_2 - y_2|.$$

3. Given (X, d) a metric space and $Y \subseteq X$, then $(Y, d|_{Y \times Y})$ is a metric space.

For example if we take $X = \mathbb{R}^2$ and Y to be the unit circle, then we also have the metric space $(Y, d_{\text{Euclidean}}|_{Y \times Y})$.

In general it is useful to first think about the proof in $\mathbb{R}^n$, and then to see if it can be adapted to a more general metric space. This is often easier as we tend to have a better intuition for $\mathbb{R}^n$.

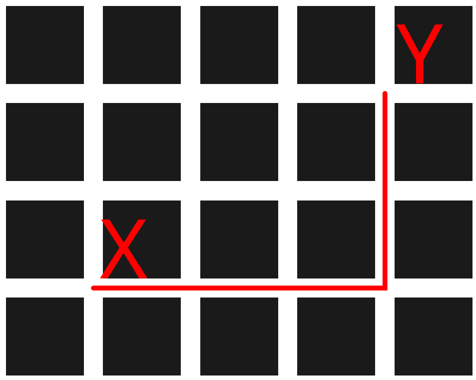

Figure 2: Manhattan (Glarus) metric visualized

Given $a < b \in \mathbb{R}$, and

$$X = \{f : [a, b] \to \mathbb{R}, f \text{ continuous }\},$$

then how can we find the distance between two functions $f, g \in X$? It can for example be defined as

$$d(f, g) = \max_{[a,b]} |f(x) - g(x)|.$$

Another distance would be

$$d(f, g) = \sqrt{\int_a^b |f(x) - g(x)|^2 dx}.$$

Lec 2 As an exercise, proof that both of these are metrics on X.

In the following, let (X, d) be a metric space. We want to define the notion of convergence and continuity in X.

Definition 1.6: Sequence

Given a set X, we call a map $x : \mathbb{N} \to X$ a SEQUENCE in X. Instead of writing $x(n)$, we write x_n as the n-th term of the sequence $\in X$. To denote the full sequence we write $(x_n)_{n\geq 0}$.

Definition 1.7: Convergent Sequence

We say $(x_n)_{n\geq 0}$ has limit $x \in X$ if $d(x_n, x) \to 0$ as real numbers. x is called the LIMIT of the sequence, and we write $\lim_{n\to\infty} x_n = x$.

Another way to phrase this is that for every $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have $d(x_n, x) < \varepsilon$.

Notice that the limit has to be in the set unlike in Analysis I, as otherwise the metric is not defined.

Lemma 1.8: Uniqueness of the limit

Given a sequence $(x_n)_{n\geq 0}$ such that $x_n \to x$ and $x_n \to y$, then $x = y$.

Proof. Assume by contradiction that $x \neq y$ and let $3\varepsilon = d(x, y) > 0$. By the definition of the limit, $\exists N_x$ such that $d(x_n, x) < \varepsilon \forall n \geq N_x$, and $\exists N_y$ such that $d(x_n, y) < \varepsilon \forall n \geq N_y$.

Let $N = \max(N_x, N_y)$, then for all $n \geq N$ we have

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

by triangle inequality, which is a contradiction. $\square$

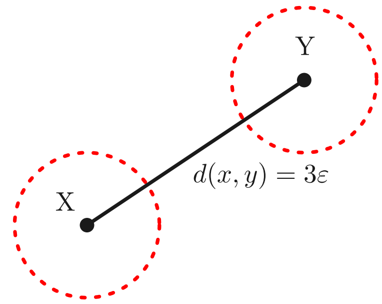

Figure 3: Uniqueness of the limit

Definition 1.9: Subsequence

Given a sequence $(x_n)_{n\geq 0}$, we define a SUBSEQUENCE as a sequence of the form

$$(x_{f(k)})_{k\geq 0},$$

where $f : \mathbb{N} \to \mathbb{N}$ is a strictly increasing function.

Usually, we will write x_{n_k} instead of $x_{f(k)}$ for the subsequence.

Definition 1.10: Accumulation Point

Given $Y \subset X$, we say that $y \in X$ is an ACCUMULATION POINT of Y if there exists a sequence $(y_n)_{n\geq 0}$ in Y such that $y_n \to y$.

Given $(x_n)_{n\geq 0}$ as a sequence of X we say that x is an ACCUMULATION POINT of $(x_n)_{n\geq 0}$ if there exists a subsequence $(x_{n_k})_{k\geq 0}$ such that $x_{n_k} \to x$.

Lemma 1.11:

Given a sequence $(x_n)_{n\geq 0}$. The sequence converges to x if and only if $\forall (x_{n_k})_{k\geq 0}$ we have $x_{n_k} \to x$.

Proof. We rewrite the statement to $x_n \to x \Leftrightarrow (x_{n_k}) \to y \Rightarrow y = x$.
$\Leftarrow$: Taking the particular subsequence $x_{n_k} = x_n$, we have $x_n \to y$, and thus $y = x$.

$\Rightarrow$: Let $f : \mathbb{N} \to \mathbb{N}$ be a strictly increasing function. Since $x_n \to x$, for every $\varepsilon > 0$ there exists N such that for all $n \geq N$ we have $d(x_n, x) < \varepsilon$. Since $f(n)$ is increasing, $f(n) \geq n$, thus for all $n \geq N$ we have $d(x_{f(n)}, x) < \varepsilon$, which means that $x_{f(n)} \to x$. $\square$

Lemma 1.12:

Given a sequence $(x_n)_{n\geq 0}$. Then $x_n \to x$ if and only if every subsequence $(x_{n_k})_{k\geq 0}$ has a subsequence $(x_{n_{k_l}})_{l\geq 0}$ such that $x_{n_{k_l}} \to x$.

Proof. See later... We will do it later try first at home. $\square$

Definition 1.13: Cauchy Sequence

We say that a sequence $(x_n)_{n\geq 0}$ is a CAUCHY SEQUENCE if $\forall \varepsilon > 0$, $\exists N$ such that $d(x_n, x_m) < \varepsilon$ whenever $n, m \geq N$.

Definition 1.14: Complete Metric Space

We say that a metric space (X, d) is COMPLETE if every Cauchy sequence in X converges (to a limit in X).

Example 1.15:

1. $(\mathbb{R}, d_{\text{Euclidean}})$ is a complete metric space.

Theorem 1.16:

$(\mathbb{R}^n, d_{\text{Euclidean}})$ is complete.

Lemma 1.17:

Given a sequence $(x_m)_{m\geq 0} \subset \mathbb{R}^n$, then $x_m \to x \in \mathbb{R}^n$ if and only if $x_{m,i} \to x_i$ for all $i = 1, \ldots, n$, where $x_m = (x_{m,1}, \ldots, x_{m,n})$ and $x = (x_1, \ldots, x_n)$.

Proof. $\Rightarrow$: Since $x_m \to x$, $\forall \varepsilon > 0 \exists N$ such that $\|x_n - x\| < \varepsilon$ for all $n \geq N$. In particular,

$$|x_{m,i} - x_i| \leq \sqrt{\sum_{j=1}^{n} |x_{m,j} - x_j|^2} = \|x_m - x\| < \varepsilon.$$

$\Leftarrow$: Since $x_{m,i} \to x_i$ for all $i = 1, \ldots, n$, $\forall \varepsilon > 0 \exists N_i$ such that $|x_{m,i} - x_i| < \frac{\varepsilon}{\sqrt{n}}$ for all $m \geq N_i$. Let $N = \max(N_1, \ldots, N_n)$, then for all $m \geq N$ we have

$$\|x_m - x\| = \sqrt{\sum_{i=1}^{n} |x_{m,i} - x_i|^2} < \sqrt{\sum_{i=1}^{n} \frac{\varepsilon^2}{n}} = \varepsilon.$$

□

Proof. [Proof of Theorem 1.16] Given $(x_m)_{m\geq 0}$ Cauchy, let us show $\exists x$ such that $x_m \to x$.

By Lemma 1.17, it suffices to show that $\forall i = 1, \ldots, n$ we have that $(x_{m,i})_{m\geq 0}$ is Cauchy in $\mathbb{R}$.

But since Cauchy sequences in $\mathbb{R}$ converge, $\exists x_i$ such that $x_{m,i} \to x_i$ for all $i = 1, \ldots, n$. Thus by Lemma 1.17, we have $x_m \to x$, where $x = (x_1, \ldots, x_n)$.

□

Tip 1.18:

An often seen counter example is the discrete metric space, where $d(x, y) = 1$ if $x \neq y$ and 0 otherwise.

Lec 3

1.1 Topology of metric spaces

We would like to define the notion of open sets in a metric space, which will allow us to define continuity and other topological properties of metric spaces.

Definition 1.19: **Open Ball**

Define the OPEN BALL of radius $r > 0$ centered at $x \in X$ as

$$B_r(x) = B(x, r) := \{y \in X : d(x, y) < r\}.$$

Definition 1.20: **Open and Closed Sets**

We say that $U \subset X$ is OPEN if for every $x \in U, \exists r > 0$ such that $B(x, r) \subset U$.

$A \subset X$ is CLOSED, if $X \setminus A$ is open.

Exercise 1.21:

Show that $(0, 1) \times (0, 1) \subset \mathbb{R}^2$ is open.

Show that $[0, 1] \times [0, 1] \subset \mathbb{R}^2$ is closed.

The topology associated to d is

$$\mathcal{T} = \{U \subset X : U \text{ is open}\}.$$

Lemma 1.22:

Arbitrary unions of open sets are open.

$$\{U_i\}_{i\in\mathcal{I}}, U_i \subset X \text{ open} \Rightarrow \bigcup_{i\in\mathcal{I}} U_i \text{ is open.}$$

Finite intersections of open sets are open.

$$U_1, \ldots, U_k \subset X \text{ open} \Rightarrow \bigcap_{i=1}^{k} U_i \text{ is open.}$$

Proof. Take $x \in \bigcup_{i\in I} U_i$, then $x \in U_j$ for some $j \in I$. Since U_j is open, $\exists r > 0$ such that $B(x, r) \subset U_j \subset \bigcup_{i\in I} U_i$, thus $\bigcup_{i\in I} U_i$ is open.

Take $x \in \bigcap_{i=1}^{k} U_i$, then $x \in U_i$ for all $i = 1, \ldots, k$. Since U_i is open, $\exists r_i > 0$ such that $B(x, r_i) \subset U_i$ for all $i = 1, \ldots, k$. Let $r = \min(r_1, \ldots, r_k)$, then $B(x, r) \subset B(x, r_i) \subset U_i$ for all $i = 1, \ldots, k$, thus $\bigcap_{i=1}^{k} U_i$ is open. □

Lemma 1.23:

Finite unions of closed sets are closed.

Arbitrary intersections of closed sets are closed.

Proof. Notice that $X \setminus \bigcup_{i=1}^{k} A_i = \bigcap_{i=1}^{k} (X \setminus A_i)$, and $X \setminus A_i$ is open for all $i = 1, \ldots, k$, thus $\bigcup_{i=1}^{k} A_i$ is closed. □

Example 1.24: **Finite is Important**

Take $(\mathbb{R}, d_{\text{Euclidean}})$ and $U_k = (-\frac{1}{k}, \frac{1}{k})$. Then

$$\bigcap_{k=1}^{\infty} U_k = \{0\}.$$

Which is not open.

Definition 1.25: **Interior, Closure and Boundary**

Given $\Omega \subset X$, we define

- $\text{int}\Omega = \Omega^\circ = \{U \subset \Omega \mid U \text{ is open }\}$ as the INTERIOR of Ω,
- $\overline{\Omega} = \{x \in X \mid \exists (x_n)_{n\geq 0} \subset \Omega, x_n \to x\}$ as the CLOSURE of Ω, and
- $\partial\Omega = \overline{\Omega} \setminus \Omega^\circ$ as the BOUNDARY of Ω.

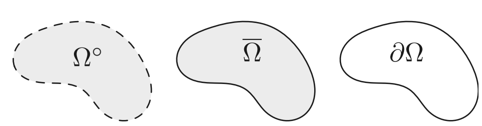

Figure 4: Interior, closure and boundary of a set Ω

Lemma 1.26:

$U \subset X$ is open if and only if whenever $(x_n)_{n\geq 0} \subset X$ such that $x_n \to x \in U$, then $x_n \in U$ eventually.

$A \subset X$ is closed if and only if whenever $(x_n)_{n\geq 0} \subset A$ such that $x_n \to x$, then $x \in A$.

Proof. 1, $\Rightarrow$: Take $x \in U$. By definition of open set, $\exists r > 0$ such that $B(x,r) \subset U$. Since $x_n \to x$, $\exists N$ such that $x_n \in B(x,r) \forall n \geq N$. (Since this is equivalent to $d(x_n, x) < r$ for all $n \geq N$). Thus $x_n \in U$ eventually.

1, $\Leftarrow$: We argue by contraposition. Since U is not open, $\exists x \in U$ such that $\forall r > 0, B(x,r) \not\subset U$. This is the same as saying $\exists x_r \in B(x,r) \cap X \setminus U$.

Taking $r = \frac{1}{n}$ and calling $x_n = x_{\frac{1}{n}} \to x$. Since $x_n \in X \setminus U$ for all $n \geq 0$, we have that $x_n \to x \in U$ but $x_n \notin U$ for all $n \geq 0$, contradicting that $x_n \in U$ eventually.

2. Exercise □

Exercise 1.27:

Given (X,d) complete, $A \subset X$ closed, show that (A,d) is complete.

Definition 1.28: Continuity

Given $f : X \to Y$ with (X, d_X) and (Y, d_Y) metric spaces, we say that f is CONTINUOUS if one of the following 3 equivalent properties hold:

- Epsilon-Delta Continuity: $\forall x \in X, \forall \varepsilon > 0, \exists \delta > 0$ such that:

$$\forall x' \in X, d_X(x,x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon.$$

Which is equivalent to saying that

$$f(B(x,\delta)) \subset B(f(x), \varepsilon).$$

- Sequential Continuity: $(x_n)_{n\geq 0} \subset X$,

$$x_n \to x \Rightarrow f(x_n) \to f(x).$$

- topological continuity: $\forall U \subset Y$ open, $f^{-1}(U)$ is open.

Proposition 1.29:

The three definitions of continuity are equivalent.

Proof. $1 \Rightarrow 2$: Assume $f : X \to Y$ is continuous in the epsilon-delta sense. Let $(x_n)_{n\geq 0} \subset X$ such that $x_n \to x$. Given $\varepsilon > 0$, by continuity, $\exists \delta > 0$ such that $f(B(x,\delta)) \subset B(f(x), \varepsilon)$. Since $x_n \to x$, $\exists N$ such that $x_n \in B(x,\delta)$ for all $n \geq N$. Thus $f(x_n) \in B(f(x), \varepsilon)$ for all $n \geq N$, which means that $f(x_n) \to f(x)$.

$2 \Rightarrow 3$: Assume f is not topologically continuous. Then $\exists U \subset Y$ open such that $f^{-1}(U)$ is not open. Hence $\exists x \in f^{-1}(U)$ and a sequence $(x_n)_{n\geq 0} \subset X \setminus f^{-1}(U)$ such that $x_n \to x$. Then $f(x) \in U$ but $f(x_n) \notin U$ for all n. But U is open so $f(x_n) \not\to f(x)$, contradicting sequential continuity.

Lec 4

$3 \Rightarrow 1$: Let $x \in X$ and $\varepsilon > 0$. The preimage $f^{-1}(B(f(x), \varepsilon))$ contains the point x and is open as $B(f(x), \varepsilon)$ is open and f is topologically continuous. Thus, $\exists \delta > 0$ such that $B(x,\delta) \subset f^{-1}(B(f(x), \varepsilon))$. Hence, f is $\varepsilon - \delta$ continuous. □

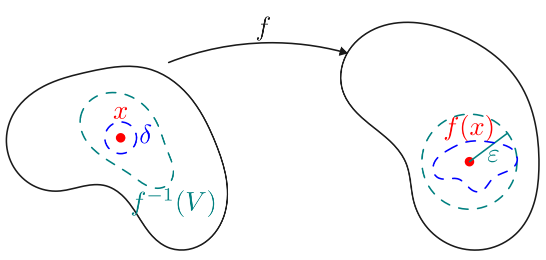

Figure 5: Topological continuity implies $\varepsilon - \delta$ continuity

Definition 1.30: Uniform and Lipschitz Continuity

Let (X, d_x) and (Y, d_Y) be metric spaces. We say that $f : X \to Y$ is UNIFORMLY CONTINUOUS if $\forall \varepsilon > 0, \exists \delta > 0$ such that

$$d_X(x,x') < \delta \Rightarrow d_Y(f(x), f(x')) < \varepsilon.$$

We say that f is LIPSCHITZ CONTINUOUS if $\exists L > 0$ such that

$$d_Y(f(x), f(x')) \leq L d_X(x,x') \forall x, x' \in X.$$

The constant L is called the LIPSCHITZ CONSTANT of f.

Example 1.31:

Fix $x_0 \in X$ and define $f := d(x, x_0)$. Then f is 1-Lipschitz.

Proof. Notice that $Y = [0, \infty)$ with the Euclidean metric. Then, for all $x, y \in X$ we have

$$d(f(x), f(y)) = |f(x) - f(y)| = |d(x,x_0) - d(y,x_0)| \leq d(x,y).$$

□

Theorem 1.32: Banach Fixed Point Theorem

Let (X,d) be a complete metric space and $T : X \to X$ λ-Lipschitz with $\lambda \in (0,1)$ (sometimes called a CONTRACTION). Then, T has a unique FIXED POINT ($\exists! x \in X$ such that $T(x) = x$).

Proof. Fix $x_0 \in X$ and define $x_1 = T(x_0), x_2 = T(x_1), \ldots, x_n = T(x_{n-1}), \ldots$. We will show that $(x_n)_{n\geq 0}$ is a Cauchy sequence, and thus converges to some $x \in X$.

Since T is a contraction,

$$\begin{aligned} d(x_{n+1}, x_n) &= d(T(x_n), T(x_{n-1})) \\ &\leq \lambda \cdot d(x_n, x_{n-1}) \leq \ldots \\ &\leq \lambda^n \cdot d(x_1, x_0). \end{aligned}$$

Since the distance is symmetric, w.l.o.g assume $m < n$. Then, by triangle inequality,

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

All terms besides λ^m are constants, and $\lambda^m \to 0$ as $m \to \infty$, thus $(x_n)_{n\geq 0}$ is a Cauchy sequence.

Since X is complete, $\exists x \in X$ such that $x = \lim_{n\to\infty} x_n$. Since T is continuous,

$$T(x) = \lim_{n\to\infty} T(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

So x is indeed a fixed point.

For uniqueness, suppose x, y are two fixed points. Then,

$$d(x,y) = d(T(x), T(y)) \leq \lambda \cdot d(x,y) < d(x,y).$$

□

We now like to extend our definition for compactness from Analysis I to metric spaces. To that extent, we need the following definition.

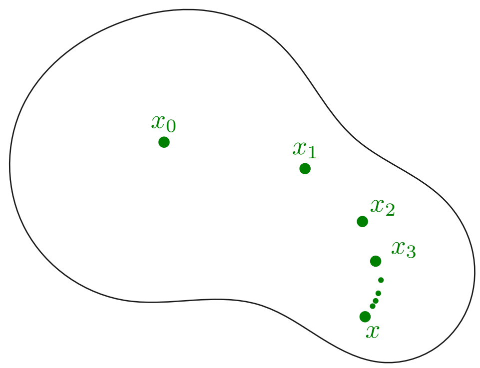

Figure 6: Proof of the Banach Fixed Point Theorem

Definition 1.33: Cover

Given a set X and $E \subset X$, we say that $\mathcal{U} = \{U_i, i \in I\}$ is a COVER of E if $E \subset \bigcup \mathcal{U} = \bigcup_{i \in I} U_i$.

If $V \subset \mathcal{U}$ is still a cover of E, we call V a SUBCOVER.

If $\mathcal{U}$ is a collection of open sets, we call it an OPEN COVER.

Definition 1.34: Compactness

A set $K \subset X$ is called

(1) SEQUENTIALLY COMPACT if for every sequence $(x_n)_{n\geq 0} \subset K$, there exists a convergent subsequence $(x_{n_k})_{k\geq 0}$ such that $x_{n_k} \to x \in K$.

(2) TOPOLOGICALLY COMPACT if for every open cover $\mathcal{U}$ of K, there exists a finite subcover $V \subset \mathcal{U}$ of K.

Example 1.35: Analysis I

Bolzano-Weierstrass: Any closed interval $[a, b] \subset \mathbb{R}$ is sequentially compact.

$[0, 1] \cap \mathbb{Q}$ is not sequentially compact. It is also not topologically compact, as $\mathbb{Q} \subset \{x_n | n \geq 0\}$ and we consider the open balls $\mathcal{U} = \{B(x_n, 2^{-n-10^3})\}$ then the total length of the intervals is $4 \cdot 2^{-10^3}$, which is less than 1, thus we cannot cover $[0, 1] \cap \mathbb{Q}$ with a finite number of intervals.

Proposition 1.36:

The two definitions of compactness are equivalent.

We will now show that (1) $\Rightarrow$ (2), and we will show the converse later.

Proof. Assume $K \subset X$ is sequentially compact. Let $\{U_i\}$ be an open cover. Thus $\forall x \in K, \exists U_i$ open, such that $x \in U_i$. Given $x \in K$, let

$$r(x) = \min\{\sup\{r > 0 : B(x, r) \subset U_i \in \mathcal{U}\}, 1\}.$$

Given $x \in K$, select U_i, such that $B\left(x, \frac{r(x)}{2}\right) \subset U_{i(x)}$[1]

Pick any $x_0 \in K$ and define

$$\mathcal{V} := \{\underbrace{U_i(x_0)}_{=U_0}, \underbrace{U_i(x_1)}_{=U_1}, \dots\},$$

where $x_1 \in K \setminus U_0$, $x_2 \in K \setminus (U_0 \cup U_1)$, and so on.

[1] We select $\frac{r}{2}$ in case the Supremum is not attained.

Doing so, unless I find a finite subcover, I will produce a sequence $x_n \in K \setminus \bigcup_{k=0}^{n-1} U_{i(x_k)}$. By sequential compactness, this sequence has a subsequence x_{n_l} such that $x_{n_l} \to x \in K$ with $r(x) > 0$. By construction $r(x_n) \to 0$! But $B_{\frac{r(x)}{2}} \subset U_{i(x)}$. Thus, for l large enough, $x_{n_l} \in U_{i(x)}$, contradicting the construction of the sequence. □ Lec 5

For the converse, the proof is a bit shorter.

Proof. Given $(x_n)_{n\geq 0} \subset K$, we want to show that there exists a subsequence $(x_{n_k})_{k\geq 0}$ such that $x_{n_k} \to x \in K$.

Assume by contradiction, $\forall x \in K$, x is not an accumulation point of $(x_n)_{n\geq 0}$ so $\forall x \in K, \exists \epsilon(x) > 0$ such that (x_n) visits $B(x, \epsilon(x))$ only finitely many times. Thus, $x_n \in K \setminus B(x, \varepsilon(x)) \forall n \geq N(x)$.

Define now

$$\mathcal{U} = \{B(x, \varepsilon(x)) | x \in K\}.$$

Since K is topologically compact,

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

This would imply that our sequence only has finitely many terms, contradicting the fact that it is a sequence. □

Corollary 1.37:

If $K \subset X$ is compact, then

1. K is closed,
2. K is complete,
3. If $A \subset X$ is closed, $K \cap A$ is compact.

Proposition 1.38:

If $f : X \to Y$ is continuous and $K \subset X$ is compact, then $f(K)$ is compact.

Proof. The goal is to show that $f(K)$ is topologically compact. Given $\mathcal{V}$ as an open cover of $f(K)$, we have that $\mathcal{U} = \{f^{-1}(V) | V \in \mathcal{V}\}$ is an open cover of K. Since K is compact, there exists a finite subcover $f^{-1}(V_1), \dots, f^{-1}(V_n)$ of K. Thus $V_1, \dots, V_n$ is a finite subcover of $f(K)$, and thus $f(K)$ is compact. □

Theorem 1.39:

Given $f : X \to \mathbb{R}$ continuous, $K \subset X$ compact, such that $\sup\{f(x)|x \in K\}$ and $\inf\{f(x)|x \in K\}$ are finite, then $\exists t \in K$ such that $f(t) = \sup\{f(x)|x \in K\}$

Proof. By definition of the supremum, $\exists (x_n)_{n\geq 0} \subset K$ such that $f(x_n) \to \sup\{f(x)|x \in K\}$ as $n \to \infty$. Since K is compact, $\exists (x_{n_k})_{k\geq 0}$ such that $x_{n_k} \to t \in K$. By continuity of f, $f(x_{n_k}) \to f(t)$ as $k \to \infty$. Thus, $f(t) = \sup\{f(x)|x \in K\}$. □

Since in this course, we will mostly work in $\mathbb{R}^n$, we will now show the following theorem.

Theorem 1.40: Heine-Borel

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

Note

We call a set $K \subset \mathbb{R}^n$ BOUNDED if $\exists M > 0$ such that

$$B(0, M) \supset K.$$

Proof. $\Rightarrow$: K is closed by corollary 1.37.

If K were unbounded, then $\exists (x_N)_{N\geq 0} \subset K$ such that $\|X_N\| \geq N \forall N$. Any subsequence of $(x_N)_{N\geq 0}$ is also unbounded. Indeed, by triangle inequality

$$x_{N_k} \to x \Leftrightarrow d(x_{N_k}, x) \to 0 \Rightarrow \|x_{N_k}\| < \|x\| + \|x_{N_k} - x\| \to \|x\|.$$

Thus, K is not sequentially compact, contradicting the fact that K is compact.

$\Leftarrow$: Our goal will be to show, that given $N \in \mathbb{N}$, $[-N,N]^n \subset \mathbb{R}^n$ is compact. This is sufficient, since if $K \subset \mathbb{R}^n$ is closed and bounded, then $K \subset B(0,N) \subset [-N,N]^n$ for some $N \in \mathbb{N}$, and thus K is a closed subset of a compact set, and thus compact by corollary 1.37.

We want to reduce the problem to Bolzano-Weierstrass in $\mathbb{R}$. Given $(x_k) \subset [-N,N]^n$, we can write $x_k = (x_{k,1}, \ldots, x_{k,n})$ where $x_{k,i} \in [-N,N]$.

1. Look at the sequence $(x_{k,1})_{k\geq 0} \subset [-N,N]$. By Bolzano-Weierstrass[2], there exists a increasing sequence $(k_m^{(1)})_{m\geq 0}$ in $\mathbb{N}$ such that

$$x_{k_m^{(1)},1} \to x_1 \in [-N,N].$$

2. Look at the sequence $(x_{k_m^{(1)},2})_{m\geq 0} \subset [-N,N]$. By Bolzano-Weierstrass, there exists a increasing sequence $(k_m^{(2)})_{m\geq 0}$ in $\mathbb{N}$ such that

$$x_{k_m^{(2)},2} \to x_2 \in [-N,N].$$

Since $(k_m^{(2)})_{m\geq 0}$ is a subsequence of $(k_m^{(1)})_{m\geq 0}$, we also have

$$x_{k_m^{(2)},1} \to x_1 \in [-N,N].$$

We continue this process, in each step making one new index converge, and keeping the previous ones converging. After n steps, we have an increasing sequence $(k_m^{(n)})_{m\geq 0}$ in $\mathbb{N}$ such that

$$x_{k_m^{(n)},i} \to x_i \in [-N,N] \quad \forall i = 1, \ldots, n.$$

Let $x = (x_1, \ldots, x_n)$, then by Lemma 1.17,

$$x_{k_m^{(n)}} \to x \in [-N,N]^n.$$

So K is sequentially compact. □

Tip 1.41:

The idea to split the proof into n times applying an argument from $\mathbb{R}$ is a very typical idea in Analysis II.

We now want to talk about CONNECTEDNESS. As a preliminary, in any metric space, $X, \varnothing$ are always open and closed (sometimes called clopen).

Question

Given $X = \mathcal{S}^2 = \{x \in \mathbb{R}^3 | \|x\| = 1\}$, equipped with the Euclidean metric restricted to $\mathcal{S}^2$.

Is it possible to write $X = U \cup V$ sich that $U \cap V = \varnothing$ and U, V are open in X and non-empty?

Definition 1.42: Connectedness

Given a metric space (X,d), we say that $A \subset X$ (nonempty) is DISSCONNECTED if $\exists U, V$ open, disjoint, such that $A \subset U \cup V$ and $A \cap U, A \cap V$ are non-empty.

Lec 6

We say that A is CONNECTED if it is not disconnected.

In other words, disconnected means that we have an open cover of A with two disjoint sets.

[2] Any sequence in a compact interval has a convergent subsequence

Tip 1.43:

Working with connectedness is usually done by contraposition, since the definition is an existence statement.

Proposition 1.44: Connected subsets of $\mathbb{R}$

$E \subset \mathbb{R}$ is connected if and only if E is an interval.

Notice that E is an interval if and only if

$$\forall x < y \in E, [x,y] \subset E.$$

Proof. $\Rightarrow$: By contraposition, assume E is not an interval. Then, $\exists x < y \in E$ such that $[x,y] \not\subset E$. So $\exists z \in [x,y]$ such that $z \notin E$.

Define $U = (-\infty, z)$ and $V = (z, \infty)$. Then, U, V are open, disjoint, and $E \subset U \cup V$. Moreover, $E \cap U$ and $E \cap V$ are non-empty since $x \in E \cap U$ and $y \in E \cap V$. Thus, E is disconnected.

$\Leftarrow$: By contraposition, assume E is disconnected. Then, $\exists U, V$ open, disjoint, such that $E \subset U \cup V$ and $E \cap U, E \cap V$ are non-empty.

Thus, pick $x \in E \cap U$ and $y \in E \cap V$ such that $x < y$[3]. Consider the supremum of the following set:

$$t* = \sup\{t \geq x \mid [x,t] \subset U\}.$$

This supremum is well defined since $t* < y$ because $[x,y] \not\subset U$.

From this we also find that $t* \in \mathbb{R} \setminus V$ since V is open and disjoint from U. But also $t* \in \mathbb{R} \setminus U$ since if $t* \in U$, then by openness of U, $\exists \varepsilon > 0$ such that $B(t*, \varepsilon) \subset U$. Thus, $[x, t* + \varepsilon] \subset U$, contradicting the definition of $t*$.

So $t* \in \mathbb{R} \setminus (U \cup V) \subset \mathbb{R} \setminus E$, and thus $[x,y] \not\subset E$. So E is not an interval. □

Proposition 1.45:

Given $(X, d_X), (Y, d_Y)$ metric spaces, $f : X \to Y$ continuous and $E \subset X$ connected, then $f(E)$ is connected.

Proof. By contraposition, assume $f(E)$ is disconnected. Then, $\exists V_1, V_2$ open, disjoint, such that $f(E) \subset V_1 \cup V_2$ and $f(E) \cap V_1, f(E) \cap V_2$ are non-empty.

Define now $U_i = f^{-1}(V_i)$, which are open since f is continuous. Then, U_1, U_2 are open, disjoint, and $E \subset U_1 \cup U_2$. Also, they are non-empty since $f(E) \cap V_i$ are non-empty. Thus, E is disconnected. □

Corollary 1.46: Intermediate Value Theorem

Given a connected metric space (X,d) and $f : X \to \mathbb{R}$ continuous, such that $f(x) = a \leq f(y) = b$. Then $\exists t \in X$ such that $f(t) = c$ for all $c \in [a,b]$.

Proof. Skipped...

Since X is connected, $f(X)$ is connected, and thus an interval. Since $a, b \in f(X)$, we have $[a,b] \subset f(X)$, and thus $\exists t \in X$ such that $f(t) = c$. □

Definition 1.47: Curve

Given a metric space (X,d), a CURVE or PATH in X is a map $\gamma : [0,1] \to X$ which is continuous.

$\gamma(0)$ is called the STARTING POINT of γ, and $\gamma(1)$ is called the ENDPOINT of γ.

γ is called CLOSED or LOOP if $\gamma(0) = \gamma(1)$.

[3] Can be assumed w.l.o.g.

Definition 1.48: Path Connectedness

Given a metric space (X, d), we call $E \subset X$ path connected if $\forall x, y \in E, \exists$ a path joining x and y, i.e. $\exists \gamma : [0,1] \to E$ continuous such that $\gamma(0) = x$ and $\gamma(1) = y$.

Proposition 1.49:

Given (X, d), E path connected, then E is connected.

Proof. Assume by contraposition that E is disconnected. Then, $\exists U_1, U_2$ open, disjoint, such that $E \subset U_1 \cup U_2$ and $\exists x_i \in E \cap U_i$.

Assume by contradiction, $\exists \gamma : [0,1] \to E$ continuous such that $\gamma(0) = x_1$ and $\gamma(1) = x_2$. A let $V_i = \gamma^{-1}(U_i)$, which are open since γ is continuous and disjoint. So $[0,1]$ is disconnected, contradicting the fact that $[0,1]$ is connected. □

The converse is not true, as the following example shows.

Example 1.50: Topologist's Sine Curve

Consider $(\mathbb{R}^2, d_{\text{Eucl}})$ and the set

$$E = \{0\} \times [-1,1] \cup \{(t, \sin(\frac{1}{t}))|t > 0\}.$$

This set is connected. However it is not path connected since intuitively, if we want to connect $(0,0)$ to $(1, \sin(1))$, we need to go through infinitely many oscillations of the sine curve, which is not possible with a continuous path.

Theorem 1.51:

In $(\mathbb{R}^n, d_{\text{Eucl}})$, given $U \subset \mathbb{R}^n$ open. U is connected if and only if U is path connected.

We first like to define the composition of paths. Given $\gamma_1 : [0,1] \to X$ and $\gamma_2 : [0,1] \to X$ such that $\gamma_1(0) = \gamma_2(0)$, we can define $\gamma_1^* = \gamma_1(1-t)$ as the reverse of γ_1. Then, the composition is defined as

$$\gamma_3(t) = \begin{cases} \gamma_1^*(2t) & t \in [0, \frac{1}{2}] \\ \gamma_2(2t-1) & t \in [\frac{1}{2}, 1] \end{cases}.$$

Proof. $\Leftarrow$: By Proposition 1.49.

$\Rightarrow$: We want to show that we can pick $x_0 \in U$ and join it with any other point $x \in U$ with a path. This is enough since if we can join x_0 to x and x_0 to y, then we can compose the paths to join x and y.

Define a set $G \subset U$ as

$$G := \{x \in U \mid \exists \text{ path } \gamma : [0,1] \to U : \gamma(0) = x_0, \gamma(1) = x\}.$$

We will now proof that G is open and that $U \setminus G$ is open, which since $x_0 \in G$ and U is connected, will imply that $G = U$.

The key observation is that for every $x \in U$, $\exists B_r(x) \subset U$ for some $r > 0$ since U is open. So $y \in B_r(x) \in G$ if and only if $x \in G$. This is because the map $t \in [0,1] \mapsto \gamma(t) = (1-t)x + ty$ is a path joining x and y.

This immediately implies that G is open, since if $x \in G$, then $B_r(x) \subset G$.

Moreover, if $x \in U \setminus G$, then $B_r(x) \subset U \setminus G$ since if $y \in B_r(x)$, then $x \in G$ if and only if $y \in G$. Thus, $U \setminus G$ is open. □

Proposition 1.52: Continuity on compact sets

Suppose $(X, d_X), (Y, d_Y)$ are metric spaces, $f : X \to Y$ is continuous, and $K \subset X$ is compact. Then, $f|_K : K \to Y$ is uniformly continuous.

Proof. Let $\varepsilon > 0$. By usual continuity of f, $\forall x \in K, \exists \delta_x > 0$ such that

$$f(B(x, \delta_x)) \subset B\left(f(x), \frac{\varepsilon}{2}\right).$$

Consider the open cover $\mathcal{U} = \{B(x, \frac{\delta_x}{2}) | x \in K\}$ of K. Since K is compact, there exists a finite subcover $B(x_1, \frac{\delta_{x_1}}{2}), \dots, B(x_n, \frac{\delta_{x_n}}{2})$ of K. Let $\delta = \min\{\frac{\delta_{x_i}}{2} | i = 1, \dots, n\}$. We will show that this δ works for uniform continuity.

If $x \in K$ and $y \in B(x, \delta)$, then

$$x \in B(x_i, \frac{\delta_{x_i}}{2}) \text{ for some } i \in \{1, \dots, n\}.$$

So

$$d(x,y) < \delta \leq \frac{\delta_{x_i}}{2} \Rightarrow d(x_i, y) < \delta(x)..$$

But then

$$f(B(x,\delta)) \subset f(B(x_i, \delta_{x_i})) \subset B\left(f(x_i), \frac{\varepsilon}{2}\right) \subset B(f(x), \varepsilon).$$

□

1.2 Normed Vector Spaces

Lec 7

Before starting with differentiating functions of several variables, we need to add some structure to our metric spaces.

Definition 1.53: Normed Vector Space over $\mathbb{R}$

Let V be a VS over $\mathbb{R}$. A map $\|\cdot\| : V \to [0, \infty)$ is called a NORM if $\forall v, u \in V$ and $\alpha \in \mathbb{R}$, the following hold:

1. Definite: $\|v\| = 0$ if and only if $v = 0$.
2. Homogeneous: $\|\alpha v\| = |\alpha| \cdot \|v\|$
3. Triangle inequality: $\|v + u\| \leq \|v\| + \|u\|$.

Example 1.54: Examples on $\mathbb{R}^n$

$\mathbb{R}^n$ with the Euclidean norm $\|x\|_{\text{Eucl}} = \sqrt{\sum_{i=1}^n x_i^2}$ is a normed vector space.

$\mathbb{R}^n$ with the p-norm $\|x\|_p = (\sum |x_i|^p)^{\frac{1}{p}}$ for $p \geq 1$ is a normed vector space.

$\mathbb{R}^n$ with the ∞-norm $\|x\|_\infty = \max |x_i|$ is a normed vector space.

From now on, we will write $|\cdot|$ instead of $\|\cdot\|$ when we use the Euclidean norm.

Definition 1.55: Hilbert-Schmidt norm

Let $M_{m \times n}(\mathbb{R})$ be the set of $m \times n$ matrices with real entries. We define the HILBERT-SCHMIDT NORM on $M_{m \times n}(\mathbb{R})$ as

$$\|M\|_2 = \sqrt{\operatorname{tr}(M^T M)} = \sqrt{\sum_{i=1}^n |Me_i|^2}.$$

To see the last equality, we square the Hilbert-Schmidt norm, we get

$$\|M\|_2^2 = \operatorname{tr}(M^T M) = \sum_{i=1}^{n}\sum_{j=1}^{m} M_i^j M_i^j.$$

But also

$$[Me_i]^j = M_i^j \Rightarrow |Me_i|^2 = \sum_{j=1}^{m} M_i^j M_i^j.$$

Lemma 1.56:

For every $x \in \mathbb{R}^n$ and $M : \mathbb{R}^n \to \mathbb{R}^m$ linear,

$$|Mx| \leq \|M\|_2 \cdot |x|.$$

Proof. We can write $x = \sum_{i=1}^{n} x_i e_i$. Thus, we compute

$$\begin{aligned}|Mx|^2 &= \sum_{i=1}^{m}[(Mx)^j]^2 = |M(\sum_{i=1}^{n} x_i e_i)|^2 \\ &= \left|\sum_{i=1}^{n} x_i M e_i\right|^2 \\ &\leq \left(\sum_{i=1}^{n} |x_i M e_i|\right)^2 = \left(\sum_{i=1}^{n} |x_i||Me_i|\right)^2 \\ &\leq |x|^2 \cdot \left(\sum_{i=1}^{n} |Me_i|^2\right) = |x|^2 \cdot \|M\|_2^2.\end{aligned}$$

□

Example 1.57: Norms on function spaces

Given $V = \{f : [a,b] \to \mathbb{R} \mid f \text{ continuous}\}$ we have the L^p-NORM for $p \geq 1$

$$\|f\|_{L^p} := \left(\int_a^b |f(t)|^p dt\right)^{\frac{1}{p}}.$$

And the SUPREMUM NORM

$$\|f\|_{L^\infty} := \sup_{t \in [a,b]} |f(t)|.$$

Proposition 1.58: Norm implies metric

Every normed vector space is a metric space with the metric defined as

$$d(x,y) = \|x - y\|.$$

Proof. d is definite because $\|x-y\| = 0$ if and only if $x - y = 0$ if and only if $x = y$.

d is symmetric because $\|x - y\| = \|(-1)(y - x)\| = \|y - x\|$.

d satisfies the triangle inequality, because

$$\|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\|.$$

□

Definition 1.59: Inner Product Vector Space

Let V be a VS over $\mathbb{R}$. A INNER PRODUCT is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ such that $\forall u, v, w \in V$ and $\alpha \in \mathbb{R}$:

1. Symmetry: $\langle u, v\rangle = \langle v, u\rangle$.
2. Bilinearity: $\langle \alpha u + \beta v, w\rangle = \alpha\langle u, w\rangle + \beta\langle v, w\rangle$.
3. Definite: $\langle v, v\rangle \geq 0$ and $\langle v, v\rangle = 0 \Leftrightarrow v = 0$.

Lemma 1.60:

Let V be an inner product vector space. Then, $\|\cdot\| : V \to [0, \infty)$ defined as

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

is a norm on V. So every inner product vector space is a normed vector space.

Proof. Exercise Sheet and Linear Algebra: Uses Cauchy-Schwarz inequality,

$$\langle v, w\rangle \leq \sqrt{\langle v, v\rangle} \cdot \sqrt{\langle w, w\rangle}.$$

□

Going back to the beginning of the lecture, we now know that $\mathbb{R}^n$ is an inner product vector space with the inner product defined as $\langle x, y\rangle = \sum_{i=1}^{n} x_i y_i$, and thus a normed vector space and a metric space.

From now on, we will by default work with the Euclidean norm and metric on $\mathbb{R}^n$. The most central definitions are the limit of a function on an open set $U \subset \mathbb{R}^n$

$$y = \lim_{x \to x_0} f(x) \Leftrightarrow \forall (x_k) \subset U \text{with } x_k \to x_0, f(x_k) \to y.$$

and the continuity of a function

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

One useful trick to compute a limit of a function $\mathbb{R}^2 \to \mathbb{R}$ as $r \to 0$ is to use polar coordinates. We can write $x = (r\cos\theta, r\sin\theta)$, and thus for example

$$\lim_{\to 0} \frac{xy}{\sqrt{x^2 + y^2}} = \lim_{r \to 0} \frac{r^2 \cos\theta \sin\theta}{r} = \lim_{r \to 0} r\cos\theta\sin\theta = 0.$$

Lec 8

2 Multidimensional Differentiation

We start with the most important definition of the course.

Definition 2.1: Differential

Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$. We say that f is DIFFERENTIABLE at $x_0 \in U$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{x \to 0} \frac{f(x_0 + x) - f(x_0) - L(x)}{|x|} = 0.$$

L is called the DIFFERENTIAL of f at x_0, and is denoted by Df_{x_0} or $Df(x_0)$.

Claim 2.2:

If $n = m = 1$, then $L(x) = f'(x_0)x$ where $f'(x_0)$ is the usual derivative of f at x_0.

Solution. We know $f'(x_0) = \lim_{x \to 0} \frac{f(x_0+x)-f(x_0)}{x}$, so we can write

$$\begin{aligned}
&\lim_{x \to 0} \frac{f(x_0 + x) - f(x_0) - f'(x_0)x}{|x|} \\
&= \lim_{x \to 0} \frac{f(x_0 + x) - f(x_0)}{x} - f'(x_0) \\
&= f'(x_0) - f'(x_0) = 0.
\end{aligned}$$

For some intuition consider the following examples.

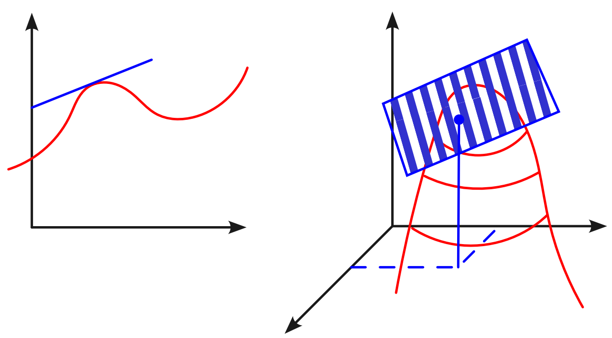

Figure 7: Differential of a function $\mathbb{R} \to \mathbb{R}$ and of a function $\mathbb{R}^2 \to \mathbb{R}$.

The equation for the slope in the first case would be

$$y = f(x_0) + f'(x_0)(x - x_0).$$

Whereas the equation for the tangent plane in the second case would be

$$y = f(x_0) + Df_{x_0}(x - x_0).$$

Lemma 2.3: Differential Componentwise

Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$ and $x_0 \in U$. Then f is differentiable at x_0 if and only if $f_i : U \to \mathbb{R}$ is differentiable at x_0 $\forall i = 1, \ldots, m$, where f_i is the i-th component of f. Moreover,

$$Df_{x_0}(x) = \begin{pmatrix} Df_{1,x_0}(x) \\ \vdots \\ Df_{m,x_0}(x) \end{pmatrix}.$$

This lemma allows us to reduce proofs to the case $m = 1$, which is usually easier to work with.

Proof. Recall the definition of the derivative of f at x_0:

$$\lim_{x \to 0} \frac{f(x_0 + x) - f(x_0) - L(x)}{|x|} = 0.$$

This can converge if and only if $\forall i = 1, \ldots, m$,

$$\lim_{x \to 0} \frac{f_i(x_0 + x) - f_i(x_0) - L_i(x)}{|x|} = 0.$$

But this is equivalent to f_i being differentiable at x_0 with differential L_i for all $i = 1, \ldots, m$. □

A convenient notation is the following

Definition 2.4: Big and Little O

Given $f, g : U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $x_0 \in U$, we say that

$$f(x) = O(g(x)) \text{ as } x \to x_0 \Leftrightarrow \lim_{x \to 0} \frac{|f(x)|}{|g(x)|} < \infty.$$

Similarly, we say that

$$f(x) = o(g(x)) \text{ as } x \to x_0 \Leftrightarrow \lim_{x \to 0} \frac{|f(x)|}{|g(x)|} = 0.$$

With this notation, f is differentiable at x_0 if and only if

$$R(x) = f(x_0 + x) - f(x_0) - Df_{x_0}(x) = o(|x|).$$

Definition 2.5: Directional derivatives

Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$ and $x_0 \in U$. Given $v \in \mathbb{R}^n$, we define the DIRECTIONAL DERIVATIVE of f at x_0 along v as

$$\partial_v f(x_0) = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t},$$

provided the limit exists.

Proposition 2.6:

Given $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$ and $x_0 \in U$. f is differentiable at $x_0 \in U$ implies that $\partial_v f(x_0)$ exists for all $v \in \mathbb{R}^n$, and

$$\partial_v f(x_0) = Df_{x_0}(v).$$

Proof. We can write

$$f(x_0 + x) - f(x_0) - Df_{x_0}(x) = o(|x|) \text{ as } x \to 0.$$

Thus for any sequence going to 0, we have that

$$\frac{f(x_0 + x) - f(x_0) - Df_{x_0}(x)}{|x|} \to 0.$$

In particular, for the sequence $x = tv$ with $t \to 0$, we have

$$\frac{f(x_0 + tv) - f(x_0) - Df_{x_0}(tv)}{|tv|} \to 0.$$

Since Df_{x_0} is linear, we can write $Df_{x_0}(tv) = tDf_{x_0}(v)$, and since $|tv| = |t||v|$, with $|v|$ finite, we can write

$$\underbrace{\frac{f(x_0 + tv) - f(x_0)}{t}}_{=\partial_v f(x_0)} - Df_{x_0}(v) \to 0.$$

□

Definition 2.7: Partial derivatives

Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$ and $x_0 \in U$. Given $i \in \{1, \dots, n\}$, we define the PARTIAL DERIVATIVE of f at x_0 along the i-th coordinate as

$$\partial_{e_i} f(x_0) =: \frac{\partial f}{\partial x_i}(x_0) = \partial_i f(x_0).$$

Notice that the partial derivatives are just the directional derivatives along the canonical basis vectors.

Example 2.8:

Let $n = 3, m = 1$ and $f(x_1, x_2, x_3) = \exp(x_2)[1 + x_1 x_3]$. Then, compute the partial derivative $\partial_1 f(0)$ and all partial derivatives at an arbitrary point $x = (x_1, x_2, x_3)$.

Solution. By definition

$$\begin{aligned}\partial_1 f(0,0,0) &= \lim_{t\to 0} \frac{f(0+t,0,0) - f(0,0,0)}{t} \\ &= \lim_{t\to 0} \frac{\exp(0)[1 + t\cdot 0] - \exp(0)[1 + 0 \cdot 0]}{t} \\ &= \lim_{t\to 0} \frac{1-1}{t} = 0.\end{aligned}$$

For the general case, we can define $g(t) = f(x_1 + t, x_2, x_3)$, so that $\partial_1 f(x) = g'(0)$. We can compute

$$\begin{aligned}\partial_1 f(x_1, x_2, x_3) &= \exp(x_2) x_3 \\ \partial_2 f(x_1, x_2, x_3) &= \exp(x_2)(1 + x_1 x_3) \\ \partial_3 f(x_1, x_2, x_3) &= \exp(x_2) x_1.\end{aligned}$$

In principle, the partial derivative can be calculated by fixing all the coordinates except the one we want to differentiate with respect to, and then applying the usual derivative in one variable.

The trick can be extended to a directional derivative along an arbitrary direction v

$$\partial_v f(x_0) = g'(0) \text{ where } g(t) = f(x_0 + tv).$$

Given $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$. When $\partial_i f(x_0)$ exists for all $x_0 \in U$ then we define the function

$$\partial_i f : U \to \mathbb{R}^m, x \mapsto \partial_i f(x).$$

Theorem 2.9: Sufficient condition for differentiability

Let $U \subset \mathbb{R}^n$ open, $f : U \to \mathbb{R}^m$. If $\partial_i f(x)$ exists and is continuous for every $i = 1, \dots, n$, and every $x \in U$, then f is differentiable at every $x \in U$.

Moreover,

$$Df_{x_0}(x) = \big(\partial_1 f(x_0), \dots, \partial_n f(x_0)\big)\, x.$$

More explicitly, if we write $x = (x_1, \dots, x_n)$, then

$$Df_{x_0}(x) = \begin{pmatrix} \partial_1 f_1(x_0) & \dots & \partial_n f_1(x_0) \\ \vdots & \ddots & \vdots \\ \partial_1 f_m(x_0) & \dots & \partial_n f_m(x_0) \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Proof. Fix $x_0 \in U$ and take $\varepsilon > 0$ such that $\{x \mid |x_i - x_{0,i}| < \varepsilon\} \subset U$. Take $x \in \mathbb{R}^n$ and define $x^{(k)} = x_0 + \sum_{i=1}^k x_i e_i$, so that $x^{(0)} = x_0$ and $x^{(n)} = x_0 + x$. Then, we can write

$$\begin{aligned}f(x_0 + x) - f(x_0) &= \sum_{k=1}^n f(x^{(k)}) - f(x^{(k-1)}) \\ &= \sum_{k=1}^n \partial_k f(y^{(k)}) x_k.\end{aligned}$$

We want to show the last equality by using the mean value theorem on a function $g_k(t) = f(x^{(k-1)} + te_k)$, so that $g_k(0) = f(x^{(k-1)})$ and $g_k(x_k) = f(x^{(k)})$.

By Lemma 2.3, we can assume $m = 1$ and thus $g_k : \mathbb{R} \to \mathbb{R}$, so that by the mean value theorem, $\exists \xi_k \in [x^{(k-1)}, x^{(k)}]$ such that

$$g_k(x_k) - g_k(0) = g_k'(\xi_k) x_k = \partial_k f(x^{(k-1)} + \xi_k e_k) x_k.$$

Letting $y^{(k)} = x^{(k-1)} + \xi_k e_k$, we get the desired equality.

From this, we can write

$$f(x_0 + x) - f(x_0) = \sum_{k=1}^n \partial_k f(y^{(k)}) x_k$$

.

But $\sum_{i=1}^{k-1} x_i e_i \to 0$ as $x \to 0$, and since $\xi_k \in [0, x_k]$, we have $\xi_k e_k \to 0$ as $x \to 0$. Thus, $y^{(k)} \to x_0$ as $x \to 0$.

$$\begin{aligned}f(x_0 + x) - f(x_0) &= \sum_{k=1}^n \partial_k f(y^{(k)}) x_k \\ &= \sum_{k=1}^n (\partial_k f(x_0) x_k + o(1)) \cdot |x|\end{aligned}$$

.

Thus, we can write

$$f(x_0 + x) - f(x_0) = \sum_{k=1}^n \partial_k f(x_0) x_k + o(|x|).$$

Which means that f is differentiable at x_0 with differential

$$Df_{x_0}(x) = \big(\partial_1 f(x_0), \dots, \partial_n f(x_0)\big)\, x.$$

□

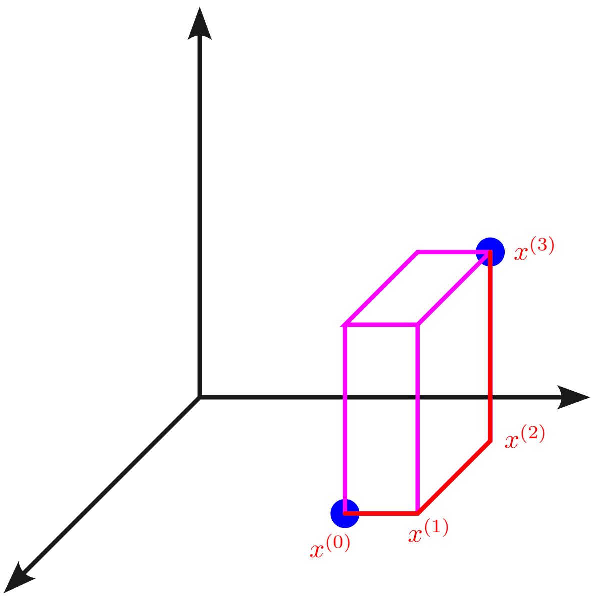

Figure 8: Definition of the $x^{(i)}$s.

Index