

CHAPTER 1

Metric Spaces

1. Definition and examples

Metric spaces generalize and clarify the notion of distance in the real line. The definitions will provide us with a useful tool for more general applications of the notion of distance:

DEFINITION 1.1. *A metric space is given by a set X and a distance function $d : X \times X \rightarrow \mathbb{R}$ such that*

i) *(Positivity) For all $x, y \in X$*

$$0 \leq d(x, y) .$$

ii) *(Non-degenerated) For all $x, y \in X$*

$$0 = d(x, y) \iff x = y .$$

iii) *(Symmetry) For all $x, y \in X$*

$$d(x, y) = d(y, x)$$

iv) *(Triangle inequality) For all $x, y, z \in X$*

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

Examples:

i) $X = \mathbb{R}$, $d(x, y) = |x - y|$.

ii) $X = \mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, $x = (x_1, x_2)$, $y = (y_1, y_2)$

$$d_1(x, y) = |x_1 - y_1| + |x_2 - y_2| .$$

iii) $X = \mathbb{R}^2$, $x = (x_1, x_2)$, $y = (y_1, y_2)$

$$d_2(x, y) = (|x_1 - y_1|^2 + |x_2 - y_2|^2)^{\frac{1}{2}} .$$

iv) Let $X = \{p_1, p_2, p_3\}$ and

$$d(p_1, p_2) = d(p_2, p_1) = 1 ,$$

$$d(p_1, p_3) = d(p_3, p_1) = 2 ,$$

$$d(p_2, p_3) = d(p_3, p_2) = 3.$$

Can you find a triangle (p_1, p_2, p_3) in the plane with these distances?

v) Let $X = \{p_1, p_2, p_3\}$ and

$$d(p_1, p_2) = d(p_2, p_1) = 1,$$

$$d(p_1, p_3) = d(p_3, p_1) = 2,$$

$$d(p_2, p_3) = d(p_3, p_2) = 4.$$

Can you find a triangle (p_1, p_2, p_3) in the plane with these distances?

vi) The French railway metric (Chicago suburb metric) on $X = \mathbb{R}^2$ is defined as follows: Let $x_0 = (0, 0)$ be the origin, then

$$d_{SNCF}(x, y) = \begin{cases} d_2(x, y) & \text{if there exists a } t \in \mathbb{R} \text{ such that } x_1 = ty_1 \\ & \text{and } x_2 = ty_2 \\ d_2(x, x_0) + d_2(x_0, y) & \text{else} \end{cases}.$$

Exercise: Show that the railroad metric satisfies the triangle inequality.

It is by no means trivial to show that d_2 satisfies the triangle inequality. In the following we write $0 = (0, \dots, 0)$ for the origin in \mathbb{R}^n .

CS LEMMA 1.2. *Let $x, y \in \mathbb{R}^n$, then*

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n |y_i|^2 \right)^{\frac{1}{2}}$$

LEMMA 1.3. *On \mathbb{R}^n the metric*

$$d_2(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}}$$

satisfies the triangle inequality.

PROOF. Let $x, y, z \in \mathbb{R}^n$. Then we deduce from Lemma CS 1.2

$$\begin{aligned} d(x, y)^2 &= \sum_{i=1}^n |x_i - y_i|^2 = \sum_{i=1}^n |(x_i - z_i) - (y_i - z_i)|^2 \\ &= \sum_{i=1}^n |(x_i - z_i)|^2 - 2 \sum_{i=1}^n (x_i - z_i)(y_i - z_i) + \sum_{i=1}^n |y_i - z_i|^2 \\ &\leq d(x, z)^2 + 2d(x, y)d(y, z) + d(y, z)^2 \end{aligned}$$

$$= (d(x, z) + d(y, z))^2.$$

Hence,

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

and the assertion is proved. ■

More examples:

(1) Let n be a prime number. On \mathbb{Z} we define

$$d_n(x, y) = n^{-\max\{m \in \mathbb{N} : n^m \text{ divides } x-y\}}.$$

The n -adic metric satisfies a stronger triangle inequality

$$d_n(x, y) \leq \max\{d_n(x, z), d_n(z, y)\}.$$

(2) Let $1 \leq p < \infty$. Then

$$d_p(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}}$$

defines a metric on \mathbb{R}^n .

(3) For $p = \infty$

$$d_\infty(x, y) = \max_{i=1, \dots, n} |x_i - y_i|$$

also defines a metric on \mathbb{R}^n .

Project 1: Let $1 < p, q < \infty$ such that $1/p + 1/q = 1$. Show Minkowski's inequality.

Mink (1.1)
$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

holds for all $x, y > 0$. **Hint:** the function $f(x) = -\ln x$ is convex on $(0, \infty)$.

PROOF OF THE TRIANGLE INEQUALITY FOR d_p . The triangle inequality for $p = 1$ is obvious. We will first show

mink2 (1.2)
$$\left| \sum_{i=1}^n x_i y_i \right| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}$$

whenever $\frac{1}{p} + \frac{1}{q} = 1$. Let $t > 0$. We first observe that

$$\begin{aligned} \left| \sum_{i=1}^n x_i y_i \right| &= \sum_{i=1}^n |t x_i| |t^{-1} y_i| \leq \sum_{i=1}^n \frac{1}{p} |t x_i|^p + \frac{1}{q} |t^{-1} y_i|^q \\ &= \frac{t^p}{p} \sum_{i=1}^n |x_i|^p + \frac{t^{-q}}{q} \sum_{i=1}^n |y_i|^q. \end{aligned}$$

What is best choice of t ? Make

$$t^p \sum_{i=1}^n |x_i|^p = t^{-q} \sum_{i=1}^n |y_i|^q$$

i.e.

$$t^{p+q} = \frac{\sum_{i=1}^n |y_i|^q}{\sum_{i=1}^n |x_i|^p}.$$

This yields

$$\begin{aligned} \left| \sum_{i=1}^n x_i y_i \right| &\leq t^p \sum_{i=1}^n |x_i|^p = \frac{\left(\sum_{i=1}^n |y_i|^q \right)^{\frac{p}{p+q}}}{\left(\sum_{i=1}^n |x_i|^p \right)^{\frac{p}{p+q}}} \sum_{i=1}^n |x_i|^p \\ &= \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}} \left(\sum_{i=1}^n |x_i|^p \right)^{1-\frac{1}{q}} \end{aligned}$$

Now, we proof the triangle inequality. Let $x = (x_i)$, (y_i) and $z = (z_i)$ in \mathbb{R}^d . Then we apply mink2 (1.2)

$$\begin{aligned} d_p(x, y)^p &= \sum_{i=1}^d |x_i - y_i|^p \leq \sum_{i=1}^d |x_i - y_i|^{p-1} (|x_i - z_i| + |z_i - y_i|) \\ &\leq \sum_{i=1}^d |x_i - y_i|^{p-1} |x_i - z_i| + \sum_{i=1}^d |x_i - y_i|^{p-1} |z_i - y_i| \\ &\leq \left(\sum_{i=1}^d (|x_i - y_i|^{p-1})^q \right)^{\frac{1}{q}} \left(\left(\sum_{i=1}^d |z_i - x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^d |z_i - y_i|^p \right)^{\frac{1}{p}} \right). \end{aligned}$$

However, $1 = 1/p + 1/q$ implies $p - 1 = p/q$ and thus $q(p - 1) = p$. Hence we get

$$d_p(x, y)^p \leq d_p(x, y)^{p-1} (d_p(x, z) + d_p(z, y)).$$

If $x \neq y$ we may divide and deduce the assertion. ■

2. Excursion: Convex functions

DEFINITION 2.1. Let I be an interval. A function $f : I \rightarrow \mathbb{R}$ is called convex if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

holds for all $x, y \in I$, $0 < \lambda < 1$.

LEMMA 2.2. Let $f : [a, b] \rightarrow \mathbb{R}$ be continuous, differentiable on (a, b) such that f' is increasing. Then f is convex.

PROOF. Let $x \in [a, b]$. We will show that

$$g(z) = \frac{f(y + z) - f(y)}{z}$$

is monotone increasing on $(0, b - x)$. Indeed, by the fundamental theorem and change of variables we deduce for $z_1 < z_2$ and $\lambda = \frac{z_1}{z_2}$ ($s = \lambda t$, $ds = \lambda dt$)

$$\begin{aligned} g(z_1) &= \int_0^{z_1} f'(s) \frac{ds}{z_1} = \int_0^{z_2} f'(\lambda t) \frac{\lambda dt}{z_1} = \int_0^{z_2} f'(\lambda t) \frac{dt}{z_2} \\ &\leq \int_0^{z_2} f'(t) \frac{dt}{z_2} = g(z_2). \end{aligned}$$

Now, we fix $y < x$ and $u = \lambda x + (1 - \lambda)y = y + \lambda(x - y)$, $z_1 = \lambda(x - y)$, $z_2 = x - y$. Then, we get

$$\frac{f(y + z) - f(y)}{\lambda(x - y)} \leq \frac{f(x) - f(y)}{(x - y)}.$$

This implies

$$f(\lambda x + (1 - \lambda)y) \leq f(y) + \lambda(f(x) - f(y)) = \lambda f(x) + (1 - \lambda)f(y). \quad \blacksquare$$

PROOF OF ^{Mink}1.1. Let $x, y > 0$. Since $-\ln x$ is convex we have

$$-\ln\left(\frac{1}{p}x^p + \frac{1}{q}y^q\right) \leq \frac{1}{p}(-\ln x^p) + \frac{1}{q}(-\ln y^q).$$

This shows by the monotonicity of \exp that

$$\frac{1}{p}x^p + \frac{1}{q}y^q \geq e^{\ln x + \ln y} = xy.$$

Minkowski's inequality is proved. \blacksquare

3. Continuous functions between metric spaces

Continuous functions ‘preserve’ properties of metric spaces and allow to describe deformation of one metric space into another. There are three different (but equivalent) ways of defining continuity, the ε - δ -criterion, the sequence criterion and the topological criterion. Each of them is interesting in its own right.

DEFINITION 3.1. *Let (X, d) and (Y, d') be metric spaces. A map $f : X \rightarrow Y$ is called continuous if for every $x \in X$ and $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$\boxed{\text{edelt}} \quad (3.1) \quad d(x, y) < \delta \implies d'(f(x), f(y)) < \varepsilon .$$

Let us use the notation

$$B(x, \delta) = \{y : d(x, y) < \delta\} .$$

For a subset $A \subset X$, we also use the notation

$$f(A) = \{f(x) : x \in A\} .$$

Similarly, for $B \subset Y$

$$f^{-1}(B) = \{x \in X : f(x) \in B\} .$$

Then $\boxed{\text{edelt}}$ (3.1) means

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

Or in a very non-formal way

f maps small balls into small balls .

Our aim is to prove a criterion for continuity in terms of so called open sets. This criterion illustrates simultaneously the role of open sets and its interaction with continuity and has a genuinely geometric flavor.

DEFINITION 3.2. *A subset O of a metric space is called open if*

$$\forall x \in O : \exists \delta > 0 : B(x, \delta) \subset O .$$

Examples:

$$O = (-1, 1), O = \mathbb{R}, O = (-1, 1) \times (-2, 2)$$

are open in \mathbb{R} , (\mathbb{R}^2, d_2) respectively.

REMARK 3.3. *The sets $B(x, \varepsilon)$, $x \in X$, $\varepsilon > 0$ are open.*

PROPOSITION 3.4. *Let (X, d) , (Y, d') be metric spaces and $f : X \rightarrow Y$ be a map. f is continuous iff $f^{-1}(O)$ is open for all open subsets $O \subset Y$.*

PROOF. \Rightarrow : We assume that f is continuous and O is open. Let $x \in f^{-1}(O)$, i.e. $f(x) \in O$. Since O is open, there exists an $\varepsilon > 0$ such that $B(f(x), \varepsilon) \subset O$. By continuity, there exists a $\delta > 0$ such that

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

Therefore

$$B(x, \delta) \subset f^{-1}(O) .$$

Since $x \in f^{-1}(O)$ was arbitrary, we deduce that $f^{-1}(O)$ is open.

\Leftarrow : Let $x \in X$ and $\varepsilon > 0$. Let us show that

$$B(f(x), \varepsilon)$$

is a on open subset of (Y, d') . Indeed, let $y \in B(f(x), \varepsilon)$ define $\varepsilon' = \varepsilon - d'(y, f(x))$.

Let $z \in Y$ such that

$$d(z, y) < \varepsilon'$$

then

$$d(f(x), z) \leq d(f(x), y) + d(y, z) < d(f(x), y) + \varepsilon - d'(y, f(x)) = \varepsilon .$$

Thus

$$B(y, \varepsilon - d'(f(x), y)) \subset B(f(x), \varepsilon) .$$

By the assumption, we see that $f^{-1}(B(f(x), \varepsilon))$ is an open set. Since $x \in f^{-1}(B(f(x), \varepsilon))$, we can find a $\delta > 0$ such that

$$B(x, \delta) \subset f^{-1}(B(f(x), \varepsilon)) .$$

Hence, for all \tilde{x} with $d(x, \tilde{x}) < \delta$, we have

$$d'(f(x), f(\tilde{x})) < \varepsilon .$$

The assertion is proved. ■

Examples:

- (1) Let (X, d) be a metric space and $x_0 \in X$ be a point , then $f(x) = d(x, x_0)$ is continuous. Indeed, the triangle inequality implies

$$d(d(x, x_0), d(d(y, 0))) = |d(x, x_0) - d(y, x_0)| \leq d(x, y)$$

This easily implies the assertion.

- (2) On \mathbb{R}^n with the standard euclidean metric $d = d_2$, the function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defined by $f(x) = d(x, 0)x$ is continuous.

(3) (Exercise) The function $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $f(x) = (\cos(x_1), \sin(x_2), \cos(x_1))$ is continuous.

DEFINITION 3.5. Let (X, d) , (Y, d') be a metric space. The space $C(X, Y)$ is the set of all continuous functions from X to Y . Let $x_0 \in X$ be a point. Then

$$C_b(X, Y) = \{f : X \rightarrow Y : f \text{ is continuous and } \sup_{x \in X} d'(f(x), f(x_0)) < \infty\}$$

is the subset of bounded continuous functions.

PROPOSITION 3.6. Let (X, d) , (Y, d') be metric spaces and $x_0 \in X$. Then $C_b(X, Y)$ equipped with

$$d(f, g) = \sup_{x \in X} d'(f(x), g(x))$$

is a metric space.

Problem: Show that d is not well-defined on $C(\mathbb{R}, \mathbb{R})$.

Proof: $d(f, g) = 0$ if and only if $f(x) = g(x)$ for all $x \in X$. This means $f = g$. Let us show that d is well-defined. Indeed, if $f, g \in C_b(X, Y)$. Then

$$\begin{aligned} \sup_x d'(f(x), g(x)) &\leq \sup_x d'(f(x), f(x_0)) + d'(f(x_0), g(x_0)) + d'(g(x_0), g(x)) \\ &\leq \sup_x d'(f(x), f(x_0)) + d'(f(x_0), g(x_0)) + \sup_x d(g(x_0), g(x)) \end{aligned}$$

is finite. Let h be a third function and $x \in X$. Then

$$d'(f(x), g(x)) \leq d'(f(x), h(x)) + d(h(x), g(x)) \leq d(f, h) + d(h, g).$$

Taking the supremum yields the assertion. ■

alg

PROPOSITION 3.7. Let (X, d) be a metric space. Then $C(X, \mathbb{R})$ is closed under (pointwise-) sums, products and multiplication with real numbers. ($C(X, \mathbb{R})$ is an algebra over \mathbb{R}).

REMARK 3.8. Let $X = \mathbb{N}$ and $d(x, y) = 1$ if $x \neq y$ and $d(x, y) = 0$ for $x = y$. (This is called the discrete metric). Then $C(X, \mathbb{R})$ is an infinite dimensional vector space.

PROOF OF ^{alg}3.7. Let $f, g \in C(X, \mathbb{R})$ be continuous and $x \in X$. Consider $x' \in X$. Then

$$\begin{aligned} fg(x) - fg(y) &= f(x)g(x) - f(y)g(y) = (f(x) - f(y))g(x) + f(y)(g(x) - g(y)) \\ &= (f(x) - f(y))g(x) + f(x)(g(x) - g(y)) + (f(y) - f(x))(g(x) - g(y)). \end{aligned}$$

Let $\varepsilon > 0$ and $\tilde{\varepsilon} = \min\{\varepsilon, 1\}$. We may choose $\delta_1 > 0$ such that

$$d(f(x), f(y))(1 + |g(x)|) < \frac{\tilde{\varepsilon}}{3}$$

holds for all $d(x, y) < \delta_1$. Similarly, we may choose $\delta_2 > 0$ such that

$$d(g(x), g(y))(1 + |f(x)|) < \frac{\tilde{\varepsilon}}{3}.$$

Let $\delta = \min(\delta_1, \delta_2)$ and $d(x, y) < \delta$. Then we deduce that

$$d(fg(x), fg(y)) = |fg(x) - fg(y)| < \frac{\tilde{\varepsilon}}{3} + \frac{\tilde{\varepsilon}}{3} + \frac{\tilde{\varepsilon}^2}{9} < \tilde{\varepsilon} \leq \varepsilon.$$

Thus fg is again continuous. The other assertions are easier. ■

COROLLARY 3.9. *The polynomials on \mathbb{R} are continuous.*

LEMMA 3.10. *Let $1 \leq p \leq \infty$ and $x, y \in \mathbb{R}^n$, then*

$$\frac{1}{n^{\frac{1}{p}}} d_p(x, y) \leq d_\infty(x, y) \leq d_p(x, y).$$

PROOF. The last inequality is obvious. For the first one, we consider $x, y \in \mathbb{R}^n$ and $1 \leq p < \infty$, then by estimating every element in the sum against the maximum

$$d_p(x, y)^p = \sum_{i=1}^n |x_i - y_i|^p \leq n \max\{|x_i - y_i|^p\}.$$

Taking the p -th root, we deduce the assertion. ■

COROLLARY 3.11. *Let $1 \leq p, q \leq \infty$, then the identity map $id : (\mathbb{R}^n, d_p) \rightarrow (\mathbb{R}^n, d_q)$ is continuous.*

PROOF. We have for all $x \in \mathbb{R}^n$ and $\varepsilon > 0$

$$B_{d_p}(x, \frac{\varepsilon}{n}) \subset B_{d_q}(x, \varepsilon).$$

This easily implies the assertion. ■

COROLLARY 3.12. *The metrics d_p define the same open sets on \mathbb{R}^n .*

DEFINITION 3.13. *Let (X, d) be a metric space. We say that a sequence (x_n) converges to x if for all $\varepsilon > 0$ there exists n_0 such that for $n > n_0$ we have*

$$d(x_n, x_0) < \varepsilon.$$

In this case we write

$$\lim_n x_n = x$$

or more explicitly

$$d - \lim_n x_n = x.$$

A sequence (x_n) is convergent, if there exists $x \in X$ with $\lim_n x_n = x$.

Examples: $d_2 - \lim_n \frac{1}{n} = 0$, $dd_3 - \lim_n 3^n = 0$. (What axioms of the natural numbers are involved?).

PROPOSITION 3.14. *Let (X, d) , (Y, d') be metric spaces and $f : X \rightarrow Y$ be a map. Then f is continuous if for every convergent sequence (x_n) in X*

$$\lim_n f(x_n) = f(\lim_n x_n).$$

Proof: \Rightarrow : Let $x = \lim_n x_n$ and $\varepsilon > 0$, then there exists a $\delta > 0$ such that

$$d(y, x) < \delta \Rightarrow d'(f(y), f(x)) < \varepsilon.$$

Let $n_0 \in \mathbb{N}$ be such that

$$d(x_n, x) < \delta$$

for all $n > n_0$, then

$$d'(f(x_n), f(x)) < \varepsilon$$

for all $n > n_0$. Hence

$$\lim_n f(x_n) = f(x).$$

\Leftarrow Let $x \in X$ and assume in the contrary that

$$\exists \varepsilon > 0 \forall \delta > 0 \exists y : d(y, x) < \delta \text{ and } d'(f(y), f(x)) \geq \varepsilon.$$

Applying these successively for all $\delta = \frac{1}{k}$, we find a sequence (x_k) such that

$$d(x_k, x) < \frac{1}{k} \quad \text{and} \quad d'(f(x_k), f(x)) \geq \varepsilon'.$$

and thus

$$\lim_k x_k = x.$$

By assumption, we have

$$\lim_k f(x_k) = f(x).$$

Hence, there exists a k_0 such that for all $k > k_0$

$$d(f(x_k), f(x)) < \varepsilon.$$

a contradiction. ■

4. Complete metric spaces and completion

Complete metric spaces are crucial in understanding existence of solutions to many equations. Complete spaces are also important in understanding spaces of integrable functions. We will review basic properties here and show the existence of a completion.

We will say that a sequence in a metric space is a Cauchy sequence if for every $\varepsilon > 0$ there exists $n_0 \in \mathbb{N}$ such that

$$d(x_n, x_m) < \varepsilon$$

for all $n, m > n_0$.

DEFINITION 4.1. *A metric space (X, d) is called complete, if every Cauchy sequence converges.*

model **PROPOSITION 4.2.** *The space (\mathbb{R}^2, d_1) is complete.*

Proof: Let x_n be a Cauchy sequence in (\mathbb{R}^2, d_1) . Then $x_n = (x_n(1), x_n(2))$ is a sequence of pairs.

Claim: The sequences $(x_n(1))_{n \in \mathbb{N}}$ and $(x_n(2))_{n \in \mathbb{N}}$ are Cauchy sequences.

Indeed, let $\varepsilon > 0$, then there exists an n_0 such that

$$d_1(x_n, x_m) < \varepsilon$$

for all $n, m > n_0$. In particular, we have

$$|x_n(1) - x_m(1)| \leq |x_n(1) - x_m(1)| + |x_n(2) - x_m(2)| \leq d_1(x_n, x_m) < \varepsilon$$

for all $n, m > n_0$ and

$$|x_n(2) - x_m(2)| \leq |x_n(1) - x_m(1)| + |x_n(2) - x_m(2)| \leq d_1(x_n, x_m) < \varepsilon.$$

Therefore, $(x_n(1))$ and $(x_n(2))$ are Cauchy.

Since \mathbb{R} is complete, we can find $x(1)$ and $x(2)$ such that

$$\lim_n x_n(1) = x(1) \quad \text{and} \quad \lim_n x_n(2) = x(2).$$

Claim: $\lim_n x_n = (x(1), x(2))$.

Indeed, Let $\varepsilon > 0$ and choose n_1 such that

$$|x_n(1) - x(1)| < \frac{\varepsilon}{2}$$

for all $n > n_1$. Choose n_2 such that

$$|x_n(2) - x(2)| < \frac{\varepsilon}{2}$$

for all $n > n_2$. Set $n_0 = \max\{n_1, n_2\}$, then for every $n > n_0$, we have

$$d_1(x_n, (x(1), x(2))) = |x_n(1) - x(1)| + |x_n(2) - x(2)| < \varepsilon$$

Thus

$$\lim_n x_n = x$$

and the assertion is proved. ■

Examples:

- (1) Let $X = \mathbb{R} \setminus \{0\}$ and $d(x, y) = |x - y|$, then (X, d) is not complete. The sequences $(\frac{1}{n})$ is Cauchy and does not converge.
- (2) Let p be a prime number. On the set of integers, we define

$$d_p(z, w) = p^{-n},$$

where $n = \max\{n : p^n \text{ divides } (z - w)\}$. This satisfies the triangle inequality. The sequence (x_n) given by $x_n = p + p^2 + \dots + p^n$ is a non convergent Cauchy sequence.

rncomp

THEOREM 4.3. *Let $n \in \mathbb{N}$. The space (\mathbb{R}^n, d_2) is a complete metric space.*

PROOF. Similar as in Proposition [4.2](#) using the following Lemma. ■

LEMMA 4.4. *Let $x, y \in \mathbb{R}^n$, then*

$$d_2(x, y) \leq \sum_{i=1}^n |x_i - y_i|.$$

PROOF. We proof this by induction on $n \in \mathbb{N}$. The case $n = 1$ is obvious. Assume the assertion is true for n and let $x, y \in \mathbb{R}^{n+1}$. We define the element $z = (x_1, \dots, x_n, y_{n+1})$, then we deduce from the triangle inequality

$$\begin{aligned} d_2(x, y) &\leq d_2(x, z) + d_2(z, y) \\ &= \left(\sum_{i=1}^{n+1} |x_i - z_i|^2 \right)^{\frac{1}{2}} + \left(\sum_{i=1}^{n+1} |z_i - y_i|^2 \right)^{\frac{1}{2}} \\ &= |x_{n+1} - y_{n+1}| + \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}}. \end{aligned}$$

To apply the induction hypothesis, we define $\tilde{x} = (x_1, \dots, x_n)$ and $\tilde{y} = (y_1, \dots, y_n)$. Then the induction hypothesis yields

$$\left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}} = d_2(\tilde{x}, \tilde{y}) \leq \sum_{i=1}^n |x_i - y_i|.$$

Hence,

$$\begin{aligned} d_2(x, y) &\leq |x_{n+1} - y_{n+1}| + \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}} \\ &\leq |x_n - y_n| + \sum_{i=1}^n |x_i - y_i| \\ &= \sum_{i=1}^{n+1} |x_i - y_i|. \end{aligned}$$

The assertion is proved. ■

DEFINITION 4.5. A subset $C \subset X$ is called closed if $X \setminus C$ is open.

complete1

PROPOSITION 4.6. Let C be closed subset of a complete metric space (X, d) , then $(C, d|_{C \times C})$ is complete.

PROOF. Let $(x_n) \subset C$ be Cauchy sequence. Since X is complete, there exists $x \in X$ such that

$$x = \lim_n x_n.$$

We have to show $x \in C$. Assume $x \notin C$. Then there exists a $\delta > 0$ such that $B(x, \delta) \subset X \setminus C$. By definition of the limit there exists n_0 such that $d(x_n, x) < \delta$ for all $n > n_0$. Set $n = n_0 + 1$. Then $d(x_n, x) < \delta$ implies $x_n \in X \setminus C$ and $x_n \in C$ by definition. This contradiction finished the proof. ■

comp1

THEOREM 4.7. Let (Y, d') be complete metric space. Let $h \in C(X, Y)$ and

$$C_h(X, Y) = \{f \in C(X, Y) : \sup_{x \in X} d'(f(x), h(x)) < \infty\}$$

Then $C_g(X, Y)$ is complete with respect to

$$d(f, g) = \sup_{x \in X} d'(f(x), g(x)).$$

PROOF. Let $(f_n) \subset C_h(X, Y)$ be Cauchy sequence. This means that for every $\varepsilon > 0$ there exists an n_0 such that

$$\text{eqq} \quad (4.1) \quad \sup_{x \in X} d'(f_n(x), f_m(x)) < \frac{\varepsilon}{2}.$$

In particular, for fixed $x \in X$, $f_n(x)$ is Cauchy. Therefore $f(x) := \lim_m f_m(x)$ is a well-defined element in Y . We fix $n > n_0$ and consider $m \geq n_0$ such that

$$d'(f_m(x), f(x)) \leq \frac{\varepsilon}{3}.$$

This implies

$$d'(f_n(x), f(x)) \leq d'(f_n(x), f_m(x)) + d'(f_m(x), f(x)) \leq \frac{5}{6}\varepsilon$$

for all $x \in X$. In particular,

$$\boxed{\text{eqq2}} \quad (4.2) \quad \sup_{n \geq n_0} \sup_{x \in X} d'(f_n(x), f(x)) \leq \frac{5}{6}\varepsilon.$$

Let us show that f is continuous. Let $z \in X$ and $\varepsilon > 0$. Choose n_0 according to [\(4.1\)](#). Choose $n = n_0 + 1$. Let $\delta > 0$ such that $d(x, y) < \delta$ implies $d'(f_n(x), f_n(y)) < \varepsilon$. Then, we have

$$d'(f(x), f(y)) \leq d'(f(x), f_n(x)) + d'(f_n(x), f_n(y)) + d'(f_n(y), f(y)) < 3\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, we see that f is continuous. Moreover, [\(4.2\)](#) implies that f_n converges to f . Finally, [\(4.2\)](#) for $\varepsilon = 1$ implies that

$$\sup_x d(f(x), h(x)) \leq \sup_x d(f(x), f_n(x)) + \sup_x d(f_n(x), h(x)) < \infty$$

implies that $f \in C_h(X, Y)$. ■

DEFINITION 4.8. Let (X, d) be a metric space and $C \subset X$. $O \subset X$ is called *sense* if for ever $x \in C$ and $\varepsilon > 0$ $B(x, \varepsilon) \cap O \neq \emptyset$.

DEFINITION 4.9. Let $O \subset X$ be a subset. Then

$$\bar{O} = \bigcap_{O \subset C, C \text{ closed}} C$$

is called the *closure*.

dens **LEMMA 4.10.** O is dense in \bar{O} and \bar{O} is closed.

PROOF. Let $x \in \bar{O}$. Assume $B(x, \varepsilon) \cap O = \emptyset$. Then $C = X \setminus B(x, \varepsilon)$ contains O . Thus

$$\bar{O} \subset C.$$

This implies that $x \notin \bar{O}$, a contradiction. Now, we show that \bar{O} is closed. Indeed, let $y \notin \bar{O}$. Then there has to be a closed set C such that $O \subset C$ but $y \notin C$. This means $y \in X \setminus C$ which is open. Hence there exists $\delta > 0$ such that

$$B(y, \delta) \subset X \setminus C$$

By definition every element $z \in B(y, \delta)$ does not belong to \bar{O} . This means $B(y, \delta) \subset X \setminus \bar{O}$. ■

comp2

THEOREM 4.11. *Let (X, d) be a non-empty metric space. For every $x \in X$ we define*

$$f_x(y) = d(x, y).$$

Let $x_0 \in X$. The map $f : X \rightarrow C_{f_{x_0}}(X, \mathbb{R})$ satisfies the following properties.

- i) $d(f(x), f(y)) = d(x, y)$,
- (1) *The closure $C = \overline{f(X)}$ is complete,*
- (2) *$f(X)$ is dense in the closure $C = \overline{f(X)}$.*

PROOF. Let $x, y \in X$ and $z \in X$. Then the ‘converse triangle’ inequality implies

$$|f_x(z) - f_y(z)| = |d(x, z) - d(y, z)| \leq d(x, y).$$

Moreover,

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

Therefore $f_x \in C_{f_{x_0}}(X, \mathbb{R})$ for every $x \in X$ and

$$d(f_x, f_y) \leq d(x, y).$$

However,

$$d(f_x, f_y) \geq |f_x(x) - f_y(x)| = |0 - d(y, x)| = d(y, x).$$

This shows i). According to Proposition [4.6](#) ^{complete1} and Theorem [4.7](#) ^{comp1}, we see that C is complete. According to Lemma [4.10](#) ^{dens}, we deduce that $f(X)$ is dense in C . ■

Project: On $C([0, 1])$ we define

$$d_1(f, g) = \int |f(s) - g(s)| ds.$$

Show that $(C([0, 1]), d_1)$ is not complete.

Project: In the literature you can find another description of the completion of a metric space. Find it and describe it.

5. Unique extension of densely defined uniformly continuous functions

In this section we will show that the completion C constructed in Theorem [4.11](#) ^{comp2} is unique (in some sense). This is based on a simple observation—the unique extension. This principle is very often used in analysis.

DEFINITION 5.1. *Let (X, d) , (Y, d') be metric spaces. A function $f : X \rightarrow Y$ is called uniform continuous if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$d(x, y) < \delta \quad \Rightarrow \quad d'(f(x), f(y)) < \varepsilon.$$

u-ext

PROPOSITION 5.2. *Let $O \subset C$ be a dense set and $f : O \rightarrow Y$ be uniformly continuous function with values in a complete metric space. Then there exists a unique continuous function $\tilde{f} : C \rightarrow Y$ such that $\tilde{f}(x) = f(x)$ for all $x \in O$.*

PROOF. Let $x \in X$. Since $B(x, \frac{1}{n}) \cap O$ is not empty, we may find $(x_n) \subset O$ such that $\lim_n x_n = x$. We try to define

$$\tilde{f}(x) = \lim_n f(x_n).$$

Let us show that this is well-defined. So we consider another Cauchy sequence (x'_n) such that $\lim_n x'_n = x$. Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that

$$d'(f(x'), y) < \varepsilon$$

holds for $d(x', y) < \delta$. We may find n_0 such that

$$d(x_n, x) < \frac{\delta}{2}$$

and

$$d(x'_n, x) < \frac{\delta}{2}$$

holds for all $n, n' > n_0$. Thus

$$d'(f(x'_n), f(x_n)) < \varepsilon.$$

This argument also shows that $(f(x_n))$ is Cauchy and hence $\tilde{f}(x)$ is well-defined. If $x \in O$, we may choose for (x_n) the constant sequence $x_n = x$ and hence $\tilde{f}(x) = f(x)$. Now, we want to show that \tilde{f} is uniformly continuous. Indeed, let $\varepsilon > 0$, then there exists $\delta > 0$ such that $d(x', y') < \delta$ implies

$$d(f(x'), f(y')) < \frac{\varepsilon}{2}.$$

Given $x, y \in C$ with $d(x, y) < \delta$, we may find (x_n) converging to x and (y_n) converging to y such that

$$d(x_n, x) < \frac{\delta - d(x, y)}{2}.$$

Thus for all $n \in \mathbb{N}$ we have

$$d(x_n, y_n) \leq d(x, y) + d(x_n, x) + d(y_n, y) < \delta.$$

This implies

$$d(f(x), f(y)) = \lim_n d(f(x_n), f(y_n)) < \frac{\varepsilon}{2}.$$

This shows that \tilde{f} is uniformly continuous. If g is another continuous function such that $g(x) = f(x)$ holds for elements $x \in O$, then we may choose a Cauchy sequence (x_n) converging to x and get

$$g(x) = \lim_n g(x_n) = \lim_n f(x_n) = f(x). \quad \blacksquare$$

Example If $f : (0, 1] \rightarrow \mathbb{R}$ is uniformly continuous, then f is bounded (why). $f(x) = 1/x$ is not uniformly continuous.

THEOREM 5.3. *The completion of a metric space is unique. More precisely, let C be the set constructed in Theorem ^{comp2}4.11. Let C' be a complete metric space and $\iota' : X \rightarrow C'$ be uniformly continuous with uniformly continuous inverse $\iota'^{-1} : \iota'(X) \rightarrow X$ such that $\iota'(X)$ is dense. Then there is a bijective, bicontinuous map $u : C \rightarrow C'$ such that $u(\iota(x)) = \iota'(x)$.*

PROOF. The map $\iota'\iota^{-1} : \iota(X) \rightarrow C'$ is uniformly continuous and hence admits a unique continuous extension $u : C \rightarrow C'$. Also $\iota'^{-1} : \iota'(X) \rightarrow X$ admits a unique extension $v : C' \rightarrow C$. Note that $vu : C \rightarrow C$ is an extension of the map $vu(\iota(x)) = \iota(x)$. Thus there is only one extension, namely the identity. This shows $vu = id$. Similarly $uv = id$. Thus $v = u^{-1}$ and u is bijective and bi-continuous. \blacksquare

Project: Find the completion of (\mathbb{Z}, d_3) .

6. A famous example

In this section we want to identify the completion of $C([0, 1])$ with respect to

$$d_1(f, g) = \int_0^1 |f(t) - g(t)| dt.$$

We will also use the function

$$I(f) = \int_0^1 f(t) dt$$

defined by the Riemann integral.

LEMMA 6.1. *I is uniformly continuous.*

PROOF. It suffices to show that

$$|I(f)| \leq \int_0^1 |f(t)| dt.$$

(This implies that I is 1-Lipschitz, i.e.

$$|I(f) - I(g)| \leq d_1(f, g) .)$$

We define $f^+ = \max\{f, 0\}$ and $f^- = \max\{-f, 0\}$. Then we have

$$I(f) = I(f^+) - I(f^-) \leq I(f^+) + I(f^-) = I(|f|) = \int |f(t)| dt .$$

Similarly, we may show that

$$I(f) = I(f^+) - I(f^-) \geq -I(f^+) - I(f^-) = -I(|f|) = -\int |f(t)| dt .$$

The assertion follows. ■

The characteristic function is given by

$$1_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} .$$

LEMMA 6.2. $I(1_{[a,b]}) = b - a$.

PROOF. We only consider $[a, b] = [0, b]$. For $2/n \leq b$ we define

$$f_n(t) = \begin{cases} nt & \text{if } t \leq \frac{1}{n} \\ 1 & \text{if } \frac{1}{n} \leq t \leq b - \frac{1}{n} \\ n(b-t) & \end{cases} .$$

Then we deduce that for $m \geq n$ we have

$$d_1(f_n, f_m) = 2 \int_0^{\frac{1}{m}} mt dt + \frac{1}{n} - \frac{1}{m} - \int_0^{\frac{1}{n}} ntdt = 2\left(\frac{1}{2m} + \frac{1}{n} - \frac{1}{m} - \frac{1}{2n}\right) = \frac{1}{n} - \frac{1}{m} .$$

Thus (f_n) is Cauchy and

$$\lim_n \int f_n(t) dt = \lim_n b - \frac{1}{n} = b .$$

For general a we simply shift. ■

In the following we denote the length of an interval by

$$|[a, b]| = b - a .$$

chy LEMMA 6.3. Let f be a continuous positive function on $[0, 1]$, $\lambda > 0$ and $\varepsilon > 0$. Then there exists intervals J_1, \dots, J_m such that

$$\{t : f(t) > \lambda\} \subset J_1 \cup \dots \cup J_m$$

and

$$\sum_k |J_k| \leq \frac{1}{\lambda - 2\varepsilon} \int f(t) dt.$$

PROOF. Let $\varepsilon < \frac{\lambda}{2}$. Since f is uniformly continuous, we may find $n \in \mathbb{N}$ such that $|x - y| \leq 1/n$ implies

$$|f(x) - f(y)| < \varepsilon.$$

We define $x_k = k/n$. Let $S = \{k : f(x_k) > \lambda - \varepsilon\}$. Let $t \in [0, 1]$ such that $f(t) > \lambda$. We consider $k = \lfloor tn \rfloor$. Then $x_k - t \leq \frac{1}{n}$ and hence

$$f(x_k) > f(t) - \varepsilon > \lambda - \varepsilon.$$

Therefore

$$\{t : f(t) > \lambda\} \subset \bigcup_{k \in S} [x_k, x_{k+1}].$$

However, $f(x_k) > \lambda - \varepsilon$ implies $f(t) > \lambda - 2\varepsilon$ for all $t \in [x_k, x_{k+1}]$. By the definition of the lower sums we deduce

$$\int_0^1 f(t) dt \geq \sum_{k \in S} (\lambda - 2\varepsilon) |J_k|.$$

Since $\lambda - 2\varepsilon > 0$, we deduce the assertion. ■

DEFINITION 6.4. i) A subset $A \subset [0, 1]$ is said to have measure 0 if for every $\varepsilon > 0$ there exists a sequence (J_k) of intervals such that

$$A \subset \bigcup_k J_k \quad \text{and} \quad \sum_k |J_k| < \varepsilon.$$

ii) A sequence (f_n) converges almost everywhere (a.e.) to a function f if there exists a set A of measure 0 such that

$$\lim_n f_n(t) = f(t)$$

for all $t \in A^c = [0, 1] \setminus A$.

limit1 PROPOSITION 6.5. Let (f_n) be a d_1 -Cauchy sequence. Then there exists a subsequence (n_k) and a function f such that f_{n_k} converges to f a.e.

PROOF. We may choose (n_k) such that

$$d_1(f_{n_k}, f_{n_{k+1}}) \leq 6^{-k}.$$

Let us denote $g_k = f_{n_k}$. We apply Lemma 6.3 to $\lambda = 2^{-k}$ and $\varepsilon = \frac{2^{-k}-3^{-k}}{2}$ and find intervals $J_1^k, \dots, J_{m(k)}^k$ such that

$$\sum_l |J_l^k| \leq \frac{6^{-k}}{2^{-k} - 2\varepsilon} = 2^{-k}$$

and

$$\{t \in [0, 1] : |g_k(t) - g_{k+1}(t)| > 2^{-k}\} \subset \bigcup_l J_l^k.$$

We define

$$A_k = \bigcup_{n \geq k} \bigcup_l J_l^n$$

and

$$A = \bigcap_k A_k.$$

Then, we see that $A \subset A_k$ and

$$\sum_{n \geq k} \sum_l |J_l^n| \leq \sum_{n \geq k} 2^{-n} = 2^{1-k}.$$

Thus shows that A has measure 0. For $t \notin A$, we may find k such that for all $n \geq k$ we have $t \notin \bigcup_l J_l^n$. This implies

$$|g_n(t) - g_{n+1}(t)| \leq 2^{-n}$$

for all $n \geq k$. In particular, $(g_k(t))$ is Cauchy for all $t \in A^c$. We may define

$$f(t) = \begin{cases} \lim_k g_k(t) & \text{if } t \notin A \\ 0 & \text{else} \end{cases}.$$

Then (g_k) converges to f almost everywhere. ■

This leads us to define the set of possible limits.

$$\mathcal{L} = \{f : [0, 1] \rightarrow \mathbb{R} : \exists (f_n) \subset C[0, 1], f_n \text{ converges to } f \text{ a.e.}\}$$

on \mathcal{L} we define the equivalence relation $f \sim g$ if there exists a set A of measure 0 such that $f1_{A^c} = g1_{A^c}$.

Exercise: Show that \sim is an equivalence relation.

We define

$$L = \mathcal{L} / \sim.$$

For a function $f \in \mathcal{L}$ we define the equivalence class $[f] = \{g : g \sim f\}$. In the following we denote by X the completion of $C[0, 1]$ with respect to the d_1 -metric. Our main theorem is the following.

inj THEOREM 6.6. *There is an injective map $\iota : X \rightarrow L$ such that*

$$\iota(x) = [f]$$

holds whenever (f_n) is a Cauchy sequence converging to x (with respect to d_1) and converging to f . a.e. Moreover, ι can be extended to $\iota(X)$.

Problem: Give a description of $\iota(X_1)$. This is done in the real analysis course (441=540).

We need some more preparation.

uniqueext LEMMA 6.7. *Let $A = \bigcup_k J_k$ the union of intervals.*

i) *Let $f \in C[0, 1]$, then $f1_A \in X$, $f1_{A^c} \in X$ and*

$$d_1(f1_A, g1_A) \leq d_1(f, g)$$

and

$$d_1(f1_{A^c}, g1_{A^c}) \leq d_1(f, g).$$

ii) *There is are continuous maps $m_A : X \rightarrow X$, $m_{A^c} : X \rightarrow X$ such that $m_A(f) = f1_A$ and $m_{A^c}(f) = f1_{A^c}$ for $f \in C[0, 1]$.*

iii) *There is a Lipschitz map $add : X \times X \rightarrow X$ such that $add(f, g) = f + g$.*

iv) *$add(m_A(x), m_{A^c}(x)) = x$ for all $x \in X$.*

v) *$d_1(f1_{A^c}, 0) \leq \sup_{t \in A^c} |f(t)|$.*

PROOF. We will start with i) for $A = [a, b]$. We use the functions Let f_n defined for $[0, b - a]$ and define $g_n(t) = f_n(t - a)$. Then we see that for every $f \in C[0, 1]$ we have

$$\begin{aligned} d_1(f f_n, f f_m) &= \int_a^b |f(t)(f_n(t) - f_m(t))| dt \leq \sup_t |f(t)| d(f_n, f_m) \\ &\leq \sup_t |f(t)| d_1(f_n, f_m) \leq \sup_t |f(t)| \left(\frac{1}{n} - \frac{1}{m} \right). \end{aligned}$$

Thus $(f f_n)$ is Cauchy. We denote the limit by $f1_{[a,b]}$. (Moreover, $f f_n$ converges pointwise to $f1_{[a,b]}$.) Now, we observe that $|f_n(t)| \leq 1$ and hence

$$d_1(f1_{[a,b]}, g1_{[a,b]}) = \lim_n d_1(f f_n, g f_n) = \lim_n \int_0^1 |f_n(t)(f(t) - g(t))| dt$$

$$\leq \int_0^1 |f(t) - g(t)| dt = d_1(f, g).$$

By Proposition [5.2](#) we find a map $u_A : X \rightarrow X$ such that $u_A(f) = f1_A$ and

$$d_1(u_A(x), u_A(y)) \leq d_1(x, y).$$

Now, we will prove iii). The metric on $X \times X$ is given by

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

Now, we consider $add : C[0, 1] \times C[0, 1] \rightarrow C[0, 1]$ and want to show that add is uniformly continuous. Indeed, elementary properties of the integral imply

$$\begin{aligned} d_1(add(f, g), add(f', g')) &= \int_0^1 |(f + g) - (f' + g')| dt \\ &\leq \int_0^1 |f - f'| dt + \int_0^1 |g - g'| dt = d((f, g), (f', g')). \end{aligned}$$

Thus Proposition [5.2](#) implies the assertion iii). In the next step we prove i) for $A = J_1 \cup \dots \cup J_n$. The key observation here is that we can find new intervals J'_1, \dots, J'_m such that the J'_l only overlap in one point and

$$A = \bigcup_l J'_l.$$

Therefore, we may define

$$u_A(x) = \sum_{l=1}^m 1_{J'_l} x = add(1_{J'_1} x, add(1_{J'_2} x, \dots add(1_{J'_{m-1}} x, 1_{J'_m} x) \dots)).$$

Being a composition of continuous function that is continuous. Moreover, for every l we may consider the sequence of function f_n^l constructed for the interval J'_l . The function

$$f_n(t) = \sum_{l=1}^m f_n^l(t)$$

is positive, continuous and vanishes in the overlapping endpoints. Hence $0 \leq f_n(t) \leq 1$ and the argument from above shows

$$d_1(f_n f, f_n g) \leq d_1(f, g).$$

This yields

$$d_1(f1_A, g1_A) = \lim_f d_1(f f_n, g f_n) \leq d_1(f, g).$$

Then we define

$$f1_{A^c} = f - f1_A = \lim_n \text{add}(f, -ff_n) = \lim_n f(1 - f_n).$$

Since $0 \leq 1 - f_n \leq 1$ we also prove as above that

$$d_1(f1_{A^c}, g1_{A^c}) \leq d_1(f, g).$$

Therefore i) is proved for A being a finite union of intervals. Let us show iv) for this particular case. Indeed,

$$\text{add}(f1_A, f1_{A^c}) = \lim_n \text{add}(ff_n, f(1 - f_n)) = \lim_n ff_n + f(1 - f_n) = f.$$

We need an additional estimate for showing the general case:

$$\boxed{\text{8-cont}} \quad (6.1) \quad d_1(f1_A, 0) \leq \sup_{t \in A} |f(t)| \sum_k |J_k|.$$

Indeed, inductively we may choose the non-overlapping J'_l in groups $J'_1, \dots, J'_{l_1}, J'_{l_1+1}, \dots, J'_{l_2}, \dots$ such

$$\sum_{l=l_{k-1}+1}^{l_{k+1}} |J'_l| \leq |J_k|.$$

Then, we have

$$\begin{aligned} d_1(f1_A, 0) &= \lim_n \int_0^1 |f_n f(t)| dt \leq \sup_{t \in A} |f(t)| \lim_n \int_0^1 f_n(t) dt \\ &= \sup_{t \in A} |f(t)| \sum_l |J'_l| \leq \sup_{t \in A} |f(t)| \sum_k |J_k|. \end{aligned}$$

Now, we consider the general case $A = \bigcup_k J_k$. We define $A_n = \bigcup_{k \leq n} J_k$. We want to show that $f1_{A_n}$ is Cauchy. For this we choose non-overlapping intervals $J'_{l_{k-1}+1}, \dots, J'_{l_k} \subset J_k$. Then, we deduce from $\boxed{\text{8-cont}}$ (6.1) that for $n \leq m$

$$\boxed{\text{chy3}} \quad (6.2) \quad d_1(f1_{A_n}, f1_{A_m}) \leq \sup_t |f(t)| \sum_{k=n+1}^m \sum_{l=l_{k-1}+1}^{l_k} |J'_l| \leq \sup_t |f(t)| \sum_{k=n+1}^m |J_k|.$$

Thus we may define $f1_A = \lim_n f1_{A_n}$. Again, we have

$$d_1(f1_A, g1_A) = \lim_n d_1(f1_{A_n}, g1_{A_n}) \leq d_1(f, g).$$

By the unique extension principle, we find a Lipschitz map $u_A : X \rightarrow X$ such that $u_A(f) = f1_A$. We use the unique extension principle to define $-x$ and then define $u_{A^c}(x) = \text{add}(x, -u_A(x))$. For $f, g \in C[0, 1]$ we have

$$d_1(u_{A^c}(f), u_{A^c}(g)) = \lim_n d_1(f1_{A_n^c}, g1_{A_n^c}) \leq d_1(f, g).$$

Thus by unique extension this also holds for $x, y \in X$. Finally, we note that for $f \in C[0, 1]$

$$\text{add}(f1_A, f1_{A^c}) = \text{add}(f1_A, \text{add}(f, -f1_A)) = \lim_n \text{add}(f1_{A_n}, \text{add}(f, -f1_{A_n})) = f.$$

Indeed, the equality holds for every $n \in \mathbb{N}$. Now, we will prove v). We may assume that

$$A = \bigcup_k J_k$$

such that the J_k 's are non-overlapping. We define

$$A_n = \bigcup_{k \leq n} J_k.$$

Let $\varepsilon > 0$ and $\delta > 0$ such that

$$|t - s| < \delta \quad \Rightarrow \quad |f(t) - f(s)| < \frac{\varepsilon}{2}.$$

Now, we consider $x \in A$ such that

$$d(x, A^c) = \inf_{y \in B} |x - y| \geq \delta$$

This means that

$$B(x, \delta) = \bigcup_k J_k \cap B(x, \delta).$$

This implies

$$\lim_n |B(x, \delta) \cap A_n^c| = 0.$$

Moreover, we can find a maximal family x_1, \dots, x_m of such points such that $d(x, A^c) \geq \delta$ implies $d(x, x_j) < \delta$ for some j . Then, we may choose n large enough such that

$$\sup_t |f(t)| \sum_{j=1}^m |B(x_j, \delta) \cap A_n^c| \leq \frac{\varepsilon}{2}.$$

Now, we define $D = \bigcup_j \bar{B}(x_j, \delta)$. Since $A_n^c \cap D^c$ is again a collection of intervals we see that

$$\begin{aligned} d_1(f1_{A_n^c}, 0) &\leq d_1(f1_{A_n^c \cap D}, 0) + d_1(f1_{A_n^c \cap D^c}, 0) \\ &\leq \sup_t |f(t)| \sum_{j=1}^m |B(x_j, \frac{\delta}{2}) \cap A_n^c| + \sup_{t \in A_n^c \cap D^c} |f(t)|. \end{aligned}$$

Now, we consider $t \in A_n^c \cap D^c$. If $d(t, B) \geq \frac{\delta}{2}$, then $d(t, x_j) \leq \frac{\delta}{2}$ hence $t \in D$. Thus we may assume $d(t, B) < \frac{\delta}{2}$. Then we find $s \in B$ such that $|t - s| < \delta$ and thus

$$|f(t)| \leq \sup_{s \in B} |f(s)| + \varepsilon.$$

This implies

$$d_1(f1_{A_n^\varepsilon}, 0) \leq \frac{\varepsilon}{2} + \varepsilon + \sup_{s \in B} |f(s)|.$$

for $n \geq n_0$. The assertion is proved. ■

REMARK 6.8. For $J = [a, b]$ we have

$$I(f1_{[a,b]}) = \int_a^b f(t) dt.$$

approx LEMMA 6.9. Let (A_m) be a sequence such that

$$A_m = \bigcup_k J_k^m$$

and

$$\lim_m \sum_k |J_k^m| = 0.$$

Then

$$\lim_m d_1(x1_{A_m}, 0) = 0$$

holds for every $x \in X$.

PROOF. Let (f_n) be Cauchy sequence converging to x . Let $\varepsilon > 0$. Then, we may choose n such that

$$d_1(f_n, x) < \frac{\varepsilon}{2}.$$

Since u_{A_m} is Lipschitz, we deduce

$$d_1(f_n 1_{A_m}, u_{A_m}(x)) < \frac{\varepsilon}{2}.$$

for all $m \in \mathbb{N}$. According to [\(6.1\)](#) ^{8-cont} we find

$$d_1(f_n 1_{A_m}, 0) \leq \sup_t |f_n(t)| \sum_k |J_k^m|.$$

By assumption, we may find m_0 such that

$$d_1(f_n 1_{A_m}, 0) < \frac{\varepsilon}{2}$$

holds for all $m \geq m_0$. This implies

$$d_1(u_{A_m}(x), 0) \leq d_1(u_{A_m}(x), f_n 1_{A_m}) + d_1(f_n 1_{A_m}, 0) < \varepsilon$$

for all $m > m_0$. ■

PROOF OF THEOREM ^{inj}5.8. Let $(f_n), (g_n)$ be Cauchy such that $d_1 - \lim_n f_n = x$ and $d_1 - \lim_n g_n = y$ and f_n converges to f and g_n converges to g a.e. We consider $h'_n = f_n - g_n$ and $z = \text{add}(x, -y)$. We want to show

$$d_1(x, 0) = 0.$$

Clearly, h'_n converges to 0 almost everywhere. Passing to subsequence (h_k) we may assume that $d_1(h_k, h_{k+1}) \leq 6^{-n}$. According to the proof of Proposition ^{limit1}8.1, we find $A_k = \bigcup_l J_k^l$ such that

$$\sum_l |J_k^l| \leq 2^{1-k}$$

such that

$$|h_n(t) - h_{n+1}(t)| \leq 2^{-n}$$

for all $t \notin A_k$. By the definition of a.e. we find $B_k = \bigcup_l \tilde{J}_k^l$ such that

$$\sum_l |\tilde{J}_k^l| \leq 2^{1-k}$$

and

$$\lim_t h_n(t) = 0$$

holds for all $t \notin B_k$. Thus we define

$$C_k = \bigcup_l J_k^l \cup \bigcup_l \tilde{J}_k^l.$$

Then

$$\sum_l |J_k^l| + \sum_l |\tilde{J}_k^l| \leq 2^{2-k}$$

and for all $t \notin C_k$ we have

$$|h_n(t)| = \left| \lim_m h_n(t) - h_m(t) \right| \leq \limsup_m |h_n(t) - h_m(t)| \leq 2^{1-n}.$$

According to Lemma ^{approx}(6.9) we find a k_0 such that

$$d_1(u_{C_k}(z), 0) < \varepsilon$$

for all $k \geq k_0$. For all $n \geq k$ we deduce from Lemma ^{uniqueext}6.7 that

$$d_1(1_{C_k^c} f_n, 0) \leq 2^{1-n}.$$

Therefore

$$d_1(u_{C_k^c}(z), 0) = \lim_n d_1(u_{C_k^c}(f_n), 0) = 0.$$

Since $x = \text{add}(u_{C_k^c}(x), u_{C_k^c}(x))$, we deduce

$$d_1(z, 0) = d_1(\text{add}(u_{C_k}(z), u_{C_k^c}(z)), \text{add}(u_{C_k}(0), u_{C_k^c}(0)))$$

$$\leq d_1(u_{C_k}(z), 0) + d_1(u_{C_k^c}(z), 0) < \varepsilon .$$

Finally, we observe that

$$d_1(x, y) = d_1(\text{add}(x, -y), 0)$$

holds by unique extension. Thus $x = y$. ■

7. Closed and Compact Sets

Let (X, d) be a metric space. We will say that a subset $A \subset X$ is *closed* if $X \setminus A$ is open.

closed PROPOSITION 7.1. *Let (X, d) be a complete metric space and $C \subset X$ a subset. C is closed iff every Cauchy sequence in C converges to an element in C .*

Proof: Let us assume C is closed and that (x_n) is a Cauchy sequence with elements in C . Let $x = \lim_n x_n$ be the limit and assume $x \notin C$. Since $X \setminus C$ is open

$$B(x, \varepsilon) \subset X \setminus C$$

for some $\varepsilon > 0$. Then there exists an n_0 such that $d(x_n, x) < \varepsilon$ for $n > n_0$. In particular,

$$x_{n_0+1} \in B(x, \varepsilon)$$

and thus $x_{n_0+1} \notin C$, a contradiction.

Now, we assume that every Cauchy sequence with values in C converges to an element in C . If $X \setminus C$ is not open, then there exists an $x \notin C$ and no $\varepsilon > 0$ such that

$$B(x, \varepsilon) \subset X \setminus C .$$

I.e. for every $n \in \mathbb{N}$, we can find $x_n \in C$ such that

$$d(x, x_n) < \frac{1}{n} .$$

Hence, $\lim x_n = x \in C$ but $x \notin C$, contradiction. ■

The most important notion in this class is the notion of compact sets. We will say that a subset $C \subset X$ is *compact* if For every collection (O_i) of open sets such that

$$C \subset \bigcup_i O_i = \{x \in X \mid \exists_{i \in I} x \in O_i\}$$

There exists $n \in \mathbb{N}$ and i_1, \dots, i_n such that

$$C \subset O_{i_1} \cup \dots \cup O_{i_n} .$$

In other words

Every open cover of C has a finite subcover .

DEFINITION 7.2. Let $X \subset \bigcup O_i$ be an open cover. Then we say that (V_j) is an open subcover if

$$X \subset \bigcup_j V_j$$

all the V_j are open and for every j there exists an i such that

$$V_j \subset O_i.$$

It is impossible to explain the importance of ‘compactness’ right away. But we can say that there would be no discipline ‘Analysis’ without compactness. The most clarifying idea is contained in the following example.

PROPOSITION 7.3. The set $[0, 1] \subset \mathbb{R}$ is compact.

Proof: Let $[0, 1] \subset \bigcup_i O_i$. For every $x \in [0, 1]$ there exists an $i \in I$ such that

$$x \in O_i.$$

Since O_i is open, we can find $\varepsilon > 0$ such that

$$x \in B(x, \varepsilon) \subset O_i.$$

Using the axiom of choice, we find a function ε_x and i_x such that

$$x \in B(x, \varepsilon_x) \subset O_{i_x}.$$

Let us define the relation $x \preceq y$ if $x < y$ and

$$y - x \leq \varepsilon_x + \varepsilon_y.$$

The crucial point here is to define

$$S = \{x \in [0, 1] \mid \exists x_1, \dots, x_n : \frac{1}{2} \preceq x_1 \preceq \dots \preceq x_n \preceq x\}.$$

We claim a) $\sup S \in S$ and b) $\sup S = 1$.

Ad a): Let $y = \sup S \in [0, 1]$. Then there exists an $x \in S$ with

$$y - \varepsilon_y < x \leq y.$$

Then obviously $x \preceq y$. Since $x \in S$, we can find

$$\frac{1}{2} \preceq x_1 \preceq \dots \preceq x_n \preceq x \preceq y.$$

Thus $y \in S$.

Ad b): Assume $y = \sup S < 1$. Let $0 < \delta = \min(\varepsilon_y, 1 - y)$. Then

$$y + \delta - y = \delta \leq \varepsilon_y + \varepsilon_{y+\delta}.$$

By a), we find

$$\frac{1}{2} \preceq x_1 \preceq \cdots \preceq x_n \preceq y \preceq y + \delta$$

and thus $y + \delta \in S$. Contradiction to the definition of the supremum. Assertion a) and b) are proved.

Therefore we conclude $1 \in S$ and thus find x_1, \dots, x_n such that

$$\frac{1}{2} \preceq x_1 \preceq \cdots \preceq x_n \preceq 1.$$

Let $x_0 = \frac{1}{2}$ and $x_{n+1} = 1$, then by definition of \preceq , we have

$$[x_j, x_{j+1}] \subset B(x_j, \varepsilon_{x_j}) \cup B(x_{j+1}, \varepsilon_{x_{j+1}}) \subset O_{i_{x_j}} \cup O_{i_{x_{j+1}}}$$

for $j = 0, \dots, n$. Thus, we deduce

$$\left[\frac{1}{2}, 1\right] \subset \bigcup_{j=0}^n [x_j, x_{j+1}] \subset \bigcup_{j=0}^{n+1} O_{i_{x_j}}.$$

Doing the same trick with $[0, \frac{1}{2}]$, we find

$$[0, 1] \subset \bigcup_{j=0}^{m+1} O_{i_{x'_j}} \cup \bigcup_{j=0}^{n+1} O_{i_{x_j}}$$

and we have found our finite subcover. ■

subcom

PROPOSITION 7.4. Let $B \subset X$ be closed set and $C \subset X$ be a compact set, then

$$B \cap C$$

is compact

Proof: Let $B \cap C \subset \bigcup O_i$ be an open cover. then

$$C \subset (X \setminus B) \cup \bigcup_i O_i$$

is an open cover for C , hence we can find a finite subcover

$$C \subset (X \setminus B) \cup O_{i_1} \cup \cdots \cup O_{i_n}.$$

Thus

$$B \cap C \subset O_{i_1} \cup \cdots \cup O_{i_n}$$

is a finite subcover. ■

ccc LEMMA 7.5. Let (X, d) be a metric space and $D \subset X$ be a countable dense set in X , then for every subset $C \subset X$ and every open cover

$$C \subset \bigcup_i O_i$$

we can find a countable subcover of balls.

Proof: Let us enumerate D as $D = \{d_n \mid n \in \mathbb{N}\}$. Let $x \in C$ and find $i \in I$ and $\varepsilon > 0$ such that

$$x \in B(x, \varepsilon) \subset O_i.$$

Let $k > \frac{2}{\varepsilon}$. By density, we can find an $n \in \mathbb{N}$ such that

$$d(x, d_n) < \frac{1}{2k}.$$

Then

$$x \in B(d_n, \frac{1}{2k}) \subset B(x, \frac{1}{k}) \subset B(x, \varepsilon) \subset O_i.$$

Let us define

$$M = \{(n, k) \mid \exists_{i \in I} B(d_n, \frac{1}{2k}) \subset O_i\}.$$

Then $M \subset \mathbb{N}^2$ is countable and hence there exists a map $\phi : \mathbb{N} \rightarrow M$ which is surjective (=onto). Hence for $V_m = B(d_{\phi_1(m)}, \frac{1}{2\phi_2(m)})$, ϕ_1, ϕ_2 the 2 components of ϕ we have

$$C \subset \bigcup_m V_m$$

and (V_m) is a countable subcover of balls of the original cover (O_i) . ■

main THEOREM 7.6. Let (X, d) be a metric space. Let $C \subset X$ be a subset. Then the following are equivalent

- i) a) Every Cauchy sequence of elements in C converges to a limit in C .
- b) For every $\varepsilon > 0$ there exists points $x_1, \dots, x_n \in X$ such that

$$C \subset B(x_1, \varepsilon) \cup \dots \cup B(x_n, \varepsilon).$$

- ii) Every sequence in C has a convergent subsequence.
- iii) C is compact.

Proof: $i) \Rightarrow ii)$. Let (x_n) be a sequence. Inductively, we will construct infinite subset $A_1 \supset A_2 \supset A_3 \cdots$ and y_1, y_2, y_3, \dots in X such that

$$\forall_{l \in A_j} : d(x_l, y_j) < 2^{-j-1}.$$

Put $A_0 = \mathbb{N}$. Let us assume $A_1 \supset A_2 \supset \cdots \supset A_n$ and y_1, \dots, y_n have been constructed. We put $\varepsilon = 2^{-n-2}$ and apply condition *i)b)* to find z_1, \dots, z_m such that

$$C \subset B(z_1, \varepsilon) \cup \cdots \cup B(z_m, \varepsilon).$$

We claim that there must be a $1 \leq k \leq m$ such that

$$A_n(k) = \{l \in A_n \mid x_n \in B(z_k, \varepsilon)\}$$

has infinitely many elements. Indeed, we have

$$A_n(1) \cup \cdots \cup A_n(m) = A_n.$$

If they were all finite, then a finite union of finite sets would have finitely many elements. However A_n is infinite. Contradiction! Thus, we can find a k with $A_n(k)$ infinite and put $A_{n+1} = A_n(k)$ and $y_{n+1} = z_k$. So the inductive procedure is finished. Now, we can find an increasing sequence (n_j) such that $n_j \in A_j$ and deduce

$$d(x_{n_j}, x_{n_{j+1}}) \leq d(x_{n_j}, y_j) + d(y_j, x_{n_{j+1}}) < \frac{1}{2}2^{-j} + \frac{1}{2}2^{-j} = 2^{-j}$$

because $n_j \in A_j$ and $n_{j+1} \in A_{j+1} \subset A_j$. Thus (x_{n_j}) is Cauchy. Indeed, by induction, we deduce for $j < m$ that

$$\begin{aligned} d(x_{n_j}, x_{n_m}) &\leq d(x_{n_j}, x_{n_{j+1}}) + d(x_{n_{j+1}}, x_{n_{j+2}}) \cdots d(x_{n_{m-1}}, x_{n_m}) \\ &\leq 2^{-j} \sum_{k=0}^{m-1} 2^{-k} = 2^{1-j}. \end{aligned}$$

This easily implies the Cauchy sequence condition. By a) it converges to some $x \in C$. We got our convergent subsequence.

*ii) \Rightarrow *iii)**: We will first show *ii) \Rightarrow *i)b)**. Indeed, let $\varepsilon > 0$ and assume for all $n \in \mathbb{N}$, $y_1, \dots, y_n \in C$ we may find

$$x(n, y_1, \dots, y_n) \in C \setminus (B(y_1, \varepsilon) \cup \cdots \cup B(y_n, \varepsilon)).$$

Then we define $x_1 \in C$ and find $x_2 \in C \setminus B(x_1, \varepsilon)$. Then we find $x_3 \in C \setminus B(x_1, \varepsilon) \cup B(x_2, \varepsilon)$. Thus inductively we find $x_n \in C$ such that

$$d(x_n, x_k) \geq \varepsilon$$

for all $1 \leq k \leq n$. It is easily seen that (x_n) has no convergent subsequence. Thus *i)b)* is showed (with points in C). For every $\varepsilon_k = \frac{1}{k}$ we find these points $y_1^k, \dots, y_{m(k)}^k \in C$ such that

$$C \subset B(y_1^k, \frac{1}{k}) \cup \cdots \cup B(y_{m(k)}^k, \frac{1}{k}).$$

Then, we see that $D = \{y_j^k : k \in \mathbb{N}, 1 \leq j \leq m(k)\}$ is dense in C . Therefore, we may work with the closure $\tilde{X} = \bar{D}$ and show that C is compact in \tilde{X} . (It will then be automatically compact in X). By Lemma ^{ccc}7.5, we may assume that

$$C \subset \bigcup_k O_k$$

and O_k 's open. If we can find an n such that

$$C \subset O_1 \cup \cdots \cup O_n$$

the assertion is proved. Assume that is not the case and choose for every $n \in \mathbb{N}$ an $x_n \in C \setminus O_1 \cup \cdots \cup O_n$. According to the assumption, we have a convergent subsequence, i.e. $\lim_k x_{n_k} = x \in C$. Then $x \in O_{n_0}$ for some n_0 and there exists a $\varepsilon > 0$ such that

$$B(x, \varepsilon) \subset O_{n_0}.$$

By convergence, we find a k_0 such that $d(x, x_{n_k}) < \varepsilon$ for all $k > k_0$. In particular, we find a $k > k_0$ such that $n_k > n_0$. Thus

$$x_{n_k} \in B(x, \varepsilon) \subset O_{n_0} \subset O_1 \cup \cdots \cup O_{n_k}.$$

Contradicting the choice of the (x_n) 's. We are done.

iii) \Rightarrow i)b) Let $\varepsilon > 0$ and then

$$C \subset \bigcup_{x \in C} B(x, \varepsilon).$$

thus a finite subcover yields *b*).

iii) \Rightarrow i)a) Let (x_n) be a Cauchy sequence. Assume it is not converging to some element $x \in C$. This means

$$\boxed{\text{ccc}} \quad (7.1) \quad \forall x \in C \exists \varepsilon(x) > 0 \forall n_0 \exists n > n_0 d(x_n, x) > \varepsilon.$$

Then

$$C \subset \bigcup_{x \in C} B(x, \frac{\varepsilon(x)}{2}).$$

Let

$$C \subset B(y_1, \frac{\varepsilon(y_1)}{2}) \cup \cdots \cup B(y_m, \frac{\varepsilon(y_m)}{2})$$

be a finite subcover (compactness). Then there exists at least one $1 \leq k \leq m$ such that

$$A_k = \{n \in \mathbb{N} \mid d(x_n, y_k) < \frac{\varepsilon(y_k)}{2}\}$$

is infinite. Fix that k and apply the Cauchy criterion to find n_0 such that

$$d(x_n, x_{n'}) < \frac{\varepsilon(y_k)}{2}$$

for all $n, n' > n_0$. By $\overbrace{(7.1)}^{\text{CCC}}$, we can find an $n > n_0$ such that

$$d(x_n, y_k) > \varepsilon(y_k).$$

Since A_k is infinite, we can find an $n' > n_0$ in A_k thus

$$\begin{aligned} \varepsilon(y_k) &< d(x_{n'}, y_k) \leq d(x_n, x_{n'}) + d(x_n, y_k) \\ &< \frac{\varepsilon(y_k)}{2} + \frac{\varepsilon(y_k)}{2} = \varepsilon(y_k). \end{aligned}$$

A contradiction. Thus the Cauchy sequence has to converge to some point in C . ■

COROLLARY 7.7. *Every interval $[a, b] \subset \mathbb{R}$ with $a < b \in \mathbb{R}$ is compact*

Proof: It is easy to see that $X \setminus [a, b]$ is open. Hence, by Proposition $\overbrace{7.1}^{\text{closed}}$ $[a, b]$ is complete, i.e. i)a) is satisfied. Given $\varepsilon > 0$, we can find $k > \frac{1}{\varepsilon}$. For $m > k(b-a)$ we derive

$$[a, b] \subset \bigcup_{j=0}^m B(a + \frac{j}{k}, \varepsilon).$$

Thus the Theorem $\overbrace{7.6}^{\text{main}}$ applies. ■

cube

LEMMA 7.8. *Let $r > 0$ and $n \in \mathbb{N}$, the set $C_r = [-r, r]^n$ is compact.*

Proof: Let $x \notin C_r$, then there exists an index $j \in \{1, \dots, n\}$ such that $|x_j| > r$. Let $\varepsilon = |x_j| - r$ and $y \in \mathbb{R}^n$ such that

$$\max_{i=1, \dots, n} |x_i - y_i| < \varepsilon,$$

then

$$|y_j| = |y_j - x_j + x_j| \geq |x_j| - |y_j - x_j| > |x_j| - \varepsilon = r.$$

thus $y \notin C_r$. Hence, C_r is closed and according to Proposition $\overbrace{4.3}^{\text{rncomp}}$, we deduce that C_r is complete.

For $n = 1$ and $\varepsilon > 0$, we have seen above that for $k > \frac{1}{\varepsilon}$ and $m > \frac{2r}{k}$

$$[-r, r] \subset \bigcup_{j=0}^m B(-r + \frac{j}{k}, \varepsilon).$$

Therefore

$$[-r, r]^n \subset \bigcup_{j_1, \dots, j_n=0, \dots, m} B_\infty((-r + \frac{j_1}{k}, \dots, -r + \frac{j_n}{k}), \varepsilon).$$

Thus i)a) and i)b) are satisfied and the Theorem $\overbrace{7.6}^{\text{main}}$ implies the assertion (The separable dense subset is \mathbb{Q}^n .) ■

THEOREM 7.9. *Let $C \subset \mathbb{R}^n$ be a subset. The following are equivalent*

- 1) C is compact.
- 2) C is closed and there exists an r such that

$$C \subset B(0, R) .$$

(That is C is bounded.)

Proof: 2) \Rightarrow 1) Let

$$C \subset B(0, R) \subset [-R, R]^n$$

be a closed set. Since $[-R, R]^n$ is compact, we deduce from Proposition ^{subcom}7.4 that C is compact as well.

1) \Rightarrow 2) Let C subset \mathbb{R}^n be a compact set. According to Theorem ^{main}7.6 i)b), we find

$$C \subset B(x_1, 1) \cup \cdots \cup B(x_m, 1)$$

thus for $r = \max_{i=1, \dots, m} (d(x_i, 0) + 1)$ we have

$$C \subset B(0, r) .$$

Moreover, by Theorem ^{main}7.6 i)a) and Proposition ^{closed}7.1, we deduce that C is closed. ■

We will now discuss one of the most important applications.

THEOREM 7.10. *Let (X, d) be a compact metric space and $f : X \rightarrow \mathbb{R}$ be a continuous function. Then there exists $x_0 \in X$ such that*

$$f(x_0) = \sup\{f(x) : x \in X\} .$$

PROOF. Let us first assume

$$A = \{f(x) : x \in X\}$$

is bounded and $s = \sup A$. For every $n \in \mathbb{N}$, we know that $s - \frac{1}{n}$ is no upper bound. Hence there $x_n \in X$ such that

$$s \geq f(x_n) > s - \frac{1}{n} .$$

Let (n_k) be such that $\lim_k x_{n_k} = x \in X$. Then we deduce from continuity that

$$f(x) = \lim_k f(x_{n_k}) \geq \lim_k s - \frac{1}{n_k} = s .$$

By definition of s we find $f(x) = s$. Now, we show that A is bounded. Indeed, if note we find $x_n \in X$ such that $f(x_n) \geq n$. Again we find a convergent subsequence

(x_{n_k}) . Since $f(x_{n_k})$ is convergent it is bounded. We assume (f_{n_k}) is bounded above by $m \in \mathbb{N}$. Choosing $k \geq m + 1$ we get

$$m \geq f(x_{n_k}) \geq n_k > n_m \geq m .$$

This contradiction shows that A is bounded and hence the first argument applies. ■

8. C(K)

For a metric space X , we denote by $C(X)$ the space of continuous functions with values in the real numbers.

simple LEMMA 8.1. *Let (X, d) be a metric space and $f, g \in C(X)$ and $t \in \mathbb{R}$, then*

- i) $f + tg$ defined by $f + tg(x) = f(x) + tg(x)$, $x \in X$, is in $C(X)$.
- ii) fg defined by $fg(x) = f(x)g(x)$, $x \in X$, is in $C(X)$.
- iii) Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a continuous function, then $h \circ f$ defined by $h \circ f(x) = h(f(x))$ is continuous.
- iv) Let (f_n) be a sequence of continuous functions such that for every $\varepsilon > 0$ there exists n_0 such that for all $n, m > n_0$

$$\sup_{x \in X} |f_n(x) - f_m(x)| \leq \varepsilon,$$

Then there is a continuous function $f : X \rightarrow \mathbb{R}$ such that

$$f(x) = \lim_n f_n(x).$$

Proof: *iii)* Let us assume that (f_n) is a sequence as above. Clearly for all $x \in X$, we see that

$$f(x) = \lim_n f_n(x)$$

exists. We have to show that f is continuous. For let $x \in X$ and $\varepsilon > 0$. Let n_0 be chosen according such that for $n, m > n_0$

$$\sup_{x \in X} |f_n(x) - f_m(x)| \leq \frac{\varepsilon}{3}.$$

In particular, for all $y \in X$ and for $n = n_0 + 1$, we deduce

limit1 (8.1) $|f(y) - f_n(y)| = \lim_m |f_m(y) - f_n(y)| \leq \frac{\varepsilon}{3}.$

Since, f_n is continuous (at x), we can find $\delta > 0$ such that $d(x, y) < \delta$ implies

$$|f_n(x) - f_n(y)| < \frac{\varepsilon}{3}.$$

Thus we get for those y

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

Hence f is continuous and the assertion is proved. ■

THEOREM 8.2. *Let K be a compact metric space and $f : K \rightarrow \mathbb{R}$ be a continuous function. Then there exists an $x_0 \in K$ such that*

$$f(x_0) = \sup\{f(x) \mid x \in K\}.$$

Proof: Note that in the proof of the main Theorem the existence of a countable dense subset has only been used to prove $ii) \Rightarrow iii)$. Thus, we still have that every sequence in a compact space has a convergent subsequence. If $\sup\{f(x) : x \in K\} = \infty$, we can find a sequence (x_n) such that

$$f(x_n) > n$$

for all $n \in \mathbb{N}$. Let us consider this case first. Let (x_{n_k}) be a convergent subsequence with

$$x = \lim_k x_{n_k} \in K.$$

Then we can find an $\varepsilon > 0$ such that

$$d(x, y) < \varepsilon \quad \Rightarrow \quad |f(x) - f(y)| \leq 1$$

Then there exists an k_0 such that $d(x_{n_k}, x) < \varepsilon$ for $k > k_0$. Let k_1 be such that $k_1 > |f(x)| + 2$, then we deduce for some $k_2 > \max\{k_1, k_2\}$

$$|f(x)| + 1 \leq k_2 \leq f(x_{k_2}) \leq f(x) + |f(x) - f(x_{k_2})| < |f(x)| + 1.$$

a contradiction. Thus $\sup\{f(x) : x \in K\} < \infty$. For every $\varepsilon > 0$, there exists an $x(\varepsilon) \in K$ such that

$$f(x(\varepsilon)) > \sup\{f(x) : x \in K\}.$$

Call the supremum *sup*. We get a sequence (x_n) such that

$$f(x_n) \leq \text{sup} \leq f(x_n) + \frac{1}{n}$$

Let (x_{n_k}) be a convergent subsequence with

$$x = \lim_k x_{n_k} \in K.$$

By continuity of f , we deduce

$$f(x) = \lim_k f(x_{n_k}) = \text{sup}.$$

The assertion is proved. ■

COROLLARY 8.3. Let K be a compact set and $f : K \rightarrow \mathbb{R}$ be a continuous function, then

$$\sup\{f(x) : x \in K\}$$

is finite.

COROLLARY 8.4. The space $C(K)$ equipped with

$$d(f, g) = \sup_{x \in K} |f(x) - g(x)|$$

is a complete metric space.

Proof: Let us first observe that for $f, g \in C(K)$ the map $|f - g|$ is continuous and thus

$$d(f, g)$$

is a real number. Clearly, d is symmetric and $f = g$, i.e. $f(x) = g(x)$ for all $x \in K$ iff $d(f, g) = 0$. The triangle inequality is obvious. Indeed, let $f, g, h \in C(K)$. Then

$$\begin{aligned} \sup_{x \in K} |f(x) - g(x)| &\leq \sup_{x \in K} |f(x) - h(x) + h(x) - g(x)| \\ &\leq \sup_{x \in K} (|f(x) - h(x)| + |h(x) - g(x)|) \\ &\leq \sup_{x \in K} |f(x) - h(x)| + \sup_{x \in K} |h(x) - g(x)| \\ &= d(f, h) + d(h, g). \end{aligned}$$

Given a Cauchy sequence (f_n) , we apply Lemma ^{simple}8.1 to obtain a continuous limit function f . According to ^{limit1}(8.1), we see that

$$\lim_{n \rightarrow \infty} d(f, f_n) = 0.$$

Hence, f_n converges to f . (Details: Exercise.) ■

Motivated by this result, we will say that a sequence of functions (f_n) converges uniformly to f (on X) if

$$\forall \varepsilon > 0 \exists n_0 \forall n > n_0 \forall x \in X |f_n(x) - f(x)| < \varepsilon.$$

This opposed to the pointwise convergence

$$\forall x \in X \forall \varepsilon > 0 \exists n_0 \forall n > n_0 |f_n(x) - f(x)| < \varepsilon.$$

Example: The functions $f_n(x) = x^n$ converge pointwise to $f(x) = 0$ on $[0, 1]$

However, the following remarkable result allows us to show that under suitable circumstances the weaker pointwise convergence implies uniform convergence.

THEOREM 8.5. *Let K be a compact space. (f_n) a sequence of continuous functions on K converging pointwise to the continuous function f on K such that for all $x \in K$ the sequence $f_n(x)$ is increasing. Then (f_n) converges uniformly to f .*

Proof: Let $\varepsilon > 0$. Then, we can find for every $x \in K$ an $n(x)$ such that

$$f(x) < f_{n(x)} + \frac{\varepsilon}{3}.$$

Since $f, f_{n(x)}$ are continuous, we can find $\delta(x) > 0$ such that

$$\delta(x, y) < \delta(x) \quad \Rightarrow \quad (|f(x) - f(y)| < \frac{\varepsilon}{3} \quad \& \quad |f_{n(x)}(x) - f_{n(x)}(y)| < \frac{\varepsilon}{3}.$$

Then we have an open cover

$$X \subset \bigcup_{x \in K} B(x, \frac{\delta(x)}{2})$$

by compactness can find a finite subcover

$$X \subset B(x_1, \delta(x_1)) \cup \dots \cup B(x_m, \delta(x_m)).$$

Let $n(\varepsilon) = \max\{n(x_1), \dots, n(x_m)\}$. Let $y \in K$ and find an $1 \leq i \leq m$ such that $d(x_i, y) < \delta(x_i)$. Then, we have for $x = x_i$

$$\begin{aligned} f(y) &\leq f(x) + \frac{\varepsilon}{3} < f_{n(x)} + \frac{2\varepsilon}{3} \\ &\leq f_n(x) + \frac{2\varepsilon}{3} \leq f_n(y) + \varepsilon. \end{aligned}$$

By monotonicity, we deduce for all $m > n$

$$f(y) < f_m(y) + \varepsilon \leq f(y) + \varepsilon.$$

The assertion is proved. ■

Let us state a further important application of compactness.

THEOREM 8.6. *Let $K \subset (X, d)$ be a compact subset and $f : K \rightarrow (Y, d)$ be a continuous function. Then for every $\varepsilon > 0$ there exists an $\delta > 0$ such that for all $x, y \in K$*

$$d(x, y) < \delta \quad \Rightarrow \quad d(f(x), f(y)) < \varepsilon.$$

Proof: For every $x \in K$, we can find $\delta(x)$ such that

$$d(x, y) < \delta \quad \Rightarrow \quad d(f(x), f(y)) < \frac{\varepsilon}{2}.$$

Then the open

$$X \subset \bigcup_{x \in X} B(x, \frac{\delta(x)}{2})$$

has a finite subcover

$$X \subset B(x_1, \frac{\delta(x_1)}{2}) \cup \cdots \cup B(x_m, \frac{\delta(x_m)}{2}).$$

Let $\delta = \min_{i=1, \dots, m} \{\frac{\delta(x_i)}{2}\}$ and consider $x, y \in K$ such that $d(x, y) < \delta$. Then there exists $1 \leq i \leq m$ such that $d(x, x_i) < \frac{\delta(x_i)}{2}$. Hence,

$$d(y, x_i) \leq d(y, x) + d(x, x_i) < \delta(x_i)$$

and thus

$$d(f(x), f(y)) \leq d(f(x), f(x_i)) + d(f(x_i), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

The assertion is proved. ■

DEFINITION 8.7. *A function satisfies the assertion of the previous theorem is called uniformly continuous.*

If time remains, we will show:

THEOREM 8.8. *The closure of the polynomials are dense in $C([a, b])$.*

Proof: *ii*) : Let $x \in X$ and $0 < \varepsilon < 1$, then there exists a $\varepsilon_f > 0$ such that

$$d(x, y) < \varepsilon_f \quad \Rightarrow \quad d(f(x), f(y)) < \frac{\varepsilon}{3(1 + |g(x)|)} < 1,$$

and $\varepsilon_g > 0$ such that

$$d(x, y) < \varepsilon_g \quad \Rightarrow \quad d(g(x), g(y)) < \frac{\varepsilon}{2(1 + |g(x)|)} < 1.$$

Let $y \in X$ such that $d(y, x) < \min\{\varepsilon_f, \varepsilon_g, 1\} = \delta$. Then we deduce from $\varepsilon < 1$

$$\begin{aligned} |fg(x) - fg(y)| &\leq |f(x)||g(x) - g(y)| + |f(x)g(y) - f(y)g(y)| \\ &\leq |f(x)||g(x) - g(y)| + |f(x) - f(y)||g(x)| + |f(x) - f(y)||g(x) - g(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon^2}{9} < \varepsilon. \end{aligned}$$

9. Some aspects of topology

DEFINITION 9.1. A topological space is given by a tuple (X, \mathcal{O}) where X is a set $\mathcal{O} \subset P(X)$ is a collection of set such that

- i) $\emptyset \in \mathcal{O}$,
- ii) If $O_1, O_2 \in \mathcal{O}$, then $O_1 \cap O_2 \in \mathcal{O}$,
- iii) If $(O_i)_i \subset \mathcal{O}$, then $\bigcup_i O_i \in \mathcal{O}$.

A basis \mathcal{B} of a topology is a collection of set such that $O \in \mathcal{O}$ if and only if for every $x \in O$ there are $B_1, \dots, B_m \in \mathcal{B}$ such that

$$B_1 \cap \dots \cap B_m \in O.$$

EXAMPLE 9.2. Let (X, d) be a metric space. Then

$$\mathcal{O} = \{O \subset X : O \text{ open}\}$$

defines a topology which is Hausdorff, i.e. for $x \neq y$ we find O_x and O_y in \mathcal{O} such that

$$\emptyset = O_x \cap O_y.$$

EXAMPLE 9.3. Let $X = \mathbb{R} \cup \{\infty\}$. Then

$$\mathcal{O} = \{O : O \subset \mathbb{R} \text{ } O \text{ open}\} \cup \{K^c \cup \{\infty\} : K \text{ } K \text{ compact}\}$$

defines a topology.

DEFINITION 9.4. A function $f : (X, \mathcal{O}_X) \rightarrow (Y, \mathcal{O}_Y)$ is continuous if

$$f^{-1}(O) \in \mathcal{O}_X$$

for all $O \in \mathcal{O}_Y$. A sequence (x_n) converges to x if for every open set $O \in \mathcal{O}$ containing x there exists n_0 such that

$$x_n \in O$$

holds for all $n > n_0$. The definition of compactness is the same as for metric spaces.

REMARK 9.5. There exists a Let $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. The map $\phi : S^1 \rightarrow \mathbb{R} \cup \{\infty\}$ defined by $\phi(e^{2\pi it}) = \frac{\sin(\pi t)}{\cos(\pi t)}$ yields a homeomorphism between S^1 and $\mathbb{R} \cup \infty$, i.e. ϕ is bijective and $\phi(O)$ open if and only if O is open.

EXAMPLE 9.6. Let $X = \{(x_n) : x_n \in \mathbb{R}\}$. Define

$$d((x_n), (y_n)) = \sum_n 2^{-n} \frac{|x_n - y_n|}{1 + |x_n - y_n|}.$$

Then $O \subset X$ is open if and only if for every $(x_n) \in O$ there exists n_0 and $\varepsilon > 0$ such that if

$$|y_n - x_n| < \varepsilon$$

holds for all $n \leq n_0$, then $(y_n) \in O$.

Let X be a set. On the space $F(X, \mathbb{R}) = \mathbb{R}^X$ of arbitrary functions, we define for $x \in X$ and $O \subset \mathbb{R}$

$$B_x^O = \{g : g(x) \in O\}.$$

Let \mathcal{O} be defined as the collection of sets such that $O \in \mathcal{O}$ if and only if for every $f \in O$ there exists x_1, \dots, x_m and O_1, \dots, O_m such that

$$f \in B_{x_1}^{O_1} \cap \dots \cap B_{x_m}^{O_m} \subset O.$$

PROPOSITION 9.7. *Let $f \in F(X, \mathbb{R})$ and (f_n) be a sequence of functions. Then $\lim_n f_n = f$ with respect to the topology above if and only if $\lim_n f_n(x) = f(x)$ holds for all $x \in X$.*

PROOF. Let $x \in X$. We define $\delta_x : F(X, \mathbb{R}) \rightarrow \mathbb{R}$ by $\delta_x(f) = f(x)$. By definition we see that δ_x is continuous. Thus $f = \lim_n f_n$ implies

$$f(x) = \delta_x(f) = \lim_n \delta_x(f_n) = \lim_n f_n(x).$$

Conversely, we assume that $\lim_n f_n(x) = f(x)$ for all $x \in \mathbb{R}$. Let O be an open set such that $f \in O$. By definition there are $x_1, \dots, x_m \in X$ and O_1, \dots, O_m subsets of \mathbb{R} such that

$$f \in B_{x_1}^{O_1} \cap \dots \cap B_{x_m}^{O_m} \subset O.$$

Thus we may choose n_0 such that

$$f_n(x_i) \in O_i$$

for all $i = 1, \dots, m$ and $n > n_0$. This implies

$$f_n \in O$$

for all $n > n_0$. ■

tych

THEOREM 9.8. *(Tychonov) Let $C \subset F(\mathbb{R})$ be a closed set. Then C is compact if and only if all the sets $\delta_x(C)$ are compact.*

We need a little preparation. We say that a collection (F_i) has the finite intersection property (fip) if

$$\emptyset \neq \bigcap_{S \subset I} F_i$$

holds for every finite subset $S \subset I$.

LEMMA 9.9. *Let $C \subset X$ be a set. Then C is compact if for every family (F_i) of closed sets such that $(C \cap F_i)$ has fip, then*

$$\emptyset \neq C \cap \bigcap_i F_i.$$

PROOF. Homework ■

Easy part of proof of Theorem ^{tych}9.8. Since δ_x is continuous, we see that for compact C we must have $\delta_x(C)$ compact. The converse is more involved. The shortest proof I know uses ultrafilters. We skip it. ■

DEFINITION 9.10. *A set $E \subset X$ of a topological space is called connected, if for all open sets O_1 and O_2*

$$E \subset O_1 \cup O_2 \quad \text{and} \quad E \cap O_1 \cap O_2 = \emptyset$$

implies $E \subset O_1$ or $E \subset O_2$.

LEMMA 9.11. *A subset E in \mathbb{R} is connected if and only if E is an interval.*

PROOF. I is called an interval if $a < c < b$ and $a, b \in I$ implies $c \in I$. If E is not an interval we find $a, b \in I$ and $a < c < b$ such that $c \notin I$. We define $O_1 = (-\infty, c)$ and $O_2 = (c, \infty)$. The condition for connectedness are then violated. Conversely, we assume that I is an interval. Let O_1 and O_2 be open sets such that $I \subset O_1 \cup O_2$ and $O_1 \cap I \neq \emptyset$ and $O_2 \cap I \neq \emptyset$. We may assume that $a < b$ and $a \in O_1$, $b \in O_2$. We define

$$c = \sup\{t : a \leq t \leq b, \forall_{a \leq s \leq t} s \in O_1\}.$$

If $c = b$, then $O_1 \cap O_2 \neq \emptyset$. If $c < b$, then $c \notin O_1$ because if it were we could use the fact that O_1 is open and find $\delta > 0$ such that $c + \delta$ satisfies the conditions and $c + \delta \leq c$. The same argument shows that $c \neq a$. Thus $c \in O_2$. Since O_2 is open we know that there exists a $\delta > 0$ such that $c > s > c - \delta$ implies $s \in O_2$. However, by the definition of the supremum we find $s \in O_1$ between $c - \delta$ and c . This shows that $O_1 \cap O_2$ is not empty. ■

The following lemma follows immediately from the definition.

LEMMA 9.12. *Let $f : X \rightarrow Y$ be a continuous function and $E \subset X$ a connected set. Then $f(E)$ is connected.*

COROLLARY 9.13 (Mean-value theorem for continuous functions). *Let $I \subset \mathbb{R}$ be an interval and $f : I \rightarrow \mathbb{R}$ be continuous. Then $f(I)$ is an interval.*

10. Picard-iteration

THEOREM 10.1. *Let $\emptyset \neq X$ be a complete metric space and $T : X \rightarrow X$ Lipschitz map with constant less than 1. Then T has a unique fixpoint.*

PROOF. Since $T : X \rightarrow X$ has Lipschitz map with constant $c < 1$ we know that

$$d(T(x), T(y)) \leq cd(x, y)$$

holds for all $x, y \in X$. We define inductively $T^0 = id$ and $T^{n+1} = T \circ T^n$. Let $x_0 \in X$. We consider the sequence $(x_n)_{n \geq 0}$ defined by $x_n = T^n(x_0)$. Note that

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq cd(x_n, x_{n-1}) \leq \cdots \leq c^n d(x_1, x_0).$$

Using the geometric series, we deduce

$$\begin{aligned} d(x_{n+m}, x_n) &\leq \sum_{k=1}^{m-1} d(x_{n+k}, x_{n+k-1}) \leq \sum_{k=1}^{m-1} c^{n+k-1} d(x_1, x_0) \\ &\leq c^n d(x_1, x_0) \sum_{k=0}^{\infty} c^k = c^n d(x_1, x_0) \frac{1}{1-c}. \end{aligned}$$

Thus (x_n) is Cauchy. Since X is complete $x = \lim_n x_n$ exists. Then we note that the continuity of $d : X \times X \rightarrow \mathbb{R}$ implies

$$d(T(x), x) = d(\lim_n T(x_n), \lim_n x_{n+1}) = \lim_n d(x_{n+1}, x_n) = 0.$$

We deduce $T(x) = x$. If x' is another fixpoint we we have

$$d(x, x') = d(T(x), T(x')) \leq cd(x, x')$$

Thus $(1-c)d(x, x') \leq 0$ implies $d(x, x') = 0$ and hence $x = x'$. ■

EXAMPLE 10.2. *The function $f(x) = 1 - x$ has a unique fixpoint. However, the iterates $x_n = f^n(x)$ only converge for $x = \frac{1}{2}$.*

THEOREM 10.3. *Let $\phi : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ be a continuous and $L > 0$ such that*

$$|\phi(s, t) - \phi(s, r)| \leq L|t - r|.$$

Let $y_0 \in \mathbb{R}$ such that $|\phi(s - x_0, y_0)| \leq Ce^{C|s|}$ for some constant $C > 0$. Then there exists a differentiable function $f : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$f'(x) = \phi(x, f(x)) \quad \text{and} \quad f(x_0) = y_0 .$$

PROOF. We will show the assertion for $x_0 = 0$. Let $\alpha > \max(L, C)$. We define X_α to be the set of continuous functions such that $\sup_{t \in \mathbb{R}} e^{-\alpha|t|} |f(t) - y_0| < \infty$. We define the distance function

$$d(f, g) = \sup_t e^{-\alpha|t|} |f(t) - g(t)| .$$

Note that if $f \in X_\alpha$ and $d(f, g) < \infty$, then $g \in X_\alpha$. It is easily checked that X_α is a complete metric space (see section 3 for a similar argument). We define

$$T(f)(t) = y_0 + \int_0^t \phi(s, f(s)) ds .$$

Since ϕ and f are continuous the function $g(s) = \phi(s, f(s))$ is also continuous. In particular, $T(f)$ is a continuous function. The key calculation here is the Lipschitz constant of T : Let f and g be continuous functions such that $d(f, g) < \infty$. Then we have

$$\begin{aligned} |T(f)(t) - T(g)(t)| &= \left| \int_0^t (\phi(s, f(s)) - \phi(s, g(s))) ds \right| \\ &\leq \int_0^t |\phi(s, f(s)) - \phi(s, g(s))| ds \\ &\leq L \int_0^t |f(s) - g(s)| ds \\ &\leq L \int_0^t e^{\alpha|s|} ds d(f, g) \\ &\leq \frac{L}{\alpha} (e^{\alpha|s|} - 1) d(f, g) \leq \frac{L}{\alpha} e^{\alpha|s|} d(f, g) . \end{aligned}$$

We define $c = L/\alpha < 1$ and deduce that

$$d(T(f), T(g)) \leq cd(f, g) .$$

Now we consider $f_0(t) = y_0$. Then, we have

$$d(T(f), f_0) \leq d(T(f), T(f_0)) + d(T(f_0), f_0) \leq cd(f, f_0) + d(T(f_0), f_0) .$$

We observe that $\alpha > C$ implies

$$d(T(f_0), f_0) \leq \sup_t e^{-\alpha|t|} \left| \int_0^t \phi(s, y_0) ds \right| \leq C \sup_t e^{-\alpha|t|} \frac{e^{C|t|} - 1}{C} \leq 1.$$

Therefore $T(X_\alpha) \subset X_\alpha$. By the contraction principle we find a unique $f \in X_\alpha$ such that $T(f) = f$. This means

$$f(t) = y_0 + \int_0^t \phi(s, f(s)) ds.$$

In particular, $f(0) = y_0$. The fundamental theorem of calculus implies $f'(t) = \phi(t, f(t))$. ■

REMARK 10.4. Let us discuss the condition $\phi(s, y_0) \leq Ce^{C|s|}$. We can restrict our attention to the interval $[-2, 2]$. Since ϕ is continuous we may find C such that $|\phi(s, y_0)| \leq C$ and are in business. The solution is differentiable in $(-1, 1)$. For the uniqueness in the space continuously differentiable functions extending to $[-2, 2]$ we may assume that g satisfies $g(x) = \phi(x, g(x))$. Then we determine \tilde{C} such that $|g(x)| \leq \tilde{C}$ and repeat the proof for some α large enough. Since then $g \in X_\alpha[-2, 2]$ it has to coincide with the solution provided by the Theorem. Finally, how do we show existence and uniqueness on \mathbb{R} ? We find a unique solution on $[-2, 2]$. Then we repeat the process on $[-1, 3]$ with the value $y_1 = f(1)$ found by the first solution. Our starting function f_0 is defined by $f_0(t) = f(t)$ for $-1 \leq t \leq 1$ and $f_0(t) = f(2)$ for $t \geq 2$. We find a unique solution on $[-1, 3]$ coinciding with the original on $[-1, 2]$. Then we continue and have a solution on $(-2, \infty)$. Since the solution are locally unique, they are unique as solution on \mathbb{R} . The example $\phi(s, t) = e^{s^2} 2s$ with solution $f(x) = y_0 + e^{x^2}$ shows that we can not always assume that the solution is in some X_α .

We consider $\phi(s, r) = r$ and $y_0 = 1$. Then the iterations are given by

$$\begin{aligned} f_0(t) &= 1, \\ f_1(t) &= 1 + \int_0^t ds = 1 + t \\ f_2(t) &= 1 + \int_0^t (1 + s) ds = 1 + t + \frac{t^2}{2} \end{aligned}$$

$$f_3(t) = 1 + \int_0^t \left(1 + s + \frac{s^2}{2}\right) ds = 1 + t + \frac{t^2}{2} + \frac{t^3}{3!}$$

We obtain the Taylor series

$$f_n(t) = \sum_{k=0}^n \frac{t^k}{k!}$$

By our theorem we know that for every $\alpha > 1$

$$\limsup_n \sup_{t \geq 0} e^{-\alpha t} |f_n(t) - e^t| = 0.$$

This is not true for $\alpha = 1$.

What happens if we want to solve

$$f'(t) = f(t)^2?$$

The function $\phi(r) = r^2$ is not Lipschitz on \mathbb{R} . We have to use a local variant.

THEOREM 10.5. *Let I and J be compact intervals with midpoints x_0 and y_0 . Let $\phi : I \times J \rightarrow \mathbb{R}$ be a continuous function such that*

$$|\phi(s, r) - \phi(s, t)| \leq L|r - t|$$

for some constant L . Then there exists a $h > 0$ such that

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

has a unique solution on $(x_0 - h, x_0 + h)$.

PROOF. We assume again $x_0 = 0$, $I = [-a, a]$ and $J = [y_0 - b, y_0 + b]$. We choose $\alpha > L$. Let $h > 0$ such that $[x_0 - h, x_0 + h] \subset I$. We define the subset C of $X_\alpha[x_0 - h, x_0 + h]$ as

$$C = \{f : f(I) \subset J\}.$$

Again T is defined as

$$T(f)(t) = y_0 + \int_0^t \phi(s, f(s)) ds.$$

Let $f \in C$. Then we have

$$\begin{aligned} |T(f)(t) - y_0| &= \left| \int_0^t \phi(s, f(s)) ds \right| \\ &\leq \left| \int_0^t (\phi(s, f(s)) - \phi(s, y_0)) ds \right| + \left| \int_0^t \phi(s, y_0) ds \right| \end{aligned}$$

$$\begin{aligned} &\leq \left| \int_0^t L|f(s) - y_0| ds \right| + |t| \sup_{s \in I} |\phi(s, y_0)| \\ &\leq h(Lb + \sup_{s \in I} |\phi(s, y_0)|). \end{aligned}$$

Hence for $h \leq \frac{b}{Lb + \sup_{s \in I} |\phi(s, y_0)|}$ we have $T(C) \subset C$. Since C is a closed subset of $X_\alpha[-h, h]$ we may apply the contraction principle. ■

For the example $f'(x) = f(x)^2$ with $f(0) = 1$ we find

$$f_0(t) = 1,$$

$$f_1(t) = 1 + \int_0^t ds = 1 + t$$

$$f_2(t) = 1 + \int_0^t (1 + s)^2 ds = 1 + \int_0^t (1 + 2s + s^2) ds = 1 + t + t^2 + \frac{t^3}{3}$$

$$\begin{aligned} f_3(t) &= 1 + \int_0^t (1 + s + s^2 + \frac{s^3}{3})^2 ds = \\ &= 1 + \int_0^t (1 + s^2 + s^4 + \frac{s^6}{9} + 2s + 2s^2 + 2\frac{s^3}{3} + \dots) ds \\ &= 1 + t + t^2 + t^3 + \dots \end{aligned}$$

The solution is $f(t) = \sum_{k=0}^{\infty} t^k = \frac{1}{1-t}$. Of course this can not be extended to \mathbb{R} and uniform convergence only works on compact subsets of $(-1, 1)$.

11. Linear differential equations

In the vector-valued situation, Picard's iteration works equally well.

PROPOSITION 11.1. *Let I be an interval and X be a Banach space. Let $x_0 \in I$ and $y_0 \in X$ and $\delta > 0$. Let $\phi : I \times \bar{B}(y_0, \delta) \rightarrow X$ be a continuous function such that*

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

Then there exists an h such that

$$f'(t) = \phi(t, f(t)) \quad f(x_0) = y_0$$

has a unique solution among continuous functions $f : (x_0 - h, x_0 + h) \rightarrow B(y_0, \delta)$.

PROOF. Define

$$T(f)(t) = y_0 + \int_0^t \phi(s, f(s)) ds$$

Let $\alpha > L$ and $h > 0$. We consider the Banach space

$$Y = \{f : [x_0 - h, x_0 + h] \rightarrow X : f \text{ continuous}\}$$

with the norm

$$\|f\| = \sup_{|s| \leq h} e^{-\alpha|s|} \|f(x_0 + s)\|.$$

We consider the closed subset C defined by

$$C = \{f : f([x_0 - h, x_0 + h]) \subset \bar{B}(y_0, \delta)\}.$$

With the appropriate choice of h we find $T(C) \subset C$. The Banach contraction principle yields a fixpoint $T(f) = f$ which satisfies

$$f'(t) = \phi(t, f(t)).$$

For any solution $g'(t) = \phi(t, g(t))$ with $g(x_0) = y_0$ we have

$$y_0 + \int_{x_0}^t \phi(s, g(s)) ds = y_0 + \int_{x_0}^t g'(s) ds = g(t).$$

Thus we have uniqueness for continuous g with values in $\bar{B}(y_0, \delta)$. ■

Let us recall the usual trick for transforming DE with linear coefficients into systems: We are given the differential equation

$$y^n(x) = \sum_{k=0}^{n-1} a_k y^k(x)$$

with conditions $y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$. Then we introduce the new variables

$$y_0 = y, y_1 = y', \dots, y_{n-1} = y^{(n-1)}.$$

This leads to the matrix valued equation

$$\vec{y}' = A(\vec{y}), \quad \vec{y}(x_0) = (y_0, \dots, y_{n-1})$$

where

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & \\ 0 & 0 & 1 & 0 & \dots \\ \vdots & & & & \\ 0 & 0 & \dots & 0 & 1 \\ a_0 & a_1 & \dots & a_{n-2} & a_{n-1} \end{pmatrix}.$$

PROPOSITION 11.2. *Let X be a Banach space and $A : X \rightarrow X$ be linear map. The differential equation*

$$f'(t) = A(t) f(t), \quad f(0) = y_0$$

has the (unique) solution

$$f(t) = e^{tA} y_0.$$

PROOF. We consider the power series

$$g(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$$

with values in $L(X, X)$. By Lemma 7.10, we deduce

$$g'(t) = \sum_{k=1}^{\infty} \frac{t^{k-1}}{(k-1)!} A^k = Ag(t).$$

Therefore the chain rule implies that

$$f(t) = g(t)y_0 = e_{y_0}(g(t))$$

satisfies

$$f'(t) = e_{y_0}(g'(t)) = g'(t)y_0 = A(g(t)y_0) = A(f(t)).$$

The condition $f(0) = y_0$ is obvious. (Note that uniqueness follows from the Picard iteration method.) ■

LEMMA 11.3. Let A be a complex $n \times n$ matrix and U invertible such that $A = UBU^{-1}$. Then for every power series $f(t) = \sum_k a_k t^k$ converging on \mathbb{R} we have

$$f(A) = Uf(B)U^{-1}.$$

PROOF. Since power series are uniformly converging it suffice to show the assertion for polynomials. By linearity we only have to consider $f(t) = t^k$. However,

$$A^k = [UBU^{-1}]^k = \underbrace{UBU^{-1}UB \cdots U^{-1}UBU^{-1}}_{k\text{-terms}} = UB^kU^{-1}. \quad \blacksquare$$

LEMMA 11.4. Let B an $n \times n$ -Jordan block for the eigenvalue λ . Then

$$e^{tB} = e^{t\lambda} \left(\sum_{k=0}^n \frac{t^k}{k!} \sum_{j=1, \dots, n-1-k} e_{j, j+k+1} \right)$$

where e_{ij} are the matrix units.

PROOF. We may write $B = \lambda 1 + C$ where

$$C = \sum_{j=1}^{n-1} e_{j, j+1}$$

is a sum of matrices with only one entry. Note that $e_{lr}e_{st} = \delta_{rs}e_{lt}$. Then we get

$$\begin{aligned} C^m &= \left(\sum_{j=1}^{n-1} e_{j, j+1} \right)^m \\ &= \sum_{j_1, \dots, j_m=1}^{n-1} e_{j_1, j_1+1} \cdots e_{j_m, j_m+1} \\ &= \sum_{j=1, \dots, n-1, j+m \leq n-1}^m e_{j, j+1} e_{j+1, j+2} \cdots e_{j+m, j+m+1} \\ &= \sum_{j=1, \dots, n-1-m}^m e_{j, j+m+1}. \end{aligned}$$

In particular, $C^n = 0$. This yields

$$\begin{aligned} e^{tC} &= \sum_{k=0}^{\infty} \frac{t^k C^k}{k!} \\ &= \sum_{k=0}^n \frac{t^k}{k!} \sum_{j=1, \dots, n-1-k} e_{j, j+k+1}. \end{aligned}$$

Finally we note that that for commuting matrices $e^{A+B} = e^A e^B$ follows from the properties of the binomial coefficients (as for scalars). \blacksquare

An example will illustrate the procedure

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then, we have

$$C^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Since $C^3 = 0$, we get

$$e^{tC} = \begin{pmatrix} 0 & t & \frac{t^2}{2} \\ 0 & 0 & t \\ 0 & 0 & 0 \end{pmatrix}.$$

Thus the monomials $\frac{t^k}{k!}$ ‘move’ in the shifted diagonal.

CHAPTER 2

Topological vector spaces

1. Vector spaces

In the following F is a field.

DEFINITION 1.1. A vector space V over F is given by a commutative group $(V, +, 0)$ and a map $m : F \times V \rightarrow V$ such that

$$m(\lambda, x + y) = m(\lambda, x) + m(\lambda, y) .$$

DEFINITION 1.2. Let V be a vector space. A system $S \subset V$ is said to be linear independent, if for every finite family (λ_s) of scalars

$$\sum_{s \in S} \lambda_s s = 0$$

implies $\lambda_s = 0$ for all $s \in S$.

EXAMPLE 1.3. $C[0, 1]$ is a vector space. The polynomials $\{p_k : k \geq 0\}$ are independent ($p_k(t) = t^k$).

EXAMPLE 1.4. The space $F(\mathbb{R}) = \mathbb{R}^{\mathbb{R}}$ is a vector space over \mathbb{R} .

EXAMPLE 1.5. Let I be an index set. Then $F^I = \{f : I \rightarrow F : f \text{ function}\}$ is a vector space over F .

EXAMPLE 1.6. \mathbb{R} is a vector space over \mathbb{Q} .

DEFINITION 1.7. A subspace W of V is a subset $W \subset V$ such that $x, y \in W$ and $\lambda \in F$ implies

$$x + \lambda y \in W .$$

EXAMPLE 1.8. Let I be an index set. Consider $F(I) \subset F^I$ defined by

$$F(I) = \{f : I \rightarrow F : \exists_{s \subset I \text{ finite}} (i \notin S \Rightarrow f(i) = 0)\} .$$

$F(I)$ is called the free vector space over I .

PROPOSITION 1.9. Let $W \subset V$ be a subspace. Define $x \sim y$ by $x - y \in W$. Then \sim is an equivalence relation and V/\sim is a vector space. This space is denoted by V/W .

PROOF. Obviously, \sim is an equivalence relation. We define the operations

$$[x] + \lambda[y] = [x + \lambda y].$$

If $x - x' \in W$ and $y - y' \in W$, then $x - x' + \lambda(y - y') \in W$. Thus V/\sim is a vector space. ■

DEFINITION 1.10. Let V and W be vector spaces. The direct sum $V \oplus W$ is the vector space $V \times W$ with the operation

$$(x, y) + \lambda(x', y') = (x + \lambda x', y + \lambda y').$$

DEFINITION 1.11. 1) Let $S \subset V$. Then the span of S is defined as

$$\text{span}(S) = \left\{ \sum_{s \in S'} \lambda_s s : \lambda_s \in F, S' \subset S \text{ finite} \right\}.$$

2) $B \subset V$ is called a basis if B is linear independent and $\text{span}(B) = V$.

b0 THEOREM 1.12. Every vector space has a basis.

PROOF. Consider the collection L of linear independent subsets of V . We say that $S_1 \leq S_2$ if $S_1 \subset S_2$. Let $(S_i)_{i \in I}$ be a chain, i.e. for every i and $j \in I$ there exists $k \in I$ such that $S_i \subset S_k$ and $S_j \subset S_k$. We claim that $S = \bigcup_{i \in I} S_i$ is linear independent. For let $(\lambda_s)_{s \in S'}$ be finite family such that

$$\sum_s \lambda_s s = 0.$$

For every $s \in S'$ we find $i(s)$ such that $s \in S_{i(s)}$. Since S' is finite we may find $k \in I$ such that $S_{i(s)} \subset S_k$ for all $s \in S'$. Hence, $\{s : S'\} \subset S_k$. By linear independence, we deduce $\lambda_s = 0$ for all $s \in S'$. By Zorn's Lemma, we find a maximal element B in L . This means for no $x \in V$ $B \cup \{x\} \in L$. We claim that $V = \text{span}(B)$. Let us show that $x \notin \text{span}(B)$ implies that $B \cup \{x\}$ is linear independent. Indeed, let λ_x and $(\lambda_s)_{s \in S}$, $S \subset B$ finite such that

$$\lambda_x x + \sum_s \lambda_s s = 0.$$

If $\lambda_x = 0$, we deduce $0 = \sum_s \lambda_s s$ and hence $\lambda_s = 0$ for all $s \in S$. If $\lambda_x \neq 0$ we find

$$x = \sum_s -\lambda_s / \lambda_x s.$$

Since $x \notin \text{span}(B)$ we deduce $\lambda_x = 0$ and the assertion follows. ■

EXAMPLE 1.13. $\{1, \sqrt{2}\}$ is linear independent over \mathbb{Q} . Let B be a basis for \mathbb{R} over \mathbb{Q} . Then we find an injective map

$$\Phi : \mathbb{R} \rightarrow \bigcup_{S' \subset B \text{ finite}} \mathbb{Q}^{S'}$$

Since $\mathbb{Q}^{S'}$ is countable and for an infinite B set the collection of all finite subsets of B has the same cardinality as B , we deduce that B has the same cardinality as \mathbb{R} .

2. Linear transformations

DEFINITION 2.1. Let V and W vector spaces over F . A map $T : V \rightarrow W$ is called linear, if

$$T(x + \lambda y) = T(x) + \lambda T(y)$$

holds for all $x, y \in V$. A linear map T is called an isomorphism if T is bijective.

REMARK 2.2. 1) If $T(V)$ is a subspace of W .

2) If $T : V \rightarrow W$ is bijective, then T^{-1} is linear.

3) If $T : V \rightarrow W$ and $S : W \rightarrow Z$ are linear maps, then the composition $ST : V \rightarrow Z$ is linear.

EXAMPLE 2.3. Let $\phi : I \rightarrow J$ be a map. Then

$$T_\phi : F(J) \rightarrow F(I), T_\phi(f) = f \circ \phi$$

is a linear map. If ϕ is bijective, then T_ϕ is an isomorphism.

THEOREM 2.4. Every vector space is isomorphic to a free vector space.

PROOF. Let B be a basis. We define $T : F(B) \rightarrow V$ by

$$T(f) = \sum_{b \in B} f(b)b.$$

Note that $T(f)$ is well-defined because only finitely many coordinates are non-zero. It is easily checked that T is an isomorphism. ■

PROPOSITION 2.5. Let V be a vector space and W be a subspace. Then V isomorphic to $W \oplus V/W$.

PROOF. Let B be a basis of V/W . Let $f : B \rightarrow V$ such that $f(b) \in b$ holds for all $b \in B$. Let us show that $\{f(b) : b \in B\}$ is linearly independent. Indeed, if we assume

$$0 = \sum_b \lambda_b f(b)$$

then $0 = [0] = \sum_b \lambda_b [f(b)] = \sum_b \lambda_b b$. Hence $\alpha_b = 0$. Therefore, we may define $T_2 : V/W \rightarrow V$ by

$$T_2\left(\sum_b \lambda_b b\right) = \sum_b \lambda_b f(b).$$

Then we define $T : W \oplus V/W \rightarrow V$ by

$$T(w, x) = w + T_2(x).$$

Obviously, T is linear. Let us show that T is injective. This is to show $T(w, x) = 0$ implies $w = 0$ and $x = 0$. Let $x = \sum_b \lambda_b f(b)$. $0 = T(w, x)$ implies $T_2(x) \in W$. Hence

$$0 = \sum_b \lambda_b [f(b)] = \sum_b \lambda_b b = x.$$

Thus $x = 0$, $T_2(x) = 0$ and $w = 0$. Let $y \in V$. We write $[y] = \sum_b \lambda_b b$. Then we have

$$\left[y - \sum_b \lambda_b f(b)\right] = [y] - \sum_b \lambda_b [f(b)] = [y] - \sum_b \lambda_b b = 0.$$

Thus $y \sim \sum_b \lambda_b f(b)$. We get

$$T\left(y - \sum_b \lambda_b f(b), [y]\right) = y.$$

Hence T is bijective. ■

DEFINITION 2.6. Let $T : V \rightarrow W$. We define the kernel

$$\ker(T) = \{x \in V : T(x) = 0\}$$

and $\text{rg}(V) = T(V)$ the range.

PROPOSITION 2.7. Let $V \rightarrow W$ be a linear map and $q : V \rightarrow V/\ker(T)$ be the quotient map $q(x) = [x]$. There exists a unique injective linear map $\hat{T} : V/\ker(T) \rightarrow W$ such that $q\hat{T} = T$.

PROOF. We have to show that $\hat{T}([x]) = T(x)$ is well-defined. However $x - x' \in \ker(T)$ implies $T(x) - T(x') = T(x - x') = 0$. Now, we assume $\hat{T}([x_1]) = \hat{T}([x_2])$. Then $T(x_1) = T(x_2)$. This implies $T(x_1 - x_2) = 0$. In particular, $x_1 \sim x_2$. ■

DEFINITION 2.8. We define $\dim(V)$ to be the smallest cardinal number given by a basis. It is easy to show that

$$\dim(V \oplus W) = \dim(V) + \dim(W).$$

COROLLARY 2.9. *Let $T : V \rightarrow W$. Then*

$$\dim(\ker(T)) + \dim(\operatorname{rg}(T)) = \dim(V).$$

PROOF. Since $\hat{T} : V/\ker(T) \rightarrow \operatorname{rg}(T)$ is an isomorphism, we see that

$$V \cong \ker(T) \oplus V/\ker(T) \cong \ker(T) \oplus \operatorname{rg}(T).$$

The formula for the dimensions follows. ■

COROLLARY 2.10. *Let V be a finite dimensional vector space and $T : V \rightarrow V$. Then T is injective if and only if T is surjective.*

PROOF. T is injective iff $\dim(\ker(T)) = 0$ iff $\dim(\operatorname{rg}(T)) = n$. ■

COROLLARY 2.11. *Let W_1 and W_2 be finite dimensional subspaces of a vector space V . Then*

$$\dim(W_1) + \dim(W_2) = \dim(W_1 + W_2) + \dim(W_1 \cap W_2).$$

PROOF. W.l.o.g. we may assume $V = W_1 + W_2$. Let us first assume that $W_1 \cap W_2 = \{0\}$. Then the map $\iota : W_2 \rightarrow V/W_1$ is bijective and hence

$$\boxed{\text{k0}} \quad (2.1) \quad \dim(W_1 + W_2) = \dim(V/W_1) + \dim(W_1) = \dim(W_2) + \dim(W_1).$$

Now, we consider the general case. Since $W_1 \cap W_2$ is a subspace of V , W_1 and W_2 we have

$$\begin{aligned} \dim(W_1 + W_2) &= \dim(W_1 + W_2/W_1 \cap W_2) + \dim(W_1 \cap W_2), \\ \dim(W_1) &= \dim(W_1/W_1 \cap W_2) + \dim(W_1 \cap W_2), \\ \dim(W_2) &= \dim(W_2/W_1 \cap W_2) + \dim(W_1 \cap W_2). \end{aligned}$$

The assertion follows from

$$\dim(W_1 + W_2/W_1 \cap W_2) = \dim(W_1/W_1 \cap W_2) + \dim(W_2/W_1 \cap W_2)$$

which is particular of our preliminary observation because $W_1/W_1 \cap W_2 \cap W_2/W_1 \cap W_2 = \{0\}$. ■

DEFINITION 2.12. *Let B be a basis for V and $S \subset V$. Then we find scalars $\lambda_{b,s}$ such that*

$$s = \sum_{b \in B} \lambda_{s,b} b.$$

The matrix $m_{S,B} = [\lambda_{s,b}]_{s \in S, b \in B}$ is called transition matrix.

LEMMA 2.13. *Let B and S be a basis for V with transition matrices $m_{B,S}$ and $m_{S,B}$. Then*

$$id = m_{S,B}m_{B,S} \quad \text{and} \quad id = m_{B,S}m_{S,B}.$$

PROOF. Let $b \in B$. Then we have

$$\begin{aligned} b &= \sum_{s \in S} \mu_{b,s} s = \sum_{s \in S} \mu_{b,s} \sum_{b' \in B} \lambda_{s,b'} b' \\ &= \sum_{b' \neq b} \left(\sum_s \mu_{b',s} \lambda_{s,b} \right) b' + \left(\sum_s \mu_{b,s} \lambda_{s,b} \right) b \end{aligned}$$

By linear independence, we get

$$\sum_s \mu_{b',s} \lambda_{s,b} = \delta_{b',b}.$$

This shows $m_{S,B}m_{B,S} = id$. Starting with elements in s yields the first assertion. ■

Let $T : V \rightarrow W$ be a linear map and C a basis for V and B be a basis for W . We may write

$$T(c) = \sum_{b \in B} \lambda_{c,b} b.$$

Then $m_{C,B}^T = [\lambda_{b,c}]$ is the matrix associated to T (with respect to C and B).

LEMMA 2.14. *Let B and S be a matrix for V and C, D be a basis for W . Then*

$$m_{D,S}^T = m_{D,C} m_{C,B}^T m_{B,S}.$$

PROOF. Assume $D = C$ first. Then

$$T(c) = \sum_{b \in B} \lambda_{c,b} b = \sum_{b \in B} \lambda_{c,b} \sum_{s \in S} \mu_{b,s} s.$$

This yields

$$m_{C,S}^T = m_{C,B}^T m_{B,S}.$$

The equation

$$m_{D,B}^T = m_{D,C} m_{C,B}^T$$

is proved similarly. ■

COROLLARY 2.15. *Let $T : V \rightarrow V$ a linear map. Let B and S be a basis. Then*

$$m_{S,S}^T = m_{S,B} m_{B,B}^T m_{B,S} = m_{B,S}^{-1} m_{B,B}^T m_{B,S}.$$

3. Determinant and adjacent matrix

We recall that on the space of $n \times n$ matrices over F the determinant is given by

$$\det(A) = \sum_{\pi \in S_n} (-1)^{I(\pi)} \prod_{i=1}^n a_{i\pi(i)}.$$

Here S_n is the space of all permutation of the set $\{1, \dots, n\}$ and

$$I(\pi) = \#\{i < j : \pi(i) > \pi(j)\}$$

is the number of inversions. It is easily checked that \det is multi-linear, antisymmetric and satisfies $\det(I) = 1$. Using the Gauss-elimination method one can then show that $\det(A) \neq 0$ if and only if A is invertible. We will use a different approach.

DEFINITION 3.1. *Given a matrix $A = (a_{ij})$ we denote by B_{ij} the matrix obtained by deleting the i and j -th column. We define*

$$b_{ij} = (-1)^{i+j} \det(X_{ji})$$

Then $B = \text{adj}(A)$ is called the adjacent matrix to A

LEMMA 3.2. $A \text{adj}(A) = \det(A)I$.

PROOF. Consider $C = A \text{adj}(A)$. Then we deduce from the well-known determinant expansion

$$c_{ii} = \sum_{j=1}^n a_{ij} (-1)^{i+j} \det(X_{ij}) = \det(A).$$

For $i \neq k$ we get

$$c_{ik} = \sum_{j=1}^n a_{ij} (-1)^{i+j} \det(X_{kj}) = \det(\tilde{A}_{ik})$$

Here \tilde{A} repeats the i -th column in the k -column. Thus $\det(\tilde{A}_{ik}) = 0$ by antisymmetry. ■

COROLLARY 3.3. *If $\det(A) \neq 0$, then $A^{-1} = \det(A)^{-1} \text{adj}(A)$.*

LEMMA 3.4. *Let F be \mathbb{R} or \mathbb{C} and A_j be $m \times m$ matrices such that*

$$\sum_{k=0}^n A_k \lambda^k = 0$$

for all $\lambda \in F$. Then $A_0 = A_1 = \dots = A_n = 0$.

PROOF. Proof by induction $n = 0$. By assumption $A_0 = 0$.

We assume the assertion is true n . Let $\sum_{k=0}^{n+1} A_k \lambda^k = 0$. Inserting $\lambda = 0$ we get $A_0 = 0$ and hence

$$\lambda \left(\sum_{k=0}^n A_{k+1} \lambda^k \right) = 0.$$

This for $\lambda \neq 0$ we must have

$$\sum_{k=0}^n A_{k+1} \lambda^k = 0$$

By continuity this also holds for $\lambda = 0$ and the induction hypothesis implies $A_1 = A_2 = \dots = A_{n+1} = 0$. ■

For a polynomial $p(\lambda) = \sum_{k=0}^n a_k \lambda^k$ and a matrix A we define

$$p(A) = \sum_{k=0}^n a_k A^k.$$

THEOREM 3.5. *Let A be an $m \times m$ matrix and*

$$p_A(\lambda) = \det(\lambda I - A)$$

the characteristic polynomial. Then

$$p_A(A) = 0.$$

PROOF. We know that

$$(\lambda I - A) \operatorname{adj}(\lambda I - A) = \det(\lambda I - A) I = p_A(\lambda) I.$$

We write

$$\operatorname{adj}(\lambda I - A) = \sum_{k=0}^{n-1} B_k \lambda^k$$

Let us write

$$p_A(\lambda) = \sum_{k=0}^n a_k \lambda^k.$$

Here $a_n = 1$. This gives

$$\begin{aligned} \sum_{k=0}^n a_k \lambda^k I &= p_A(\lambda) I = (\lambda I - A) \operatorname{adj}(\lambda I - A) \\ &= (\lambda I - A) \sum_{k=0}^{n-1} B_k \lambda^k \\ &= -AB_0 + \sum_{k=1}^n (B_{k-1} - AB_k) \lambda^k. \end{aligned}$$

Comparing coefficients we get $B_{n-1} = I$, $-AB_0 = a_0$ and

$$B_{k-1} - AB_k = a_k I$$

for $k = 1, \dots, n-2$. Multiplying with A^{k-1} we deduce

$$A^k B_{k-1} - A^{k+1} B_k = a_k A^k.$$

Therefore

$$\sum_{k=0}^n a_k A^k = A^n - AB_0 + \sum_{k=0}^{n-1} (A^{k-1} B_{k-1} - A^k B_k) = A^n - A^n = 0. \quad \blacksquare$$

DEFINITION 3.6. *Let us consider the ideal*

$$I_A = \{p \in \mathbb{C}[X] : p(A) = 0\}$$

of polynomials. Since the integral domain of polynomials admits a factorization algorithm there exists a polynomial m_A with minimal degree and leading coefficients 1 such that

$$I_A = \mathbb{C}[X]m_A.$$

In particular, the minimal polynomial divides the characteristic polynomial p_A .

4. Eigenvalues and eigenvectors

DEFINITION 4.1. *Let $T : V \rightarrow V$ be a linear map. A number $\lambda \in F$ is called an eigenvalue if there exists $v \neq 0$ such that*

$$T(v) = \lambda v$$

In this case v is called eigenvector. The space K_λ of eigenvectors is given by

$$E_\lambda = \ker(T - \lambda \text{id}).$$

EXAMPLE 4.2. On $F^{\mathbb{N}}$ we define

$$T(f)(n) = f(n+1).$$

Then $T(f) = \lambda f$ implies

$$\lambda f(n) = f(n+1)$$

Therefore every eigenvector for $T(f) = \lambda f$ is given by $f(n) = \lambda^n f(0)$ with $f(0) \neq 0$. Note that for $\lambda \neq 0$ the function $f_\lambda(n) = \lambda^n$ does not belong to the free vector space $F(\mathbb{N})$. For $\lambda = 0$ we must have $0 = f(2) = f(3) = \dots$. Hence on f given

by $f(1) = 1$ and 0 else is an eigenvector for T . We may also consider the subspace $\ell_2 \subset \mathbb{R}^{\mathbb{N}}$ given by

$$\ell_2 = \left\{ f : \mathbb{N} \rightarrow \mathbb{R} : \sum_n |f(n)|^2 < \infty \right\}.$$

Then $f_\lambda \in \ell_2$ if and only if $|\lambda| < 1$.

LEMMA 4.3. *Let $F \in \{\mathbb{C}, \mathbb{R}\}$, $A \in M_n(F)$ and λ an eigenvalue. Then $m_A(\lambda) = 0$.*

PROOF. Let $v \neq 0$ such that $A(v) = \lambda v$. Then $A^k v = \lambda^k v$ implies

$$0 = m(A)v = m(\lambda)v.$$

Thus $m(\lambda) = 0$. ■

eigen

PROPOSITION 4.4. *Let V be a complex finite dimensional vector space of positive dimension. Let $T : V \rightarrow V$ be a linear map. Then T has an eigenvalue.*

PROOF. After fixing a basis, we may associate with T a matrix A . Then $p_A(\lambda) = \det(\lambda I - A)$ is a polynomial with leading coefficient $\lambda^{\dim(V)}$. Thus $\dim(V) > 0$ implies with the fundamental theorem that there exists λ with $p_A(\lambda) = 0$. This implies $\det(\lambda I - A) = 0$ and hence there exists $0 \neq v \in \ker(\lambda I - A)$. Using the transition matrix, we deduce that T has an eigenvector. ■

REMARK 4.5. *The eigenvalues are exactly the roots of the characteristic polynomial. Indeed, λ is an eigenvalue iff $\ker(\lambda I - A) \neq 0$ iff $\det(\lambda I - A) = 0$.*

5. Jordan normal form

DEFINITION 5.1. *We say that A is similar to B if there exists an invertible map S such that $A = S^{-1}BS$.*

THEOREM 5.2. *Let $A \in M_n(\mathbb{C})$ a complex matrix with*

$$p_A(x) = (x - \lambda_1)^{r_1} \cdots (x - \lambda_p)^{r_p}$$

and

$$m_A(x) = (x - \lambda_1)^{s_1} \cdots (x - \lambda_p)^{s_p}.$$

Then A is similar to a block matrix B with blocks

$$B_i = \begin{pmatrix} \lambda_i & 1 & 0 & 0 \cdots \\ 0 & \lambda_i & 1 & 0 \cdots \\ \vdots & & & \\ 0 & \cdots & 0 & \lambda_i & 1 \\ 0 & \cdots & 0 & 0 & \lambda_i \end{pmatrix}$$

Here not more than s_i blocks occur.

DEFINITION 5.3. Let $T : V \rightarrow V$ be a linear map. We define

$$\rho(T) = \{\lambda \in \mathbb{C} : (\lambda I - T)^{-1} \text{ exists}\}$$

and the spectrum $\sigma(T) = \mathbb{C}\rho(T)$.

REMARK 5.4. $A \in M_n(\mathbb{C})$. The $\lambda \in \sigma(A)$ iff $p_A(\lambda) = 0$ if $m_A(\lambda) = 0$.

LEMMA 5.5. Let $A \in M_n(\mathbb{C})$. Let $\lambda \in \sigma(A)$. Then the sequence of subspaces
Consider

$$M^j = \ker(A - \lambda)^j$$

is ordered by inclusion. Moreover, there exists a minimal k such that $M^k = M^{k+1}$.

PROOF. Consider $d_j = \dim(M^j)$. Then d_j are integers and (d_j) is bounded by n . Thus $d = \lim_j d_j$ converges. Using $\varepsilon = \frac{1}{2}$ we see that for some $k \geq k_0$ we must have $d_j = d$. Hence, we define $k = \min\{j : d_j = d\}$. ■

DEFINITION 5.6. For two subspaces V and W of \mathbb{C}^n we write $V \oplus W = \mathbb{C}^n$ if $V + W = \mathbb{C}^n$ and $V \cap W = \{0\}$.

dec LEMMA 5.7. $M^k \oplus \text{rg}(A - \lambda)^k = \mathbb{C}^n$ and $\text{rg}(A - \lambda)^j = \text{rg}(A - \lambda)^k$ for all $j \geq k$.

PROOF. Note that

$$n = \dim(\ker(\lambda - A)^j) + \dim(\text{rg}(\lambda - A)^j).$$

Moreover, the sequence $W_j = \text{rg}((A - \lambda)^j)$ is decreasing. Thus $d_j = d_k$ for all $j \geq k$ implies $\dim(\text{rg}(\lambda - A)^j) = \dim(\text{rg}(\lambda - A)^k)$ for all $j \geq k$. Thus $W_j = W_k$.

Now, let $v \in \mathbb{C}^n$. Define $w = (A - \lambda)^k v$. Note that $(A - \lambda)(W_k) = (A - \lambda)^k(A - \lambda)(\mathbb{C}^n) \subset W_k$. Moreover, $\dim(A - \lambda)(W_k) = \dim W_{k+1} = \dim W_k$ implies that $\lambda I - A$ is injective on W_k . Hence $(A - \lambda)^k$ is injective and surjective when restricted to W_k . Therefore we find $v_0 \in W_k$ such that $(A - \lambda)^k(v_0) = w$. Equivalently $v_0 - v \in \ker((A - \lambda)^k)$. Hence

$$v = v - v_0 + v_0 \in M^k + W_k.$$

This implies

$$\mathbb{C}^n = M^k + W_k$$

Using the dimension formula we must have $\dim(M^k \cap W_k) = 0$. ■

In the following we use $\sigma(A) = \{\lambda_1, \dots, \lambda_p\}$

$$M_i = \bigcup_l \ker((A - \lambda)^l) \quad , \quad W_i = \bigcap_l \operatorname{rg}((A - \lambda)^l) .$$

We denote by k_i the smallest integer from Lemma [5.7](#).

inj LEMMA 5.8. $i \neq j$ implies $M_i \subset W_j$.

PROOF. Let us show that $(A - \lambda_j)$ leaves M_i invariant. Indeed, let $x \in M_i$. Then

$$(A - \lambda_j)(x) = (\lambda_j - \lambda_i)(x) + (A - \lambda_i)(x) \in M_i .$$

Recall that for $x \in \ker((A - \lambda)^k)$ we know that

$$(A - \lambda)^k(\lambda_i I - A)x = \lambda(\lambda_i I - A)(A - \lambda_i)^k(x) = 0 .$$

Thus we get $(\lambda_j - A)(M_i) \subset M_i$. Now, we want to show

$$\ker(\lambda_j - A) \cap M_i = \{0\} .$$

Indeed, let $x \in \ker((\lambda_j - A)) \cap M_i$. Then we have

$$\begin{aligned} (\lambda_j - \lambda_i)^k(x) &= ((A - \lambda_i) - (A - \lambda_j))^k(x) \\ &= (A - \lambda_i)^k(x) + \sum_{l=1}^{k-1} \binom{k}{l} (-1)^l (A - \lambda_i)^{k-l} (A - \lambda_j)^l(x) = 0 . \end{aligned}$$

Therefore $(A - \lambda_j)$ is an isomorphism when restricted to M_i . Thus for every $x \in M_i$ we may find $y \in M_i$ such that

$$x = (A - \lambda_j)^k(y) \in W_i .$$

The assertion is proved. ■

LEMMA 5.9. $\mathbb{C}^n = M_1 \oplus M_2 \oplus \dots \oplus M_p$.

PROOF. We have

$$\mathbb{C}^n = M_1 \oplus W_1$$

Every element $x \in W_1$ can be written in a unique way $x = z + y$, $z \in W_2$, $y \in M_2$. Since $y \in M_2$, we get $z \in W_2 \cap W_1$:

$$W_1 = W_1 \cap W_2 \oplus M_2 .$$

Hence

$$\mathbb{C}^n = M_1 \oplus M_2 \oplus W_1 \cap W_2 .$$

Since $M_3 \subset W_1 \cap W_2$, we continue and get

$$\mathbb{C}^n = M_1 \oplus \cdots \oplus M_p \oplus W_1 \cap \cdots \cap W_p.$$

Let $W = \bigcap_i W_i$. Note that

$$(A - \lambda_j)(W_i) \subset [(\lambda_j - \lambda_i)W_i + (A - \lambda_i)(W_i)] \subset W_i.$$

Therefore $(A - \lambda_j)$ maps W into W for all $j = 1, \dots, p$. In particular $A(W) \subset W$. Let us denote the induced linear map on W by T . If λ is an eigenvalue for T , then λ is an eigenvalue for A . Thus $\sigma(T) \subset \{\lambda_1, \dots, \lambda_p\}$. However, for every $\lambda = \lambda_i$ every eigenvector x is contained in $\ker((A - \lambda_i)^{k_i})$. This yields $x \in M_i \cap W \subset M_i \cap W = \{0\}$. Therefore we have shown that T has no eigenvalues. According to Proposition [4.4](#) ~~eigen~~ we must have $\dim(W) = 0$. ■

LEMMA 5.10. *The dimensions $k_i = \dim(M_i)$ coincide with degree s_i corresponding to λ_i in the minimal polynomial.*

PROOF. Let us recall that $M_i = \ker((A - \lambda_i)^{k_i})$ and hence

$$(A - \lambda_i)^{k_i}(M_i) = 0.$$

Since \mathbb{C}^n is a direct sum of the M_i 's and the $(A - \lambda_i)^{k_i}$'s commute we get

$$(A - \lambda_1)^{k_1} \cdots (A - \lambda_p)^{k_p}(\mathbb{C}^n) = \{0\}.$$

This implies that the minimal polynomial m_A divides $q(x) = \prod_{i=1}^p (\lambda_i - x)$. Thus $s_i \leq k_i$ for all $i = 1, \dots, p$. Suppose there exists i such that $s_i < k_i$. Then there exists an $0 \neq x \in M_i$ such that $(A - \lambda_i)^{s_i}(x) \neq 0$. We have shown in Lemma [5.8](#) ~~in j~~ that for every $j \neq i$ the map $(A - \lambda_j)$ maps M_i to M_i and is injective. Therefore $\prod_{j \neq i} (A - \lambda_j)^{s_j}$ is injective on M_i . We get

$$m_A(A)(x) = \left[\prod_{j \neq i} (A - \lambda_j)^{s_j} \right] (A - \lambda_i)^{s_i} x \neq 0.$$

Thus $m_A(A) \neq 0$ and this contradiction concludes the proof. ■

The next result concludes our proof of the existence of the Jordan normal form.

PROPOSITION 5.11. *Let $\lambda_i \in \sigma(A)$ and for $j = 1, \dots, k = s_i$ we define*

$$\Delta_j = \dim(\ker(A - \lambda_i)^j) - \dim(\ker(A - \lambda_i)^{j-1})$$

and $\partial_k = \Delta_k$ and

$$\partial_j = \Delta_j - \Delta_{j+1}.$$

Then the restriction of A to M_i is similar to a direct sum of ∂_j Jordan blocks of length j .

PROOF. In the following $\lambda = \lambda_i$ and $k = s_i$. We consider $M = \ker((A - \lambda)^k)$. We decompose

$$M = \ker((A - \lambda)^{k-1}) \oplus \ker((A - \lambda)^k) / \ker((A - \lambda)^{k-1})$$

Let $[b_1], \dots, [b_m]$ be a basis for the quotient space. Note that $m = \Delta_k = \partial_k$. Let us show that

$$S = \{(A - \lambda)^j(b_i) : 0 \leq j < k, 1 \leq i \leq m\}$$

is a system of linear independent vectors. Indeed, assume

$$\sum_{ij} a_{ij}(A - \lambda)^j(b_i) = 0$$

Since $(A - \lambda)^{k-1}(A - \lambda)^j = 0$ for $j \geq 1$, we get

$$\sum_{i=1}^m a_{i0}(A - \lambda)^{k-1}(b_i) = 0.$$

This implies $\sum_i a_{i0}b_i \in \ker((A - \lambda)^{k-1})$. By linear independence of the $[b_1], \dots, [b_m]$ we deduce $a_{i0} = 0$ for $i = 1, \dots, m$. Similarly, we assume

$$\sum_{i=1}^m \sum_{j=1}^{k-1} a_{ij}(A - \lambda)^j(b_i) = 0$$

and deduce $a_{11}, \dots, a_{m1} = 0$. Inductively, we find $a_{ij} = 0$. For a fixed $1 \leq i \leq m$ we consider

$$x_j = (A - \lambda)^j(b_i).$$

Note that

$$A(x_j) = (A - \lambda)(x_j) + \lambda x_j = x_{j+1} + \lambda x_j$$

for $j = 0, \dots, k-1$ but $(A - \lambda)(x_{k-1}) = (A - \lambda)^k(b_i) = 0$. Hence x_{k-1} is an eigenvector. We see that A leaves

$$F_i = \text{span}\{x_0, \dots, x_{k-1}\}$$

invariant. In this basis we finally find the Jordan block

$$A|_{F_i} \cong \begin{pmatrix} \lambda & 1 & 0 & 0 \cdots \\ 0 & \lambda & 1 & 0 \cdots \\ \vdots & & & \\ 0 & \cdots & 0 & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix}.$$

Using the reversed order $v_{(i-1)m+j} = x_{k-j}$ we obtain (finally) a Jordan block and a basis $(v_j)_{1 \leq j \leq mk}$ for m blocks of length k . Now, we proceed inductively and consider $\ker((A - \lambda)^{k-1}) / \ker((A - \lambda)^{k-1})$. Then vectors $c_1, \dots, c_m = [(A - \lambda)(b_i)]$ are already linearly independent. Thus we may complete this system with linearly independent vectors $[B_1], \dots, [B_l]$ where

$$l = \Delta_{k-1} - \Delta_k = \partial k.$$

The descendants of the B_i form l Jordan normal blocks of size $k - 1$. In general, we have

$$\partial_j = \Delta_j - (\partial_k + \cdots + \partial_{j+1})$$

many Jordan blocks of size j . For example, we have

$$\partial_k + \partial_{k-1} = \Delta_k + \Delta_{k-1} - \Delta_k = \Delta_{k-1}.$$

By induction, we get

$$\partial_k + \cdots + \partial_{j+1} = \Delta_{j+1}.$$

Our claim is proved. ■

REMARK 5.12. *The JNF is uniquely determined by the dimensions ∂_j . It is unique up to permutation of the blocks.*

COROLLARY 5.13. $\Delta_k \leq \Delta_j$ for $j = 1, \dots, k$ and

$$k_i = \min\{j : \dim(\ker(A - \lambda_i)^j) = \dim(\ker(A - \lambda_i)^{j+1})\}.$$

PROOF. In fact we have seen that for b_1, \dots, b_m such that $[b_i] = \ker((A - \lambda)^k) / \ker((A - \lambda)^{k-1})$ are linearly independent we have $(A - \lambda)^{k-j}(b_i)$ are linearly independent. This yields $\Delta_k \leq \Delta_k$. Since $\Delta_k > 1$ by definition, we deduce $d_{j+1} > d_j$ for all $j = 1, \dots, k - 1$. ■

Let us consider an example

$$A = \begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 0 & 2 & 0 & -1 & -1 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

We have $\chi_A(x) = \det(A - x)^2 = (2 - x)^5$. Thus the eigen value 2. Consider

$$(A - 2)^2 = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = 0.$$

The minimal polynomial is $m_A(x) = (2 - x)^2$. The kernel of $(A - 2)$ is the $\text{span}\{e_1, e_2 - e_3, e_4 - e_5\}$. Thus A is similar to

$$B = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix}.$$

In fact I constructed A as

$$A = S^{-1}BS$$

where

$$S = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

6. Topology and vector spaces

DEFINITION 6.1. A topological group is a group G with a topology τ such that $\cdot : G \times G \rightarrow G$, $\cdot(g, h) = gh$ and $I : G \rightarrow G$, $I(g) = g^{-1}$ is continuous.

EXAMPLE 6.2. (1) Let G be group and $d(g, h) = \begin{cases} 1 & \text{if } g \neq h \\ 0 & \text{else} \end{cases}$. This induces the discrete metric. In the induced topology every set is open and hence G is a topological group.

(2) Let us consider $\mathbb{R}^{\mathbb{R}}$ with the topology of pointwise convergence. Then $(\mathbb{R}^{\mathbb{R}}, +)$ is a commutative topological group.

DEFINITION 6.3. A topological vector space is a vector space over $K \in \{\mathbb{R}, \mathbb{C}\}$ with a topology on V such that $(V, +)$ is a topological group and $\cdot : K \times V \rightarrow V$ is continuous.

Why topology? For differentiation. We need even more.

DEFINITION 6.4. V be vector spaces with a metric. Let $\Omega \subset V$ be an open set and $f : \Omega \rightarrow \mathbb{R}$ be a map. f is called differentiable at $x_0 \in \Omega$ if there exists a continuous linear map $T : V \rightarrow \mathbb{R}$ such that

$$\lim_{d(x, x_0) \rightarrow 0} \frac{|f(x) - f(x_0) + T(x - x_0)|}{d(x, x_0)} = 0.$$

REMARK 6.5. Let $V = \mathbb{R}(\mathbb{N})$ equipped with the pointwise topology. Then a linear map $T : V \rightarrow \mathbb{R}$ is continuous if and only if $\sum_k |T(e_k)| < \infty$. Indeed, $C = \{(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n) : \varepsilon \in \{-1, 0, 1\}\} \cap V$ is compact. Thus

$$\sum_k |T(e_k)| = \sup_{x \in C} |T(x)|.$$

Differentiation is usually done in normed vector spaces.

DEFINITION 6.6. $(V, \|\cdot\|)$ is called a normed vector space if V is a vector space and $\|\cdot\| : V \rightarrow [0, \infty)$ satisfies.

- i) $\|x\| = 0 \Leftrightarrow x = 0$,
- ii) $\|\lambda x\| = |\lambda| \|x\|$,
- iii) $\|x + y\| \leq \|x\| + \|y\|$,

for all $x, y \in V$, $\lambda \in K$. The associated metric on $(V, \|\cdot\|)$ is defined by

$$d_{\|\cdot\|}(x, y) = \|x - y\|.$$

REMARK 6.7. For a normed vector space $(V, +)$ is a topological group.

LEMMA 6.8. A normed vector space $(V, \|\cdot\|)$ is complete if and only if every absolutely convergent series is convergent.

PROOF. Let us assume that V is complete and that

$$\sum_n \|x_n\| < \infty.$$

Using the Cauchy criterion in \mathbb{R} , we find for every $\varepsilon > 0$ a natural number n_0 such that for $m \geq n \geq n_0$

$$\sum_{k=n+1}^m \|x_k\| < \varepsilon.$$

This shows that $y_n = \sum_{k=1}^n x_k$ satisfies

$$\|y_m - y_n\| = \left\| \sum_{k=n+1}^m x_k \right\| \leq \sum_{k=n+1}^m \|x_k\| < \varepsilon.$$

Thus (y_n) is Cauchy. Since V is complete we find a limit $y = \lim_n y_n$. For the converse we assume that (y_n) is Cauchy. Passing to a subsequence (if necessary) we may assume $\|y_{n+1} - y_n\| = d(y_{n+1}, y_n) < 2^{-n}$. Then the series $x_0 = y_0$, $x_n = y_n - y_{n-1}$ satisfies

$$\sum_n \|x_n\| < \infty.$$

By assumption, the partial sums

$$z_n = x_0 + x_1 - x_0 + x_2 - x_1 + \cdots + x_n - x_{n-1} = x_n$$

converge. Thus (x_n) is convergent. ■

DEFINITION 6.9. V and W be normed vector spaces. Let $\Omega \subset V$ be an open set and $f : \Omega \rightarrow W$ be a map. f is called differentiable at $x_0 \in \Omega$ if there exists a linear map T such that

$$\lim_{\|v\| \rightarrow 0} \frac{\|f(x_0 + v) - f(x_0) + T(v)\|}{\|v\|} = 0.$$

REMARK 6.10. If f is differentiable at x_0 , then f is continuous at x_0 .

Funny, the derivative is a linear map! In order to understand what it means to be continuously differentiable we need a norm on $L(X, Y)$.

PROPOSITION 6.11. Let X be a normed space and Y be a Banach space. We define $L(X, Y)$ as the space of map $T : X \rightarrow Y$ which are linear, i.e.

$$T(x + \lambda y) = T(x) + \lambda T(y).$$

and continuous. The norm on $L(X, Y)$ is given by

$$\|T\|_{op} = \sup_{\|x\| \leq 1} \|T(x)\|.$$

Then $L(X, Y)$ is a Banach space.

PROOF. Let us first show that a linear map $T : X \rightarrow Y$ is continuous iff $\|T\| < \infty$. Indeed, if $\|T\|$ is finite, then

$$\|T(x) - T(y)\| = \|T(x - y)\| \leq \|T\|_{op} \|x - y\|$$

holds for all $x, y \in V$. Thus T is Lipschitz and thus continuous. For the converse, we assume that T is continuous. Then $T^{-1}(B(0, 1))$ is open and henceforth contains $B(0, \varepsilon)$ for some $\varepsilon > 0$. Now let $\|x\| \leq 1$ and $0 < \delta < \varepsilon$. Then $\|(\varepsilon - \delta)x\| < \varepsilon$ and hence

$$\|T(x)\| = (\varepsilon - \delta)^{-1} \|T(\varepsilon - \delta)x\| < (\varepsilon - \delta)^{-1}.$$

This shows that $\|T\|_{op} \leq (\varepsilon - \delta)^{-1}$ for every $\delta > 0$ and thus $\|T\|_{op} \leq \varepsilon^{-1}$. Now, we observe that $\|\cdot\|_{op}$ is a norm. We only check the triangle inequality. Indeed,

$$\begin{aligned} \|T + S\|_{op} &= \sup_{\|x\| \leq 1} \|(T + S)(x)\| = \sup_{\|x\| \leq 1} \|T(x) + S(x)\| \leq \sup_{\|x\| \leq 1} \|T(x)\| + \|S(x)\| \\ &\leq \|T\|_{op} + \|S\|_{op}. \end{aligned}$$

Finally we have to show that $L(X, Y)$ is complete. Let (T_n) be a Cauchy sequence of linear maps. For fixed $x \in X$, we have

$$\|T_n(x) - T_m(x)\| \leq \|T_n - T_m\| \|x\|.$$

Thus $(T_n(x))$ is Cauchy and we may define

$$T(x) = \lim_n T_n(x).$$

Then we have

$$T(x + \lambda y) = \lim_n T_n(x + \lambda y) = \lim_n T_n(x) + \lambda T_n(y) = T(x) + \lambda T(y).$$

Thus T is linear. Let us show that

$$\boxed{\text{op}} \quad (6.1) \quad \lim_n \|T - T_n\|_{op} = 0.$$

Indeed, let $x \in X$ with $\|x\| \leq 1$. Then we have

$$\begin{aligned} \|T(x) - T_n(x)\| &= \left\| \lim_m T_m(x) - T_n(x) \right\| \leq \limsup_{m \geq n} \|T_m(x) - T_n(x)\| \\ &\leq \sup_{m \geq n} \|T_m - T_n\| \|x\| \leq \sup_{m \geq n} \|T_m - T_n\|. \end{aligned}$$

In particular $\|T\|_{op} \leq \|T - T_1\|_{op} + \|T_1\|_{op}$ is finite and T is continuous. Moreover, $\lim_n d(T, T_n) = 0$ implies that $\lim_n T_n = T$. ■

PROPOSITION 6.12. (*Chain rule*) $\Omega \subset V$ open, $\tilde{\Omega} \subset W$ open. $f : \Omega \rightarrow W$, $g : \tilde{\Omega} \rightarrow Z$, $x_0 \in \omega$, $y_0 = f(x_0) \in \tilde{\Omega}$. If f is differentiable at x_0 and g is differentiable at y_0 , then $g \circ f$ is differentiable and

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

PROOF. Let us introduce the error functions

$$\varepsilon_f(v) = f(x_0 + v) - f(x_0) - T(v)$$

and

$$\varepsilon_g(w) = g(y_0 + w) - g(y_0) - S(w).$$

Note that

$$\lim_{\|v\| \rightarrow 0} \frac{\|\varepsilon_f(v)\|}{\|v\|} = 0 = \lim_{\|w\| \rightarrow 0} \frac{\|\varepsilon_g(w)\|}{\|w\|}.$$

For $v \in V$ we introduce $w = f(x_0 + v) - f(x_0) = T(v) + \varepsilon_f(v)$. Then we have

$$\begin{aligned} g(f(x_0 + v)) - g(f(x_0)) - R(v) &= g(f(x_0) + w) - g(f(x_0)) - R(v) \\ &= S(w) + \varepsilon_g(w) - R(v) = S(\varepsilon_f(v)) + \varepsilon_g(w). \end{aligned}$$

Since S is continuous, we have

$$\lim_{\|v\| \rightarrow 0} \frac{\|S(\varepsilon_f(v))\|}{\|v\|} \leq \|S\| \frac{\|\varepsilon_f(v)\|}{\|v\|} = 0.$$

Moreover, for every $\varepsilon > 0$ there exists a $\gamma > 0$ such that $\|w\| < \gamma$ implies

$$\|\varepsilon_g(w)\| < \frac{\varepsilon}{1 + \|T\|} \|w\|.$$

Since f is continuous, there exists a $\delta > 0$ such that $\|v\| < \delta$ implies $\|f(x_0 + v) - f(x_0)\| < \gamma$. Moreover, by making δ smaller we can also assume

$$\|\varepsilon_f(v)\| \leq \|v\|.$$

Then, we get

$$\|w\| = \|f(x_0 + v) - f(x_0)\| = \|T(v) + \varepsilon(v)\| \leq \|T\|\|v\| + \|v\|.$$

Thus we get

$$\|\varepsilon_g(w)\| < \frac{\varepsilon}{1 + \|T\|} \|w\| \leq \varepsilon\|v\|.$$

for all $\|v\| < \delta$. ■

EXAMPLE 6.13. *Let u be continuously differentiable function in two variables. We are looking for the derivative of*

$$h(t) = \int_0^t u(t, s) ds.$$

Solution: We define $g(t, r) = \int_0^t u(r, s) ds$. The derivative is given by the gradient

$$\begin{aligned} T(x, y) &= \nabla g(t, r) \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \frac{\partial g}{\partial t}(t, r)x + \frac{\partial g}{\partial r}(t, r)y = u(r, t)x + \left(\int_0^t \frac{\partial u}{\partial r} u(r, s) ds \right) y \end{aligned}$$

We define $f(t) = (t, t)$. The derivative is $f'(t) = \begin{pmatrix} 1 & 1 \end{pmatrix}$. Hence, we get

$$h'(t) = F'(t, t)f'(t) = u(t, t) + \int_0^t \frac{\partial u}{\partial r} u(t, s) ds.$$

REMARK 6.14. *Let $T : V \rightarrow W$ be a linear map. Then $T'(x) = T$ for all $x \in V$.*

EXAMPLE 6.15. *Consider $\det : M_n(\mathbb{R}) \rightarrow \mathbb{R}$. Then*

$$(\det)'(A)(B) = \sum_{k=1}^n \det(A(e_1), \dots, B(e_k), \dots, A(e_n)).$$

Here $B(e_j)$ is the j -th column and for $k = 1, \dots, m$ we replace the k 'th column of B by the k -th column of A . Moreover, the derivative of the function

$$g(t) = \det(1 + tA)$$

is given by

$$g'(0) = \operatorname{tr}(A) = \sum_{i=1}^n a_{ii}.$$

Indeed, we note

$$\begin{aligned} \det(A + B) &= \sum_{\sigma} \varepsilon(\sigma) \prod_{i=1}^n (A + B)_{i, \sigma(i)} \\ &= \sum_{\sigma} \varepsilon(\sigma) \prod_{i=1}^n (a_{i, \sigma(i)} + b_{i, \sigma(i)}) \\ &= \det(A) + \sum_{\sigma} \varepsilon(\sigma) \sum_{k=1}^n b_{k, \sigma(k)} \prod_{i \neq k} a_{i, \sigma(i)} \\ &\quad + \text{higher monomials in the coefficient } b_{ij}. \end{aligned}$$

This yields the first assertion. For the second that if $a_{ij} = \delta_{ij}$ only the term $\sigma = id$ survives and we get

$$\det(1 + tA) = 1 + t \operatorname{tr}(A) + \text{higher monomials in } t.$$

7. Taylor formula

We will first consider the Taylor formula for functions with values in \mathbb{R} . Recall the scalar Taylor formula

THEOREM 7.1. (*Taylor formula with integral remainder*) Let $f : I \rightarrow \mathbb{R}$ a $(n + 1)$ -times continuously differentiable. Then

$$f(x_0 + t) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} t^k + R_n(x_0, t)$$

where

$$R_n(x_0, t) = \frac{1}{n!} \int_0^t f^{(n+1)}(x_0 + s)(t - s)^n ds.$$

LEMMA 7.2. Let $f : \Omega \rightarrow \mathbb{R}$ be n times differentiable in Ω . Let $v \in V$ such that $[x_0, x_0 + v] = \{x_0 + tv : 0 \leq t \leq v\} \subset \Omega$. Define $h(t) = x_0 + tv$. Then the scalar function $g(t) = f(x_0 + tv)$ satisfies

$$g^{(n)}(t) = f^{(n)}(x_0 + tv, \dots, x_0 + tv).$$

PROOF. $n = 1$: By the chain rule we know that

$$g'(t) = f'(x_0 + tv)(v).$$

Now, we consider $f' : \Omega \rightarrow L(V, \mathbb{R})$ and the function $F : \Omega \rightarrow \mathbb{R}$ defined by

$$F(x) = f'(x)(v)$$

We apply the chain rule again and get

$$g''(t) = \frac{d}{dt}F(x_0 + tv) = F'(x_0 + tv)(v).$$

In order to calculate this derivative we write $F(x) = e_v(f'(x))$, where $e_v : L(V, \mathbb{R}) \rightarrow \mathbb{R}$ is given by $e_v(T) = T(v)$. Since e_v is linear, we deduce from the chain rule

$$F'(x)(w) = e_v \circ f''(x) = f''(x)(v)(w).$$

The general case is proved by induction following the same arguments. ■

tay1

COROLLARY 7.3. Let $\Omega \subset V$ be an open set and $f : \Omega \rightarrow \mathbb{R}$ be $(n + 1)$ -times continuously differentiable. Let $x_0 \in \Omega$ and $v \in V$ such $[x_0, x_0 + v] \subset \Omega$. Then

$$f(x_0 + v) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} \underbrace{(v, \dots, v)}_{k \text{ times}} + R_n(x_0, v)$$

where

$$R_n(x_0, v) = \frac{1}{n!} \int_0^1 f^{(n+1)}(x_0 + sv, \dots, x_0 + sv)(t - s)^n ds.$$

Now, we want to consider function $f : I \rightarrow X$ where X is a Banach space.

fund

LEMMA 7.4. Let $f : [a, b] \rightarrow X$ be continuous function. For a partition $\pi = \{a = x_0, \dots, x_n = b\}$ and ξ_1, \dots, ξ_n with $\xi_i \in [x_{i-1}, x_i]$ the Riemann sum is given by

$$S(\pi, \xi) = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

Then

$$\int_a^b f(s) ds = \lim_{\text{mesh}(\pi) \rightarrow 0} S(\pi, \xi)$$

exists. Moreover,

$$F(t) = \int_a^t f(s) ds$$

satisfies $F'(t) = f(t)$.

PROOF. Since f is uniformly continuous, we can find for $\varepsilon > 0$ a $\delta > 0$ such that $|t - s| < \delta$ implies

$$\|f(t) - f(s)\| < \varepsilon.$$

We consider two partitions such that π_1 and π_2 with $\text{mesh}(\pi_1) < \delta$ and $\text{mesh}(\pi_2) < \delta$. Let $\pi = \pi_1 \cup \pi_2$. Let us assume that π_1 has $n + 1$ points and $\xi = (\xi_1, \dots, \xi_n)$ is an intermediate vector (i.e. such that $\xi_i \in (x_{i-1}^1, x_i^1)$). We assume that π has $m + 1$ points η is an intermediate vector for π . Let us agree to use the ξ_i 's for every interval of π which is contained in $[x_{i-1}, x_i]$. Then, we get

$$\begin{aligned} \|S(\pi_1, \xi) - S(\pi, \eta)\| &= \left\| \sum_{i=1}^n [f(x_i)(x_i - x_{i-1}) - \sum_{[y_{j-1}, y_j] \subset [x_{i-1}, x_i]} f(\eta_j)(y_j - y_{j-1})] \right\| \\ &\leq \sum_{i=1}^n \sum_{[y_{j-1}, y_j] \subset [x_{i-1}, x_i]} \|f(\xi_i) - f(\eta_j)\| (y_j - y_{j-1}) \leq \varepsilon(b - a). \end{aligned}$$

Thus we get

$$\|S(\pi_1, \xi) - S(\pi_2, \tilde{\xi})\| \leq 2\varepsilon(b - a).$$

This proves the first assertion. Let us also note the immediate consequences

$$\left\| \int_a^b f(s) ds \right\| \leq \int_a^b \|f(s)\| ds$$

and

$$\int_a^b (f(s) + g(s)) ds = \int_a^b f(s) ds + \int_a^b g(s) ds.$$

For the prove of the second assertion, we observe that

$$\begin{aligned} \left\| \int_a^{b+t} f(s) ds - \int_a^b f(s) ds - f(b)t \right\| &= \left\| \int_b^{b+t} (f(s) - f(b)) ds \right\| \\ &\leq \int_b^{b+t} \|f(s) - f(b)\| ds. \end{aligned}$$

Given $\varepsilon > 0$ we may find $\delta > 0$ such that $|s| < \delta$ implies $\|f(s) - f(b)\| < \varepsilon$. This yields

$$\|\varepsilon(t)\| = \left\| \int_a^{b+t} f(s) ds - \int_a^b f(s) ds - f(b)t \right\| \leq \varepsilon|t|.$$

The assertion follows. ■

LEMMA 7.5. Let $[a, b] \subset \bigcup_{x \in [a, b]} B(x, \delta_x)$ be an open cover for $[a, b]$. Then there exists a partition $\pi = \{a = t_0 < t_1 < \dots < t_m = b\}$ such that for every $i = 1, \dots, m$ we find $x_i \in [t_i, t_{i+1}]$ and $\max\{|x_i - t_i|, |t_{i+1} - x_i|\} < \delta_{x_i}$.

PROOF. Let us denote by S the set of all points $s \in [a, b]$ such that there are $t_0 = a < t_1 < \cdots < t_m$ with $x_i \in [t_{i-1}, t_i]$, $\max\{|x_i - t_i|, |t_{i+1} - x_i|\} < \delta_{x_i}$ and $x_{m+1} \in [t_m, y]$ such that

$$\max\{|x_{m+1} - t_m|, |y - x_{m+1}|\} < \delta_{x_{m+1}}.$$

Note that S is not empty because $a \in S$. Let $s = \sup S$. We claim that $s \in S$. Indeed, let us find $y \in S$ such that $y > s - \delta_s$. Then we find $t_0 = a < \cdots < t_m < y$ and $t_m \leq x_{m+1} \leq y$. We may define $t_{m+1} = y$ and $x_{m+2} = s$. Also, we must have $s = b$. Indeed, assume $s < b$. $s \in S$ implies that $s - x_{m+1} < \delta_{x_{m+1}}$. Let $\rho < \delta_{x_{m+1}} - (s - x_{m+1})$ such that also $s + \rho < b$. Then $s + \rho \in S$ yields a contradiction. Thus $s = b$ and the assertion is proved. ■

fund2 LEMMA 7.6. *Let $f : [a, b] \rightarrow X$ be continuously differentiable. Then*

$$\int_a^b f(s) ds = f(b) - f(a).$$

PROOF. By continuity it suffices to assume that f is differentiable on an open subset of $[a, b]$. Let $\varepsilon > 0$ and $\delta > 0$. For every $x \in [a, b]$ we may find $\delta_x < \delta$ such that

$$|t - x| < 2\delta_x \Rightarrow \|f(t) - f(x) - f'(x)(t - x)\| \leq \varepsilon|t - x|.$$

Then we have $[a, b] \subset B(x, \delta_x)$. By compactness, we may find a finite subset ξ_1, \dots, ξ_n such that

$$[a, b] \subset \bigcup_{i=1}^n B(\xi_i, \delta_{\xi_i}).$$

Inductively we rename the sequence such that $a \in B(\xi_1, \delta_{\xi_1})$ and define $x_1 = \xi_1 + \frac{3}{2}\delta_{\xi_1}$. Then we find $\xi_2 \neq \xi_1$ such that $x_1 \in B(\xi_2, \delta_{\xi_2})$. Continuing in this way we find a partition $x_0 = a, x_1, \dots, x_n$ and $\xi_i \in [x_{i-1}, x_i]$ such that $x_i, x_{i+1} \in B(\xi_i, \delta_{\xi_i})$. This yields

$$\begin{aligned} \|f(b) - f(a) - S(\pi, \xi)\| &= \left\| \sum_{i=1}^n f(x_{i+1}) - f(x_i) - f'(\xi_i)(x_{i+1} - x_i) \right\| \\ &\leq \sum_{i=1}^n \|f(x_{i+1}) - f(\xi_i) - f'(\xi_i)(x_{i+1} - \xi_i)\| \\ &\quad + \|f(\xi_i) - f(\xi_{i-1}) - f'(\xi_i)(\xi_i - x_{i-1})\| \\ &\leq \sum_{i=1}^n \varepsilon(|x_{i+1} - \xi_i| + |\xi_i - x_{i-1}|) \leq \varepsilon(b - a). \end{aligned}$$

Thus for δ small enough we find

$$\begin{aligned} \|f(b) - f(a) - \int_a^b f'(s)ds\| &\leq \|f(b) - f(a) - S(\pi, \xi)\| + \left\| \int_a^b f'(s)ds - S(\pi, \xi) \right\| \\ &< \varepsilon(b-a) + \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we deduce the assertion. ■

COROLLARY 7.7. *Let X be a Banach space and $f : (a, b) \rightarrow X$ be $(n+1)$ times continuously differentiable. Then*

$$f(x_0 + t) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} t^k + R_n(x_0, t)$$

where

$$R_n(x_0, t) = \frac{1}{n!} \int_0^t f^{(n+1)}(x_0 + s)(t-s)^n ds.$$

PROOF. We use induction on n . For $n = 1$ we deduce from Lemma [fund2 7.6](#) that

$$f(x_0 + t) - f(x_0) = \int_0^t f'(x_0 + s)ds.$$

For $n = 2$ we consider

$$F(t) = f'(x_0 + t)t.$$

Then we have

$$F'(t) = f''(x_0 + t)t + f'(x_0 + t).$$

This yields

$$\begin{aligned} f(x_0 + t) - f(x_0) &= \int_0^t f'(x_0 + s)ds = \int_0^t (F'(s) - f''(x_0 + s)s)ds \\ &= F(t) - F(0) - \int_0^t f''(x_0 + s)sds = f'(x_0 + t)t - \int_0^t f''(x_0 + s)sds \\ &= \int_0^t f''(x_0 + s)(t-s)ds. \end{aligned}$$

For general n it is best to use integration by parts (justified as above):

$$\begin{aligned} &\frac{1}{(k-1)!} \int_0^t f^{(k)}(x_0 + s)(t-s)^{k-1}ds \\ &\leq \left[-\frac{1}{k!} f^{(k)}(x_0 + s)(t-s)^k \right]_0^t + \frac{1}{k!} \int_0^t f^{(k+1)}(x_0 + s)(t-s)^k ds. \\ &= \frac{f^{(k)}(x_0)t^k}{k!} + \frac{1}{k!} \int_0^t f^{(k+1)}(x_0 + s)(t-s)^k ds \end{aligned}$$

Iterating yields assertion. ■

COROLLARY 7.8. *Let V be a normed space and X be a Banach space. Let $\Omega \subset V$ be open, $f : \Omega \rightarrow X$ be $(n + 1)$ -times (continuously) differentiable. Let $x_0 \in \Omega$ such that $B(x_0, \delta) \subset \Omega$. Then*

$$f(x_0 + v) = \sum_{k=0}^n \frac{1}{k!} f^{(k)}(x_0)(v, \dots, v) + R_n(x_0, v)$$

such that

$$\|R_n(x_0, v)\| \leq \frac{1}{(n + 1)!} \sup_{y \in B(x_0, \delta)} \|f^{(n+1)}(y)\| \|v\|^{n+1}$$

holds for all $\|v\| \leq \delta$.

PROOF. Since Ω is open we may $\delta > 0$ such that $B(x_0, \delta) \subset \Omega$. Let $\|v\| \leq \delta$ and consider the function $g(t) = f(x_0 + tv)$. Then g is $(n + 1)$ times continuously differentiable and we get

$$g(1) = \sum_{k=0}^n \frac{g^{(k)}(0)}{k!} + R_n(0, 1).$$

As in Corollary [7.3](#) ^{tay1} we find

$$g^k(t) = f^{(k)}(x_0 + tv)(v, \dots, v).$$

Therefore, we get

$$\begin{aligned} \|R_n(0, 1)\| &= \left\| \frac{1}{n!} \int_0^1 g^{(n+1)}(s)(1-s)^n ds \right\| \\ &= \left\| \frac{1}{n!} \int_0^1 f^{(n+1)}(x_0 + sv)(v, \dots, v)(1-s)^n ds \right\| \\ &\leq \frac{1}{n!} \sup_{y \in B(x_0, \delta)} \|f^{(n+1)}(y)\| \|v\|^{n+1} \int_0^1 (1-s)^n ds \\ &= \frac{1}{(n + 1)!} \sup_{y \in B(x_0, \delta)} \|f^{(n+1)}(y)\| \|v\|^{n+1}. \end{aligned}$$

This is exactly the estimate for the remainder claimed in the assertion. ■

A power series with values in a Banach space X is given by

$$f(t) = \sum_{k=0}^{\infty} x_k (t - t_0)^k$$

where $x_k \in X$. We will focus on $t_0 = 0$. The radius of convergence is given by

$$R = \left(\limsup_k \|x_k\|^{\frac{1}{k}} \right)^{-1}.$$

REMARK 7.9. Let $0 \leq r < u < R$. Then for $n \geq n_0$

$$\sum_{k \geq n} |t|^k \|x_k\| \leq \frac{\left(\frac{r}{u}\right)^n}{1 - r/u}.$$

Thus f is a convergent sum of continuous functions and hence continuous.

pow LEMMA 7.10. On $(-R, R)$ f is infinitely often differentiable and

$$f^{(n)}(t) = \sum_{k=n}^{\infty} \frac{k!}{n!} x_k (t - t_0)^{k-n}$$

PROOF. We assume $t_0 = 0$. Let $0 < r < R$ and $|t| < r$. Let $|s| < r$. Let $r < u$. Then we find n_0 such that $\|x_k\| \leq u^{-k}$. First we observe that for $k \geq 2$

$$\begin{aligned} |s^k - t^k - kt^{k-1}(s-t)| &= \left| \int_t^s k(v^{k-1} - t^{k-1})dv \right| \\ &= \left| \int_t^s k(k-1) \int_t^v w^{k-2}dw dv \right| \leq k(k-1)r^{k-2} \int_t^s \int_t^w dw dv \\ &= \frac{k(k-1)}{2} r^{k-2} |s-t|^2 \end{aligned}$$

Then we get for all $n \geq n_0$

$$\begin{aligned} &\left\| \sum_{k \geq n} s^k x_k - \sum_{k \geq n} t^k x_k - \sum_{k \geq n} t^{k-1} k x_k (s-t) \right\| \\ &= \left\| \sum_{k \geq n} (s^k - t^k - kt^{k-1}(s-t)) x_k \right\| \\ &\leq \sum_{k \geq n} |s^k - t^k - kt^{k-1}(s-t)| \|x_k\| \\ &\leq \sum_{k \geq n} |t^k - s^k - kt^{k-1}(t-s)| u^{-k} \\ &\leq \sum_{k \geq n} \frac{k(k-1)}{2} r^{k-2} |s-t|^2 u^{-k} \\ &\leq \frac{|s-t|^2}{2} \sum_{k \geq n_0} k(k-1) \left(\frac{r}{u}\right)^k. \end{aligned}$$

Note that the sum on the right hand side is convergent. This yields

$$\left(\sum_{k \geq n} t^k x_k \right)' = \sum_{k \geq n} t^{k-1} k x_k.$$

Differentiating the polynomial $p(t) = \sum_{k=0}^{n_0} t^k x_k$ is no threat and the assertion follows for $n = 1$. Induction yields the assertion in full generality. ■