TOPOLOGIE ET CALCUL DIFFÉRENTIEL. I. TOPOLOGIE

DMITRY CHELKAK, DMA ENS 2020

[Last corrected: October 19, 2021]

This is the first part of informal lecture notes for the L3 course *Topology and Differential Calculus* at the DMA ENS, Paris, given in the Fall 2020 term.

Contents

September 21, 2020

1. Metric and topological spaces: basics

Definition 1.1. *Let E be a set. A function $d : E \times E \to \mathbb{R}_+ = [0, +\infty)$ is called a* metric *(or a* distance function*) on E if*

(i) $d(x, y) = 0$ *if and only if* $x = y$*;*
(ii) $d(x, y) = d(y, x)$ *for all* $x, y \in E$*;*
(iii) $d(x, z) \le d(x, y) + d(y, z)$ *for all* $x, y, z \in E$ (triangle inequality).

The pair (E, d) (or just E for shortness, once d is fixed) is called a metric space.

Let us mention several examples of metric spaces:

(0) $E = \mathbb{R}$ or $\mathbb{R}^N$, the Euclidean distance: $d(x, y) := \big(\sum_{k=1}^n |y_k - x_k|^2\big)^{1/2}$. More generally, for each $p \ge 1$ the function $d_p(x, y) := \big(\sum_{k=1}^n |y_k - x_k|^p\big)^{1/p}$ is a distance. (For $p \ne 1, 2$, the triangle inequality (iii) is not fully trivial; it is called *Minkowski's inequality.*) Also, for $p = +\infty$ one can naturally extend this definition by setting $d_\infty(x, y) := \max_{k=1,\ldots,n} |y_k - x_k|$.

(1) $E = G$, a connected (finite or infinite) graph; $d(u, v)$ is the minimal number of edges required to go from u to v. (A common illustration was to think about points of G as cities and edges as flights linking them; then $d(u, v)$ is the minimal number of flights required to go from u to v. In the current circumstances, this is not a useful model.)

(2) Without airplanes: E – a certain landscape, $t(x, y)$ – time required to walk from x to y, thus $t(x, z) \le t(x, y) + t(y, z)$. Note that a priori t is not symmetric (i.e., (ii) fails): to go uphill is typically longer than downhill; one can fix this by declaring $d(x, y) := \max\{t(x, y), t(y, x)\}$ and the triangle inequality is still there.

(3) $E = 2^{\mathbb{R}^N}$, the set of all non-empty subsets of $\mathbb{R}^N$, the *Hausdorff distance:*

$$d_{\mathrm{H}}(X, Y) := \inf\{\varepsilon > 0 : Y \subset \textstyle\bigcup_{x \in X} B(x, \varepsilon) \text{ and } X \subset \bigcup_{y \in Y} B(y, \varepsilon)\},$$

where $B(x, \varepsilon) := \{y : d(x, y) < \varepsilon\}$ is the open ball in $\mathbb{R}^N$. This is symmetric by construction and (iii) is also easy to check. However, there is a problem with (i): one can have $d_{\mathrm{H}}(X, Y) = 0$ for $X \ne Y$: e.g., think about the open ball $X = B(x, r)$ and the closed ball $\overline{B}(x, r) := \{y : d(x, y) \le r\}$ of the same radius. To avoid this problem, one can, e.g., replace E by the set of all *compact* subsets of $\mathbb{R}^N$; see TD.

What can one do with/in metric spaces?

(a) to speak about *convergence* of (sequences of) points $x_n \in E$: by definition, $x_n \to x$ as $n \to \infty$ if and only if $d(x_n, x) \to 0$ as $n \to \infty$;
(b) to speak about *continuous mappings*: $f : E_1 \to E_2$ is continuous if and only if, the convergence of points $x_n \to x$ in E_1 as $n \to \infty$ implies the convergence of their images $f(x_n) \to f(x)$ in E_2 as $n \to \infty$;
(c) e.g., to generalize – <u>under</u> <u>certain</u> <u>assumptions</u> <u>on</u> <u>E</u> – standard theorems on continuous functions $f : [0, 1] \to \mathbb{R}$. E.g., continuous functions $f : E \to \mathbb{R}$ attain their extremal values provided that E is <u>compact</u>; they also attain all intermediate values provided that E is <u>connected</u>; etc.

One of the important ideas behind such a generalization is to consider continuous mappings defined on more complicated/interesting objects than numbers (e.g., functions on functions, which are traditionally called *functionals*) and to be able to apply general theorems (extremal value, fixed point, etc)

Certainly, all these notions should be properly defined and the 'standard' proofs re-developed/translated to the context of abstract metric spaces. This is one of the things that we will be doing below. Actually, it turns out that there exists even a (strictly) more general framework of *topological spaces* in which the notions like compactness, connectedness etc make perfect sense. In order to motivate their definition, we need more notation, which directly generalize the same notation for subsets of $\mathbb{R}$ or $\mathbb{R}^N$:

- Let (E,d) be a metric space, $x \in E$ and $r \geq 0$. Denote

  $$B(x,r) := \{y \in E : d(x,y) < r\}, \quad \overline{B}(x,r) := \{y \in E : d(x,y) \leq r\}.$$

  For shortness, sometimes we will also use the notation $B_r(x) = B(x,r)$ and $\overline{B}_r(x) = \overline{B}_r(x)$ for the same sets, which are called *open* and *closed balls* (of radius r with the center at x), respectively.
- A set $U \subset E$ is called *open* if for each $x \in U$ there exists $r = r_x > 0$ such that $B(x,r_x) \subset U$. (Note that in this case we have $U = \bigcup_{x \in U} B(x,r_x)$.)
- A set $F \subset E$ is called *closed* if for each $x \notin F$ there exists $r > 0$ such that $B(x,r) \cap F = \emptyset$. (Trivially, F is closed if and only if $E \setminus F$ is open).
- For a subset $X \subset E$, we define its *interior* $\overset{\circ}{X} := \bigcup_{U \subset X;\ U-\text{open}} U$;

  its *closure* $\overline{X} := \bigcap_{F \supset X;\ F-\text{closed}}$ and the *boundary* $\partial X := \overline{X} \setminus \overset{\circ}{X}$. Note that the interior of X is the maximal (under inclusion) open subset of X while the closure of X is the minimal closed set containing X.
- *Warning:* it can be that $\overline{B_r(x)} \neq \overline{B}_r(x)$ (only $\overline{B_r(x)} \subset \overline{B}_r(x)$ is OK).

It is easy to see that the notion of continuous mappings between metric spaces relies not upon the metric d but on the concept of open sets:

Lemma 1.2. *Let (E_1,d_1) and (E_2,d_2) be metric spaces. A mapping $f : E_1 \to E_2$ is continuous (as defined in (b) above: $x_n \to x$ in E_1 implies $f(x_n) \to f(x)$ in E_2) if and only if the following property holds:*

- *for each open set $U \subset E_2$ its pre-image $f^{-1}(U)$ is an open set in E_1.*

Passing to complements $f^{-1}(E_2 \setminus U) = E_1 \setminus f^{-1}(U))$ one can also reformulate the continuity of $f : E_1 \to E_2$ by requiring that

- $f^{-1}(F) \subset E_1$ is closed for each closed $F \subset E_2$.

Proof. "$\Downarrow$" Let f be continuous, $U \subset E_2$ open, and $x \in f^{-1}(U)$. Since U is open, we can find $r > 0$ such that $B(f(x),r) \subset U$. Assume that $B(x,2^{-n}) \not\subset f^{-1}(U)$ for all $n \in \mathbb{N}$, which means that there exists $x_n \notin f^{-1}(U)$ such that $d_1(x,x_n) < 2^{-n}$. By continuity, $d_2(f(x_n),f(x)) \to 0$ as $n \to \infty$. In particular, $f(x_n) \in B(f(x),r) \subset U$ for large enough n and hence $x_n \in f^{-1}(U)$, a contradiction.
"$\Uparrow$" Let $x_n \to x$ in E_1 as $n \to \infty$. For each $\varepsilon > 0$ the set $f^{-1}(B(f(x),\varepsilon)$ is open and thus there exists $\delta = \delta(x,\varepsilon) > 0$ such that $B(x,\delta) \subset f^{-1}(B(f(x),\varepsilon)$ or, equivalently, $f(B(x,\delta)) \subset B(f(x),\varepsilon)$. Since $d_1(x_n,x) \to 0$ as $n \to \infty$, we have $x_n \in B(x,\delta)$ for all $n \geq N_0$, and hence $d_2(f(x),f(x_n)) < \varepsilon$ for all $n \geq N_0$. $\square$

- Two metrics d', d'' defined on the same set E are called *equivalent* if for each $x \in E$ and $r > 0$ there exist $r', r'' > 0$ such that $B'(x, r') \subset B''(x, r)$ and $B''(x, r'') \subset B'(x, r)$. In other words, two metrics are equivalent if and only if they define the same collection of open subsets of E. In view of Lemma 1.2, the notion of continuous mappings $f : E_1 \to E_2$ remains unchanged if one replaces a metric in E_1 or in E_2 by an equivalent one.

Definition 1.3. *Let E be a set and $\mathcal{O} \subset 2^E$ be a collection of subsets of E satisfying the following assumptions:*

(i) *$\emptyset \in \mathcal{O}$ and $E \in \mathcal{O}$;*
(ii) *if $\{U_\alpha\}_{\alpha \in A} \subset \mathcal{O}$, then $\bigcup_{\alpha \in A} U_\alpha \in \mathcal{O}$ (note that A can be infinite);*
(iii) *if $U, V \in \mathcal{O}$, then $U \cap V \in \mathcal{O}$ (by induction, the same holds for all* finite *intersections of sets from $\mathcal{O}$).*

Then $\mathcal{O}$ is called a topology *on E; the sets $U \in \mathcal{O}$ are called* open *sets; and the pair $(E, \mathcal{O})$ (or simply E for shortness if $\mathcal{O}$ is fixed) is called a* topological space.

The interior $\overset{\circ}{X}$, closure $\overline{X}$ and the boundary ∂X of a set $X \subset E$ are defined as above; note that we relied only upon the notion of open sets there.

Definition 1.4. *A mapping $f : E_1 \to E_2$ between topological spaces is called* continuous *if for each open set $U \subset E_2$ its pre-image $f^{-1}(U)$ is an open set in E_1.*

- The terminology is consistent: let (E, d) be a metric space and $\mathcal{O}$ denote the collection of open sets in E as defined above. Then, the properties (i)–(iii) are straightforward (to prove (iii), note that if $x \in U \cap V$, then $B(x, r_1) \subset U$ and $B(x, r_2) \subset V$ imply $B(x, \min\{r_1, r_2\} \subset U \cap V)$.
- Vice versa, given a topology $\mathcal{O}$ on E one says that $\mathcal{O}$ is *metrizable* if there exists a metric d on E such that the collection of open sets constructed out of d coincides with $\mathcal{O}$.
- Not all topologies are metrizable! In particular, all metrizable topologies satisfy the following property (called *axiom T_2* or *Hausdorff space*): if $x \neq y$, then there exist $U_x, U_y \in \mathcal{O}$ such that $x \in U_x$, $y \in U_y$ and $U_x \cap U_y = \emptyset$. (Indeed, we can take $U_x = B(x, \frac{1}{2}d(x, y))$, $U_y = B(y, \frac{1}{2}d(x, y))$.) However, this property fails in the examples given below.

Non-metrizable topological spaces:

(4) *Co-finite topology* (a toy 'model' for the next example): E - infinite set; $U \in \mathcal{O}$ if and only if $E \smallsetminus U$ is finite or if $U = \emptyset$. Equivalently, a set $F \subset E$ is closed if and only if it is finite or $F = E$. One can easily see that the Hausdorff property (T_2 axiom) does not hold.

(5) *Zariski's topology:* $E = k^n$; k is a field ($\mathbb{C}$ or $\mathbb{R}$ or $\mathbb{F}_{p^n}$ or ...) and $n \in \mathbb{N}$. A set $F \subset k^n$ is called closed if there exists a (possible infinite, no restriction) family of polynomials in n variables $\{P_\alpha\}_{\alpha \in A} \subset k[x_1, \ldots, k_n]$ such that $F = \{x \in k^n : P_\alpha(x_1, \ldots, x_n) = 0 \text{ for all } \alpha \in A\}$. The property (ii) is trivial, the property (iii) follows by considering the family of products $\{P_\alpha Q_\beta\}_{\alpha \in A, \beta \in B}$:

$$\begin{aligned} & P_\alpha(x) Q_\beta(y) = 0 \text{ for all } \alpha \in A \text{ and } \beta \in B \\ & \Leftrightarrow\ (P_\alpha(x) = 0 \text{ for all } \alpha \in A) \text{ or } (Q_\beta(y) = 0 \text{ for all } \beta \in B). \end{aligned}$$

(In other words, if $E \smallsetminus U$ is defined by $\{P_\alpha\}_{\alpha \in A}$ and $E \smallsetminus V$ by $\{Q_\beta\}_{\beta \in B}$, then $E \smallsetminus (U \cap V)$ is defined by $\{P_\alpha Q_\beta\}_{\alpha \in A, \beta \in B}$.)

September 23, 2020

Reminder. We have the following (consistent!) definitions:

<u>Metric</u> <u>spaces</u> ($d : E \times E \to \mathbb{R}_+$):

- Convergent sequences in E:
  $x_n \to x$ iff $d_E(x_n, x) \to 0$.
- Continuous mappings $f : E_1 \to E_2$:
  $x_n \to x \;\Rightarrow\; f(x_n) \to f(x)$.

- $U \subset E$ is open if for each $x \in U$ there exists $r_x > 0$ s.t. $B(x, r_x) \subset U$; note that $U = \bigcup_{x \in U} B(x, r_x)$.

<u>Topological</u> <u>spaces</u> ($\mathcal{O} \subset 2^E$):

- Continuous mappings $f : E_1 \to E_2$:
  $U \in \mathcal{O}_{E_2} \;\Rightarrow\; f^{-1}(U) \in \mathcal{O}_{E_1}$.

It is easy to see that, if $f : E_1 \to E_2$ and $g : E_2 \to E_3$ are continuous, then their composition $f \circ g : E_1 \to E_3$ is also continuous.
(Indeed, $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U))$.

Recall also the definitions of the interior and the closure of a set $X \subset E$:

$$\begin{aligned} \overset{\circ}{X} &= \bigcup\nolimits_{U\,-\,\text{open: } U \subset X} U &&= \{x \in E : \exists U_x \in \mathcal{O} \text{ such that } x \in U_x \subset X\}; \\ \overline{X} &= \bigcap\nolimits_{F\,-\,\text{closed: } X \subset F} F &&= \{x \in E : \forall U_x \in \mathcal{O}\ (x \in U_x \Rightarrow U_x \cap X \neq \emptyset)\}. \end{aligned}$$

Moreover, in the metric setup it is enough to consider open balls $B(x, r)$ instead of generic neighborhoods $U_x \in \mathcal{O}$ in the second column.

---

2. Topological spaces: more notions

Definition 2.1. *A bijection $f : E_1 \to E_2$ between topological spaces is called a* homeomorphism *if both f and f^{-1} are continuous. Topological spaces E_1 and E_2 are called* homeomorphic *if there exists a homeomorphism $f : E_1 \to E_2$.*

Note that

- 'to be homeomorphic' is an equivalence relation on topological spaces;
- if a certain 'topological' (i.e., formulated only in terms of open sets and notions derived out of $\mathcal{O}$) property holds for E_1, it also holds for E_2;
- a topological space is metrizable if and only if it is homeomorphic to a metric space. Below we sometimes (a bit inaccurately) say 'E is metric space' instead of 'E is a metrizable topological space'.

Further, note that the intersection $\mathcal{O}' \cap \mathcal{O}''$ of two topologies on the same E is again a topology on E. Let us introduce more terminology.

Definition 2.2. *Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two topologies defined on the same E such that $\mathcal{O}_1 \subset \mathcal{O}_2$, then one says that $\mathcal{O}_2$ is* finer *than $\mathcal{O}_1$ and that $\mathcal{O}_1$ is* coarser *than $\mathcal{O}_2$. The extreme cases $\mathcal{O} = \{\emptyset, E\}$ and $\mathcal{O} = 2^E$ are called the* trivial *and the* discrete *topologies, respectively.*

(The intuition behind the names is that we 'distinguish' points in E from each other using open sets: the larger $\mathcal{O}$ is, the 'finer' structure of E we can speak about.)

Definition 2.3. *A subset $\mathcal{B} \subset \mathcal{O}$ is called a* base *of the topology $\mathcal{O}$ if for each $U \in \mathcal{O}$ there exists a subset $\{B_\alpha\}_{\alpha \in A} \subset \mathcal{B}$ such that $U = \bigcup_{\alpha \in A} U_\alpha$. Equivalently, for each $U \in \mathcal{O}$ and $x \in U$ there exists $B \in \mathcal{B}$ such that $x \in B \subset U$.*

An instructive example is the collection $\mathcal{B} := \{B(x,r)\}_{x\in E, r>0}$ in a metric space. Instead of describing the full collection $\mathcal{O}$ of open sets it is often easier to prescribe a base $\mathcal{B}$ of $\mathcal{O}$. However, *not* all collections $\mathcal{B} \subset 2^E$ can serve as a base of a topology.

Proposition 2.4. *(i) Let $\mathcal{B} \subset \mathcal{O}$ be a base of $\mathcal{O}$. Then, $\bigcup_{B\subset\mathcal{B}} B = E$ and*

$$\textit{for each } B', B'' \in \mathcal{B} \textit{ and } x \in B' \cap B'' \textit{ there exists } B \in \mathcal{B} \textit{ s.t. } x \in B \subset B' \cap B''.$$

(ii) Vice versa, assume that $\mathcal{B} \subset 2^E$ satisfies the two conditions given above. Then, $\mathcal{O} := \{\bigcup_{B\in\mathcal{A}} B\}_{\mathcal{A}\subset\mathcal{B}}$ is a topology on E. This topology is called the topology generated *(or* defined*) by $\mathcal{B}$ and is the minimal topology on E that contains $\mathcal{B}$.*

Proof. (i) Both properties directly follow from the definition of a base of a topology (and since $E \in \mathcal{O}$ and $B_1 \cap B_2 \in \mathcal{O}$).

(ii) It is clear that $\emptyset, E \in \mathcal{O}$ and that unions of sets from $\mathcal{O}$ also belong to $\mathcal{O}$. To prove that $U' \cap U'' \in \mathcal{O}$ for $U', U'' \in \mathcal{O}$, note that

$$\big(\bigcup_{B'\in\mathcal{A}'} B'\big) \cap \big(\bigcup_{B''\in\mathcal{A}''} B''\big) \ = \ \bigcup_{B'\in\mathcal{A}', B''\in\mathcal{A}''} (B' \cap B'').$$

For each $x \in U' \cap U''$ one can find $B'_x \in \mathcal{A}'$ and $B''_x \in \mathcal{A}''$ such that $x \in B'_x \cap B''_x$, and then use the assumption on $\mathcal{B}$ to find $B_x \in \mathcal{B}$ such that $x \in B_x \subset B'_x \cap B''_x$. Clearly, $B_x \subset U' \cap U''$ and hence $U' \cap U'' = \bigcup_{x\in U'\cap U''} B_x \in \mathcal{O}$ as required. $\square$

Definition 2.5. *$\mathcal{B}_x \subset \mathcal{O}$ is called a* local base *of $\mathcal{O}$ at $x \in E$ if $x \in B_x$ for all $B_x \in \mathcal{B}_x$ and for each $U \in \mathcal{O}$ s.t. $x \in U$ there exists $B_x \in \mathcal{B}_x$ s.t. $x \in B_x \in \mathcal{B}_x$.*

Clearly, if $\{\mathcal{B}_x\}_{x\in E}$ is a collection of local bases of $\mathcal{O}$, then $\mathcal{B} := \cup_{x\in E}\mathcal{B}_x$ is a base of $\mathcal{O}$. Also, if $B'_x, B''_x \in \mathcal{B}_x$ then there exists $B_x \in \mathcal{B}_x$ such that $B_x \subset B'_x \cap B''_x$. However, the latter property is *not* enough to guarantee that $\bigcup_{x\in E} \mathcal{B}_x$ is a base of a topology. (The actual condition to check is that for each $B'_{x'} \in \mathcal{B}_{x'}$, $B''_{x''} \in \mathcal{B}_{x''}$ and $x \in B'_{x'} \cap B''_{x''}$ there exists $B_x \in \mathcal{B}_x$ such that $B_x \subset B'_{x'} \cap B''_{x''}$.)

Definition 2.6. *A topological space $(E, \mathcal{O})$ is called*

- first-countable *if $\mathcal{O}$ has an at most countable local base $\mathcal{B}_x$ at each $x \in E$;*
- second-countable *if $\mathcal{O}$ has an at most countable base $\mathcal{O}$.*

It is easy to see that all metric spaces (and so all metrizable topologies, this is an example of a 'formal' inaccuracy mentioned above) are first-countable: indeed, one can take a collection of open balls $\mathcal{B}_x := \{B(x, 2^{-n})\}_{n\in\mathbb{N}}$ as a local base at x.

Definition 2.7. *A topological space E is called* separable *if there exist an at most countable subset $X \subset E$ s.t. $\overline{X} = E$ (such X are called* everywhere dense *in E).*

Warning: There is also a totally different notion '*separated* space' = Hausdorff = T_2 space; see above. A (personal) practical advice in order to avoid a confusion is not to use the latter name, saying Hausdorff (or T_2, depending whether you talk with a 'generic' mathematician or with a topologist) instead.

Proposition 2.8. *A metric space is second-countable if and only if it is separable.*

Proof. "$\Rightarrow$" Let $\mathcal{B} = \{B_k\}_{k\in\mathbb{N}}$ be a (at most) countable base of the topology on E. Choose a point $x_k \in B_k$ and let $X := \{x_k\}_{k\in\mathbb{N}}$. If we had $x \notin \overline{X}$, this would imply the existence of an open set U s.t. $x \in U \subset E \smallsetminus \overline{X}$ and $B_k \in \mathcal{B}$ such that $x \in B_k \subset U$, which is a contradiction since $B_k \cap X \ni x_k$.

"$\Leftarrow$" Let $X = \{x_k\}_{k\in\mathbb{N}}$ be such that $\overline{X} = E$ and consider $\mathcal{B} := \{B(x_k, 2^{-n})\}_{k,n\in\mathbb{N}}$. For each open set U and a point $x \in U$ there exists $m \in \mathbb{N}$ such that $B(x, 2^{-m}) \subset U$ and $k \in \mathbb{N}$ such that $d(x, x_k) \leq 2^{-(m+2)}$. Then, $x \in B(x_k, 2^{-(m+1)}) \subset U$. $\square$

Let us now reformulate several standard 'metric' notions (e.g., take $E = \mathbb{R}$ as a motivating example) in a 'topological' way.

Definition 2.9. *A set $X \subset E$ is called*

- everywhere dense *(in E) if $\overline{X} = E$;*
- nowhere dense *(in E) if $\operatorname{Int} \overline{X} = \emptyset$,*

where $\operatorname{Int}(Y) := \overset{\circ}{Y}$ is the alternative notation for the interior of a set $Y \subset E$.

Note that the latter condition is much stronger than to require $\operatorname{Int} X = \emptyset$: e.g., if $X := \mathbb{Q} \subset E := \mathbb{R}$, then $\operatorname{Int} X = \emptyset$ but $\mathbb{Q}$ is actually *everywhere* dense in $\mathbb{R}$. A good example of a nowhere dense set is the (standard or '$\frac{1}{3}$') Cantor set (see TD) $\mathrm{C} := [0,1] \smallsetminus \big((\frac{1}{3}, \frac{2}{3}) \cup (\frac{1}{9}, \frac{2}{9}) \cup (\frac{7}{9}, \frac{8}{9}) \cup \ldots\big)$ or similar sets (called Cantor-like nowadays; actually Cantor introduced C as a particular example in a larger family). Note that such sets can have (strictly) *positive* Lebesgue measure (='length') provided that the size of removed intervals decays fast enough. Equivalently,

- X is everywhere dense iff for each open non-empty set U one has $U \cap X \neq \emptyset$.
- X is nowhere dense iff each open non-empty set U contains an open non-empty subset V such that $V \cap X = \emptyset$.

(To prove the second statement: if X is nowhere dense and, for some $U \neq \emptyset$, such V does not exist, then $x \in \overline{X}$ for all $x \in U$ (i.e., $U \subset \overline{X}$) and hence $U \subset \operatorname{Int} \overline{X}$. Vice versa, if $\operatorname{Int} \overline{X} \neq \emptyset$, then there exists a non-empty open set U such that $U \subset \overline{X}$. At the same time, $V \subset E \smallsetminus X$ implies that $V \subset E \smallsetminus \overline{X}$, a contradiction.)

Definition 2.10. *Given $X \subset E$ and $x \in E$, one says that x is*

- *an* interior *point of X if $x \in \overset{\circ}{X}$;*
- *an* adherent *point of X if $x \in \overline{X}$. Further, such a point is*
  - *an* isolated *point of X if $x \in X$ and there exists a neighborhood $U_x \in \mathcal{O}$ of x such that $U_x \cap X = \{x\}$;*
  - *an* accumulation *point of X if $(U_x \cap X) \smallsetminus \{x\} \neq \emptyset$ for all $x \in U_x \in \mathcal{O}$.*

Note that the set $X' \subset \overline{X}$ of accumulation points of X is closed (as there exists an open set $U = \bigcup U_{x_\alpha}$ such that $U \cap X$ is the set of all isolated points x_α of X).

Détour. Actually, Cantor(1845–1918) pointed out the importance of a point-topology when studying (150 years ago, the paper is published in 1872) the following question on Fourier(1768–1830) series:

> assume that numbers a_n, b_n, $n \in \mathbb{N}$, satisfy the following condition: $\sum_{n \in \mathbb{N}} (a_n \cos(nx) + b_n \sin(nx)) = 0$ (the convergence is understood pointwise) for all x on $[0, 2\pi]$ except a certain ('small') set X. Does this assumption imply that all the coefficients a_n, b_n vanish?

The result of Cantor provides the affirmative answer under the following condition: there exists $k \in \mathbb{N}$ such that $X^{(k)} = \emptyset$, where $X^{(k+1)} := (X^{(k)})'$ and $X^{(0)} := X$. In a certain sense, this can be considered as the starting point of the development of the topology; it is worth noting that for 40-50 years not all mathematicians were fascinated by such considerations and that Cantor himself felt unhappy, in particular because of those controversies.

The word 'closed' can be viewed as 'closed under the operation of adding accumulation points to a given set' and the word 'open' appeared later [? 1910s, suggested by Carathéodory ?] as a conventional antonym.

---

3. Normed vector spaces: definitions, examples

We now discuss a situation when E is a vector space.

Definition 3.1. *A function* $\|\cdot\| : V \to \mathbb{R}_+$ *defined on a* vector space E *(over* $\mathbb{R}$ *or* $\mathbb{C}$*) is called a* norm *if it satisfies the following conditions:*

(i) $\|x\| = 0$ *if and only if* $x = 0$;
(ii) $\|\alpha x\| = |\alpha| \cdot \|x\|$ *for all scalars* α *and all* $x \in E$;
(iii) $\|x + y\| \leq \|x\| + \|y\|$ *for all* $x, y \in E$.

A vector space equipped with a norm is called a normed vector space*.*

It is easy to see that normed vector spaces can be viewed as metric (and, further, as topological) ones: the distance function can be introduced as $d(x, y) := \|x - y\|$. We already saw such an example when discussing the space $\mathbb{R}^N$ equipped with norms $\|\cdot\|_p$, where $p \in [1, +\infty]$.

Definition 3.2. *Let* k *be a field (or, more generally, an integral domain). A function* $|\cdot| : k \to \mathbb{R}_+$ *is called an absolute value if it satisfies the same conditions (i)–(iii) as above, where (ii) should be now read as*

(ii') $|xy| = |x| \cdot |y|$ *for all* $x, y \in k$.

Again, $d(x, y) := |y - x|$ can be taken as a distance function of k, which provides it a structure of a metric space. Given k, one can ask to describe the set of all possible absolute values defined on k. If $k = \mathbb{Q}$, *Ostrowski's theorem* (see also TD) claims that there are only two interesting options: if $|\cdot| : \mathbb{Q} \to \mathbb{R}_+$ is an absolute value, then it is

- either trivial, i.e., $|r| = 1$ for all $r \neq 0$;
- or *equivalent* to the Euclidean absolute value: there exists $\gamma \geq 1$ such that $|r| = |-r| = r^\gamma$ for all $r \in \mathbb{Q}_+$.
- or *equivalent* to the *p-adic* absolute value: there exist a prime integer p and $\gamma > 0$ such that $|r| = p^{-\gamma \max\{k \in \mathbb{Z} :\, r = p^k \cdot n_1/n_2, n_1 \in \mathbb{Z}, n_2 \in \mathbb{Z} \smallsetminus p\mathbb{Z}\}}$ for all $r \in \mathbb{Q}^*$.

Note that the parameter γ is irrelevant from the topological perspective. Also, the p-adic absolute value satisfies a stronger version of the triangle inequality (iii):

(iii') $|q + r| \leq \max\{|q|, |r|\}$.

Such absolute values (satisfying the stronger condition (iii')) are called *ultrametric* or *non-Archimedean*. If (iii') fails, then the absolute value is called *Archimedean*.

September 28, 2020

Examples of vector spaces:

(0) The spaces $\mathbb{R}^N$ (or $\mathbb{C}^N$, one can similarly replace $\mathbb{R}$ by $\mathbb{C}$ in the forthcoming examples). As already mentioned above,
  - $\|x\|_p := \big(\sum_{k=1}^N |x_k|^p\big)^{1/p}$, where $1 \leq p < +\infty$,
  - and $\|x\|_\infty := \max_{k=1,\dots,N} |x_k|$,

  are norms in this space (recall that the triangle inequality for $p \neq 1, 2, \infty$ is not fully trivial and is called the Minkowski inequality). All these norms are equivalent in the following sense: $\|x\|_\infty \leq \|x\|_p \leq N\|x\|_\infty$ for all p.

We can consider *sequences* $x = (x_n)_{k\in\mathbb{N}}$ instead of finite-dimensional vectors.

(1) Denote

$$\ell^p := \{x = (x_n)_{n\in\mathbb{N}} : \|x\|_p := \big(\textstyle\sum_{n\in\mathbb{N}} |x_n|^p\big)^{1/p} < +\infty\}$$

for $1 \le p \le \infty$; with a usual modification $\|x\|_\infty := \sup_{n\in\mathbb{N}} |x_n|$. This is a normed vector space (where we again rely upon the Minkowski inequality, now for sequences: e.g., it can be obtained from a finite-dimensional version by passing to the limit $N \to \infty$).

- It is worth noting that, contrary to the finite-dimensional case, the spaces ℓ^p are pairwise different *even as sets* ($\ell^p \subsetneq \ell^q$ if $p < q$), not speaking about the norms: the sequence $x_n = n^{-1/r}$ with $p < r < q$ belongs to ℓ^q but does not belong to ℓ^p.
- *Exercise:* if $p < +\infty$, then the space ℓ^p is separable.

  Proof: Consider a countable set

  $$X := \{(r_0, \ldots, r_{N-1}, 0, 0, \ldots),\ N \in \mathbb{N},\ r_k \in \mathbb{Q}\}$$

  of finite sequences composed of rational numbers. We claim that X is dense in ℓ^p. Indeed, for each $x \in \ell^p$ and each $\varepsilon > 0$ one can find $N \in \mathbb{N}$ such that $\sum_{n\ge N} |x_n|^p < \frac{1}{2}\varepsilon^p$ and, further, for each $n = 0, \ldots, N-1$, find $r_n \in \mathbb{Q}$ such that $|x_n - r_n|^p < \frac{1}{2N}\varepsilon^p$. We have constructed an element $r \in X$ such that $\|x - r\|_p < \varepsilon$. $\square$
- However, the space ℓ^∞ is not separable. To show that, for each subset $S \subset \mathbb{N}$ of indices, let the indicator sequence $y_S \in \ell^\infty$ be defined as $(y_S)_n = 1$ if $n \in S$ and $(y_S)_n = 0$ if $n \notin S$. Note that
  (a) $\|y_S - y_{S'}\|_\infty = 1$ provided that $S \ne S'$;
  (b) the set $Y := \{y_S,\ S \subset \mathbb{N}\}$ is uncountable (since the set $2^{\mathbb{N}}$ of all subsets of $\mathbb{N}$ is uncountable).

  If ℓ^∞ was separable and $X = \{x_k\}_{k\in\mathbb{N}} \subset \ell^\infty$ was a countable dense subset, then for each $S \subset \mathbb{N}$ there would exist $k = k(S) \in \mathbb{N}$ such that $\|y_S - x_k\| < \frac{1}{2}$. In such a situation, the mapping $S \mapsto k(S)$ would be injective due to (i), which leads to a contradiction with (ii).

Even more conceptual is to consider vector spaces of *functions*.

(2) For instance, let

$$E = C([0,1], \mathbb{R}) \ := \ \{f : [0,1] \to \mathbb{R} \text{ s.t. } f \text{ is continuous}\},$$

equipped with the *supremum* norm: $\|f\|_C := \|f\|_\infty = \max_{x\in[0,1]} |f(z)|$ (One can replace sup by max since we work with continuous functions on a segment.) This space (better to say, its generalization $C(K, E)$, where K is a compact and E is a complete normed vector space) is both very natural and important; we will often come back to its properties in what follows.

(2') The space of k times differentiable functions on $[0,1]$ is defined in a similar way:

$$C^k([0,1], \mathbb{R}) \ := \ \left\{ f : [0,1] \to \mathbb{R} \text{ s.t. } \begin{array}{l} f \text{ is } k \text{ times differentiable,} \\ f^{(k)} \text{ is continuous on } [0,1] \end{array} \right\},$$

equipped with the norm

$$\|f\|_{C^k} \ := \ \max_{x\in[0,1],\ m=0,\ldots,k} |f^{(m)}(x)|.$$

As in all other metric spaces, a local base of the topology at $f \in C^k$ is given by the sets

$$U_f^{(k,n)} := \{g : \max_{x\in[0,1]} |g^{(m)}(x) - f^{(m)}(x)| < 2^{-n} \text{ for all } m \le k\}, \qquad (3.1)$$

with $n \in \mathbb{N}$ (recall that k is fixed). As usual, instead of taking the maximum over all $m = 0, \ldots, k$, one could also take the sum (or, more generally, any other $\|\cdot\|_p$ norm in $\mathbb{R}^k$), this does not make any difference; cf. example (0).

(3) In principle, the space $C([0,1],\mathbb{R})$ could be viewed as something resembling the space ℓ^∞ of bounded sequences (actually, this is not a good analogy: to have a proper analogue of ℓ^∞ one should drop the continuity assumption; thus obtained space is called $L^\infty([0,1])$ and is much bigger than $C([0,1])$). It is also natural to introduce other norms, similar to those in ℓ^p:

$$\|f\|_p := \big(\textstyle\int_0^1 |f(x)|^p dx\big)^{1/p}, \quad 1 \le p < +\infty.$$

At least on the same *set* of continuous functions the integrals are well-defined and we obtain another normed vector spaces. Though we have not defined the relevant terminology yet, let us nevertheless mention that

- The normed vector space $(C([0,1],\mathbb{R}), \|\cdot\|_p)$ is not complete. To have a complete space one needs to enlarge the set $C([0,1],\mathbb{R})$ to *'all'* functions f on $[0,1]$ such that $\int_0^1 |f(x)|^p dx < +\infty$. This raises a deep question: for which functions on $[0,1]$ there is a reasonable definition of the integral? This is one of the subjects of the course *'Intégration et probabilités'*. Eventually, one defines an analogue of the space ℓ^p which is called $L^p([0,1])$. We will return to this discussion later.

Let us now discuss a somewhat similar example to (2) – the vector space C^∞ of *infinitely* differentiable functions. However, note that C^∞ is not a normed space.

(4) Denote

$$C^\infty([0,1],\mathbb{R}) \ := \ \big\{f : [0,1] \to \mathbb{R} \text{ s.t. } f \in C^k([0,1],\mathbb{R}) \text{ for all } k \in \mathbb{N}\big\},$$

equipped with a topology defined by the local bases $\{U_f^{(k,n)}\}_{k,n\in\mathbb{N}}$ (see (3.1)), note that k is *not* fixed now and runs over $\mathbb{N}$ similarly to n. Certainly, one should check that the collection $\bigcup_{f\in C^\infty([0,1],\mathbb{R})}\{U_f^{(k,n)}\}_{k,n\in\mathbb{N}}$ can serve as a (global) base of a certain topology (see Proposition 2.4). To this end, assume that $f \in U_{f_1}^{(k_1,n_1)} \cap U_{f_2}^{(k_2,n_2)}$ and let $k := \max\{k_1,k_2\}$. Further, let

$$2^{-n} < 2^{-n_1} - \|f - f_1\|_{C^{k_1}} \quad \text{and} \quad 2^{-n} < 2^{-n_2} - \|f - f_2\|_{C^{k_2}},$$

which implies that $U_f^{(n,k)} \subset U_{f_1}^{(n_1,k)} \subset U_{f_1}^{(n_1,k_1)}$ and similarly for $U_{f_2}^{(n_2,k_2)}$.

Note that the topology defined above, by construction, is first-countable. The following lemma describes the convergent sequences in this topology:

Lemma 3.3. *The convergence $f_m \to f$, $m \to \infty$, in the space $C^\infty([0,1],\mathbb{R})$ is equivalent to say that $f_m^{(k)} \to f^{(k)}$ in $C([0,1],\mathbb{R})$ for all $k \in \mathbb{N}$.*

Proof. '⇒' This is a triviality since, for each $k \in \mathbb{N}$, one can consider neighborhoods $U_f^{(n,k)}$ of f with *this* k, which are open both in C^k and C^∞.
'⇐' This is also a triviality: to prove that $f_m \to f$ in C^∞, it is enough to consider open neighborhoods U_f from the local base at f. By construction, each such a neighborhood is an open set in C^k for *some* $k \in \mathbb{N}$. □

It is easy to see that the construction of the topology given in (4) is pretty general. Let us mention one more example of the same kind (see also TD).

(4') Schwartz's space $\mathcal{S}(\mathbb{R})$ of rapidly decreasing smooth functions on $\mathbb{R}$:

$$\mathcal{S}(\mathbb{R}) \;:=\; \left\{ f:\mathbb{R}\to\mathbb{R} \text{ s.t. } \begin{array}{l} f\in C^k(\mathbb{R}) \text{ for all } k\in\mathbb{N} \text{ and} \\ \sup_{x\in\mathbb{R}}(|x|^s+1)|f^{(k)}(x)| < +\infty \text{ for all } s,k\in\mathbb{N} \end{array} \right\}.$$

(An example of such a function is $P(x)e^{-x^2}$, where $P(x)$ is a polynomial.) To define the topology one can, for instance, use local bases

$$V_f^{m,n} \;:=\; \left\{ g\in\mathcal{S}(\mathbb{R}) : \max_{(s,k):s+k\le m} \sup_{x\in\mathbb{R}}(|x|^s+1)|g^{(k)}(x)-f^{(k)}(x)| < 2^{-n} \right\}$$

(The proof of the fact that these local bases can be used to define a topology mimics the case of $C^\infty([0,1])$.) Similarly to Lemma 3.3, the convergence $f_m\to f$ in the topology of the space $\mathcal{S}(\mathbb{R})$ is equivalent to the convergence $\sup_{x\in\mathbb{R}}(|x|^s+1)|f_m^{(k)}-f^{(k)}(x)|\to 0$ as $m\to\infty$ for all $s,k\in\mathbb{N}$.

A similar construction can be done for all *countable* collections of norms $\|\cdot\|^{(n)}$ defined on the same vector space E. (More generally, it is enough to require that $\|\cdot\|^{(n)}:E\to\mathbb{R}_+$ is a *semi-norm* – i.e. satisfies conditions (ii) and (iii) but not necessarily (i) in Definition 3.1 – provided that $\|x\|^{(n)}=0$ for *all* n implies $x=0$.)

4. Sequential continuity and separation axioms

In examples (4) and (4') given above we described both the (base of the) topology and the convergent sequences in this topology. One can wonder whether to know the latter (i.e., how the convergent sequences look like) is enough to know the former (i.e., what are open/closed sets and what are continuous mappings).

Below we assume that E_1,E_2 are *topological* spaces and start with a few preliminary comments/observations:

- Recall that $x_n\to x$ iff for each open neighborhood U_x of x (i.e., each $U_x\in\mathcal{O}$ such that $x\in U_x$) there exists $N\in\mathbb{N}$ such that $x_n\in U_x$ for all $n\ge N$. This can be understood as the continuity property of the mapping $\mathbb{N}\cup\{*\}\to E$, where the set $\mathbb{N}\cup\{*\}$ can be viewed as a metric (and hence topological) space if we define $d(n,m):=|\frac{1}{n+1}-\frac{1}{m+1}|$ and $d(n,*):=\frac{1}{n+1}$ (i.e., if we 'represent' this set as a subset $\{\frac{1}{n+1},n\in\mathbb{N}\}\cup\{0\}$ of the *metric* space $\mathbb{R}$).
- In particular, if $f:E_1\to E_2$ is continuous, then the convergence $x_n\to x_*$ (in E_1) implies the convergence $f(x_n)\to f(x_*)$ (in E_2).
- However, requiring that $x_n\to x\Rightarrow f(x_n)\to f(x)$ does *not* imply that f is a continuous mapping.
  - A counterexample: $E=\mathbb{R}$ equipped with the co-countable topology (F is closed iff it is finite or countable or $F=E$). In this topology, $x_n\to x_*$ as $n\to\infty$ if and only if the sequence x_n stabilizes at x_* (i.e., if $x_n=x_*$ for all sufficiently large n). Therefore, *all* mappings $f:E\to E$ are sequentially continuous: if x_n stabilizes, so do $f(x_n)$. However, there exist non-continuous mappings: e.g., $x\mapsto\operatorname{sign}(x)$.
- **Warning:** In topological spaces, a limit of a sequence maybe *not* unique.
  - Example: in the co-finite topology the convergence $x_n\to x_*$ means the following: if $x_n=y$ for infinitely many indices n, then $y=x_*$. In particular, a sequence with no repetitions converges to *each* $x_*\in E$.

(*Proof:* Let $x_* \in U$, where U is open in the co-finite topology (i.e., equals to E without finitely many 'exceptional' points). The statement 'there exists N such that $x_n \in U$ for all $n \geq N$' is equivalent to say that these exceptional points appear only finitely many times in (x_n).)

- *Exercise*: give a similar description of convergent sequences in the Zariski topology.

Definition 4.1. *A mapping $f : E_1 \to E_2$ is called* sequentially continuous *if*

$$x_n \to x_* \text{ as } n \to \infty \text{ (in } E_1) \quad \Rightarrow \quad f(x_n) \to f(x_*) \text{ as } n \to \infty \text{ (in } E_2).$$

Recall that all continuous mappings are sequentially continuous but not vice versa.

Definition 4.2. *A topological space E_1 is called a* Fréchet–Urysohn *space if the following property holds: for each $X \subset E$ and each point $x \in \overline{X}$ there exists a sequence of points $x_n \in X$ such that $x_n \to x$ as $n \to \infty$.*

Proposition 4.3. *Assume that E_1 is a Fréchet–Urysohn space. Then, a mapping $f : E_1 \to E_2$ is continuous if (and only if) it is sequentially continuous.*

Proof. Given $F \subset E_2$ closed and $x \in \overline{f^{-1}(F)}$, use the Fréchet–Urysohn property to find a sequence $x_n \to x$ with $x_n \in f^{-1}(F)$ (which is equivalent to $f(x_n) \in F$). Due to the sequential continuity of f, we have $f(x_n) \to f(x)$ and hence $f(x) \in F$ since F is closed. Therefore, $x \in f^{-1}(F)$, which means that the set $f^{-1}(F)$ is closed. □

Lemma 4.4. *All first-countable (and hence all metrizable) topological spaces are Fréchet–Urysohn spaces.*

Proof. Let $x \in \overline{X}$ and $\{U_x^{(n)}\}_{n\in\mathbb{N}}$ be a *countable* local base at x. Further, denote $V_x^{(n)} := \bigcap_{k=0}^{n} U_x^{(n)}$, note that all these sets are open as they are finite intersections of open sets. Since $x \in \overline{X}$, for each $n \in \mathbb{N}$ we can find a point $x_n \in V_x^{(n)} \cap X$. It is easy to see that $x_n \to x$ as $n \to \infty$: indeed, each open neighborhood U_x of x contains a local base element $U_x^{(N)}$ and $x_n \in U_x^{(N)}$ for all $n \geq N$ by construction. □

- To summarize, in first-countable spaces (a) the notion of convergent sequenced defines the topology and (b) the class of continuous mappings coincides with sequentially continuous ones.

Let us now discuss properties of topological spaces known under the name **Separation axioms.** The full list goes from (T_0) to (T_6) including a few half-integers (with (T_k) being *strictly* weaker than (T_m) provided that $k < m$); we mention only the most important ones.

(T_0) (Kolmogorov's axiom) For each $x, y \in E$ s.t. $x \neq y$ the following holds: there exists an open neighborhood $x \in U_x \in \mathcal{O}$ of x such that $y \notin U_x$ *or* there exists an pone neighborhood $y \in U_y \in \mathcal{O}$ of y such that $x \notin U_y$.

Remark: this property allows to distinguish points in the space E: if $x \neq y$ do not satisfy this condition, then each open set $U \in \mathcal{O}$ either contains both x, y or none of them. For instance, consider the topology $\mathcal{O} = \{\emptyset, \{1\}, \{2,3\}, \{1,2,3\}\}$ on the three-element set $E = \{1,2,3\}$. The points 2 and 3 are not distinguishable: $\mathcal{O}$ can be viewed as the discrete topology on the *two*-element set $E/(2 \sim 3)$.

(T_1) (Fréchet's axiom) For each $x, y \in E$ s.t. $x \neq y$ there exists an open neighborhood $y \in U_y \in \mathcal{O}$ such that $x \notin U_y$.

Remark: This is equivalent to say that $\{x\}$ is a closed set for each $x \in E$. (Indeed, the (T_1) axiom implies that $E \setminus \{x\} = \bigcup_{y \neq x} U_y$ is an open set and thus $\{x\}$ is closed. Vice versa, if $\{x\}$ is closed, then one can take $U_y := E \setminus \{x\}$ for all $y \neq x$.)

Note that both co-finite and Zariski topologies *satisfy* this axiom (and do not satisfy the next one, which has been already mentioned above).

(T_2) (=Hausdorff=separated spaces) For each $x, y \in E$ s.t. $x \neq y$ there exists open neighborhoods $U_x, U_y \in \mathcal{O}$ such that $U_x \cap U_y = \emptyset$.

Remark: This property, in particular, guarantees the uniqueness of limits of sequences: if both $x_n \to x$ and $x_n \to y$ as $n \to \infty$ for $y \neq x$, then finding neighborhoods such that $U_x \cap U_y = \emptyset$ leads to a contradiction.

(T_4) (normal Hausdorff spaces) E satisfies (T_1) and for each closed sets F_0, F_1 s.t. $F_0 \cap F_1 = \emptyset$ there exists open sets $U_0 \supset F_0$ and $U_1 \supset F_1$ s.t. $U_0 \cap U_1 = \emptyset$.

(T_6) (perfectly normal Hausdorff spaces) E satisfies (T_1) and for each closed sets F_0, F_1 s.t. $F_0 \cap F_1 = \emptyset$ there exists a continuous function $f : E \to [0, 1]$ such that $F_0 = f^{-1}(\{0\})$ and $F_1 = f^{-1}(\{1\})$.

Remark: To see that $(T_6) \Rightarrow (T_4)$, take $U_0 := f^{-1}(\,[0, \frac{1}{2}[\,)$ and $U_1 := f^{-1}(\,]\frac{1}{2}, 1]\,)$. Vice versa, *Urysohn's lemma* (see TD) ensures that, if a topological space E satisfies (T_4) and closed sets $F_0, F_1 \subset E$ are such that $F_0 \cap F_1 = \emptyset$, then there exists a continuous function $f : E \to [0, 1]$ such that $f(x) = 0$ for all $x \in F_0$ and $f(x) = 1$ for all $x \in F_1$. However, note that this statement is strictly *weaker* than (T_6) as we only know that $F_0 \subset f^{-1}(\{0\})$ and not $F_0 = f^{-1}(\{0\})$ (and similarly for F_1).

We already mentioned above that metric spaces are Hausdorff. In fact, a stronger statement holds:

Proposition 4.5. *Metric (and hence metrizable topological) spaces are perfectly normal Hausdorff, i.e., satisfy the separation axiom* (T_6).

We need a lemma. Provided E is a metric space and $X, Y \subset E$, define

$$d(x, Y) := \inf_{y \in Y} d(x, y), \qquad d(X, Y) := \inf_{x \in X, y \in Y} d(x, y).$$

Lemma 4.6. *$F \subset E$ is closed iff the following holds: $(d(x, F) = 0 \Rightarrow x \in F)$.*

(However, note that $d(F_0, F_1) = 0$ does *not* imply $F_0 \cap F_1 \neq \emptyset$: e.g., consider two closed sets $F_0 := \{(x, y) \in \mathbb{R}^2 : y \leq 0\}$ and $F_1 := \{(x, y) \in \mathbb{R}^2 : y \geq |x|^{-1}\}$ in $\mathbb{R}^2$.)

Proof of Lemma 4.6. '$\Rightarrow$': Let F be closed and $d(x, F) = 0$, which means that there exist $x_n \in F$ such that $d(x_n, x) \leq 2^{-n}$. This gives $x_n \to x$ and $x \in \overline{F} = F$.

'$\Leftarrow$': Since metric spaces are first countable and hence Fréchet–Urysohn, for each $x \in \overline{F}$ there exists a sequence $x_n \in F$ such that $x_n \to x$ (i.e., $d(x_n, x) \to 0$). This implies $d(x, F) = 0$ and hence $x \in F$ due to the assumption. □

Proof of Proposition 4.5. The case $F_0 = F_1 = \emptyset$ is trivial: one can take $f(x) := \frac{1}{2}$. If $F_0 \neq \emptyset$ but $F_1 = \emptyset$, then the function $f(x) := d(x, F_0)/(1 + d(x, F_0))$ works. Finally, if both $F_0, F_1 \neq \emptyset$, then we can take $f(x) := d(x, F_0)/(d(x, F_0) + d(x, F_1))$; note that the denominator never vanishes due to Lemma 4.6 and $F_0 \cap F_1 = \emptyset$. □

September 30, 2020

5. How to introduce a topological space? Subspaces, product spaces, quotient spaces, 'final' and 'initial' topologies

In this section we discuss a few 'standard' ways to construct a topological space starting from other topological spaces.

5.0. **Détour: defining topologies via mappings.** Before going to concrete definitions, let us briefly discuss a more general idea of describing topologies via classes of mappings that we want to be continuous. (Let us emphasize that definitions given in the preceding paragraphs as well as those given after this discussion go in the opposite direction: given topologies in E_1 and E_2 one defines what are continuous mappings from E_1 to E_2.)

Reminder: in metric spaces we have the following characterizing property of closed sets: F is closed if and only if ($x \in F \Leftrightarrow d(x,F)=0$). In particular,

a subset F of a metric (or metrizable topological) space E is closed iff there exists a continuous mapping $f : E \to \mathbb{R}$ s.t. $F = f^{-1}(\{0\})$.

Indeed, the sets $f^{-1}(\{0\})$ are closed by definition of continuous mappings and for each closed set $F \subset E$ one can take $f := d(\cdot, F)$. (Q: Btw, why is $d(\cdot, F)$ continuous? A: It follows from the triangle inequality that $|d(x,F)-d(x',F)| \le d(x,x')$ by taking the infimum over $y \in F$ in $d(x',y) \le d(x,y)+d(x,x')$.)

Note that this description is pretty similar to the definition of the *Zariski topology* on k^n, recall that the closed sets in this topology are characterized as follows:

$F \subset k^n$ is closed iff there exists a *family* $\{P_\alpha\}_{\alpha \in A} \subset k[x_1,\ldots,x_n]$ such that $F = \{x \in k^n : P_\alpha(x_1,\ldots,x_n) = 0 \text{ for all } \alpha \in A\}$.

The only difference between the two situations is that we have a *single* mapping f in the former (metric) setup and a *family* P_α of polynomials in the latter (Zariski).

- A priori, we do not require any restriction on the set of indices A: it could be infinite (uncountable, etc): this is because we want the intersection of closed sets $P_\alpha^{-1}(\{0\})$ to be closed, no matter what the set of indices A is.
- However (see the course *Algébre 2*), provided that k is a field (or, more generally, a *Noetherian ring*: e.g., $k = \mathbb{Z}$ is OK), it is not that hard to show that, for each family $\{P_\alpha\}_{\alpha \in A}$ of polynomials in n variables with coefficients in k, there exists a *finite* subfamily $\{P_{\alpha_s}\}_{s=1,\ldots,m}$ such that

$$P_\alpha(x_1,\ldots,x_n) = 0 \text{ for all } \alpha \in A \ \Leftrightarrow P_{\alpha_s}(x_1,\ldots,x_n) = 0 \text{ for all } s = 1,\ldots,m.$$

  This statement (to be precise, the fact that the ring $k[x_1,\ldots,x_n]$ of polynomials is Noetherian, see *'Algébre 2'*) allows to rewrite the definition of closed sets in the Zarisky topology as follows:

  a set $F \subset k^n$ is closed iff there exists a polynomial mapping $P = (P_{\alpha_s})_{s=1,\ldots,m} : k^n \to k^m$ such that $F = P^{-1}(\{0\})$.
- Since compositions of polynomials are again polynomials, one easily see that they are all continuous in this topology: if a closed set $F \subset k^n$ is defined by $P : k^n \to k^m$ (i.e., if $F = P^{-1}(\{0\})$) and $Q : k^s \to k^n$ is algebraic, then $Q^{-1}(F) = (P \circ Q)^{-1}(\{0\})$ and hence $Q^{-1}(F)$ is closed.

To summarize the preceding discussion: if we start with the class of polynomial mappings $P : k^n \to k^m$ (which is closed under compositions) and want to introduce topologies in these spaces so that all these mappings are continuous, then the Zariski topology is very natural: it is the *coarsest* – i.e. having as less open (or, equivalently, closed) sets – (T_1) topology in which all P's are continuous.

The construction explained above (classes of 'nice' mappings between certain sets $\rightsquigarrow$ coarsest topologies that make these mappings continuous) is very general but we will <u>not</u> develop it further.

Let us now move back to a more pedestrian level.

5.1. **Subspaces of topological spaces.**

Definition 5.1. *Let $(E, \mathcal{O})$ be a topological space and $E' \subset E$ be a subset of this space (<u>not</u> necessarily open or closed). Denote*

$$\mathcal{O}' \ := \ \{U' \subset E' : \ \exists U \in \mathcal{O} \ s.t. \ U' = U \cap E'\}.$$

Then, $\mathcal{O}'$ is a topology on E', which is called the induced *(from E to E') topology.*

(The check that $\mathcal{O}'$ is a topology is straightforward: e.g., if $U'_\alpha = U_\alpha \cap E' \in \mathcal{O}'$, then $\bigcup_{\alpha \in A} U'_\alpha = \big(\bigcup_{\alpha\in A} U_\alpha\big) \cap E' \in \mathcal{O}'$; the other properties are even simple.)

- Note that the induced topology on E' is the *coarsest* topology on E' in which the inclusion $E' \hookrightarrow E$ (i.e., the mapping $x \mapsto x$) is continuous. Indeed, each set $U' \in \mathcal{O}'$ is a pre-image of an open set U in E under this inclusion and thus must be open if we want $E' \hookrightarrow E$ to be continuous.
- This topological definition is consistent with the metric setup. Namely, if E was a metric space, then one view E' as a metric space simply by restricting the distance function $d : E \times E \to \mathbb{R}_+$ to $E' \times E'$. To see that thus obtained $d' = d|_{E\times E}$ defines the same topology as the induced one from E, it is enough to check that open balls are in $\mathcal{O}'$, which is trivial.
- If $E'' \subset E' \subset E$, then the topology induced on E'' from E' is the same as the topology induced on E'' directly from E.

As a very simple example, consider $E = \mathbb{R}$ (equipped with the standard topology) and $E' = [0, 1]$ (or $E' = \mathbb{Q}$). Note that, e.g., $[0, \frac{1}{2})$ is *open* in $[0, 1]$. Below we list several properties of induced topologies. It is straightforward to see that:

- $F' \subset E'$ is closed iff there exists a closed set $F \subset E$ s.t. $F' = F \cap E$. (Indeed, $F' = F \cap E'$ is equivalent to $E' \smallsetminus F' = (E \smallsetminus F) \cap E'$.)
- If $X \subset E'$, then $\overline{X}^{E'} = \overline{X}^{E} \cap E'$, where in the left-hand side the closure of X is taken in the topology of E' whilst in the right-hand side the closure is taken in E. (Indeed, $\bigcap_{X\subset F\text{ – closed in }E} F \cap E' = \bigcap_{X\subset F'\text{ – closed in }E'} F'$.)
  - However, note that $\mathrm{Int}_{E'} X \neq E' \cap \mathrm{Int}_E X$.
    (E.g., consider $E = \mathbb{R}^2$ and $X = E' = \mathbb{R}$ in the usual topology.)
- Let $x_n, x_* \in E'$. Then, the convergence $x_n \to x$ in E' is equivalent to the same convergence in E. (Indeed, to say that $x_n \in U'_x = U_x \cap E'$ is equivalent to saying that $x_n \in U_x$.)

We now discuss which 'nice' properties of topological space survive when passing from E to its subspace E'.

(i) ◦ If E is first-countable, then so is E'.
◦ If E is second-countable, then so is E'.
(Indeed, a (local or global) countable base $\{U'_n\}_{n\in\mathbb{N}}$ of the topology $\mathcal{O}'$ in E' is given by $U'_n := U_n \cap E'$, where $\{U_n\}_{n\in\mathbb{N}}$ is a (local or global) base in E.)

(ii) ◦ If E is Hausdorff (=(T_2)=separated), then E' is also Hausdorff.
◦ The same property holds for the separation axiom (T_1).
(Indeed, take $U'_x := U_x \cap E'$ and $U'_y \cap E'$.)

- **Warning:** However, this is <u>not</u> the case for (T_4). (The catch is that, given two closed sets $F'_{0,1} \subset E'$ s.t. $F'_0 \cap F'_1 = \emptyset$, it can be not possible to find closed $F_{0,1} \subset E$ such that $F'_{0,1} = F_{0,1} \cap E'$ *and* $F_0 \cap F_1 = \emptyset$.)

- **Warning:** The fact that E is separable does <u>not</u> imply that E' is separable (though this is OK for metric spaces since in the metric setup the separability is equivalent to the second countability).

Useful reference: Steen, Lynn Arthur; Seebach, J. Arthur, *Counterexamples in topology. 2nd ed.*, New York-Heidelberg-Berlin: Springer-Verlag. XI, 244 p. (1978).

5.2. **Products of topological spaces.**

Definition 5.2. *Let $\{E_\alpha\}_\alpha \in A$ be a family (possibly infinite, uncountable etc) of topological spaces and denote*

$$E \ := \ \prod_{\alpha\in A} E_\alpha \ = \ \{x=(x_\alpha)_{\alpha\in A} \ \text{s.t.}\ x_\alpha \in E_\alpha \ \text{for all}\ \alpha\in A\}.$$

The base of the product *topology in E is given by the collection of sets*

$$\left\{ \prod_{\alpha\in A} U_\alpha : \begin{array}{l} U_\alpha \text{ is open in } E_\alpha \text{ for all } \alpha\in A, \\ U_\alpha = E_\alpha \text{ unless } \alpha\in A_0, \text{ where } A_0\subset A \text{ is a \underline{finite} set.} \end{array} \right\}. \tag{5.1}$$

Certainly, one should check that the collection (5.1) can serve as a base of a topology in E, i.e. that the assumption of Proposition 2.4 holds. Indeed,

$$\big(\prod_{\alpha\in A} U'_\alpha\big) \cap \big(\prod_{\alpha\in A} U''_\alpha\big) = \prod_{\alpha\in A}(U'_\alpha \cap U''_\alpha)$$

and, if we declare $U_\alpha := U'_\alpha \cap U''_\alpha$, then $U_\alpha = E_\alpha$ for all α except the finite set $A'_0 \cup A''_0$. Still, there is a question: why do we impose this finiteness condition in (5.1) (note that, if we drop it, then the result is again a base of a certain topology on E). The answer to this question is given by the following statement:

- The product topology is the *coarsest* topology on $E = \prod_{\alpha\in A} E_\alpha$ such that all the projections $\pi_\alpha : E \to E_\alpha$, $x = (x_\alpha)_{\alpha\in A} \mapsto x_\alpha$, $\alpha \in A$, are continuous.

(Indeed, the continuity of π_α is equivalent to say that the sets (5.1) with $A_0 = \{\alpha\}$ are open and finite intersections (which have to be open too) lead to finite A_0's.)

Lemma 5.3. *The convergence $x^{(n)} \to x$ in the product space E is equivalent to the coordinate-wise convergence, i.e., to saying that $x^{(n)}_\alpha \to x_\alpha$ for each $\alpha \in A$.*

Proof. To deduce $x^{(n)}_\alpha \to x_\alpha$ from $x^{(n)} \to x$, simply consider the open sets (5.1) with $A_0 = \{\alpha\}$. Vice versa, assume that $x^{(n)}_\alpha \to x_\alpha$ for all $\alpha \in A$. For each base set (5.1) containing x and each $\alpha \in A_0$ there exists n_α such that $x^{(n)}_\alpha \in U_\alpha$ for all $n \geq n_\alpha$ (since $x^{(n)}_\alpha \to x_\alpha \in U_\alpha$). Therefore, $x^{(n)}$ belongs to the set (5.1) for all $n \geq \max_{\alpha\in A_0} n_\alpha$; note that we take the maximum over a finite set $A_0 \subset A$. □

Example. Recall the space $C^\infty([0,1]) = C^\infty([0,1],\mathbb{R})$ discussed in Section 3; a similar consideration applies to the Schwartz space $\mathcal{S}(\mathbb{R})$ and to all other vectored spaces with the topology defined by a countable family of semi-norms (see TD). We have the inclusion

$$\begin{array}{rcl} C^\infty([0,1]) & \hookrightarrow & C([0,1]) \times C^1([0,1]) \times C^2([0,1]) \times ..., \\ f & \mapsto & (f, f, f, \ldots); \end{array}$$

in other words, $C^\infty([0,1]) = \bigcap_{k\in\mathbb{N}} C^k([0,1])$ can be identified (as a *set*) with the 'diagonal' of the Cartesian product $\prod_{k\in\mathbb{N}} C^k([0,1])$. It is easy to see that the *topology* induced on this set from the product topology in $\prod_{k\in\mathbb{N}} C^k([0,1])$ coincides with that discussed in Section 3: compare Lemma 3.3 with Lemma 5.3 and use the fact that a subspace of a countable product of first-countable spaces is first-countable; see below. (The first-countability guarantees that describing convergent *sequences* is equivalent to describing the topology.)

We will start the next lecture with discussing

- how to define the product of *metric* spaces (spoiler, see TD: if all E_α are metrizable, then the product topology in $E = \prod_{\alpha\in A} E_\alpha$ is metrizable if A is finite or countable and <u>not</u> metrizable otherwise, provided that all E_α are non-trivial, i.e., contain at least two points);
- which 'nice' topological properties of E_α are inherited by their product E.

October 05, 2020

Let us discuss which 'nice' properties of topological spaces E_α are necessarily inherited in their product $E = \prod_{\alpha\in A} E_\alpha$.

- **Separation axioms.**
  - If all E_α are Haudorff=(T_2), then E is also Hausdorff; the same holds for (T_1). (Indeed, if $x \neq y$, then $x_\alpha \neq y_\alpha$ for a certain $\alpha \in A$ and we can separate these coordinates, i.e., find open (in E_α) sets $U_{x_\alpha} \cap U_{y_\alpha} = \emptyset$ and then consider cylindric sets (5.1) with $U_{x_\beta} = U_{y_\beta} = E_\beta$, $\beta \neq \alpha$.)
  - E is *not* necessarily (T_4) if all E_α are (T_4). To see a difficulty, note that one cannot say almost anything about projections $\pi_\alpha(F_{0,1}) \subset E_\alpha$.

    Exercise:* the Sorgenfrey line ($\mathbb{R}$ equipped with the topology generated by the base $\mathcal{B} = \{[a,b[, -\infty < a < b < +\infty\})$ is a (T_4) space.

    However, the product of two Sorgenfrey lines – the so-called Sorgenfrey plane – is not (T_4): it is not hard to see that the anti-diagonal $\{(x,-x)\}$ is a closed set and ,moreover, all subsets of the anti-diagonal are closed in the Sorgenfrey plane. Choosing $F_0 = \{(x,-x), x \in \mathbb{Q}\}$ and $F_1 = \{(x,-x), x \notin \mathbb{Q}\}$ leads to a contradiction[1] with (T_4)

[1] This is not a triviality. If $F_{0,1} \subset U_{0,1}$, where $U_{0,1}$ are open disjoint sets in the Sorgenfrey plane, then for each point $(x,-x) \in F_1$, there exists $\varepsilon_x > 0$ such that $[x, x+\varepsilon_x[\times[-x, -x+\varepsilon_x[\subset U_1$ and similarly for $(x,-x) \in F_0$. However, it is easy to see that the assumption $U_0 \cap U_1 = \emptyset$ implies that each of the sets $\{x \notin \mathbb{Q} : \varepsilon_x > 2^{-n}\}$ is *nowhere dense*, a contradiction with the material that we will discuss later (see *Baire's theorem*).

- **Countability axioms.**
  - If A is at most countable and all E_α are first/second-countable, then E is also first/second-countable. (Indeed, in this case the collection of sets (5.1) with U_α being taken from a countable base (or a local base) in E_α is a *countable* union (over $A_0 \subset A$) of countable collections, each of which is a finite Cartesian product of countable bases.)
  - If A is uncountable (and provided that all E_α are nontrivial, i.e., contain a nontrivial open set $x_\alpha \in U_\alpha \subsetneq E_\alpha$), then E cannot be even first-countable at $x = (x_\alpha)_{\alpha\in A}$. Indeed, let $U^{(\alpha)} := \prod_{\beta\in A} U^{(\alpha)}_\beta$, where $U^{(\alpha)}_\beta = E_\beta$ if $\alpha \neq \beta$ and $U^{(\alpha)}_\alpha = U_\alpha$. If the topology of E had a countable base, then uncountably many of open sets $U^{(\alpha)} \subset E$ would contain the same base set V, which means that $\pi_\alpha(V) \subset U_\alpha \subsetneq E_\alpha$ for infinitely (in fact, uncountably) many indices $\alpha \in A$. This leads to a contradiction with the definition of the topology on E as V – being open in E – must contain one of the sets (5.1).
- **Separability.**
  - Again, if A is at most countable and all E_α are separable, then E is also separable. (Indeed, let $\{x^{(\alpha)}_m\}_{\alpha\in A}$ be countable dense subsets of E_α. If $A = \{1, \ldots, N\}$ is finite, then the set $\{(x^{(1)}_{m_1}, \ldots, x^{(N)}_{m_N})\}_{m_1,\ldots,m_N\in\mathbb{N}}$ is still countable and dense in E. If $A = \mathbb{N}$, then we can consider the set of sequences $x = (x^{(n)}_{m_n})_{n\in\mathbb{N}} \in E$ such that $m_n = 0$ for all $n \geq N = N(x)$. Clearly, each set (5.1) contains such a sequence since $A_0 \subset A$ is finite.)
  - **[!]**[2] A less trivial statement is that E is separable provided that $A = \mathbb{R}$ (which has the same cardinality as $2^{\mathbb{N}}$); see the footnote.
  - If the cardinality of A exceeds that of $2^{\mathbb{N}}$ and each E_α contains two non-empty open sets U_α, V_α s.t. $U_\alpha \cap V_\alpha = \emptyset$, then E is not separable.
- **Metrizability.** Let E_α be *metric* spaces with distance functions d_{E_α}.
  - If $A = \{1, \ldots, N\}$ is finite, then $d_E(x, y) := \max_{1\leq k\leq N} d_{E_k}(x_k, y_k)$ is a distance function on E, which defines the product topology on E. It is worth noting that this can be also done for *normed* vector spaces E_k: setting $\|x\|_E := \max_{1\leq k\leq N} \|x_k\|_{E_k}$ makes E a normed vector space.
  - If $A = \mathbb{N}$ is countable, then E is still a metric space: for instance, one can set $d(x, y) := \max_{k\in\mathbb{N}} \min\{d_{E_k}(x_k, y_k), 2^{-k}\}$. To see that d_E defines the product topology, one can, e.g., argue that the convergence $d_E(x^{(n)}, x) \to 0$ as $n \to \infty$ is equivalent to the fact that, for

[2]Let us show that the product of separable spaces E_α <u>is</u> separable if $A = \mathbb{R}$. Let $\{x^{(n)}_\alpha\}$ denote a dense countable set in E_α. Given a finite set $Q = \{q_1, \ldots, q_m\}$ of rational numbers (where $-\infty =: q_0 < q_1 < \ldots < q_m < q_{m+1} := +\infty$) and a multi-index $N = (n_1, \ldots, n_{m+1}) \in \mathbb{N}^{m+1}$, let

$$x^{(Q,N)}_\alpha := x^{(n_k)}_\alpha \quad \text{if } q_k \leq \alpha < q_{k+1}.$$

This is a countable collection of points in E and all sets (5.1) contain at least one point $x^{(Q,N)}_\alpha$.

On the other hand, assume that each E_α contain two non-empty disjoint sets U_α, V_α and the cardinality of A exceeds that of $2^{\mathbb{N}}$. Given $\{x^{(n)}\}_{n\in\mathbb{N}} \subset E$, let $A^{(n)} := \{\alpha \in A : x^{(n)}_\alpha \in U_\alpha\}$. Since $\mathrm{card} A > \mathrm{card}\, 2^{\mathbb{N}}$, one can find $\alpha \neq \alpha'$ such that for all $n \in \mathbb{N}$ either both $\alpha, \alpha' \in A^{(n)}$ or both $\alpha, \alpha' \notin A^{(n)}$. Then, the open set $U_\alpha \times V_{\alpha'} \times \prod_{\beta\neq\alpha,\alpha'} E_\beta$ does not contain points $x^{(n)}$.

each $k \in \mathbb{N}$ one has $d_{E_k}(x_k^{(n)}, x_k) \to 0$. (This argument relies on the first-countability of E; see above.) However, note that the countable product of normed vector spaces is <u>not</u> a normed space.

- ○ If A is uncountable (and all factors E_α are non-trivial, i.e., if each E_α contains at least two distinct points), then the space E is <u>not</u> metrizable. Indeed, as discussed above, this space is not first-countable.

5.3. **'Initial' and 'final' topologies.** Mostly for the completeness of the presentation (recall also the digression made at the beginning of the previous lecture), let us give two abstract definitions.

Definition 5.4. *Let E be a set and assume that we are given a collection of mappings $f_\alpha : E \to E_\alpha$ from E to topological spaces E_α, $\alpha \in A$. The* 'initial' *topology defined by $\{f_\alpha\}_{\alpha\in A}$ is the <u>coarsest</u> topology on E such that all f_α are continuous.*

It easy to see that this definition makes sense (i.e., that the intersection of all topologies on E such that all f_α are continuous is again a topology on E). In fact, above we already see two examples of this construction:

- The subspace topology on $E' \subset E$ is the coarsest topology in which the mapping $E' \hookrightarrow E$ is continuous.
- The product topology is the coarsest topology in which all the projections $\pi_\alpha : E = \prod_{\alpha\in A} E_\alpha \to E_\alpha$ are continuous.

As often in math, one can also consider a 'dual' definition:

Definition 5.5. *Let E be a set and assume that we are given a collection of mappings $f_\alpha : E_\alpha \to E$ from topological spaces E_α, $\alpha \in A$, to E. The* 'final' *topology defined by $\{f_\alpha\}_{\alpha\in A}$ is the <u>finest</u> topology on E such that all f_α are continuous.*

Equivalently, $U \subset E$ is open in the 'final' topology iff $f_\alpha^{-1}(U)$ is open in E_α for all $\alpha \in A$; it is straightforward to check that this condition defines a topology.

Warning: While constructing the 'initial' topology is a reasonably nice operation as can be seen from the examples discussed above, passing to the 'final' topology can be potentially dangerous even if we work with a single mapping $E_1 \to E$ and pushforward a nice topology from E_1 to E as we will see now.

5.4. **Quotient topology.** Let $\sim$ be an equivalence relation on E; denote by $[x] \subset E$ the equivalence class of x and by $E/\!\sim$ the set of all such equivalence classes.

Definition 5.6. *The quotient topology on $E/\!\sim$ is the finest topology in which the mapping $x \mapsto [x]$ is continuous. In other words, a set $U \subset E/\!\sim$ is open if and only if the set $\{x \in E : [x] \in U\}$ is open in E.*

It is easy to see that the space $E/\!\sim$ is separable provided that so is E. (Indeed, if $X \subset E$ is a dense set in E, then the set $\{[x] \in E/\!\sim, x \in X\}$ is dense in $E/\!\sim$.) However, both separation and countability properties can be lost under passing to a quotient space. We illustrate this by the following examples.

- **'Real line with two origins'.** This is an example of a Hausdorff space E and an equivalence relation $\sim$ on it leading to a non-Hausdorff space $E/\!\sim$.

  Let $E := \mathbb{R} \sqcup \mathbb{R}$ be the *disjoint union* of two copies of $\mathbb{R}$, each endowed with the standard topology. Formally, $E := \{(x,1), x \in \mathbb{R}\} \cup \{(x,2), x \in \mathbb{R}\}$ and a set $U \subset E$ is open iff both sets $\{x \in \mathbb{R} : (x,1) \in U\}$ and $\{x \in \mathbb{R} : (x,2) \in U\}$

are open. (This is a general construction, the disjoint union of other topological spaces is defined in the same way.)

Now define an equivalence relation on E by declaring $(x,1) \sim (x,2)$ for $x \neq 0$. In other words, most equivalence classes in E consist of two points and can be identified with $x \in \mathbb{R} \setminus \{0\}$ but there are also two one-point equivalence classes $0_1 := [(0,1)]$ and $0_2 := [(0,2)]$.

The local bases of the quotient topology on $E/{\sim}$ at $x \in \mathbb{R} \setminus \{0\}$ are the same as that of $\mathbb{R}$, while open neighborhoods of the point 0_1 (and similarly for 0_2) are $\{x \in \mathbb{R} \setminus \{0\} : |x| < r\} \cup \{0_1\}$. In particular, the quotient space $E/{\sim}$ is not Hausdorff: 0_1 and 0_2 cannot be separated by open sets.

- **'Bouquet of infinitely many circles'.** This is an illustration that a quotient space of $E = \mathbb{R}$ can be not first-countable. For $x, y \in \mathbb{R}$, let $x \sim y$ if *both* $x, y \in \mathbb{Z}$. In other words, most equivalence classes consist of a single point $x \in \mathbb{R} \setminus \mathbb{Z}$ while there is a single infinite equivalence class $\mathbb{Z}$, which we will denote $\mathbf{0} \in E/{\sim}$. Again, the local bases of $E/{\sim}$ at points $x \in \mathbb{R} \setminus \mathbb{Z}$ are the usual ones while the condition $\mathbf{0} \in U$ – open in $E/{\sim}$ means that for each $m \in \mathbb{Z}$ there exists $0 < r_m$ such that $(m - r_m, m) \cup (m, m + r_m) \subset U$.

  It is not hard to see that the point $\mathbf{0}$ does not admit a countable local base in $E/{\sim}$. Indeed, if this was the case and $r_m^{(n)} > 0$ were the radii corresponding to a base set $U^{(n)} \ni \mathbf{0}$, $n \in \mathbb{Z}$, then one could consider the open neighborhood
  $$\{\mathbf{0}\} \cup \bigcup_{m \in \mathbb{Z}} \big((m - \tfrac{1}{2} r_m^{(m)}, m) \cup (m, m + \tfrac{1}{2} r_m^{(m)})\big) \qquad (5.2)$$
  of $\mathbf{0}$, which does not contain any of the sets $U^{(n)}$.

Though considering the quotient topology is potentially dangerous, this is what people quite often do, as this is often the simplest way to define a topological space.

Examples. (Disclaimer: on a formal level, in each of these examples a certain technical work should be done to verify the claims; we leave this work to the reader. The primary goal of these examples is to understand the language used.)

- Consider the set $E = \mathbb{R}$ and the equivalence relation $x \sim y$ if $x - y \in \mathbb{Z}$. The quotient space $\mathbb{R}/{\sim}$, also known as $\mathbb{R}/\mathbb{Z}$, is homeomorphic to the one-dimensional circle $\mathrm{S}^1 := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$; the homeomorphism is given by $x \mapsto (\cos(2\pi x), \sin(2\pi x))$. Note that another way to get S^1 as a quotient space is to consider the segment $[0,1]$ and to declare that $0 \sim 1$:
  $$\mathbb{R}/\mathbb{Z} \ \cong\ \mathrm{S}^1 \ \cong\ [0,1]/(0 \sim 1).$$
- One can construct the two-dimensional torus in a similar way:
  $$\mathbb{C}/(\mathbb{Z} + i\mathbb{Z}) \ \cong\ \mathrm{S}^1 \times \mathrm{S}^1 \ \cong\ [0,1]^2/((x,0) \sim (x,1), (0,y) \sim (1,y)).$$
  (In the left-hand side, the equivalence relation is given by $z \sim w$ if both $\mathrm{Re}(z-w), \mathrm{Im}(z-w) \in \mathbb{Z}$. In the right-hand side all four corners of the square lie in the same equivalence class: $(0,0) \sim (0,1) \sim (1,0) \sim (1,1)$.)
- In the same spirit, one obtains
  - the cylinder $[0,1]^2/((x,0) \sim (x,1))$;
  - the *Möbius strip* $[0,1]^2/((0,y) \sim (1,1-y))$;
  - the *Klein bottle* $[0,1]^2/((x,0) \sim (x,1), (0,y) \sim (1,1-y)))$.

Remark: One can construct the quotient topology step by step: e.g., to obtain the Klein bottle one can either start with the cylinder and glue two its boundaries to each other in a twisted way or, equivalently, start with the Möbius strip and glue a half of its boundary to the other half.

- Finally, playing with identifications of the opposite sides of the square, one can also consider the space
$$[0,1]^2/((x,0)\sim(1-x,1),(0,y)\sim(1,1-y)) \;\cong\; \mathbb{R}\mathrm{P}^2, \tag{5.3}$$
the so-called real two-dimensional *projective space*. The right-hand side is defined as follows:
$$\begin{aligned}\mathbb{R}\mathrm{P}^2 \;:=\; & (\mathbb{R}^3\smallsetminus\{0\})/((x,y,z)\sim(\lambda x,\lambda y,\lambda z))\\ =\; & \mathrm{S}^2/((x,y,z)\sim(-x,-y,-z)),\end{aligned}$$
where $\mathrm{S}^2 := \{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2=1\}$ is the two-dimensional sphere. The fact that the latter quotient space is homeomoprhic to (5.3) follows by considering the mapping $(x,y)\mapsto(x,y,\sqrt{1-x^2-y^2})$ of the closed unit disc $\overline{\mathbb{D}}$ onto the upper hemisphere, noting that $\overline{\mathbb{D}}$ is homeomorphic to the square $[0,1]^2$, and checking that the identification $(x,y)\sim(-x,-y)$ on the boundary of $\overline{\mathbb{D}}$ matches the identification on the boundary of $[0,1]^2$ in (5.3).

Détour: topological manifolds and Lie groups. The torus, Klein bottle and $\mathbb{R}\mathrm{P}^2$, as well as the cylinder and the Möbius strip considered *without* their boundaries (i.e., we take the space $[0,1]\times(0,1)$ and factorize it by the equivalence relation $(0,y)\sim(1,y)$ or $(0,y)\sim(1,1-y)$, respectively) are examples of:

Definition 5.7. *A Hausdorff, second-countable topological space X is called a C^k-smooth n-dimensional* topological manifold *if there exists a collection (called an* atlas*) of open sets $U_\alpha\subset X$ (called* charts*) such that*

(i) *$X=\bigcup_{\alpha\in A}U_\alpha$ (i.e., the whole space is covered by charts);*
(ii) *each U_α is homeomorphic to $\mathbb{R}^n$ (or, equivalently, to an open ball in $\mathbb{R}^n$), let $f_\alpha : U_\alpha\to\mathbb{R}^n$ denote the corresponding homeomorphism;*
(iii) *if $U_\alpha\cap U_\beta\neq\emptyset$, then the mapping $f_\beta\circ f_\alpha^{-1} : f_\alpha(U_\alpha\cap U_\beta)\to f_\beta(U_\alpha\cap U_\beta)$ is C^k-continuous (i.e., passing from one chart of X to another is continuous); note that this condition is automatic if $k=0$.*

A manifold X is called smooth, if all $f_\beta\circ f_\alpha^{-1}$ are C^∞ mappings.

Remarks: (i) At this point we do not (at least formally) know what C^k mappings acting from $\mathbb{R}^n$ to $\mathbb{R}^n$ are. This will be discussed in the second part of the course.

(ii) At first sight, it is not clear why do we additionally require X to be Hausdorff if we also ask that each point x is a neighborhood homeomorphic to $\mathbb{R}^n$. However, the latter holds for the real line with two origins, which is *not* Hausdorff.

A trivial example of a smooth manifold is an open set $X=V$ in $\mathbb{R}^n$: by definition, it can be covered by open balls U_α and one can take $f_\alpha(x):=x$. Let us give a more interesting example of a different nature:

Let $X := \{A\in\mathbb{R}^{2\times 2} : \det A = 1\}$, which can be equivalently written as $\{(a,b,c,d) : ad-bc=1\}\subset\mathbb{R}^4$, with the subspace topology. We will start the next lecture by (re-)discussing this example.

October 07, 2020

The manifold $X = \mathrm{SL}(2,\mathbb{R}) = \{A \in \mathrm{GL}(2,\mathbb{R}) : \det A = 1\}$ is an example of a (matrix) Lie group (Sophus Lie, 1842–1899, Norwegian $\Rightarrow$ pronounced as 'Lee').

Definition 5.8. *A group G is called an (abstract)* Lie group *if G also carries the structure of a smooth manifold and both mappings $g_1, g_2 \mapsto g_1g_2$, $G \times G \to G$, and $g \to g^{-1}$, $G \to G$, are continuous. G is called a* matrix Lie group *if it is a subgroup of the group $\mathrm{GL}(N,\mathbb{R})$ (or $\mathrm{GL}(N,\mathbb{C})$) of all invertible $N \times N$ matrices.*

Indeed, on $X = \mathrm{SL}(2,\mathbb{R})$ one can consider four open sets $V_a^\pm, V_b^\pm, V_c^\pm, V_d^\pm \subset X$ defined by the conditions $\pm a > 0$, $\pm b > 0$ etc. Let $f_a^\pm : V_a^\pm \to \mathbb{R}^3$ be defined as $f_a^\pm(a,b,c,d) := (a,b,c)$, note that $(f_a^\pm)^{-1}(a,b,c) = (a,b,c,(bc+1)/a)$. This mapping is a homeomorphism of $V_a^\pm$ and $\mathbb{R}_\pm \times \mathbb{R} \times \mathbb{R} \cong \mathbb{R}^3$. (Note that we are in the metric setup, so one only needs to check that convergent sequences are mapped to convergent sequences, which is trivial.) All composition mappings like $f_b^\pm \circ (f_a^\pm)^{-1}$ are rational and thus C^∞-smooth.

• **Whitney's embedding theorem.** In the example discussed above, the three-dimensional manifold $X = \mathrm{SL}(2,\mathbb{R})$ can be viewed as (more rigorously, is homeomorphic and, moreover, *diffeomorphic* – see the second part of the course – to) a subset of $\mathbb{R}^4$, which locally looks like a graph of a function $((d = (bc+1)/a$, $c = (ad-1)/b$ etc). A natural question arises: whether this is always the case, i.e., is it true that each smooth n-dimensional manifold X can be embedded into $\mathbb{R}^m$ with large enough m in this way? The affirmative answer to this question for $m = 2n$ and C^∞-smooth manifolds is given in 1930s by Hassler Whitney(1907–1989). If (and only if) $n = 2^s$, then the bound $m = 2n$ is sharp: for such n, the projective space $\mathbb{R}\mathrm{P}^n$ cannot be embedded into $\mathbb{R}^{2n-1}$ (note a highly nontrivial interplay between 'geometric' and 'arithmetic' properties). In fact, one can replace the C^∞-smoothness assumption by C^2 (see also the second part of this course).

---

6. Compact spaces and sets

We now move back to the pedestrian level and continue discussing how the notions familiar from the real analysis are formulated for topological/metric spaces.

Definition 6.1. *One says that a topological space E satisfies the* Borel–Lebesgue property *if for each family of open sets $\{U_\alpha\}_{\alpha\in A}$ such that $E = \bigcup_{\alpha\in A} U_\alpha$ (such families are called* open covers *of E) there exists a finite set $\{\alpha_1, \dots, \alpha_n\} \subset A$ such that $E = \bigcup_{k=1}^n U_{\alpha_k}$ (then, $\{U_{\alpha_k}\}_{k=1,\dots,n}$ is called a* finite sub-cover *of E).*

Trivially (by passing to the complements), the Borel–Lebesgue property of E is equivalent to say that for each family $\{F_\alpha\}_{\alpha\in A}$ of closed sets such that $\bigcap_{\alpha\in A} F_\alpha = \emptyset$ there exists a finite subset $\{F_{\alpha_k}\}_{k=1,\dots,n}$ such that $\bigcap_{k=1,\dots,n} F_{\alpha_k} = \emptyset$.

[**!**] It so happened that the following definition is country-dependent.

Definition 6.2. *In France:*

- *A topological space E is called* quasi-compact *iff it satisfies the BL property.*
- *A quasi-compact space E is called* compact *if it is Hausdorff=(T_2).*

In the rest of the world:

- *E is called* compact *iff it satisfies the Borel–Lebesgue property.*

Definition 6.3. *A set $K \subset E$ is called (quasi-)compact if K endowed with the subspace topology is a (quasi-)compact topological space. This can be re-formulated as follows: for each open cover of K (i.e., for each collection of open – in E – sets U_α such that $K \subset \bigcup_{\alpha \in A} U_\alpha$) there exists a finite subcover: $K \subset \bigcup_{k=1,\dots,n} U_{\alpha_k}$.*

In the metric setup there is no difference between the French/English terminologies. However, this becomes relevant in algebraic geometry: *open* sets in the Zariski topology satisfy the BL property (i.e., are *quasi-compact*) and it is debatable whether the name *compact* is also appropriate for them or not.

Let us now discuss a few simple examples.

(1) The segment $[0, 1] \subset \mathbb{R}$ is compact.

(*Proof.* Let $[0, 1] \subset \bigcup_{\alpha \in A} U_\alpha$ and denote

$$t^* := \sup\{t \in [0, 1] : \text{ the segment } [0, t] \text{ admits a finite sub-cover by } U_\alpha\}$$

Note that $t^* > 0$. Indeed, there exists $\alpha \in A$ such that $0 \in U_\alpha$ and, since U_α is open, there exists $r > 0$ such that $[0, r) \subset U_\alpha$, which implies $t^* \geq r$. Also, it cannot be that $0 < t^* < 1$: in this case one uses the fact that there exists $\alpha \in A$ and $r > 0$ such that $(t^* - r, t^* + r) \subset U_\alpha$ and hence the segment $[0, t^* + \frac{1}{2}r]$ also admits a finite subcover since one can add this U_α to a finite subcover of the segment $[0, t^* - \frac{1}{2}r]$. Finally, if $t^* = 1$, then one finds $\alpha \in A$ and $r > 0$ such that $(1 - r, 1] \subset U_\alpha$ and adds this U_α to a finite subcover of the segment $[0, 1 - \frac{1}{2}r]$. □)

(2) **Compact sets in $\mathbb{R}^n$.** In fact, the following is fulfilled:

$$K \subset \mathbb{R}^n \text{ is compact} \iff K \text{ is a bounded closed set in } \mathbb{R}^n.$$

(The proof relies upon several general facts, which we will discus next.)

Proof. '⇒' The fact that K is bounded follows by considering the family $U_m := B(0, m)$ in the Borel–Lebesgue property. Trivially, $K \subset \bigcup_{m \in \mathbb{N}} B(0, m)$ and hence there exists a finite subcover of K, i.e., K is bounded. The fact that K is closed follows from a general fact that compact sets in Hausdorff(!) spaces are closed; see Lemma 6.6 below.

'⇐' We rely upon a fact that a closed subset of a compact set is also compact (see Lemma 6.5) and on the fact that the closed cubes $[-m, m]^n \subset \mathbb{R}^n$ are compact; the latter follows, e.g., from the compactness of the segment $[-m, m] \subset \mathbb{R}$ and the Tykhonov theorem (see Theorem 6.11 or 6.14). □

(3) To illustrate why *quasi*-compact sets are not necessarily closed in non-Hausdorff spaces, one can consider the co-finite topology on an infinite set E. In fact, in this topology <u>all</u> sets $K \subset E$ are compact: indeed, if $K \subset \bigcup_{\alpha \in A} U_\alpha$, then already the first open set U_{α_0} covers all elements of K except finitely many, and the remaining elements in $K \setminus U_{\alpha_0}$ obviously require only finitely many of remaining U_α's to cover them.

We now discuss basic properties of quasi-compact sets. To begin with, a finite union of quasi-compacts is quasi-compact:

Lemma 6.4. *Let $K_1, K_2 \subset E$ be quasi-compact. Then, $K_1 \cup K_2$ is quasi-compact.*

Proof. This is a triviality: if $K_1 \cup K_2 \subset \bigcup_{\alpha \in A} U_\alpha$, then one can find a finite subcover of K_1 by U_α's, a finite subcover of K_2 by U_α's, and consider the union of the two, which is a finite subcover of $K_1 \cup K_2$. □

Note that, if E is not Hausdorff, it can happen that both quasi-compacts $K_{1,2}$ *are* Hausdorff while their union $K_1 \cup K_2$ is not: e.g., one can consider the subsets $K_{1,2} := [-1,0) \cup 0_{1,2} \cup (0,1]$ of the line with two origins discussed above.

6.1. **(Quasi-)compacts and closed sets.**

Lemma 6.5. *If K is quasi-compact and $K \supset F$ is closed, then F is a quasi-compact. (The same holds for compacts: if K is Hausdorff, then so is F.)*

Proof. Indeed, if $F \subset \bigcup_{\alpha\in A} U_\alpha$, then $K = (K \smallsetminus F) \cup \bigcup_{\alpha\in A}$. As $K \smallsetminus F$ is an open set in K, one can find a finite subcover of K by $K \smallsetminus F$ and U_α's, which gives a finite sub-cover of F by U_α's as required. $\square$

Lemma 6.6. *Let E be a Hausdorff topological space (see the discussion above: one cannot remove this assumption) and $K \subset E$ be compact. Then, K is closed in E.*

Proof. Let $y \in E \smallsetminus K$. As E is Hausdorff, for each $x \in K$ one can find open neighborhoods $x \in U_x^{(y)}$ and $y \in U_y^{(x)}$ such that $U_x^{(y)} \cap U_y^{(x)} = \emptyset$. Clearly, we have $K \subset \bigcup_{x\in K} U_x^{(y)}$ and hence one can find a finite subcover $K \subset \bigcup_{k=1,\dots,n} U_{x_k}^{(y)}$. Therefore,

$$K \cap V_y = \emptyset, \quad \text{where} \quad y \in V_y := \bigcap\nolimits_{k=1,\dots,n} U_y^{(x_k)}$$

is an open neighborhood of y. This proves that $E \smallsetminus K$ is an open set. $\square$

Definition 6.7. *Let E be a Hausdorff space. A set $X \subset E$ is called* relatively compact *(in E) if there exists a compact set $K \subset E$ such that $X \subset K$. Equivalently, X is relatively compact if and only if $\overline{X}$ is compact.*

(Indeed, if $X \subset K$ and K is compact, then K must be closed due to Lemma 6.6, which implies $\overline{X} \subset K$, and hence $\overline{X}$ is compact due to Lemma 6.5.)

6.2. **(Quasi-)compacts and continuous mappings.** The next, straightforward but important, proposition demonstrates that we work with a good definition.

Proposition 6.8. *Let K be quasi-compact and $f : K \to E'$ be a continuous mapping. Then $f(K)$ is also quasi-compact.*

Proof. This is a triviality: if $f(K) \subset \bigcup_{\alpha\in A} U_\alpha$, then $K = \bigcup_{\alpha\in A} f^{-1}(U_\alpha)$ and, since K is compact, one can find a finite subcover in this open (since f is continuous) cover: $K = \bigcup_{k=1,\dots,n} f^{-1}(U_\alpha)$ and hence $f(K) \subset \bigcup_{k=1,\dots,n} U_\alpha$. $\square$

Corollary 6.9. *Let K be quasi compact and $f : K \to \mathbb{R}$ be a continuous function. Then, there exists $x_{\max} \in K$ such that $f(x_{\max}) = \sup_{x\in K} f(x) = \max_{x\in K} f(x)$.*

Proof. This follows form the fact that $f(K)$ is a compact set in $\mathbb{R}$ and hence must be bounded and closed. $\square$

Again, if K is compact (i.e., quasi-compact and Hausdorff) but E' is not Hausdorff, then there is no a priori reason why $f(K)$ should be Hausdorff. However, if we assume that E' is Hausdorff, then the following useful statement holds:

Proposition 6.10. *Let K be compact, E' be Hausdorff, and $f : K \to f(K) \subset E'$ be a continuous bijection. Then, f is a homeomorphism of K and $f(K)$ (i.e., the inverse mapping $f^{-1} : f(K) \to K$ is also continuous).*

Proof. By definition, to check the continuity of f^{-1} it is enough to prove that $f(F) = (f^{-1})^{-1}(F)$ is closed in E' if F is closed in K. This is straightforward: if $F \subset K$ is closed, then it is compact by Lemma 6.5, hence $f(F)$ is compact (Proposition 6.8) and thus closed (Lemma 6.6). □

- Note that Proposition 6.10 implies the following fact: if K is quasi-compact and Hausdorff, then one cannot introduce a coarser topology on K so that it remains Hausdorff. (Indeed, if such a Hausdorff topology $\mathcal{O}' \subsetneq \mathcal{O}$ on K existed, then one could consider the continuous (since $\mathcal{O}' \subset \mathcal{O}$) bijection $f : (K, \mathcal{O}) \to (K, \mathcal{O}')$, $x \mapsto x$, which has to be a homeomorphism.)

6.3. **Products of (quasi-)compact spaces: Tykhonov's theorem.**

The following theorem, due to Tykhonov (or Tychonoff(1906–1993); he introduced and studied the product topology – as defined above – in 1920s), is a fundamental result. Unfortunately, we have no time to prove it in this course unless in the very special case when all the topological spaces in question are metrizable (which also means that A will be restricted to be at most countable); see Theorem 6.14.

Theorem 6.11. *Let $(K_\alpha)_{\alpha\in A}$ be a family of quasi-compact topological spaces. Then (whatever the cardinality of A is), their product $\prod_{\alpha\in A} K_\alpha$ is also a quasi-compact.*

- ◦ Recall that the product of Hausdorff spaces is always Hausdorff. Thus, in the French terminology, the product of compact spaces is also compact.
- ◦ Note that even the product of *two* topological compacts $K_1 \times K_2$ is not that easy to handle as one should find a way to link open covers of $K_1 \times K_2$ by open sets $U_1 \times U_2$ with covers of K_1 and K_2; see TD.
- ◦ In fact, Tykhonov's theorem for quasi-compacts is equivalent to the axiom of choice; in particular, the proof must rely upon this axiom.
- ◦ Curiosity-type remark: however, the particular case of this theorem – products of 'French' compacts are compact – is strictly weaker than the AC.

6.4. **Compact sets in metric spaces.** Let us now discuss compact sets in metric spaces. To begin with, recall that (see TD), given $r > 0$, a subset X of a metric space K is called a *r-net* iff for each $y \in K$ there exists $x \in K$ such that $d(x, y) \le r$.

Lemma 6.12. *Let K be a metric compact and $r > 0$. Then, there exists a finite r-net in K. In particular, K is separable.*

Proof. Trivially, $K \subset \bigcup_{x\in K} B(x, r)$ and hence $K \subset \bigcup_{k=1,\dots,n} B(x_k, r)$ for certain $x_1, \dots, x_n$. By definition, the set $X = \{x_1, \dots, x_n\}$ is a r-net. To see that K is separable, take the union of all 2^{-n}-nets with $n \in \mathbb{N}$. □

The following theorem gives an equivalent definition of metric compacts.

Theorem 6.13. *Let K be a metric space. The following properties are equivalent:*

(i) *K is compact (i.e., satisfies the Borel–Lebesgue property);*

(ii) *for each sequence $(x_n)_{n\in\mathbb{N}}$, $x_n \in K$, there exists a subsequence – i.e., an infinite set $J \subset \mathbb{N}$ – and a point $x \in K$ such that $x_n \to x$ as $J \ni n \to \infty$.*

Moreover, if K is a metric compact, then the following statement (known under the name Lebesgue's lemma) holds:

(iii) *for each open cover $K = \bigcup_{\alpha\in A} U_\alpha$ there exists $r > 0$ (which does not depend on x) such that for each $x \in K$ there exists α_x s.t. $B(x, r) \subset U_{\alpha_x}$.*

October 12, 2020

We start this lecture by proving Theorem 6.13 and, as a corollary, a particular case of the Tykhonov theorem (see Theorem 6.11 and Theorem 6.14) for a (at most) countable product of compact metric spaces.

Proof of Theorem 6.13. (i) $\Rightarrow$ (ii). Assume that a sequence $(x_n)_{n\in\mathbb{N}}$ does not contain a convergent subsequence. This means that each $x\in K$ has an open neighborhood $U_x=B(x,2^{-m})$, where $m=m(x)\in\mathbb{N}$, such that the set

$$J_m=J_m(x):=\{n\in\mathbb{N}:x_n\in B(x,2^{-m})\}$$

is finite. (If this is was not the case and all the sets $J_0\supset J_1\supset J_2\supset\ldots$ were infinite, then one removes first k elements from J_k to guarantee that $\min J_{k+1}>\min J_k$ and constructs an infinite set $J_*:=\{\min J_k, k\in\mathbb{N}\}$; this would imply $x_n\to x$ as $J_*\ni n\to\infty$, a contradiction). We can now use the compactness of K to conclude that $K\subset\bigcup_{x\in K}U_x$ is actually covered by *finitely* many U_x's, each of them containing only finitely many elements of the sequence $(x_n)_{n\in\mathbb{N}}$, a contradiction.

(ii) $\Rightarrow$ (iii). On the contrary, assume that for each $n\in\mathbb{N}$ there exists a point $x_n\in K$ such that $B(x_n,2^{-n})\not\subset U_\alpha$ for all $\alpha\in A$. Due to our assumption (ii), we can then find a convergent subsequence $x_n\to x$ as $J\ni n\to\infty$. Since $x\in U_{\alpha_x}$ for some $\alpha_x\in A$, there exists $m\in\mathbb{N}$ such that $B(x,2^{-m})\subset U_{\alpha_x}$ and, therefore, $B(x_n,2^{-m-1})\subset U_{\alpha_x}$ provided that $n\in J$ is large enough, a contradiction.

(ii) $\Rightarrow$ (i). Given an open cover $K\subset\bigcup_{\alpha\in A}U_\alpha$, find $r>0$ as in (iii). If this open cover did not admit a finite sub-cover, we could inductively construct a sequence of points $x_n\in K$ and open sets U_{α_n} such that $x_n\notin\bigcup_{k=0}^{n-1}U_{\alpha_k}$ and $B(x_n,r)\subset U_{\alpha_n}$ (start with an arbitrary x_0, find α_0, then find x_1, then find α_1, etc). Clearly, the sequence $(x_n)_{n\in\mathbb{N}}$ cannot contain a convergent subsequence since, by construction, we have $d(x_n,x_k)\geq r>0$ for all $n\neq k$. $\square$

Theorem 6.14. *Let $(K_n)_{n\in\mathbb{N}}$ be a countable family of compact metrizable topological spaces. Then their product $\prod_{m\in\mathbb{N}}K_m$ is also a compact metrizable space.*

Proof. Recall that the topology on $K:=\prod_{m\in\mathbb{N}}K_m$ is first-countable (and hence Fréchet–Uryhson) and $x^{(n)}\to x$ in K if and only if $x^{(n)}_m\to x_m$ for each $m\in\mathbb{N}$.

The proof is a standard application of the *'Cantor diagonal process'* argument. To prove that K is compact, we will check the property (ii) from Theorem 6.13. Let $(x^{(n)})_{n\in\mathbb{N}}$ be a sequence in K. As K_0 is compact, there exists an infinite set (subsequence) $J_0\subset\mathbb{N}$ and a point $x^*_0\in K_0$ such that $x^{(n)}_0\to x^*_0$ as $J_0\ni n\to\infty$. Similarly, there exists an infinite set $J_1\subset J_0$ and a point $x^*_1\in K_1$ such that $x^{(n)}_1\to x^*_1$ as $J_1\ni n\to\infty$. Repeating this argument, we obtain infinite sets $J_0\supset J_1\supset\ldots$ and points $x^*_m\in K_m$ such that $x^{(n)}_m\to x^*_m$ as $J_m\ni m\to\infty$.

> Denote by $J^{(m)}_m$ the set J_m without its first m elements and let $J_*:=\{\min J^{(m)}_m,\, m\in\mathbb{N}\}$. Similarly, denote by $J^{(m)}_*$ the set J_* without its first m elements. By construction, for all $m\in\mathbb{N}$, we have $J^{(m)}_*\subset J^{(m)}_m$ and hence $x^{(n)}_m\to x^*_m$ as $J_*\ni n\to\infty$.

We have found a subsequence such that $x^{(n)}\to x^*:=(x^*_m)_{m\in\mathbb{N}}$ as $J_*\ni n\to\infty$. $\square$

6.5. **Locally compact and σ-compact spaces.** Let us now discuss a small extension of the notion of compact spaces; a typical example is an open subset of a compact metric space (see Lemma 6.17). Below we use *English* terminology: compactness=Borel-Hausdorff property (=quasi-compactness in French).

Definition 6.15. *A topological space E is called* locally compact *if for each $x \in \mathbb{E}$ there exists an open neighborhood U_x and a compact set K_x such that $x \in U_x \subset K_x$.*

For Hausdorff (and, in particular, for metric) spaces, this is the same as to require the existence of an open neighborhood U_x such that its closure $\overline{U}_x$ is compact. (Indeed, if K_x is compact, then it is closed; here we rely upon the Hausdorff property. Therefore, $\overline{U}_x \subset K_x$, which implies that $\overline{U}_x$ is also compact.) Moreover, it is enough to consider open balls $U_x = B(x, r)$.

Warning: Spaces like ℓ^p or $C([0,1],\mathbb{R})$ are <u>not</u> locally compact. To see that, recall that (e.g., see Lemma 6.12) compact sets in metric spaces must admit finite r-nets. This is not the case for closed balls $\overline{B}(0,\rho)$ if $r < \frac{1}{2}\rho$, as these balls contain countably many points x_n such that $\|x_n - x_m\| \geq \rho$ for all $n \neq m$.

Definition 6.16. *A topological space is called* σ-compact *(in French, one says* dénobrable à l'infini*) if there exist $E \supset K_n$ – compact such that $E = \bigcup_{n\in\mathbb{N}} K_n$. In other words, E is called σ-compact if it can be covered by countably many compacts.*

A simple example of a locally and σ- compact space is $\mathbb{R}$ or $\mathbb{R}^N$.

Lemma 6.17. *Let K be a metric compact and $K \supset U$ be an open subset of K. Then, U is a locally compact and σ-compact metric space.*

Proof. Denote $\phi(x) := \dot{(}x, K \smallsetminus U)$; note that $x \in U \Leftrightarrow \phi(x) > 0$ (since $K \smallsetminus U$ is a closed set). To prove the local compactness, note that for each $x \in U$ there exists $m \in \mathbb{N}$ such that $2^{-m} < \phi(x)$ and hence $\overline{B}(x, 2^{-m}) \subset U$; the closed ball $\overline{B}(x, 2^{-m}) \subset K$ is compact since K is compact. To prove the σ-compactness of U, consider the closed – and thus compact – sets $K_n := \phi^{-1}([2^{-m}, \infty))$. □

Certainly, one can always additionally require that $K_0 \subset K_1 \subset K_2 \subset$ in Definition 6.16 simply by replacing the sequence K_n by the unions $K'_n := \bigcup_{k=0}^{n} K_k$. If we also require the local compactness of E, then the following stronger property is fulfilled; see also the discussion after Definition 6.20.

Lemma 6.18. *If E is <u>both</u> locally compact and σ-compact, then there exists a sequence of compacts $K_0 \subset K_1 \subset K_2 \subset \ldots \subset E$ such that $E = \bigcup_{n\in\mathbb{N}} K_n$ and that the following holds: for each $E \supset K$ - compact, there exists n such that $K \subset K_n$.*

Proof. Let $E = \bigcup_{n\in\mathbb{N}} K_n$ and $K'_0 := K_0$. We now inductively construct the sequence K'_n as follows: consider the compact(!) set $L_n := K'_n \cup K_{n+1}$; for each point $x \in L_n$ find an open neighborhood U_x and a compact set $K_{n+1}^{(x)}$ such that $x \in U_n^{(x)} \subset K_{n+1}^{(x)}$; using the compactness of L_n find a finite(!) subcover

$$L_n \subset U_n := \bigcup_{k=1}^{m} U_n^{(x_k)}; \quad \text{denote } K'_{n+1} := \bigcup_{k=1}^{m} K_{n+1}^{(x_k)}.$$

By construction, $K'_0 \subset U_0 \subset K'_1 \subset U_1 \subset K'_2 \subset \ldots$ and $E = \bigcup_{n\in\mathbb{N}} K'_n = \bigcup_{n\in\mathbb{N}} U_n$. Since each compact set $K \subset E$ admits a finite subcover by open sets U_n, there exists $n = n(K) \in \mathbb{N}$ such that $K \subset U_n \subset K'_{n+1}$. □

6.6. **Spaces of continuous functions on (locally and σ-) compact sets.** Let us conclude this section by introducing the spaces of continuous functions on compacts (as well as those on locally and σ-compact sets – see Lemma 6.17 for the basic example), which will play a central role in what follows.

Definition 6.19. *Let K be a (topological) compact and E be a metric space. Then,*

$$C(K,E) := \{f : K \to E,\ f \text{ is continuous}\}, \quad d_{C(K,E)}(f,g) := \max_{x\in K} d_E(f(x),g(x)),$$

is a metric space. Moreover, if E is a normed vector space, then so is $C(K,E)$.

[!] Note that we put $\max_{x\in K}$ instead of $\sup_{x\in K}$: the reason is that the function $x \mapsto (f(x),g(x)) \mapsto d_E(f(x),g(x))$ is continuous on K and thus attains its maximal value since k is compact; in particular $d_{C(K,E)}(f,g) < +\infty$.

Definition 6.20. *Let U be a (topological) locally and σ-compact space and E be a metric space. Denote*

$$C(U,E) \ := \ \{f : U \to E,\ f \text{ is continuous}\}.$$

To endow $C(U,E)$ with a metrizable(!) topology, consider a sequence of compacts $K_0 \subset K_1 \subset \ldots \subset U$ from Lemma 6.18 and set

$$d_{C(U,E)}(f,g) \ := \ \max\nolimits_{m\in\mathbb{N}} \min\{d_{C(K_m,E)}(f,g), 2^{-m}\}.$$

Though the definition of the metric on $C(U,E)$ depends on the choice of the exhausting sequence of compacts K_m, the (first-countable) topology on $C(U,E)$ does not rely upon this choice and can be described as follows (see Lemma 6.18):

$$\begin{aligned} f_n \to f \text{ in } C(U,E) \quad &\Longleftrightarrow \quad \forall\, m\in\mathbb{N} \ \ f_n \to f \text{ in } C(K_m,E), \\ &\Longleftrightarrow \quad \forall\, K \subset U \text{ - compact} \ \ f_n \to f \text{ in } C(K,E). \end{aligned}$$

7. Complete metric spaces

The following definition mimics the one from the real analysis.

Definition 7.1. *Let E be a metric space. A sequence $(x_n)_{n\in\mathbb{N}}$, $x_n \in E$, is called a* Cauchy sequence *if for each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that for all $n,m \geq N$ one has $d(x_m,x_m) < \varepsilon$.*

It follows from the triangle inequality that all convergence sequences are Cauchy but not vice versa (e.g., think about $E := \mathbb{Q} \cap [0,1]$.)

Definition 7.2. *A metric space E is called* complete *if for each Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ there exists $x \in E$ such that $x_n \to x$ as $n \to \infty$.*

Warning: The completeness is <u>not</u> a topological notion: e.g., the metric spaces $(0,1)$ and $\mathbb{R}$ are homeomorphic, the latter is complete but the former is not.

Let us introduce more terminology: two metrics d_1 and d_2 on E are called

- *topologic equivalent* if they define the same topology on E;
- *metric equivalent* if $\forall \varepsilon > 0\ \exists \delta > 0$ such that $d_{1,2}(x,y) < \delta \Rightarrow d_{2,1}(x,y) < \varepsilon$.
- *Lipschitz equivalent* if $\exists c, C > 0$ such that $cd_1(x,y) \leq d_2(x,y) \leq Cd_1(x,y)$.

The notions of Cauchy sequences and of completeness are stable under *metric* equivalence. In particular, note that (a) all metrics $d_p(x,y) := \|x-y\|_p$, $p \in [1,+\infty]$, on $\mathbb{R}^N$ are Lipschitz equivalent (with $c = 1/N$, $C = N$) and that (b) e.g., the metrics $\max_{n\in\mathbb{N}} \min\{d_n(x_n,y_n), 2^{-n}\}$ and $\sum_{n\in\mathbb{N}} \min\{d_n(x_n,y_n), 2^{-n}\}$ on $\prod_{n\in\mathbb{N}} E_n$ are not only topologically but also metric (though not Lipschitz) equivalent.

October 14, 2020

We now discuss basic properties of complete metric spaces. It is easy to see that

- If E is a complete metric space and $F \subset E$ is a closed set, then F is also a complete metric space. (Indeed, each Cauchy sequence of elements of F has a limit in E and this limit has to belong to F since F is closed.)
- Vice versa, let $F \subset E$ (we do not make any assumption in E) and F – equipped with the subspace topology – is a complete metric space. Then, F is closed in E. (Indeed, if $F \ni x_n \to x \in E$ as $n \to \infty$, then the sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy and thus has a limit in F, i.e., $x \in F$.)

Lemma 7.3. *Let E be a complete metric space and $F_0 \supset F_1 \supset F_2 \supset \dots$ be a nested sequence of non-empty closed subsets of E s.t.* $\operatorname{diam} F_n := \sup_{x,y\in F_n} d(x,y) \to 0$ *as $n \to \infty$. Then, $\bigcap_{n\in\mathbb{N}} F_n = \{x\}$, for a certain point $x \in E$.*

Proof. Denote $F := \bigcap_{n\in\mathbb{N}} F_n$. Clearly, $\operatorname{diam} F \le \operatorname{diam} F_n$ for all $n \in \mathbb{N}$ and hence $\operatorname{diam} F = 0$. Therefore, F cannot contain more than one point and we only need to prove that $F \ne \emptyset$. To this end, choose a point $x_n \in F_n$. For all $m \ge n$ we have $d(x_n, x_m) \le \operatorname{diam} F_n \to 0$, hence the sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy and thus has a limit: $x_n \to x$ as $n \to \infty$. Since, for each $n \in \mathbb{N}$, the set F_n is closed and $x_m \in F_m \subset F_n$ for all $m \ge n$, we have $x \in F_n$. Therefore, $x \in F$. □

Warning: At first sight, it seems that the assumption $\operatorname{diam} F_n \to 0$ is made only in order to guarantee that the intersection $F := \bigcap_{n\in\mathbb{N}} F_n$ consists of a single point. However, without this assumption it *can* also happen that $F = \emptyset$: for instance, one can consider the closed sets $F_n = [n, +\infty) \subset \mathbb{R} = E$. In general, if F_0 is compact – e.g., if F_0 is a bounded closed set in $\mathbb{R}^n$ – then $F \ne \emptyset$, essentially due to the same argument. However, without the compactness one can have $F = \emptyset$ even if F_0 is bounded: e.g., the sets $F_n = \{e_n, e_{n+1}, e_{n+2}, \dots\} \subset \ell^2$ are closed (e.g., since they do not contain nontrivial convergent sequences) and bounded but $F = \emptyset$.

Proposition 7.4. *A metric space E is compact if and only if*[3] *E is complete and for each $r > 0$ there exists a finite r-net in E.*

Proof. '⇒'. The fact that E admits finite r-nets was discussed in Lemma 6.12, let us prove that E is complete. If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, then – due to the compactness of E – it contains a convergent subsequence: $x_n \to x$ as $J \ni x_n \to x$. This implies the convergence of the *whole* sequence since $(x_n)_{n\in\mathbb{N}}$ is Cauchy.

'⇐' Consider a sequence $(x_n)_{n\in\mathbb{N}}$. Our goal is to find a *Cauchy* subsequence of this sequence. Since E admits a finite 1-net X_1, there exists an infinite set $J_0 \subset \mathbb{N}$ such that $d(x_n, x_m) < 2$ for all $n, m \in J_1$: one simply assigns to each x_n the point from X_1 to which x_n is 1-close and uses the fact that to a certain point from X an infinite number of x_n's is assigned. Similarly – since E has a finite $\frac{1}{2} - net$ – there exists an infinite set $J_1 \subset J_0$ such that $d(x_n, x_m) < 1$ for all $n, m \in J_1$; and so on. Thus, we obtain infinite sets $J_0 \supset J_1 \supset J_2 \supset \dots$ such that $d(x_n, x_m) < 2^{-k+1}$ for all $n, m \in J_k$. Applying the diagonal process argument as above we obtain a Cauchy subsequence $(x_n)_{n\in J_*}$. Since E is complete, this subsequence has a limit. □

[3] This proposition also implies that the Lebesgue condition (iii) in Theorem 6.13 is equivalent to the compactness *provided that E contains only finitely many r-isolated points for each* $r > 0$. E.g., if $X := (x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence that does not have a limit, then one can consider an open cover $E = (E \smallsetminus X) \cup \bigcup_{n\in\mathbb{N}} B(x_n, r_n)$ where $r_n \to 0$ can be chosen so that these balls are disjoint and that for each $k \in \mathbb{N}$ one has $B(x_n, r_n) \subsetneq B(x_n, 2^{-k})$ for large enough $n \ge N(k)$.

- The metric spaces $[0,1]$ and $[0,1]^n$ are complete as they are compact. Moreover, $\mathbb{R}$ and $\mathbb{R}^n$ are also complete since each Cauchy sequence in $\mathbb{R}^n$ is bounded and thus belongs to a cube $[-R,R]^n$, which is a complete space.
- The normed vector spaces ℓ^p, $1 \le p \le +\infty$, are also complete (below we assume that $p < +\infty$, the case $p = +\infty$ is simpler). To see this, consider a Cauchy sequence $x^{(n)} \in \ell^p$, $x^{(n)} = (x^{(n)}_m)_{m\in\mathbb{N}}$ and note that for each $m \in \mathbb{N}$ there exists a limit $x^{(n)}_m \to x^*_m$ as $n \to \infty$ since $(x^{(n)}_m)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$. Moreover, for each $N \in \mathbb{N}$ we have
$$\textstyle\sum_{k=0}^N |x^*_k|^p \ \le\ \limsup_{n\to\infty} \sum_{k=0}^N |x^{(n)}_k|^p \ \le\ \limsup_{n\to\infty} \|x^{(n)}\|_p^p \ <\ +\infty,$$
which implies (by passing to the limit $N \to \infty$) that $x^* \in \ell^p$ and
$$\textstyle\sum_{k=0}^N |x^{(n)}_k - x^*_k|^p \ \le\ \limsup_{m\to\infty} \sum_{k=0}^N |x^{(n)}_k - x^{(m)}_k|^p \tag{7.1}$$
$$\le \limsup_{m\to\infty} \|x^{(n)} - x^{(m)}\|_p^p, \tag{7.2}$$
which implies (by sending $N \to \infty$ and then $n \to \infty$) that $x^{(n)} \to x^*$ in ℓ^p.

Another example is given the spaces $C(K,E)$ and $C(U,E)$ of E-valued continuous functions on a topological compact K or a locally and σ-compact space U, which were introduced in the previous lecture. Recall that $f_n \to f$ in $C(K,E)$ if and only if $f_n(x) \to f(x)$ *uniformly* in $x \in K$, while $f_n \to f$ in $C(U,E)$ if and only if $f_n \to f$ in $C(K,E)$ for each compact $K \subset U$.

Theorem 7.5. *(i) Let K be a (topological) compact and E be a complete metric space. Then, the metric space $C(K,E)$ is complete.*

(ii) Let U be locally compact and σ-compact, and E be a complete metric space. Then, the metrizable space $C(U,E)$ is complete.

Proof. (i) Let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $C(K,E)$. For each $x \in K$, we have $d_E(f_n(x), f_m(x)) \le d_{C(E,K)}(f_n, f_m)$, so the sequence $(f_n(x))_{n\in\mathbb{N}}$ is Cauchy in E and thus has a limit, which we denote by $f(x)$. Clearly,
$$\textstyle\sup_{x\in K} d_E(f_n(x), f(x)) \le \limsup_{m\to\infty} d_{C(E,K)}(f_n, f_m) \to 0 \ \text{ as } \ n \to \infty.$$
However, it remains(!) to check that $f : K \to E$ is a continuous function. This follows from the continuity of f_n and the *uniform* convergence $f_n \to f$. Indeed, given $x \in K$ and $\varepsilon > 0$ one can find $n \in \mathbb{N}$ such that $d_{C(E,K)}(f_n, f) < \frac{1}{3}\varepsilon$ and then – using the continuity of f_n – an open neighborhood $U_x \ni x$ such that $d_E(f(y), f(x)) < \frac{1}{3}\varepsilon$ for $y \in U_x$. By the triangle inequality, this implies $d_E(f(y), f(x)) < \varepsilon$ for $y \in U_x$.

(ii) The same reasoning applies: the pointwise limits $f(x)$ exist and the limit function f is continuous on U. Moreover, by (i), for each compact set $K \subset U$ the functions f_n converge to f in $C(K,E)$, which implies that $f_n \to f$ in $C(U,E)$. □

Quasi-détour. Let be a bounded[4] metric space. Denote by $\mathcal{F}(E)$ the set of all its non-empty closed subsets and by $\mathcal{K}(E) \subset \mathcal{F}(E)$ the set of all non-empty compact subsets; both equipped with the Hausdorff distance $d_{\mathrm{H}}(F_1, F_2)$, defined as follows:
$$d_{\mathrm{H}}(F_1, F_2) < r \quad \Leftrightarrow \quad \text{for each } x \in F_{1,2} \text{ there exist } y \in F_{2,1} \text{ such that } d(x,y) < r.$$
Note that $\mathcal{F}(E) = \mathcal{K}(E)$ if E is compact.

[4] In other words, we assume that $\operatorname{diam} E < +\infty$. This restriction is irrelevant: if needed, one can always replace a given metric $d(x,y)$ by $\min\{d(x,y), 1\}$, which is metrically equivalent to d.

Theorem 7.6. *(i) Let E be a complete metric space. Then, both $\mathcal{F}(E)$ and $\mathcal{K}(E)$ are complete metric spaces. (ii) Moreover, if E is compact, then so is $\mathcal{K}(E)$.*

Proof.[5] Let us first prove the implication (i) $\Rightarrow$ (ii). According to Proposition 7.4, it is enough to check that, if E has a finite r-net for each $r > 0$, then so does $\mathcal{K}(E)$. This is straightforward: if $X = \{x_1, \ldots, x_n\}$ is a r-net for E, then the set $2^X \smallsetminus \{\emptyset\}$ of all non-empty subsets of X is a r-net for $\mathcal{K}(E)$.

(i). Let E be complete and $(F_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{F}(E)$; our goal is to find a subsequence (i.e., an infinite set $J_* \subset \mathbb{N}$ in our notation) and a closed set $F_* \subset E$ such that $F_n \to F_*$ as $J_* \ni n \to \infty$. In addition, if all F_n are compact, then we should also prove that F_* is compact too. By passing to a subsequence, we can assume that $d_{\mathrm{H}}(F_n, F_m) < 2^{-n}$ for all $m \geq n$. Denote

$$F_* \ := \ \{x_* \in E : \text{ there exist } x_m \in F_m \text{ such that } x_m \to x_* \text{ as } m \to \infty\}.$$

It is not hard to see that:

- For each $x_* \in F_*$ and $n \in \mathbb{N}$ there exists $x_n \in F_n$ s.t. $d(x_n, x_*) < 2^{-n+1}$. Indeed, if $x_* = \lim_{m\to\infty} x_m$ with $x_m \in F_m$, then one can choose m large enough so that $d(x_m, x_*) < 2^{-n}$ and then find a point $x_n \in F_n$ such that $d(x_n, x_m) < 2^{-n}$ using the fact that $d_{\mathrm{H}}(F_n, F_m) < 2^{-n}$.
- For each $x_n \in F_n$ there exists $x_* \in F_*$ such that $d(x_n, x_*) < 2^{-n+1}$. Indeed, given $x_n \in F_n$ one can find $x_{n+1} \in F_{n+1}$ s.t. $d(x_n, x_{n+1}) < 2^{-n}$, then $x_{n+2} \in F_{n+2}$ such that $d(x_{n+1}, x_{n+2}) < 2^{-n-1}$ and so on. It follows from the triangle inequality that the sequence $(x_{n+m})_{m\in\mathbb{N}}$ is Cauchy and therefore there exists a limit $x_{n+m} \to x_*$ as $m \to \infty$. By definition, $x_* \in F_*$ and we have $d(x_n, x_*) < 2^{-n} + 2^{-n-1} + 2^{-n-2} + \ldots = 2^{-n+1}$.

Therefore, $d_{\mathrm{H}}(F_n, F_*) \leq 2^{-n+1} \to 0$ and we are almost done: e.g., to prove that $\mathcal{F}(E)$ is complete it remains to check that F_* is closed. This is straightforward: let $x_*^{(n)} \in F_*$ and $x_*^{(n)} \to x_*$ as $n \to \infty$. As explained above, for each $n \in \mathbb{N}$ we can find $x_n^{(n)} \in F_n$ such that $d(x_*^{(n)}, x_n^{(n)}) \leq 2^{-n}$. Then, $x_n^{(n)} \to x_*$ and hence $x_* \in F_*$.

- Finally, let us prove that F_* is compact if all F_n are compact.

This is done as follows:

- Given a sequence $x_*^{(n)} \in F_*$, consider $x_m^{(n)} \in F_m$ s.t. $d(x_m^{(n)}, x_*^{(n)}) \leq 2^{-m}$; the existence of such approximations follows from the preceding discussion.
- Use the compactness of each of F_m and the diagonal process to find $J_* \subset \mathbb{N}$ and $x_m^* \in F_m$ such that, for each $m \in \mathbb{N}$, $x_m^{(n)} \to x_m^*$ as $J_* \ni n \to \infty$.
- Note that the sequence $(x_m^*)_{m\in\mathbb{N}}$ is Cauchy: for all $k \geq m$ we have
$$d(x_m^{(n)}, x_k^{(n)}) \leq 2^{-m} + 2^{-k} \leq 2^{-m+1} \quad \Rightarrow \quad d(x_m^*, x_k^*) \leq 2^{-m+1}.$$
- Since E is a complete space, there exists a limit $x_m^* \to x_*^*$ as $m \to \infty$. Note that $d(x_m^*, x_*^*) \leq 2^{-m+1}$. In particular, $x_*^* \in F_*$.
- Thus, it remains to check that $x_*^{(n)} \to x_*^*$ as $J_* \ni n \to \infty$, i.e., that we can exchange the limits in m and in n. This follows from the fact that the convergence $x_m^{(n)} \to x_*^{(n)}$ is *uniform* in n. (Given $\varepsilon > 0$, find $m \in \mathbb{N}$ such that $2^{-m+2} < \varepsilon$ and note that $d(x_*^{(n)}, x_*^*) \leq d(x_m^{(n)}, x_m^*) + 2^{-m} + 2^{-m+1}$; then use the fact that $d(x_m^{(n)}, x_m^*) \leq 2^{-m}$ if $n \in J_*$ is large enough.) $\square$

[5]This proof was skipped during the lecture in order to save time. Also, note that the theorem is extended: for non-compact E, *two* metric spaces $\mathcal{K}(E) \subsetneq \mathcal{F}(E)$ are considered.

The next standard fact to discuss is the *Banach–Caccioppoli fixed point theorem.*

Theorem 7.7. *Let E be a complete metric space, $q < 1$, and $f : E \to E$ be a q-contraction, i.e., $d(f(x), f(y)) \le q \cdot d(x, y)$ for all $x, y \in E$. Then, there exists a unique point $x_* \in E$ such that $f(x_*) = x_*$. Moreover, for each $x_0 \in E$ the following holds: if we define inductively $x_{n+1} := f(x_n)$ for $n \in \mathbb{N}$, then $x_n \to x_*$ as $n \to \infty$.*

Proof. The uniqueness of x_* is trivial (otherwise, we would have $d(x'_*, x''_*) \le q \cdot d(x'_*, x''_*)$, a contradiction). To prove the existence, note that, for each starting point $x_0 \in E$, the sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy since

$$d(x_n, x_m) \;\le\; (q^n + q^{n+1} + \ldots + q^{m-1}) \cdot d(x_0, x_1) \;\le\; q^n(1-q)^{-1} \cdot d(x_0, x_1)$$

for all $m \ge n$. As E is complete, it has a limit x_* and one easily sees that $x_* = f(x_*)$ by passing to the limit $n \to \infty$ in the equation $x_{n+1} = f(x_n)$. $\square$

Example. Consider the metric space $\mathcal{K}([0,1])$ of all non-empty compact subsets of $[0,1]$; due to Theorem 7.6, this space is compact and hence complete. Let the mapping $f : \mathcal{K}([0,1]) \to \mathcal{K}([0,1])$ be defined as follows: $K \mapsto \frac{1}{3}K \cup (\frac{1}{3}K + \frac{2}{3})$; one can easily see that f is a $\frac{1}{3}$-contraction. Then, the '$\frac{1}{3}$-Cantor' set C can be defined as the unique non-empty compact subset of $[0,1]$ such that $f(\mathrm{C}) = \mathrm{C}$.

This construction of C admits a straightforward generalization: given a family $f_1, \ldots, f_n$ of q-contractions of a complete metric space E, one can define a q-contraction $f : K \mapsto f_1(K) \cup f_2(K) \cup \ldots f_n(K)$ on the complete metric space $\mathcal{K}(E)$ and define a compact $\mathrm{C}_{f_1,\ldots,f_n}$ as the unique fixed point of f.

E.g., the *Koch snowflake* and similar fractal sets can be defined in this manner.

Détour. Fixed point theorems – there are dozens of them – provide a useful tool of proving the existence of an object satisfying certain conditions. Note that Theorem 7.7 is a *metric* statement relying upon very strong assumptions on the mapping f: this is why its proof is so straightforward. There are also (much deeper) purely *topological* statements of that kind, the simplest of them is the following:

Theorem 7.8 (Brouwer, ∼1910)**.** *Let $\overline{B}^n$ denote the closed ball $\overline{B}(0,1) \subset \mathbb{R}^n$. Each continuous mapping $f : \overline{B}^n \to \overline{B}^n$ has a fixed point (i.e., $\exists x \in \overline{B}^n : f(x) = x$).*

This theorem first appeared in the beginning of the 20th century – the early time of the development of the topology as a subject – and can be proven by several techniques, including the combinatorial proof via the *Sperner lemma*. A conceptual generalization of this result is the *Lefschetz fixed point theorem*, in which the space $\overline{B}^n$ is replaced by a compact (triangulable) manifold X with boundary. (Clearly, not all X and f work: e.g., there are continuous mappings from a n-dimensional sphere or a torus to itself that do not have fixed points.) It turns out that there exists a numerical characteristics Λ_f, called the *Lefschetz number* of a continuous mapping $f : X \to X$, such that the condition $\Lambda_f \ne 0$ implies the existence of a fixed point of f. If $X = \overline{B}^n$, then $\Lambda_f = 1$ for all f.

Unfortunately, this material goes far beyond the scope of these lectures.

October 19, 2020

7.1. **Completion of a metric space.** If a given metric space E is not complete, then there is a canonical procedure to embed it into a complete metric space $E' \supset E$ such that E is dense in E', the latter space is called the *completion* of E.

This procedure works as follows:

- To start with, consider the (huge) set $\widetilde{E}$ of all Cauchy sequences in E.
- Introduce an equivalence relation on this set: $(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}}$ if for each $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, y_n) < \varepsilon$ for all $n \geq N$. (Exercise: provided that both sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are Cauchy – this is equivalent to saying that the sequence $(x_0, y_0, x_1, y_1, x_2, y_2, \ldots)$ is Cauchy.)
- Denote $E' := \widetilde{E}/\sim$ and introduce a metric on E' by
$$d'((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) \ := \ \lim_{n\to\infty} d(x_n, y_n).$$
(Exercise: (a) if both $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are Cauchy, then the sequence $d(x_n, y_n)$ is also Cauchy; (b) d' is a well defined metric on E'.)
- To each point $x \in E$ one can associate a trivial Cauchy sequence $x_n = x$ for all n, this is why E can be viewed as a dense subspace of E'.
- Finally, to see that E' is complete, assume that $(x^{(n)})_{n\in\mathbb{N}}$, $x^{(n)} = (x^{(n)}_m)_{m\in\mathbb{N}}$ is a Cauchy sequence of (equivalence classes of) Cauchy sequences in E. Our goal is to find a convergent (in E') subsequence in $x^{(n)}$. Passing to a subsequence, we can assume that $d'(x^{(n)}, x^{(n+1)}) < 2^{-n}$ for all $n \in \mathbb{N}$. Using the definition of d' and the fact that each sequence $(x^{(n)}_m)_{m\in\mathbb{N}}$ is Cauchy in E, we can find an increasing sequence $m_{-1} := 0 < m_0 < m_1 < m_2 < \ldots$ such that
$$d(x^{(n)}_p, x^{(n+1)}_p) \leq 2^{-n} \quad \text{and} \quad d(x^{(n)}_p, x^{(n)}_q) \leq 2^{-n} \ \text{ for all } \ q \geq p \geq m_n.$$
Now define the sequence $(x^*_m)_{m\in\mathbb{N}}$ as follows: $x^*_p := x^{(n)}_p$ if $m_{n-1} \leq p < m_n$. It is easy to see that
$$d(x^*_p, x^{(n)}_p) \leq 2^{-n+1} \quad \text{and} \quad d(x^*_p, x^*_q) \leq 5\cdot 2^{-n} \ \text{ for all } \ q \geq p \geq m_n.$$
Therefore, $(x^*_m)_{m\in\mathbb{N}}$ is a Cauchy sequence in E and $x^{(n)} \to x^*$ in E'.

Remark: Let $E \subset E''$ be a complete metric space such that E is dense in E''. Then, E'' is isometric to the completion E' of E constructed above.

Proof. Indeed, each Cauchy sequence in E must have a limit in E''; it is straightforward to check that this defines an isometric inclusion $\iota : E' \to E''$. On the other hand, since E is dense in E'', each point of E'' should appear as a limit (in E'') of points from E. Thus, $\iota(E') = E''$. □

Lemma 7.9. *Let E_1 be a complete metric space and $f : E \to E_1$ be a uniformly continuous function. Then, there exists a unique continuous function $f' : E' \to E_1$ such that $f'|_E = f$, where E' denotes the completion of E.*

Proof. This is straightforward: if $x = (x_n)_{n\in\mathbb{N}} \in \widetilde{E}$ is a Cauchy sequence in E, then the uniform continuity of f implies that the sequence $(f(x_n))_{n\in\mathbb{N}}$ is Cauchy in E_1 and thus has a limit $\widetilde{f}(x)$. This allows to correctly(!) define f' on $E' = \widetilde{E}/\sim$ by $f'([x]) := f(x)$; the continuity of f' follows from the uniform continuity of f. □

Let us now discuss several examples of this construction:

- The basic example of the completion of a metric space is $E = (\mathbb{Q}, |\cdot|)$ and $E' = \mathbb{R}$. One can also consider other metrics on $\mathbb{Q}$, notably those provided by p-adic absolute values $|\cdot|_p$. The completion of the space $(\mathbb{Q}, |\cdot|_p)$ is known under the name *p-adic numbers* $\mathbb{Q}_p$; the elements of $\mathbb{Q}_p$ can be identified with formal series $\sum_{n=n_0}^{+\infty} a_n p^n$, where $n_0 \in \mathbb{Z}$ and $a_n \in \{0, \ldots, p-1\}$.
- Another example is the set ℓ_0 of all *finite* sequences equipped with one of the norms $\|\cdot\|_p$, $1 \le p < \infty$. Completing it with respect to $\|\cdot\|_p$ one gets pairwise different complete spaces $E' = \ell^p$. (Indeed, each sequence $(x_n)_{n\in\mathbb{N}} \in \ell^p$ can be approximated by finite sequences $(x_0, \ldots, x_N, 0, 0, \ldots)$ in the metric $\|\cdot\|_p$. Therefore, ℓ_0 is dense in ℓ^p for each $p \in [1, +\infty)$.)
- However, note that the completion of ℓ_0 with respect to the metric $\|\cdot\|_\infty$ is not ℓ^∞ since ℓ_0 is not dense in ℓ^∞. In fact, this completion is
$$\ell_0^\infty := \{(x_n)_{n\in\mathbb{N}} : x_n \to 0 \text{ as } n \to \infty\} \subsetneq \ell^\infty;$$
(Exercise: check that the normed vector space $(\ell_0^\infty, \|\cdot\|_\infty)$ is complete.)
- Similarly, if one starts with the *set* $C([0,1], \mathbb{R})$ and complete it with respect to the norms $\|\cdot\|_p$, $1 \le p < +\infty$, then the family of complete(!) $L^p([0,1])$ spaces arise; see *'Intégration et proba'*. (Also, note that the space $C([0,1], \mathbb{R})$ equipped with the norm $\|\cdot\|_\infty$ is already complete 'as is'.)
- Finally, consider a complete metric space E and the set of its *finite* non-empty subsets $\mathcal{K}_0(E)$. Exercise: the completion of $(\mathcal{K}_0(E); d_{\mathrm{H}})$ is $\mathcal{K}(E)$ – the set of all non-empty *compact* subsets of E; see Theorem 7.6.

7.2. **Baire's theorem.** Recall that a subset X of a topological space E is called

- *nowhere dense* if $\operatorname{Int} \overline{X} = \emptyset$ or, equivalently, $U \not\subset \overline{X}$ for each open set $U \neq \emptyset$. This property can be further reformulated as follows: for each open set $U \neq \emptyset$ there exists an open subset $U \supset V \neq \emptyset$ such that $V \cap X = \emptyset$. (Indeed, if X is nowhere dense, then one can take $V := U \cap (E \smallsetminus \overline{X})$. Vice versa, the condition $V \subset E \smallsetminus X$ implies that $V \subset E \smallsetminus \overline{X}$ and thus $U \not\subset \overline{X}$.)
- *(everywhere) dense* if $\overline{X} = E$, i.e., if $U \cap X \neq \emptyset$ for each open set $U \neq \emptyset$.

It is easy to see that the following conditions are equivalent:

$U = E \smallsetminus F$ is open everywhere dense $\iff$ $F = E \smallsetminus U$ is closed nowhere dense.

Definition 7.10. *A set $X \subset E$ is called* maigre *if there exists a (at most) countable collection of nowhere dense sets $(F_n)_{n\in\mathbb{N}}$ such that $X \subset \bigcup_{n\in\mathbb{N}} F_n$. (The traditional English terminology is* 'first category set'. *However, the name* 'meagre' *also exists.)*

Clearly, since the closure of a nowhere dense set is also nowhere dense, in the above definition one can assume that all F_n are closed, without loss of generality.

Theorem 7.11 (Baire)**.** *Let E be a complete metric space and $X \subset E$ be a maigre set. Then, the complement $E \smallsetminus X$ is everywhere dense in E. In particular, $X \neq E$.*

Proof. Let $X \subset \bigcup_{n\in\mathbb{N}} F_n$, where F_n are nowhere dense sets. It is enough to prove that the set $E \smallsetminus \bigcup_{n\in\mathbb{N}} F_n$ is everywhere dense.

Let $U \subset E$ be an open set. Since F_0 is nowhere dense we can find an open ball $B_0 := B(x_0, r_0) \subset U$ such that $B_0 \cap F_0 = \emptyset$. Similarly, since F_1 is nowhere dense we can find an open ball $B_1 := B(x_1, r_1) \subset B(x_0, \frac{1}{2}r_0)$ such that $B_1 \cap F_1 = \emptyset$. Iterating this construction, we obtain a sequence of open balls

$$B_n = B(x_n, r_n) \subset B(x_{n-1}, \tfrac{1}{2}r_{n-1}) \ \text{ such that } \ B_n \cap F_n = \emptyset.$$

Since $\overline{B}_n \subset \overline{B}(x_{n-1}, \frac{1}{2}r_{n-1}) \subset B_{n-1}$ and $r_n \to 0$ as $n \to \infty$, it is easy to see that there exists a point $x \in E$ such that $\bigcap_{n\in\mathbb{N}} B_n = \{x\} \neq \emptyset$; see Lemma 7.3. By construction, $x \notin F_n$ for all $n \in \mathbb{N}$ and hence $V \cap (E \smallsetminus \bigcup_{n\in\mathbb{N}} F_n) \neq \emptyset$. □

It is tempting to say, at least informally, that a 'typical' element of a complete metric space cannot belong to a maigre set. However, as it is shown by the following example, one should be extremely careful when using such an informal terminology.

Illustration: Liouville numbers. Denote

$$L := \bigcap\nolimits_{m\in\mathbb{N},\, n\in\mathbb{N}} U_n^{(m)}, \qquad U_n^{(m)} := \bigcup\nolimits_{q>n,\, p\in\mathbb{Z}} \left(\left(\frac{p}{q} - \frac{1}{q^m}, \frac{p}{q} + \frac{1}{q^m} \right) \smallsetminus \left\{ \frac{p}{q} \right\} \right).$$

In other words, x is a *Liouville number* (i.e., $x \in L$) iff for each $m \in \mathbb{N}$ there exists infinitely many denominators $q \in \mathbb{N}$ such that $|x - \frac{p}{q}| < \frac{1}{q^m}$ for a certain $p = p(q)$.

- It is easy to see that $\mathbb{R} \smallsetminus L = \bigcup_{m\in\mathbb{N}} \bigcup_{n\in\mathbb{N}} F_n^{(m)}$, where $F_n^{(m)} := \mathbb{R} \smallsetminus U_n^{(m)}$, is *maigre*: indeed, each $U_n^{(m)}$ is an open everywhere dense set.
- On the other hand, L is a set of Lebesgue measure 0: if $m \geq 3$, we have
$$\mathrm{Leb}(U_n^{(m)} \cap [0,1]) \ \leq \ 2\sum\nolimits_{q>n} q^{1-m} \ = \ O(n^{2-m}) \ \to \ 0 \ \text{ as } \ n \to \infty.$$

To summarize, $\mathbb{R} = L \cup (\mathbb{R} \smallsetminus L)$ is a splitting of $\mathbb{R}$ to a set of Lebesgue measure zero and a maigre set. Each of them is 'untypical' but from very different viewpoints.

Détour[6]**.** The name 'Liouville numbers' originates from the following observation:

> if $x \in \mathbb{R}$ is a non-rational algebraic number (i.e., if there exists a polynomial $P \in \mathbb{Z}[x]$ such that $P(x) = 0$), then
> $$\left| x - \frac{p}{q} \right| \geq \frac{c(x)}{q^m}, \quad \text{for all } \ \frac{p}{q} \in \mathbb{Q}, \ \text{ where } \ c(x) > 0 \ \text{ and } \ m = \deg P.$$
> For a proof, note that $|P(\frac{p}{q})| \geq \frac{1}{q^m}$, this gives $c(x) \approx |P'(x)|^{-1}$. (Replacing P by P' or $P^{(k)}$, one can always assume that $P'(x) \neq 0$.)

In particular, the set L does not contain algebraic numbers; this allowed Joseph Liouville to give a first explicit example $\sum_{n\in\mathbb{N}} 10^{-n!} \in L$ of such a number in 1844.

In fact, a much deeper statement holds true:

Theorem 7.12 (Klaus Roth, 1955)**.** *For each non-rational algebraic number $x \in \mathbb{R}$ and each $\varepsilon > 0$ there exists a constant $c(x, \varepsilon) > 0$ such that*

$$\left| x - \frac{p}{q} \right| \ \geq \ \frac{c(x,\varepsilon)}{q^{2+\varepsilon}} \quad \textit{for all } \ \frac{p}{q} \in \mathbb{Q}.$$

(Theorem 7.12 is also referred to as the Thue–Siegel–Roth theorem to acknowledge the preceding work of Axel Thue and that of Carl Ludwig Siegel.) Note that for *each* $x \in \mathbb{R} \smallsetminus \mathbb{Q}$ there are infinitely many $\frac{p}{q} \in \mathbb{Q}$ such that $|x - \frac{p}{q}| < \frac{1}{q^2}$: such approximations are given by truncations of the continued fraction expansion of x.

---

[6]This material was not mentioned during the lecture.

A typical use of the Baire theorem is to give a 'cheap' prove of the existence of certain objects. Let us illustrate this by the following claim:

- The set $\big\{f \in C([0,1],\mathbb{R}) : \exists x \in [0,1] \text{ s.t. } f \text{ is differentiable at } x\big\}$ is maigre. Therefore, the set of *nowhere* differentiable functions is dense in $C([0,1],\mathbb{R})$.

Proof.[7] Let us consider the sets

$$F_m := \{f \in C([0,1],\mathbb{R}) : \ \exists x \in [0,1] \text{ s.t. } \forall y \in [0,1] \ |f(y)-f(x)| \le m|x-y|\}.$$

If f is differentiable at a point x, then the function $y \mapsto |f(y)-f(x)|/|y-x|$ is continuous and thus $f \in F_m$ provided that $m \ge \max_{y\in[0,1]}|f(y)-f(x)|/|y-x|$.

It is not hard to see that F_m is nowhere dense. Indeed, let us first prove that F_m is a closed set. Given a sequence $f^{(n)} \in F_n$ such that $f^{(n)} \to f$ in $C([0,1],\mathbb{R})$, let $x^{(n)} \in [0,1]$ be the corresponding points from the definition of the set F_m. By compactness, we can find a subsequence such that $x^{(n)} \to x^* \in [0,1]$ as $J \ni n \to \infty$. Now note that

$$\begin{aligned} |f^{(n)}(y) - f^{(n)}(x^*)| \ &\le \ |f^{(n)}(y) - f^{(n)}(x^{(n)})| + |f^{(n)}(x^*) - f^{(n)}(x^{(n)})| \\ &\le \ m \cdot (|y - x^{(n)}| + |x^* - x^{(n)}|) \ \to \ m|y - x^*| \end{aligned}$$

as $J \ni n \to \infty$. Therefore, $|f(y) - f(x^*)| \le m|y-x^*|$ and hence $f \in F_m$.

To prove that F_m is nowhere dense, it remains to note that each open set $U \subset C([0,1],\mathbb{R})$ contains a piece-wise linear function g. Let $B(g,\rho) \subset U$ and denote by $M < +\infty$ the maximal gradient of g. Further, let $\phi : \mathbb{R} \to [-\frac{1}{2},\frac{1}{2}]$ be defined as

$$\phi(x) = \textstyle\int_0^x \operatorname{sign}(t - \lfloor t \rfloor - \frac{1}{2})dt.$$

Then, the piece-wise linear function $\widetilde{g}(x) := g(x) + \rho\phi((M+m+1)x/\rho)$ cannot belong to F_m since, at each point on $[0,1]$, its gradient is at least $m+1$ (or at most $-m-1$). Still, we have $\widetilde{g} \in B(g,\rho) \subset U$; hence, $U \not\subset F_m = \overline{F}_m$. $\square$

October 21, 2020

8. Connected and path-connected topological spaces

Definition 8.1. *A topological space E is called*

- connected *if there exists no non-trivial pair of open sets $\emptyset \ne U_0, U_1 \subset E$ such that $U_0 \cap U_1 = \emptyset$ and $U_0 \cup U_1 = E$;*
- path-connected *if for each pair of points $x_0, x_1 \in E$ theere exists a continuous mapping $\gamma : [0,1] \to E$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. (Informally, one says that x_0 and x_1 can be joined by a continuous path γ.)*

As usual, a subset $X \subset E$ is called connected/path-connected iff it is a connected/path-connected topological space (with the subspace topology inherited from E).

Lemma 8.2. *Let $f : E \to E_1$ be a continuous mapping. (i) If E is connected, then $f(E)$ is connected. (ii) If E is path-connected, then $f(E)$ is path-connected.*

[7] The proof was skipped during the lecture due to the lack of time.

Proof. Assume that there exists two open (in the subspace topology of $f(E) \subset E_1$) sets $\emptyset \neq V_{0,1} = f(E) \setminus V_{1,0}$. Since the mapping f is continuous, their preimages $U_{0,1} := f^{-1}(V_{0,1})$ are open in E and satisfy the same conditions $\emptyset \neq U_{0,1} = E \setminus U_{1,0}$, a contradiction with the connectedness of E.

(ii) Let $y_0, y_1 \in f(E)$ and choose $x_{0,1} \in E$ such that $y_{0,1} = f(x_{0,1})$. Since E is path-connected, there exists a continuous mapping $\gamma : [0,1] \to E$ such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. Then, the mapping $f \circ \gamma : [0,1] \to f(E)$ is also continuous and satisfies $(f \circ \gamma)(0) = y_0$ and $(f \circ \gamma)(1) = y_1$. □

Lemma 8.3. *(i) The segment* $[0,1]$ *is a connected topological space.*
(ii) If a topological space E *is path-connected, then it is connected.*

Proof. (i) Let $U_0 \subset [0,1]$ be an open *and* closed set such that $0 \in U$ and let $x^* := \sup\{x \in [0,1] : [0,x] \subset U_0\}$. Since $[0,1] \setminus U_0$ is open, we must have $x^* \in U_0$. If $x^* < 1$, one easily obtains a contradiction with the fact that U_0 is open.

(ii) On the contrary, assume that $\emptyset \neq U_{0,1} = E \setminus U_{1,0}$ are both open and closed sets. Let $x_{0,1} \in U_{0,1}$ and $\gamma : [0,1] \to E$ be a continuous mapping such that $\gamma(0) = x_0$ and $\gamma(1) = x_1$. It is easy to see that $\gamma([0,1])$ is also not connected (indeed, consider the sets $U_{0,1} \cap \gamma([0,1])$), which contradicts to (i) and Lemma 8.2(i). □

Lemma 8.4. *(i) Let* $E_\alpha \subset E$ *be a connected set, for each* $\alpha \in A$. *If* $\bigcap_{\alpha \in A} E_\alpha \neq \emptyset$, *then* $\bigcup_{\alpha \in A} E_\alpha$ *is connected. (ii) The same claim holds for path-connected sets.*

Proof. (i) Let $x \in \bigcap_{\alpha \in A} E_\alpha$ and assume that non-empty disjoint open sets $U_{0,1} \subset E$ are such that $\bigcup_{\alpha \in A} E_\alpha \subset U_0 \cup U_1$. Without loss of generality, assume that $x \in U_0$. Since each E_α is connected, we should have $E_\alpha \subset U_0$. Hence, $\bigcup_{\alpha \in A} E_\alpha \subset U_0$.

(ii) Let $x_{0,1} \in E_{\alpha_{0,1}}$ and $x \in E_{\alpha_0} \cap E_{\alpha_1}$. Concatenating the paths from x_0 to x (running in E_{α_0}) and from x to x_1 (running in E_{α_1}) one obtains a required path from x_0 to x_1. More pedantically, given continuous mappings $\gamma_{0,1} : [0,1] \to E_{\alpha_{0,1}}$ such that $\gamma_{0,1}(0) = x_{0,1}$ and $\gamma_{0,1}(1) = x$, one defines

$$\gamma(t) := \gamma_0(2t) \ \text{ if } \ t \le \tfrac{1}{2} \quad \text{and} \quad \gamma(t) := \gamma_1(2-2t) \ \text{ if } \ t \ge \tfrac{1}{2}. \qquad \square$$

Lemma 8.5. *(i) Let* E_α *be a connected topological space, for each* $\alpha \in A$. *Then, the space* $\prod_{\alpha \in A} E_\alpha$ *is connected. (ii) The same claim holds for path-connected spaces.*

Proof. (i) Consider first the case of a two-element set A, e.g., $A = \{\mathrm{x}, \mathrm{y}\}$. For each $x_0 \in E_x$ and $y_0 \in E_y$ the 'cross-like' set $C_{(x_0,y_0)} := \{(x,y) : x = x_0 \text{ or } y = y_0\}$ is connected due to Lemma 8.5(i). Therefore, $E_\mathrm{x} \times E_\mathrm{y} = \bigcup_{y \in E_y} C_{(x_0,y)}$ is also connected due to the same lemma.

The case of a *finite* set A easily follows by induction. Assume now that A is infinite and that $E := \prod_{\alpha \in A} E_\alpha = U_0 \cup U_1$, where $U_{0,1}$ are non-empty disjoint open sets. By definition of the product topology, each of the sets $U_{0,1}$ contains a base set of the form (5.1). Therefore, one can find points $x_{0,1} \in U_{0,1}$ such that the set $A_0 := \{\alpha : (x_0)_\alpha \neq (x_1)_\alpha\}$ is finite. From the above consideration it follows that the 'hyperplane' set $H := \{x \in E : x_\alpha = (x_0)_\alpha \text{ for all } \alpha \notin A_0\}$ is connected: indeed, it is homeomorphic to a product of connected spaces over the finite set of indices A_0. On the other hand, both sets $U_{0,1} \cap H \neq \emptyset$ are open in H, a contradiction.

(ii) This part is straightforward. Let $x^{(0,1)} \in E$. Since each of the spaces E_α is path-connected, there exist continuous mappings $\gamma_\alpha : [0,1] \to E_\alpha$ with $\gamma_\alpha(0) = x^{(0)}_\alpha$ and $\gamma_\alpha(1) = x^{(1)}_\alpha$. It is easy to see that the mapping $\gamma = (\gamma_\alpha)_{\alpha \in A} : [0,1] \to E$ is also continuous: this follows from the continuity of all mappings $\pi_\alpha \circ \gamma = \gamma_\alpha$ and from the definition of the product topology in E; e.g., see (5.1). $\square$

Let us now discuss two examples:

- $\mathbf{E} = \mathbb{R}$**.** A subset $X \subset \mathbb{R}$ is connected if and only if it is an (open, closed or half-closed; finite or infinite) interval; i.e. iff the condition $x_0, x_1 \in X$ implies that $x \in X$ for all $x \in \mathbb{R}$ such that $x_0 < x < x_1$. In particular, in this case all connected sets are also path-connected; cf. Lemma 8.3.

  Proof. Assume that $x_{0,1} \in X$ but $x \notin X$ for a certain $x \in (x_0, x_1)$. By choosing $U_0 := (-\infty, x) \cap X$ and $U_1 := (x, +\infty) \cap X$, it is easy to see that X is not connected. Vice verse, if X is an interval, then it is obviously path-connected and hence connected due to Lemma 8.3. $\square$

  In particular, one immediately concludes that:
  - if E is a connected topological space and $f : E \to \mathbb{R}$ is continuous, then $f(E)$ is an (open or closed or half-open; finite or infinite) interval;
  - moreover, if E is a compact connected space, then $f(E) = [f_{\min}, f_{\max}]$ for some $-\infty < f_{\min} \le f_{\max} < +\infty$.
- **Topologist's sine curve**. Let $E = \mathbb{R}^2$ and consider
  $$X := \{(x, \sin \tfrac{1}{x}), x > 0\} \cup \{(0, y), y \in [-1, 1]\}.$$
  This set is(!) connected but <u>not</u> path-connected. (The former claim follows from Lemma 8.7, the latter – from the fact that each continuous mapping $\gamma : [0,1] \to \mathbb{R}^2$ s.t. $\gamma(0) = (0,0)$ and $\gamma(1) = (\pi, 0)$ is uniformly continuous and hence cannot oscillate between the lines $y = \pm 1$ infinitely many times.)

Definition 8.6. *A <u>metric</u> space E is called* chain-connected (bien enchaîné *in French) if for each $\varepsilon > 0$ and each pair of points $x_0, x_1 \in E$ there exists a finite sequence of points $x_0 = y_0, y_1, \ldots, y_{N+1} = x_1$ such that $d(y_k, y_{k+1}) < \varepsilon$.*

Lemma 8.7. *(i) If E is a connected metric space, then it is chain-connected.*
(ii) If E is a <u>compact</u> chain-connected metric space, then it is connected.

It is worth noting that one cannot withdraw the compactness assumption in (ii): e.g., the metric space $\mathbb{R} \setminus \{0\}$ is chain-connected but not connected. Another example of the same kind is the closed set $X := \{(x, y) : |y| \cdot x \ge 1\} \subset \mathbb{R}^2$.

Proof of Lemma 8.7. (i) Let $x_{0,1} \in E$ and $\varepsilon > 0$. Denote
$$U_0 := \{y : y \text{ is } \varepsilon\text{-chain connected to } x_0\}, \qquad U_1 := E \setminus U_0.$$
It is easy to see than both sets U_0 and U_1 are open in E. (Indeed, if $y \in U_0$, then $B(y, \varepsilon) \subset U_0$ as we can add one more step to the ε-chain going from x to y. Similarly, if $y \in U_1$, then $B(y, \varepsilon) \subset U_1$: otherwise, there would exist a point $z \in B(y, \varepsilon)$ such that z is ε-connected to x and we could add one more step from z to y to this chain.) Since E is connected and $U_0 \neq \emptyset$, we must have $U_0 = E$.

(ii) Assume that E is not connected, i.e., that exist open sets $\emptyset \neq U_{0,1} \subset E$ such that $U_{0,1} = E \setminus U_{1,0}$. In particular, both $U_{0,1}$ are also closed in E and hence compact since E is compact (see Lemma 6.5). This implies that
$$d(U_0, U_1) = \inf\nolimits_{x_0 \in U_0, x_1 \in U_1} d(x_0, x_1) = \min\nolimits_{x_0 \in U_0, x_1 \in U_1} d(x_0, x_1) > 0.$$
and hence $x_{0,1} \in U_{0,1}$ cannot be ε-chain connected provided that $\varepsilon < d(U_0, U_1)$. $\square$

Definition 8.8. *Let E be a topological space. The connected component of a point $x \in E$ is the maximal (under inclusion) connected set $C_x \subset E$ such that $x \in C_x$.*

Note that

- such a maximal set exists due to Lemma 8.5(i);
- the connected component C_x is a *closed* set. Indeed, it is easy to see that, if a set $C \subset E$ is connected, then its closure $\overline{C}$ is also connected. (Let $\overline{C} \subset U_0 \cup U_1$, where U_0 and U_1 are disjoint open sets in E. Since C is connected we should have either $C \cap U_0 = \emptyset$ or $C \cap U_1 = \emptyset$, which implies that either $\overline{C} \cap U_0 = \emptyset$ or $\overline{C} \cap U_1 = \emptyset$.)
- Thus, E is the union of disjoint (see Lemma 8.5) closed connected sets C_x.
- Trivially, if $x \in U \subset E$ is a both open and closed set, then $C_x \subset U$. However, note that C_x is in general *not* open.

Definition 8.9. *E is called* totally disconnected *if $C_x = \{x\}$ for all $x \in E$.*

A trivial example of a totally disconnected space is provided by the discrete topology on E. A less trivial example is the Cantor set C or any other nowhere dense subset of $\mathbb{R}$ (indeed, connected subsets of $\mathbb{R}$ are intervals and nowhere dense sets do not contain any nontrivial interval).
Exercise: prove that the space $\mathbb{Q}_p$ of p-adic numbers is also totally disconnected.

November 2, 2020

9. The space $C(K,E)$ of continuous functions on compacts

9.1. **Arzelà–Ascoli theorem.** Let K be a metric compact and E be a complete metric space. Recall the characterization of compact metric spaces provided by Proposition 7.4: K is compact if and only if it is complete and, for each $r > 0$, admits finite r-nets.

Recall also that we denote by $C(K,E)$ the vector space of all continuous function $f : K \to E$, equipped with the distance

$$d_{C(K,E)}(f,g) := \sup_{x \in K} d_E(f(x), g(x)) = \max_{x \in K} d_E(f(x), g(x)).$$

As discussed above (see Theorem 7.5), $C(K,E)$ is a complete metric space.

- If E is a complete normed space – such spaces are called *Banach spaces* (due to Stefan Banach, 1892–1945) – then $C(K,E)$ is also a Banach space as the definition of the distance $d_{C(K,E)}$ can be transformed into the norm $\|f\|_{C(K,E)}(f,g) := \max_{x \in K} \|f(x)\|_E$.
- Moreover, if $E = \mathbb{R}$ or $E = \mathbb{C}$ (or other field, which is complete with respect to a certain absolute value), then $C(E,K)$ becomes a *Banach algebra*: one can also multiply the functions and this operation satisfies $\|fg\| \le \|f\| \cdot \|g\|$.

Given a sequence of functions $f_n \in C(K,E)$, $n \in \mathbb{N}$, and $f \in C(K,E)$ one says that

- $f_n(x) \to f(x)$ *pointwise* if this convergence holds for each $x \in K$;
- $f_n \to f$ *uniformly for* $x \in K$ if $f_n \to f$ in $C(K,E)$; we will also use the notation $f_n \rightrightarrows f$ in order to avoid a confusion.

Clearly, the uniform convergence implies the pointwise one but not vice versa (unless K is a *finite* set: note that in this case $C(K,E) \cong E^{\#K}$, the pointwise convergence is the coordinate-wise while the metric in $C(K,E)$ is the ℓ^1-type metric on $E^{\#K}$). Indeed, if $x_0 \in K$ is a non-isolated point, then one, e.g., can consider functions $f_n(x) := \phi_0(n \cdot d(x, x_0))$, where $\phi_0(t) := te^{-t}$. (If $x_0, x_1, x_2, \ldots \in K$ are isolated points, then one can take $f_n(x) := 1$ if $x = x_k$, $k \geq n$, and $f_n(x) := 0$ otherwise.)

The following two facts are essentially known from the real analysis:

- **Dini's theorem.** Let $E = \mathbb{R}$ and assume that the sequence $f_n \in C(K, \mathbb{R})$ is monotone, e.g., $f_0(x) \leq f_1(x) \leq \ldots$. Let $f_n(x) \to f(x)$ pointwise (i.e., for each $x \in K$), where $f \in C(K, \mathbb{R})$. Then, this convergence is uniform.

  Proof. Given $\varepsilon > 0$, denote $U_n := \{x \in K : f_n(x) > f(x) - \varepsilon\} \subset U_{n+1}$. These sets are open (as f_n, f are continuous) and cover K. Since K is compact, there exists $n \in \mathbb{N}$ such that $K = U_n$. □

- **Heine's theorem:** Provided that K is compact, each continuous function $f : K \to E$ is *uniformly continuous:* for each $\varepsilon > 0$ there exists $\delta > 0$ such that the following holds: $d_K(x', x'') < \delta$ implies $d_E(f(x'), f(x'')) < \varepsilon$.

  Proof. For each $x \in K$ one can find an open neighborhood $U_x = B(x, 2\delta_x)$ such that $d_E(f(x'), f(x)) < \frac{1}{2}\varepsilon$ for all $x' \in U_x$). By compactness, K is covered by finitely many *twice smaller* open balls $B(x, \delta_x)$; let $\delta > 0$ be the minimum of their radii. Now, if $x' \in B(x, \delta_x)$ and $d_K(x', x'') < \delta$, then both $x', x'' \in U_x = B(x, 2\delta_x)$ and hence $d_E(f(x'), f(x'')) < 2 \cdot \frac{1}{2}\varepsilon = \varepsilon$. □

The first of the two main statements in this section is an 'if and only if' characterization of (pre)compact sets in $C(K,E)$.

Exercise: closed balls are *not* compact in $C(K,E)$ (provided that K is not finite).

Theorem 9.1 (Arzelá–Ascoli). *A set $\mathcal{F} \subset C(K,E)$ is precompact (i.e., its closure $\overline{\mathcal{F}}$ is compact) if and only if the following two conditions hold:*

(a) *for each $x \in K$ the set $\mathcal{F}(x) := \{f(x), f \in \mathcal{F}\}$ is precompact in E;*
(b) *the family of functions $f \in \mathcal{F}$ is uniformly* equicontinuous*: for each $\varepsilon > 0$ there exists $\delta > 0$ such that for all $f \in \mathcal{F}$ the following holds: $d_K(x', x'') < \delta$ implies $d_E(f(x'), f(x'')) < \varepsilon$.*

Proof. '⇒' Assume that $\overline{\mathcal{F}}$ is compact in $C(K,E)$. The condition (a) is trivial since, for each $x \in K$, the application $f \mapsto f(x)$ is continuous and thus $\{f(x), f \in \overline{\mathcal{F}}\}$ is compact in E. Thus, it remains to prove (b). Let $\varepsilon > 0$, $x \in K$ and $r > 0$. Denote

$$\mathcal{U}^{(\varepsilon)}_{x,r} := \{f \in C(K,E) : \sup_{x' \in B(x,r)} d_E(f(x'), f(x)) < \varepsilon\}.$$

It is easy to see that the set $\mathcal{U}^{\varepsilon}_{x,r}$ is open in $C(K,E)$: indeed, if

$$\sup_{x' \in B(x,r)} d_E(f(x'), f(x)) = \varepsilon - \rho, \quad \rho > 0,$$

and $d_{C(K,E)}(f,g) < \frac{1}{2}\rho$, then $\sup_{x' \in B(x,r)} d_E(g(x'), g(x)) < \varepsilon$ (i.e, $B(f, \frac{1}{2}\rho) \in \mathcal{U}^{(\varepsilon)}_{x,r}$). Clearly, for each $x \in K$ we have $\mathcal{F} \subset \bigcup_{r>0} \mathcal{U}^{(\varepsilon)}_{x,r} = C(K,E)$. Since $\mathcal{F}$ is precompact, for each $x \in K$ there exists $r_x > 0$ such that $\mathcal{F} \subset \mathcal{U}^{(\varepsilon)}_{x,r_x}$. Similarly to the proof of Heine's theorem, we now note that $K = \bigcup_{x \in K} B(x, \frac{1}{2}r_x)$ and hence – by compactness – K is covered by finitely many such balls, let r be the minimum of the corresponding radii r_x. As in Heine's theorem, it is easy to see that

$$\text{for all } f \in \mathcal{F}, \text{ if } d_K(x', x'') < \tfrac{1}{2}r, \text{ then } d_E(f(x'), f(x'')) < 2\varepsilon.$$

(Indeed, if $x' \in B(x, \frac{1}{2}r_x)$ and $d_K(x', x'') < \frac{1}{2}r$, then both $x', x'' \in B(x, r_x)$ and the claim follows since $\mathcal{F} \subset \mathcal{U}^{(\varepsilon)}_{x,r_x}$.)

'$\Leftarrow$' Consider a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathcal{F}$, our goal is to find a subsequence that converges to a continuous function f uniformly on K. Recall that a metric compact K admits finite 2^{-k}-nets $X^{(k)} := \{x^{(k)}_1, \dots, x^{(k)}_{m_k}\}$ for all $k \in \mathbb{N}$ and, in particular, is separable: the set $X := \bigcup_{k\in\mathbb{N}} X^{(k)}$ is dense in K. Using the assumption (a) for each of the (countably many) points $x^{(k)}_m \in X$ and the Cantor diagonal process, we can find a subsequence $J \ni n \to \infty$ such that the sequences $f_n(x^{(k)}_m)$ converge (in E) for all $k \in \mathbb{N}$ and $m \le m_k$. Denote

$$f(x^{(k)}_m) \ := \ \lim\nolimits_{J\ni n\to\infty} f_n(x^{(k)}_m).$$

At the moment, the function f is defined on the dense subset $X \subset K$ only. However, since the functions f_n are uniformly *equi*-continuous (see (b)), the following holds:

$$\text{if } x', x'' \in X \text{ and } d_K(x', x'') < \delta, \ \text{ then } \ d_E(f(x'), f(x'')) \le \varepsilon, \tag{9.1}$$

with the same $\delta = \delta(\varepsilon) > 0$ as in (b). Note that K can be viewed as a completion of X. Therefore, f can be extended from X to E as in Lemma 7.9: given a point $x \in K \smallsetminus X$, find a convergent sequence $X \ni x_n \to x$ and note that (9.1) implies that the limit $f(x) := \lim_{n\to\infty} f(x_n)$ exists and does not depend on the choice of the approximating sequence. Moreover, thus constructed function f satisfies the same condition (9.1) for all $x', x'' \in K$. In particular, $f \in C(K, E)$.

We are almost done: it only remains to check that $f_n \to f$ in $C(K, E)$ (as $J \ni n \to \infty$). Given ε, δ from the equicontinuity condition (b) and let $2^{-k} < \delta$. Due to the pointwise convergence $f_n(x) \to f(x)$ at each of the (finitely many) points $x^{(k)} \in X^{(k)}$ in the 2^{-k}-net, there exists $N \in \mathbb{N}$ such that $d_E(f_n(x^{(k)}), f(x^{(k)})) < \varepsilon$ for all $x^{(k)} \in X^{(k)}$ and for all $J \ni n \ge N$. Together with (9.1), this implies that

$$d_E(f_n(x), f(x)) \ \le \ d_E(f_n(x), f_n(x^{(k)})) + d(f_n(x^{(k)}), f(x^{(k)})) + d(f(x^{(k)}), f(x)) \ < 3\varepsilon$$

for all $x \in K$ and $J \ni n \ge N$, provided that $d_K(x, x^{(k)}) < 2^{-k} < \delta(\varepsilon)$. $\square$

Example. Certainly, the key condition in Theorem 9.1 is the equicontinuity. E.g., the closed(!) set of 1-Lipshitz functions

$$\{f : [0,1] \to \mathbb{R} : |f(x) - f(y)| \le |x - y| \text{ and } |f(0)| \le M\}$$

is compact in $C([0,1], \mathbb{R})$; note that, though the second condition $|f(0)| \le M$ is indispensable, it can be replaced by any other bound on the the values of f.

9.2. **Stone–Weierstrass theorem.** Consider now the Banach algebra $C(K, \mathbb{R})$. The next theorem essentially says that a *subalgebra* of this algebra is always dense unless it is not due to a trivial reason (see condition (b) below). Note that (see a discussion before Corollary 9.3) that a similar result for the algebra $C(K, \mathbb{C})$ of complex-valued functions does *not* hold 'as is': one should also require that the subalgebra in question is closed under the conjugation.

Theorem 9.2 (**Stone–Weierstrass**). *Let $\mathcal{A}$ be a subalgebra of $C(K, \mathbb{R})$ (in other words, $\alpha f + \beta g \in \mathcal{A}$ and $fg \in \mathcal{A}$ for all $f, g \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{R}$). Assume that*

(a) $\mathbb{R} \subset \mathcal{A}$, *i.e., that constant functions belong to* $\mathcal{A}$*;*
(b) $\mathcal{A}$ *distinguishes points of* K*, i.e., for each* $x, y \in K$ *such that* $x \neq y$ *there exists* $f \in \mathcal{A}$ *such that* $f(x) \neq f(y)$*.*

Then, $\mathcal{A}$ *is dense in* $C(K, \mathbb{R})$*.*

Before given the proof of this theorem, let us first discuss a few examples:

- The model case to have in mind is the algebra of polynomials $\mathbb{R}[x]$ (or, similarly, trigonometric polynomials), which is dense in $C(K;\mathbb{R})$ for all compact subsets $K \subset \mathbb{R}$.
- Clearly, the same holds for the algebra $\mathbb{R}[x_1, \ldots, x_n]$ of polynomials in n variables, on each compact set $K \subset \mathbb{R}^n$.
- However, if we replace $\mathbb{R}$ by $\mathbb{C}$, the conditions (a) and (b) are not anymore sufficient. To see this, consider the closed unit disc $\overline{\mathbb{D}} := \{z : |z| \leq 1\}$ and the algebra of polynomials $\mathbb{C}[z]$ with complex coefficients. For each $P \in \mathbb{C}[z]$ we have $\oint_{|z|=1} P(z)dz = i\int_0^{2\pi} P(e^{i\theta})e^{i\theta}d\theta = 0$. If $\mathbb{C}[z]$ were dense in $C(\overline{\mathbb{D}};\mathbb{C})$, then the same identity would hold for all continuous functions $f : \overline{\mathbb{D}} \to \mathbb{C}$ since $f \mapsto \oint_{|z|=1} f(z)dz$ is a continuous mapping $C(\overline{\mathbb{D}};\mathbb{C}) \to \mathbb{C}$. However, this is not the case: e.g., $\oint_{|z|=1} \overline{z}dz = 2\pi i$.
- A usual modification of the Stone–Weierstrass theorem for complex-values functions is the following

  Corollary 9.3. *Let* $\mathcal{A}$ *be a subalgebra of* $C(K, \mathbb{C})$ *such that, as in Theorem 9.2, (a)* $\mathcal{A}$ *contains constants, (b)* $\mathcal{A}$ *distinguish points of* K *and*
  (c) $\mathcal{A}$ *is closed under the conjugation: if* $f \in \mathcal{A}$*, then* $\overline{f} \in \mathcal{A}$*.*
  Then, $\mathcal{A}$ *is dense in* $C(K, \mathbb{C})$*.*

  We will discuss an (almost trivial) proof of Corollary 9.3 after Theorem 9.2.

Proof of Theorem 9.2. We start with two preliminary steps:

- **Lemma.** *There exists polynomials* $P_n \in \mathbb{R}[x]$ *such that* $P_n(x) \to |x|$ *as* $n \to \infty$ uniformly *for* $x \in [-1, 1]$*.*

  Proof. For instance, let $P_0(x) := 0$, $P_{n+1}(x) := P_n(x) + \frac{1}{2}(x^2 - (P_n(x))^2)$. By induction, $P_n(x) \leq P_{n+1}(x) \leq |x|$ and hence $P_n(x) \to |x|$ as $n \to \infty$ for all $x \in [-1, 1]$. The convergence is uniform due to the Dini theorem. □

- *Denote* $\mathcal{L} := \overline{\mathcal{A}}$*, the completion of* $\mathcal{A}$ *in* $C(K, \mathbb{R})$*. Then,* $\mathcal{L}$ *is a* lattice*, i.e.* $f, g \in \mathcal{L}$ *implies that* $\max\{f, g\} \in \mathcal{L}$ *and* $\min\{f, g\} \in \mathcal{L}$*.*

  Proof. Since $\max\{f, g\} = \frac{1}{2}(f+g+|f-g|)$, it is enough to prove that $f \in \overline{\mathcal{A}}$ implies that $|f| \in \overline{\mathcal{A}}$. Note that $\overline{\mathcal{A}}$ is an algebra: if $f, g \in \overline{\mathcal{A}}$ and $f_n, g_n \in \mathcal{A}$ are such that $f_n \to f$, $g_n \to g$ in $C(K, \mathbb{R})$, then $\alpha f_n + \beta g_n \to \alpha f + \beta g \in \overline{\mathcal{A}}$ and $f_n g_n \to fg \in \overline{\mathcal{A}}$. (Note that in the latter implication we use the fact that continuous functions on K are bounded.) Now let $f \in \overline{\mathcal{A}}$ and $M \in \mathbb{R}$ be such that $|f| \leq M$. For polynomials P_n provided by the above lemma, we have $\overline{\mathcal{A}} \ni MP_n(f) \to |f|$ in $C(K, \mathbb{R})$. Therefore, $|f| \in \overline{\mathcal{A}}$. □

Theorem 9.2 now follows from

Proposition 9.4. *Let $\mathcal{L} \subset C(K, \mathbb{R})$ be a lattice such that for each $x, y \in K$, $x \neq y$, and each $a, b \in \mathbb{R}$ and $\varepsilon > 0$ there exists a function $f \in \mathcal{L}$ such that $|f(x) - a| < \varepsilon$ and $|f(y) - b| < \varepsilon$. Then, $\mathcal{L}$ is dense in $C(K, \mathbb{R})$.*

Indeed, if $a = b$, then one can take $f(x) := a$ using assumption (a) and, if $a \neq b$, then one can find a function f that distinguishes points x and y (see (b)) and consider the function $b \cdot (f(\cdot) - f(x))/(f(y) - f(x)) + a \cdot (f(y) - f(\cdot))/(f(y) - f(x))$. In both cases, we find a function $f \in \mathcal{A}$ such that $f(x) = a$ and $f(y) = b$.

Thus, the proof of Theorem 9.2 is complete modulo key Proposition 9.4, which is sometimes viewed as another version of the Stone–Weierstrass theorem. □

We begin the next lecture with the proof of Proposition 9.4.

November 4, 2020

Recall that in the previous lecture we deduced the Stone–Weierstrass theorem from the following proposition, we now give its proof.

Proposition 9.4. *Let $\mathcal{L} \subset C(K, \mathbb{R})$ be a lattice such that for each $x, y \in K$, $x \neq y$, and each $a, b \in \mathbb{R}$ and $\varepsilon > 0$ there exists a function $f \in \mathcal{L}$ such that $|f(x) - a| < \varepsilon$ and $|f(y) - b| < \varepsilon$. Then, $\mathcal{L}$ is dense in $C(K, \mathbb{R})$.*

Proof. Let $f \in C(K, \mathbb{R})$ be a continuous function on K that we want to approximate (in $C(K, \mathbb{R})$, i.e., uniformly on K) by a function from $f_{\mathcal{L}} \in \mathcal{L}$ with a precision $\varepsilon > 0$. Given $x, y \in K$, let us find a function $f_x^{(y)} \in \mathcal{L}$ such that

$$|f_x^{(y)}(x) - f(x)| < \varepsilon \quad \text{and} \quad |f_x^{(y)}(y) - f(y)| < \varepsilon.$$

(Note that such a function exists due to the assumption if $x \neq y$ but also if $x = y$: take an arbitrary other point as the second point in the assumption.) Let, for a while, $y \in K$ be fixed and consider the open (in K) sets

$$x \in U_x^{(y)} \ := \ \{x' \in K : f_x^{(y)}(x') - f(x') < \varepsilon\}.$$

Since K is compact, it is covered by finitely many sets $U_{x_k}^{(y)}$, $k = 1, \ldots, m = m^{(y)}$. Let

$$f^{(y)} \ := \ \min\{f_{x_1}^{(y)}, \ldots, f_{x_m}^{(y)}\} \ \in \ \mathcal{L}.$$

By construction,

$$f^{(y)}(y) - f(y) > -\varepsilon \quad \text{and} \quad f^{(y)}(x) - f(x) < \varepsilon \ \text{ for all } x \in K.$$

We now use the same trick as above: for each $y \in K$, consider an open neighborhood

$$y \in V^{(y)} \ := \ \{y' \in K : f^{(y)}(y') - f(y') > -\varepsilon\}.$$

Since K is compact, it is covered by finitely many sets $V^{(y_k)}$, $k = 1, \ldots, m$. Denote

$$f_{\mathcal{L}} \ := \ \max\{f^{(y_1)}, \ldots, f^{(y_m)}\} \in \mathcal{L}.$$

By construction, we have $-\varepsilon < f_{\mathcal{L}}(x) - f(x) < \varepsilon$ for all $x \in K$. □

Let us also prove of a version of the Stone–Weierstrass theorem for $C(K, \mathbb{C})$; recall that in this case we additionally assume that $\mathcal{A}$ is closed under the conjugation.

Proof of Corollary 9.3. Note that $\operatorname{Im} f = \operatorname{Re}(-if)$ and let

$$\mathcal{A}_{\mathbb{R}} \ := \ \{\operatorname{Re} f,\ f \in \mathcal{A}\} \subset C(K,\mathbb{R}) \ = \ \{\operatorname{Im} f,\ f \in \mathcal{A}\}.$$

It easily follows from assumption (c) that $\mathcal{A}_{\mathbb{R}} = \mathcal{A} \cap C(K,\mathbb{R})$: indeed, if $f \in \mathcal{A}$, then $\operatorname{Re} f = \frac{1}{2}(f+\overline{f}) \in \mathcal{A}$. Therefore, $\mathcal{A}_{\mathbb{R}}$ is an *algebra* and thus, given $f \in C(K,\mathbb{C})$, one can apply Theorem 9.2 to construct approximations $f_{\mathrm{Re}}, f_{\mathrm{Im}} \in \mathcal{A}_{\mathbb{R}}$ of $\operatorname{Re} f, \operatorname{Im} f$ in $C(K,\mathbb{R})$. Then, the function $f_{\mathrm{Re}} + if_{\mathrm{Im}} \in \mathcal{A}$ approximates f in $C(K,\mathbb{C})$. □

10. Bounded linear operators in Banach spaces

Definition 10.1. *A vector space E over $\mathbb{R}$ or $\mathbb{C}$ is called*

- *a* Banach space *if it is a complete normed space.*

Two related definitions are:

- *E is called a* Fréchet space *if there is a countable family of (semi-)norms $\|\cdot\|_k : E \to \mathbb{R}_+$ such that E is complete wrt to the metric constructed out of these norms: if a sequence $(x_n)_{n\in\mathbb{N}}$ is Cauchy wrt to each of $\|\cdot\|_k$, then there exists $x \in E$ such that $\|x_n - x\|_k \to 0$ as $n \to \infty$ for each k. Clearly, Fréchet spaces are more general than Banach ones.*
- *A very particular case of Banach spaces is* Hilbert spaces*: E is called Hilbert if it carries a structure of a* scalar product *(i.e., a function $\langle\cdot,\cdot\rangle$ on $E \times E$, which is linear with respect to one of the arguments, anti-linear with respect to the other, and satisfies $\langle y,x\rangle = \overline{\langle y,x\rangle}$) such that $\langle x,x\rangle = \|x\|^2$.*

(Rhetoric question.) Why at all should one wonder about linear operators between Banach spaces? This is because we often want to study 'nice' functions $f : E \to F$ and to say that this function is differentiable at a point $x \in E$ is to say that there exists a linear mapping $A : E \to F$ such that f can be approximated by this linear mapping near x (i.e., $\|f(x') - f(x) - A(x'-x)\|_F = o(\|x'-x\|_E)$ as $x' \to x$). Comparing with the most trivial case $E = F = \mathbb{R}$, the linear mapping A now plays the role of $\lambda = f'(x) \in \mathbb{R}$, which can be understood as a mapping $t \mapsto \lambda t$, $\mathbb{R} \to \mathbb{R}$.

Definition 10.2. *Let E, F be two Banach (or just normed) vector spaces.*

- *A mapping $A : E \to F$ is called a* bounded linear operator *if A is a linear mapping (i.e., $A(\alpha x + \beta y) = \alpha Ax + \beta Ay$) and there exists $M > 0$ such that $\|Ax\|_F \le M\|x\|_E$ for all $x \in E$.*
- *Let $\mathcal{L}(E,F)$ be the vector-space of all bounded linear operators $A : E \to F$.*
- *In fact, it is easy to see that $\mathcal{L}(E,F)$ is a <u>normed</u> vector space if we set*

$$\begin{aligned}\|A\|_{\mathcal{L}(E,F)} \ &:= \ \inf\{M > 0 :\ \text{we have } \|Ax\|_F \le M\|x\|_E \text{ for all } x \in E\} \\ &= \ \sup\nolimits_{x\in E:\|x\|=1} \|Ax\|_F/\|x\|_E\,.\end{aligned}$$

(The last equality is a simple exercise.) Let us start with a list of simple comments:

- If $A : E \to F$ is a linear mapping, then

$$\begin{aligned}A \in \mathcal{L}(E,F) \ &\Leftrightarrow \ A \text{ is a Lipschitz mapping } \ \Leftrightarrow \ A \text{ is continuous} \\ &\Leftrightarrow \ A \text{ is continuous at the point } 0 \in E.\end{aligned}$$

  (Indeed, all '$\Rightarrow$' are trivial and, if A is continuous at $0 \in E$, then there exists $\delta > 0$ such that $\|x\|_E < \delta$ implies $\|Ax\|_F < 1$, so one can take $M := \delta^{-1}$.)

- In particular, a natural generalization of the notion of bounded linear operators from Banach to Fréchet spaces is continuous linear mappings.
- If F is a Banach space, then $\mathcal{L}(E,F)$ is also a Banach space.

  Proof. We need to check that $\mathcal{L}(E,F)$ is complete. Let $(A_n)_{n\in\mathbb{N}}$ be a Cauchy sequence. Then, for each $x\in E$ the sequence $(A_nx)_{n\in\mathbb{N}}$ is a Cauchy sequence in F: indeed, $\|A_nx-A_mx\|_F\le\|A_n-A_m\|_{\mathcal{L}(E,F)}\cdot\|x\|$. As we assumed that F is complete, we can define $Ax:=\lim_{n\to\infty}A_nx$. It is straightforward to check that A is a linear operator, that it is bounded and that $\|A_n-A\|_{\mathcal{L}(E,F)}\le\limsup_{m\to\infty}\|A_n-A_m\|_{\mathcal{L}(E,F)}\to 0$ as $n\to\infty$. □

- In the finite-dimensional case, the space $\mathcal{L}(\mathbb{R}^n,\mathbb{R}^m)$ is just the space of $\mathbb{R}^{n\times m}$ matrices. In particular, in this case *all* linear operators are bounded. However, note that the definition of the norm on $\mathbb{R}^{n\times m}$ *depends* on the choice of the norms in $\mathbb{R}^n$ and $\mathbb{R}^m$.
- It is easy to see that all norms on $\mathbb{R}^n$ are equivalent: if $\|\cdot\|'$ is a certain norm on $\mathbb{R}^n$, e_k are the standard basis vectors, and $x=\sum_{k=1}^n x_ke_k$, then

  $$\|x\|'\le M\cdot\|x\|_1,\ \text{ where }\ M:=\max_{k=1,\dots,n}\|e_k\|'\ \text{ and }\ \|x\|_1=\textstyle\sum_{k=1}^n|x_k|.$$

  From here, it follows that the mapping $x\mapsto\|x\|'$, $(\mathbb{R},\|\cdot\|_1)\to\mathbb{R}_+)$, is continuous and thus has a strictly positive minimum on the unit sphere, which implies the inverse inequality $\|x\|'\ge m\|x\|_1$ for a certain $m>0$.
- Thus, if E is finite-dimensional, then all norms on E are equivalent and one easily sees that *all* linear operators $A:E\to F$ are bounded:

  $$\|Ax\|_F\le\max_{k=1,\dots,n}\|Ae_k\|\cdot\|x\|_1\le M\cdot\|x\|_E,$$

  where $M:=\max_{k=1,\dots,n}\|Ae_k\|\cdot(\min\|x\|_E/\|x\|_1)^{-1}$.
- However, if E is *infinite-dimensional*, then there always exist linear mappings $A:E\to\mathbb{R}$ which are not bounded. To see this
  - choose linearly independent vectors $(e_n)_{n\in\mathbb{N}}$ in E with $\|e_n\|=1$;
  - find an algebraic basis $(f_\beta)_{\beta\in B}$ in E that contains vectors $f_0=e_0$, $f_1=e_1-2e_0$, $\dots$, $f_n=e_n-2e_{n-1}$, $\dots$ (recall that the fact that $(f_\beta)_{\beta\in B}$ is an algebraic basis means that each element of E equals to a finite linear combination of f_β's; one needs the axiom of choice here);
  - Define the mapping $A:x\to\lambda_0(x)$, where $\lambda_0(x)$ denotes the coefficient in front of $f_0=e_0$ in the representation of x in the basis $(f_\beta)_{\beta\in B}$. Then, $Ae_n=A(f_n+2f_{n-1}+\ldots+2^nf_0)=2^n$, so A is not bounded.

Definition 10.3. *Let E be a Banach space. The space $\mathcal{L}(E,\mathbb{R})$ (or $\mathcal{L}(E,\mathbb{C})$ for Banach spaces over $\mathbb{C}$) is called the* (continuous) dual to E *and is denoted by E' (or E^* in certain places other than France).*

Let us discuss several concrete examples:

(1) **Let $E=\ell^p$ with $1\le p<+\infty$. Then, $E'\cong\ell^q$, where $\frac{1}{p}+\frac{1}{q}=1$** and the isomorphism works as follows:

$$y=(y_n)_{n\in\mathbb{N}}\in\ell^q\quad\mapsto\quad\big(Y:\ x\in\ell^p\ \mapsto\ Y(x):=\textstyle\sum_{n\in\mathbb{N}}x_ny_n\big).$$

Proof. Due to the Hölder inequality, this mapping defines an inclusion of ℓ^q into $(\ell^p)'$; moreover, we have $\|Y\|_{(\ell^p)'} \le \|y\|_{\ell^q}$. Thus, we need to prove two things: the first is that actually $\|Y\| = \|y\|$ and the second is that the dual space $(\ell^p)'$ does not any contain other elements.

Assume that $p \neq 1$. Given $y \in \ell^q$, define $x_n := |y_n|^{q-1} \cdot \mathrm{sign}(y_n)$. Since $p(q-1) = q$, we have $x \in \ell^p$ and

$$\|x\|_p = (\textstyle\sum_{n\in\mathbb{N}} |y_n|^q)^{1/p} = \|y\|_q^{q-1} \quad \text{and} \quad Y(x) = \textstyle\sum_{n\in\mathbb{N}} x_n y_n = \|y\|_q^q.$$

This implies that, for $y \in \ell^q$, we have $\|Y\|_{\mathcal{L}(\ell^p,\mathbb{R})} \ge \|y\|_{\ell^q}$ (and hence '$=$').

Now let $Y \in (\ell^p)'$ be an arbitrary element of the dual space and define $y_n := Y(e_n)$, where e_n stands for the n-th standard basis vector in ℓ^p. Assume that $y \notin \ell^q$ and, as above, denote $x_n^{(N)} := |y_n^{(N)}|^{q-1} \cdot \mathrm{sign}(y_n)$, where $y^{(N)} = (y_0, \ldots, y_N, 0, 0, \ldots)$. Then,

$$Y(x^{(N)})/\|x^{(N)}\|_p \;=\; \|y^{(N)}\|_q \;\to\; \infty \;\text{ as }\; N \to \infty$$

provided that $y \notin \ell^q$, a contradiction.

Exercise: give a proof in the remaining (simpler) case $p = 1$. $\square$

(2) **However, $(\ell^\infty)' \supsetneq \ell^1$.** It is always true that one can identify E with a *subspace* of the second dual $(E')'$ by mapping $E \ni x \mapsto X \in (E')'$, where X is defined as $E' \ni A \mapsto X(A) := Ax$. Since $\ell^\infty = (\ell^1)'$, this means that $(\ell^\infty)' \supset \ell^1$. However, *not all* continuous functionals on ℓ^∞ can be obtained in this way, we will return to this question during the next lecture.

(3) The same statements hold for L^p-spaces but we do not discuss it here.

(4) **Riesz(–Markov–Kakutani)'s theorem:**
Let $E = C(K, \mathbb{R})$, where K is a metric compact. Then, E' can be identified with the space $\mathcal{M}(K)$ of all *signed Borel measures on K:* to a measure $\mu \in \mathcal{M}(K)$, one associates a functional $C(K,\mathbb{R}) \ni f \mapsto \int_K f d\mu$.
This is a rather deep result which goes beyond the scope of these lectures.

Proposition 10.4. *Let E be a Banach space and $A : E \to \mathbb{R}$ be a linear mapping. Then, $A \in \mathcal{L}(E, \mathbb{R})$ if and only if* $\mathrm{Ker}\, A$ *is a* closed *subspace of E.*

Proof. '$\Rightarrow$': this is a triviality since $\mathrm{Ker}\, A = A^{-1}(\{0\})$ and A is continuous.
'$\Leftarrow$': denote $F := \mathrm{Ker}\, A$ and let $x_0 \notin F$ (if $F = E$, then there is nothing to prove). Let $x_1 := x_0/(Ax_0)$; note that $Ax_1 = 1$ and that $d := \mathrm{dist}(x_1, F) > 0$, this is where we use the fact that F is closed. Since $x/(Ax) - x_1 \in F$ for each $x \in E$ such that $Ax \neq 0$, we have $\|x/(Ax)\| \ge d$ and hence $|Ax| \le d^{-1}\|x\|$ for all $x \in E$. $\square$

As we have seen above, in infinite dimensional Banach spaces there are (plenty of) unbounded linear functionals $A : E \to \mathbb{R}$ and hence plenty of non-closed linear subspaces. However, it is worth noting that to construct an example of a non-closed subspaces one should *not* do anything exotic: for instance, let

$$F := \{x = (x_n)_{n\in\mathbb{Z}} : \textstyle\sum_{n\in\mathbb{Z}} n^2|x_n|^2 < +\infty\} \subset E = \ell^2(\mathbb{Z}).$$

Clearly, F cannot be closed as it contains finite sequences, which are dense in ℓ^2.

(**Quasi-détour:** Applying the Fourier transform $(x_n)_{n\in\mathbb{Z}} \mapsto \sum_{n\in\mathbb{Z}} x_n e^{i\pi n t}$ – see TD=DM Toussaint – this example can be reformulated as follows: the Sobolev space H^1 on the circle is not closed in L^2.)

November 9, 2020

Let us now discuss linear operators acting from a Banach space E to itself. Recall that $\|A\|_{\mathcal{L}(E,E_1)} = \sup_{x\in E:x\neq 0}\|Ax\|/\|x\|$. It is easy to see that

$$\begin{array}{l} A\in\mathcal{L}(E,E_1), \\ B\in\mathcal{L}(E_1,E_2) \end{array} \quad\Rightarrow\quad \begin{array}{l} BA\in\mathcal{L}(E,E_2) \\ \|BA\|_{\mathcal{L}(E,E2)} \le \|A\|_{\mathcal{L}(E,E1)}\cdot\|B\|_{\mathcal{L}(E_1,E_2)}\,, \end{array}$$

where we write BA instead of $B\circ A$ (similarly to Ax instead of $A(x)$).

Definition 10.5. *Given a Banach space E, denote $\mathcal{L}(E) := \mathcal{L}(E,E)$.*

Clearly, $\mathcal{L}(E,E)$ is a *Banach algebra:* if $A_n\to A$ and $B_n\to B$ in $\mathcal{L}(E)$, then $B_nA_n\to BA$ in $\mathcal{L}(E)$ since $\|B_nA_n - BA\|\le\|B\|\cdot\|A_n-A\|+\|A_n\|\cdot\|B_n-B\|$.

Definition 10.6. *The* spectral radius $\rho(A)$ *of a bounded linear operator $A\in\mathcal{L}(E)$ is defined as*

$$\rho(A) \;:=\; \lim_{n\to\infty}\|A^n\|^{1/n} \;\le\; \|A\|. \tag{10.1}$$

- The limit in (10.1) exists as $\|A^{n+m}\|\le\|A^n\|\cdot\|A^m\|$. (This is a standard lemma: if a sequence $a_n\ge 0$ is sub-additive, i.e., $a_{n+m}\le a_n+a_m$, then the sequence a_n/n has a limit; in our case one takes $a_n := \log\|A^n\|$.)
- The spectral radius of $A\in\mathcal{L}(E)$ does not change if the norm on E is replaced by an equivalent norm. Indeed, the corresponding operator norms on $\mathcal{L}(E)$ are also equivalent, hence $(\|A^n\|')^{1/n}\le(C\|A^n\|)^{1/n}\sim\|A^n\|^{1/n}$ and vice versa. In particular, in order to speak about the spectral radius of $N\times N$ matrices (which are elements of $\mathbb{R}^{N\times N}$, there is no need to fix the choice of a norm on $\mathbb{R}^N$ (as all of them are equivalent).
- If $E=\mathbb{R}^N$ (more accurately, if E is finite-dimensional and thus isometric to $\mathbb{R}^N$ with a certain norm), then linear operators $A\in\mathcal{L}(E)\cong\mathbb{R}^{N\times N}$ can be understood via the *Jordan normal form*. It is easy to see that in this case $\rho(A)$ is the maximum of absolute values of the eigenvalues of A.
- However, in infinite-dimensional spaces there is no analogue of such a (Jordan) decomposition; bounded linear operators can have very complicated structure. To understand complications compared to the finite-dimensional case, consider the (backward) shift operator $\mathrm{S}:\ell^2\to\ell^2$ (which is even a Hilbert(!) space) defined by $(x_n)_{n\in\mathbb{N}}\mapsto(x_{n+1})_{n\in\mathbb{N}}$. This operator looks like a (transposed) 'infinite Jordan cell' corresponding to $\lambda=0$. However, it is easy to see that for each $\lambda\in\mathbb{C}$ such that $|\lambda|<1$ there exists a vector $f_\lambda=(\lambda^n)_{n\in\mathbb{N}}\in\ell^2$ such that $\mathrm{S}f_\lambda=\lambda f_\lambda$.

Détour.[8] It is easy to see that the 'transposed' operator $(x_n)_{n\in\mathbb{N}}\mapsto(0,x_0,x_1,\ldots)$ does not have any eigenvalue except $\lambda=0$. In particular, this observation means that (even in Hilbert spaces, not speaking about Banach ones) the notion of the *spectrum* of an operator $A\in\mathcal{L}(E)$ should be defined not as the set of eigenvalues but in a certain other way. The 'proper' definition is the following:

$$\mathrm{spec}(A) := \{\lambda\in\mathbb{C}:\ \text{the operator } A-\lambda I \text{ is not invertible in } \mathcal{L}(E)\}.$$

In fact, $\rho(A)=\sup\{|\lambda|:\lambda\in\mathrm{spec}(A)\}$; note that the inequality '$\ge$' follows from the next proposition while the inverse one is less trivial.

---

[8]This was skipped during the lecture

Denote by I the identity operator in $\mathcal{L}(E)$.

Proposition 10.7. *Let E be a Banach space and $A \in \mathcal{L}(E)$.*

(i) *If $\rho(A) < 1$, then the operator $\mathrm{I} - A$ is invertible in $\mathcal{L}(E)$. Moreover, the following identity holds (where the series converges in $\mathcal{L}(E)$):*
$$(\mathrm{I} - A)^{-1} = \mathrm{I} + A + A^2 + \dots .$$

(ii) *If A is invertible in $\mathcal{L}(E)$ and $\|B\| \le \|A^{-1}\|^{-1}$, then the operator $A - B$ is also invertible in $\mathcal{L}(E)$ and*
$$(A - B)^{-1} = (\mathrm{I} - A^{-1}B)^{-1}A^{-1} = \mathrm{I} + A^{-1}BA^{-1} + A^{-1}BA^{-1}BA^{-1} + \dots$$

(iii) *The set of invertible operators is an open set in $\mathcal{L}(E)$.*

Proof. (i) The series $\mathrm{I} + A + A^2 + \dots$ is Cauchy in E and hence converges since, for each $\varepsilon > 0$, we have $\|A^n\| \le (\rho(A) + \varepsilon)^n$ for large enough n. The identity $(\mathrm{I} - A)(\mathrm{I} + A + A^2 + \dots) = \mathrm{I}$ is a triviality.

(ii) We have $\rho(A^{-1}B) \le \|A^{-1}B\| \le \|A^{-1}\| \cdot \|B\| < 1$; the claim follows from (i).

(iii) This is straightforward from (ii): if A is invertible in $\mathcal{L}(E)$, then so are all operators in the open ball $B_{\mathcal{L}(E)}(A, \|A^{-1}\|^{-1})$. □

11. Hahn–Banach theorem

Theorem 11.1 (Hahn–Banach)**.** *Let E be a subspace of a normed ($\mathbb{R}$ or $\mathbb{C}$)-vector space E_1. Assume that a linear functional $A \in E' = \mathcal{L}(E, \mathbb{R}$ or $\mathbb{C})$ defined on E is such that $\|A\|_E = 1$. Then, there exists a linear functional $A_1 \in E_1'$ such that $A_1|_E = A$ and $\|A_1\| = \|A\| = 1$.*

(Note that, if $A_1\big|_E = A$, then one always has $\|A_1\| \ge \|A\| = 1$. The Hahn–Banach theorem claims that one can construct an extension of A on a bigger space E_1 such that the norm of A does not increase.)

Proof. The $\mathbb{R}$ and $\mathbb{C}$ case are not fully similar. We start with considering

Banach spaces over $\mathbb{R}$. The proof relies upon the axiom of choice, which we will in its equivalent form provided by

> *Zorn's lemma:* if $\mathcal{F}$ is a partially ordered set such that each totally ordered subset of $\mathcal{F}$ has a majorant, then there exists at least one maximal (i.e., not majorated by any other element) element in $\mathcal{F}$.

Let
$$\mathcal{F} \ := \ \{(F, A_F) : \ E \subset F \subset E_1, \ A_F|_E = A, \ \|A_F\|_{F'} = 1\},$$
where F is a linear subspace of E_1 and A_F a bounded linear functional on F; and
$$(F_1, A_{F_1}) \prec (F_2, A_{F_2}) \ \text{ if } \ F_1 \subset F_2 \text{ and } A_{F_2}|_{F_1} = A_{F_1}.$$
It is easy to see that each *totally* ordered set $\{(F_\beta, A_{F_\beta}\}_{\beta \in B}$ admits a majorant $(F, A_F) \in \mathcal{F}$, where $F := \bigcup_{\beta \in B} F_\beta$ and $A_F|_{F_\beta} := A_{F_\beta}$ (the latter definition is consistent provided that the set under consideration is totally ordered). Therefore, there are maximal elements in $\mathcal{F}$ and our goal is to prove that,
$$\text{if } (F, A_F) \text{ is a maximal element in } \mathcal{F}, \text{ then } F = E_1.$$
Assume that $F \subsetneq E_1$ and find $x_1 \in E_1 \smallsetminus F$. To get a contradiction, it remains to prove that we can extend the operator A_F from F further to the subspace

$F_1 := \{x + \lambda x_1,\ x \in F,\ \lambda \in \mathbb{R}\}$. In order words, the preceding discussion reduces the Hahn–Banach theorem to its particular case when

$$E_1 = \{x + \lambda x_1,\ x \in E,\ \lambda \in \mathbb{R}\}, \quad x_1 \notin E,\ \|x_1\| = 1.$$

(the last assertion can be added for free by scaling the vector x_1).

To prove that such a 'one-dimensional' extension is possible, we need to assign a value $Ax_1 := a_1 \in \mathbb{R}$ such that $|Ax + \lambda a_1| \le \|x + \lambda x_1\|$ for all $x \in E$ and $\lambda \in \mathbb{R}$. By scaling (and since this holds for $\lambda = 0$ by the assumption $\|A\|_{E'} = 1$), it is equivalent to say that $|Ax + a_1| \le \|x + x_1\|$ for all $x \in E$ or, equivalently to

$$-a_1 \;\in\; \textstyle\bigcap_{x\in E} [\, Ax - \|x{+}x_1\| \,,\, Ax + \|x{+}x_1\| \,] \,.$$

The intersection of close segments in the right-hand side is non-empty if and only if there segments intersect *pairwise*. (Indeed, we only need to prove that the supremum of the left-ends of these segments is less or equal than the infimum of their right-ends, which follows from the fact that each of these left-ends is less or equal to each of the right-ends.) Therefore, it is enough to prove that

$$Ax - \|x + x_1\| \le Ay + \|y + x_1\| \quad \text{for all } x, y \in E,$$

which follows from the triangle inequality: $A(x{-}y) \le \|x{-}y\| \le \|y{+}x_1\| - \|x{+}x_1\|$. The proof of the Hahn–Banach theorem for Banach spaces over $\mathbb{R}$ is complete.

Banach spaces over $\mathbb{C}$. The 'standard' proof uses a reduction to the $\mathbb{R}$ case, which has already been treated above. To this end, for a while consider both E and E_1 as *real*-linear spaces and denote

$$Rx := \mathrm{Re}(Ax) \quad \text{for} \;\; x \in E.$$

Clearly, R is a *real*-linear functional on E and $\|R\| \le \|A\| \le 1$. Let $R_1 : E_1 \to \mathbb{R}$ be a *real*-linear functional on E_1 such that $R_1|_E = R$ and $\|R_1\| = \|R\| \le 1$. Denote

$$A_1 x := R_1(x) - iR_1(ix), \quad x \in E_1.$$

Then,

- For $x \in E$ we have $A_1 x = \mathrm{Re}(Ax) - i\,\mathrm{Re}(A(ix)) = Ax$ (where we used the fact that A is complex-linear on E).
- A_1 is actually a *complex*-linear functional on E_1:
$$A_1(ix) = R_1(ix) - iR_1(-x) = i(R_1(x) - iR_1(ix)) = iA_1(x).$$
- In fact, $\|A_1\| \le \|R_1\| \le 1$: indeed, for each $x \in E_1$ there exists $\theta \in \mathbb{R}$ such that $A_1 x \in e^{-i\theta}\mathbb{R}_+$. Therefore,
$$|A_1 x| \;=\; A_1(e^{i\theta}x) \;=\; R_1(e^{i\theta}x) \;\le\; \|e^{i\theta}x\| \;=\; \|x\|.$$

The proof is complete. □

Quasi-détour: another ('geometric') proof for Banach spaces over $\mathbb{C}$. In fact, using the same strategy as in the real case, one can reduce the general statement to constructing a 'one-complex-dimensional' extension of a functional $A : E \to \mathbb{C}$, which is possible if and only if

$$\textstyle\bigcap_{x\in E} \overline{B}_{\mathbb{C}}(Ax, \|x + x_1\|) \;\neq\; \emptyset.$$

Recall that in the 'real' case, a similar intersection is non-empty if and only if the corresponding segments intersect pairwise. There exists a (much deeper) analogue of this statement for $\mathbb{C} \cong \mathbb{R}^2$ (and, more generally, $\mathbb{R}^d$), known under the name

- **Helly's theorem:** Let $\{K_\alpha\}_{\alpha\in A}$ be a family of *convex* compacts in $\mathbb{R}^d$. Assume that each $d+1$ compacts taken from this family have a non-empty intersection. Then, $\bigcap_{\alpha\in A} K_\alpha \neq \emptyset$.

Applying this theorem to the family of closed discs $\{\overline{B}_{\mathbb{C}}(Ax, \|x+x_1\|)\}_{x\in E}$, we see that it is enough to prove that, for all $x, y, z \in E$,

$$\overline{B}_{\mathbb{C}}(Ax, \|x+x_1\|) \cap \overline{B}_{\mathbb{C}}(Ay, \|y+x_1\|) \cap \overline{B}_{\mathbb{C}}(Az, \|z+x_1\|) \neq \emptyset.$$

This can be done via the following

- **Lemma:** Let $a_1, a_2, a_3 \in \mathbb{C}$ and $r_1, r_2, r_3 \in \mathbb{R}_+$. Then, the intersection of three discs $\overline{B}_{\mathbb{C}}(a_1, r_1) \cap \overline{B}_{\mathbb{C}}(a_2, r_2) \cap \overline{B}_{\mathbb{C}}(a_3, r_3)$ is non-empty if and only if the following condition holds: for all $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{C}$ such that $\lambda_1 + \lambda_2 + \lambda_3 = 0$ one has $|\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3| \leq |\lambda_1| r_1 + |\lambda_2| r_2 + |\lambda_3| r_3$.

Finally, to see that the required condition holds, it remains to note that

$$\begin{aligned} |\lambda_1 Ax + \lambda_2 Ay + \lambda_3 Az| &\leq \|\lambda_1 x + \lambda_2 y + \lambda_3 z\| \\ &= \|\lambda_1(x+x_1) + \lambda_2(y+x_1) + \lambda_3(z+x_1)\| \\ &\leq |\lambda_1|\cdot\|x+x_1\| + |\lambda_2|\cdot\|y+x_1\| + |\lambda_3|\cdot\|z+x_1\|. \quad \square \end{aligned}$$

Let us now discuss two corollaries of the Hahn–Banach theorem.

(1) Recall that we have an embedding of a Banach space into its second dual:

$$\iota : E \hookrightarrow (E')', \quad x \mapsto X : (E' \ni A \mapsto Ax \in \mathbb{R} \text{ or } \mathbb{C}).$$

Clearly, $\|X\| = \sup_{A\in E'} \|X(A)\|/\|A\| = \sup_{A\in E'} \|Ax\|/\|A\| \leq \|x\|$. The Hahn–Banach theorem implies that ι *is an isometry:* $\|X\| = \|x\|$.

Indeed, consider a linear functional $\lambda x \mapsto \lambda\|x\|$ defined on the one-dimensional subspace $\{\lambda x, \lambda \in \mathbb{R} \text{ or } \mathbb{C}\}$ of E and denote by A_x its lift onto E such that $\|A_x\| = 1$. Then, $\|X\| \geq \|X(A_x)\|/\|A_x\| = \|A_x x\| = \|x\|$.

(2) Consider the following subspace of ℓ^∞:

$$C := \{x = (x_n)_{n\in\mathbb{N}} : \text{ there exists a limit } x_* = \lim_{n\to\infty} x_n\} \subset \ell^\infty.$$

The linear functional $x \mapsto x_*$ is bounded (and has norm 1) on C. Therefore, applying the Hahn–Banach theorem, one can lift this functional to a functional $L : \ell^\infty \to \mathbb{R}$ (or $\mathbb{C}$) such that $Lx \leq \|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|$ and that $Lx = \lim_{n\to\infty} x_n$ provided that the sequence $(x_n)_{n\in\mathbb{N}}$ converges. In other words, we obtain a construction of a 'limit' of *all bounded sequences* that respects the linear structure (i.e., $L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$) and coincides the usual limit on convergent sequences.

In particular, this proves that $((\ell^1)')' = (\ell^\infty)' \supsetneq \ell^1$.

November 16, 2020

We start with a couple of further remarks on a linear functional $L \in (\ell^\infty)'$ such that $Lx = \lim_{n\to\infty} x_n$ for convergent sequences $x \in C \subset \ell^\infty$ (and $\|L\|_{(\ell^\infty)'} = 1$).

- The subspace $C \subset \ell^1$ of convergent sequences can be equivalently viewed as the space of continuous functions on the set $\overline{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$, the Alexandrov compactification of $\mathbb{N}$. Note that the topology on $\overline{\mathbb{N}}$ is metrizable, so – by general theory (Riesz's theorem) – all linear functionals on C can be viewed as signed measures on $\overline{\mathbb{N}}$. Clearly, the functional L corresponds to the Dirac measure at the point $\infty \in \mathbb{N}$ (and $\ell^1 \subset C'$ – to measures on $\mathbb{N}$).
- Functionals L obtained as above respect the additive structure but – in general – we do *not* have $L(xy) = (Lx)(Ly)$ or $L(f(x)) = f(L(x))$ for continuous functions f (where $(xy)_n := x_n y_n$ and $(f(x))_n := f(x_n)$). It is worth noting that one *can* construct such functionals (certainly, taking for granted the axiom of choice), they are known under the name *ultra-filters*.

[!] See the *partiel* problems for related concepts/questions. [!]

12. Open mapping (Banach–Schauder) theorem

Let us now discuss one more important theorem on linear operators in Banach spaces, the so-called *open mapping theorem.*

Theorem 12.1 (Banach–Schauder). *Let both E and E_1 be Banach spaces and a bounded linear operator $A \in \mathcal{L}(E, E_1)$ be surjective. Then, A is an open mapping (i.e., $A(U) \subset E_1$ is open for each open set $U \in E$).*

Proof. It is enough (due to linearity) to prove that $B_{E_1}(0, r) \subset A(B_E(0, 1))$ for a certain $r > 0$. Denote $F := \overline{A(B_E(0, 1))}$, the closure of the latter set in E_1. Since A is surjective, we have $E_1 = \bigcup_{n\in\mathbb{N}}(nF)$. Baire's theorem implies that at least one of the closed sets nF has a non-empty interior. By linearity, this means $\operatorname{Int} F \neq \emptyset$.

It is now easy to see that there exists $r > 0$ such that $B_{E_1}(0, r) \subset F$. Indeed, if $B_{E_1}(y, r) \subset F$, then $B_{E_1}(-y, r) \subset F$ and hence $B_{E_1}(0, r) \subset F$ by linearity. However, this is not enough: we now need to replace F back by $B_E(0, 1)$. To this end, apply the following iterative construction:

- Let $\|y\| < r$ and $q < 1$ be such that $\|y\| \le (1-q)^2 r$; denote $y_0 := y/(1-q)^2$.

Since $y_0 \in B_{E_1}(0, r) \subset F$, it can be approximated by points Ax_0, $x_0 \in B_E(0, 1)$, with an arbitrary precision. In particular, we can find $x_0 \in E$ such that

- $\|x_0\| < 1$ and $\|y_0 - Ax_0\| < qr$; denote $y_1 := q^{-1}(y_0 - Ax_0)$.

Since $y_1 \in B_{E_1}(0, r) \subset F$, it can be approximated by points Ax_1, $x_1 \in B_E(0, 1)$, with an arbitrary precision. Repeating this procedure, we can inductively construct a sequence of points $y_n \in E_1$, $x_n \in E$ such that

- $\|x_n\| < 1$, $\|y_n - Ax_n\| < qr$, and $y_{n+1} = q^{-1}(y_n - Ax_n)$.

By construction, we have

$$y_0 = Ax_0 + qy_1 = Ax_0 + qAx_1 + q^2 y_2 = \ldots = A(x_0 + qx_1 + \ldots + q^n x_n) + q^{n+1} y_n.$$

Denote $x_* := \sum_{n\in\mathbb{N}} q^n x_n$ (this is where we use that E is a Banach space) and note that $\|y_n\| \le qr + \|A\|$. Therefore, passing to the limit $n \to \infty$ we arrive at the equality $y_0 = Ax_*$ and hence $y = Ax$, where $x := (1-q)^2 x_*$. Finally, note that $\|x\| = (1-q)^2\|x_*\| \le (1-q)^2 \cdot (1-q)^{-1} < 1$. The proof is complete. $\square$

The open mapping theorem has a number of **corollaries:**

(1) Let $A \in \mathcal{L}(E, E_1)$ be a bijection of Banach spaces. Then, $A^{-1} \in \mathcal{L}(E_1, E)$.
Proof. This trivially follows from Theorem 12.1.

(2) Let two norms $\|\cdot\|$, $\|\cdot\|_1$ be defined on the same vector space E so that it is complete with respect to *each* of them. If $\|x\|_1 \le C\|x\|$ for some $C > 0$ and all $x \in E$, then there exists $C_1 > 0$ such that $\|x\| \le C_1\|x\|_1$ for all $x \in E$.
Proof. Consider $E = (E, \|\cdot\|)$ and $E_1 = (E, \|\cdot\|_1)$. □

(3) Let E, E_1 be Banach spaces and $A : E \to E_1$ be a linear mapping. Then, $A \in \mathcal{L}(E, E_1)$ if and only if its *graph* $G_A := \{(x, Ax), x \in E\}$ is a *closed* linear subspace of the Banach space $E \times E_1$.

Proof. If A is continuous, then G_A is closed in $E \times E_1$. Vice versa, assume that G_A is closed and define $\|x\|_A := \|x\| + \|Ax\|$ for $x \in E$. As $\|x\|_A \ge \|x\|$, in order to prove that $\|x\|_A \le C_A\|x\|$ for some $C_A > 0$, it remains to prove that E is a complete with respect to $\|\cdot\|_A$. Let a sequence $(x_n)_{n\in\mathbb{N}}$ be Cauchy with respect to $\|\cdot\|_A$. Then, both sequences $(x_n)_{n\in\mathbb{N}}$ and $(Ax_n)_{n\in\mathbb{N}}$ are Cauchy (in E and E_1, resp.) and hence converge: $x_n \to x \in E$, $Ax_n \to y \in E_1$. Since G_A is closed, we have $y = Ax$, so $\|x_n - x\|_A \to 0$. □

12.1. **Détour. Self-adjoint operators in Hilbert spaces.** Though this material goes far beyond the scope of our course, we include a *very brief* discussion of this topic, mostly because of its relevance for the mathematical formalism of the quantum mechanics.

In what follows, H is assumed to be a *separable* complex Hilbert space. In fact, all such infinite-dimensional spaces are unitary equivalent to each other: e.g., by choosing an orthonormal basis in H one can isometrically map it onto ℓ^2. However, it is useful not to reduce the discussion to a concrete space (like ℓ^2) since the definition of operators usually comes from a certain physical context.

- $A \in \mathcal{L}(H)$ is called a *bounded self-adjoint operator* in H iff it is symmetric:
$$\langle Ax, y\rangle = \langle x, Ay\rangle \quad \text{for all } x, y \in H.$$
(A finite-dimensional analogue of this notion is Hermitian matrices in $\mathbb{C}^N$.)

The simplest possible example of a bounded self-adjoint operator is the following:

- Let μ be a (σ-finite) Borel measure on $\mathbb{R}$ with a *bounded* support. Consider the space $H = H_\mu := L^2(\mathbb{R}, \mu) = \{f : \mathbb{R} \to \mathbb{C} \text{ s.t. } \int_{\mathbb{R}} |f(\lambda)^2|\mu(d\lambda) < +\infty\}$ and an operator $\Lambda_\mu : f \mapsto \lambda f$. This operator is defined on the whole space $L^2(\mathbb{R}, \mu)$ (this is where we use the fact that μ is supported on a bounded set), is bounded and symmetric:
$$\langle \Lambda_\mu f, g\rangle \;=\; \int_{\mathbb{R}} \lambda f(\lambda)\overline{g(\lambda)}\mu(d\lambda) \;=\; \langle f, \Lambda_\mu g\rangle.$$
(Note that the space $L^2(\mathbb{R}, \mu)$ is finite-dimensional if and only if μ is supported on a finite set, i.e., iff μ is a finite sum of Dirac measures. This is an example of Hermitian matrices with pairwise distinct eigenvalues.)

Clearly, one can also consider (finite or countable) products $H := \prod H_{\mu_k}$ of such $L^2(\mathbb{R}, \mu_k)$ spaces and define the linear operator Λ acting on functions $f = (f_k) \in H$ by $(\Lambda f)_k := \lambda f_k$. Then, $\Lambda \in \mathcal{L}(H)$ provided that the supports of measures μ_k are uniformly (in k) bounded. (Note that in the finite-dimensional case this allows to include multiple eigenvalues into the consideration.)

The following – remarkable and important result – says that this construction in fact represents *all* bounded self-adjoint operators.

- **Spectral theorem:** Let $A \in \mathcal{L}(H)$ be a bounded self-adjoint operator in a separable Hilbert space H. Then, there exists a (at most countable) collection of Borel measures μ_k on $\mathbb{R}$, with uniformly bounded supports, such that A is unitary equivalent to the operator Λ discussed above (i.e., $A = U^{-1}\Lambda U$, where $U : H \to \prod H_{\mu_k}$ is an isomorphism of Hilbert spaces).

However, to consider *bounded* self-adjoint is not enough, e.g., for the mathematical formulation of the quantum mechanics (but also for many other applications). As a motivation for a general notion of self-adjoint operators, let us formulate

- **Stone's theorem** Let $U : \mathbb{R} \to \mathcal{L}(H)$ be a (weakly-)continuous one-dimensional group (i.e., $U(t+s) = U(t)U(s)$ and $U(0) = \mathrm{I}$) of unitary ($\langle U(t)x, U(t)y\rangle = \langle x, y\rangle$) operators in a separable Hilbert space. Then, there exists a (not necessarily bounded!) self-adjoint operator A in H such that $U(t) = e^{itA}$ for all $t \in \mathbb{R}$.

We now give an intrinsic definition of (possibly unbounded) self-adjoint operators:

- A linear operator $A : D(A) \to H$ defined on a *dense* subspace $D(A) \subset H$ is called self-adjoint if
  - A is symmetric on $D(A)$, i.e., $\langle Ax, y\rangle = \langle x, Ay\rangle$ for all $x, y \in D(A)$;
  - the graph $G_A = \{(x, Ax), x \in D(A)\}$ is a closed subspace of $H \times H$;
  - both operators $A \pm i\mathrm{I} : D(A) \to H$ are surjective (*a posteriori,* this also implies that there exist bounded inverses $(A \pm i\mathrm{I})^{-1} \in \mathcal{L}(H)$).

The spectral theorem generalizes in the most natural way:

- **Spectral theorem (general case):** Each self-adjoint operator A in a separable Hilbert space is unitary equivalent to an operator $\Lambda = (\Lambda_{\mu_k})$ discussed above where one *drops* the assumption that the measures μ_k have bounded support and set
$$D(A) := \left\{ f = (f_k) \in \prod H_{\mu_k} : \textstyle\sum_k \int_{\mathbb{R}} (1 + |\lambda|^2)|f_k(\lambda)|^2 < +\infty \right\}.$$

In particular, the spectral theorem allows to *define* unitary operators $e^{itA} \in \mathcal{L}(H)$: their images under the equivalence of H and $\prod H_{\mu_k}$ act as $f = (f_k) \mapsto (e^{it\lambda} f_k)$. (Note that *a priori* it is not clear even that there exists $x \in D(A)$ such that $Ax \in D(A)$, not speaking about the convergence of the series $\sum_{n \in \mathbb{N}} \frac{1}{n!}(itA)^n$; however *a posteriori* it turns out that this can be done on a dense subset of H, which corresponds to functions $f_k \in L^2(\mathbb{R}, d\mu_k)$ having bounded support.) Similarly, one can use the spectral theorem to define operators $g(A)$ for other functions $g : \mathbb{R} \to \mathbb{C}$.

- Let us emphasize that, contrary to the finite-dimensional case, the *spectral measure(s)* μ_k are *not* necessarily discrete. Properties of the spectral measure(s) of a Hamiltonian (viewed as self-adjoint operators in a Hilbert space) of a given physical system are very important in the quantum mechanics.

Obviously, all that is a subject of another course.

---

13. Additional material: partiél homework

Because of the lockdown, this year the *partiél* is not an intermediate test but a week-long homework (dead-line: November 22). It will be corrected but *not graded.*

13.1. **Weak topology on ℓ^1.**

0. Recaps: $(\ell^1)' \cong \ell^\infty$. For $v \in \ell^\infty$, denote by $\phi(v)$ the linear functional acting on $u \in \ell^1$ as $\phi(v)u := \sum_{n\in\mathbb{N}} v_k u_k$.

(i) Prove that $\phi(v) \in (\ell^1)'$ for all $v \in \ell^\infty$ and that $\|\phi(v)\|_{(\ell^1)'} = \|v\|_{\ell^\infty}$.

(ii) Prove that $\phi : \ell^\infty \to (\ell^1)'$ is a bijection and hence an isometry.

1. Weak topology on ℓ^1: definition, weak closure of the unit sphere. Let $F \subset (\ell^1)'$ be a finite set. For $u \in \ell^1$ and $r > 0$, denote

$$B_F(u,r) \;:=\; \{u' \in \ell^1 : |f(u'-u)| < r \text{ for all } f \in F\}.$$

(i) Check that the collection of sets $\{B_F(u,r) : u \in \ell^1,\ r>0,\ F \text{ - finite subset of } (\ell^1)'\}$ can serve as a base of a topology. This topology is called the *weak topology on* ℓ^1.

(ii) Prove that this topology is indeed weaker=coarser that the standard (also called strong) topology in the space ℓ^1, which is defined by the norm $\|\cdot\|_1$.

Let $S := \{u \in \ell^1 : \|u\|_1 = 1\}$ be the unit sphere in ℓ^1 and denote by $\overline{S}^{\mathrm{w}}$ its closure in the *weak* topology. We want to show that

$$\overline{S}^{\mathrm{w}} = \overline{B} := \overline{B}_{\ell^1}(0,1) = \{u \in \ell^1 : \|u\|_{\ell^1} \le 1\}.$$

(Note that, trivially, S is a closed set in the strong topology, i.e., $\overline{S} = S$.)

(iii) Let $u \in \overline{B}$. Prove that each open – in the weak(!) topology – neighborhood of u contains an affine line passing through u (i.e., for each $u \in U$ - open in the weak topology, there exists $v \in \ell^1$ such that $u + v\mathbb{R} \subset U$). Conclude that $\overline{B} \subset \overline{S}^{\mathrm{w}}$.

(iv) Let $u \in \ell^1$ be such that $\|u\|_1 > 1$. Prove that there exists $v \in \ell^\infty$ such that $\|v\|_\infty = 1$ and $\phi(v)u > 1$. Conclude that $\overline{S}^{\mathrm{w}} \subset \overline{B}$.

2. Convergent sequences: Schur's property of ℓ^1. Let $u^{(n)}, u \in \ell^1$. We write $u^{(n)} \rightharpoonup u$ as $n \to \infty$ if the sequence $u^{(n)}$ converges to u in the weak topology, and keep the notation $u^{(n)} \to u$ for the usual convergence $\|u^{(n)} - u\|_1 \to 0$. Clearly, $u^{(n)} \to u$ implies that $u^{(n)} \rightharpoonup u$. We now prove that these notions *are equivalent.* (By linearity, we can assume that $u = 0$ without loss of generality.)

(i) Prove that $u^{(n)} \rightharpoonup u$ as $n \to \infty$ if and only if $f(u^{(n)}) \to f(u)$ for all $f \in (\ell^1)'$.

(ii) Recall that $(\ell^1)' \cong \ell^\infty$ and define $d(v,w) := \sum_{k\in\mathbb{N}} 2^{-k}|v_k - w_k|$ for $v, w \in \ell^\infty$. Prove that the set $\overline{B}^* := \{v \in \ell^\infty : \|v\|_\infty \le 1\}$ is compact in the metric space (ℓ^∞, d).

(iii) Let $v^{(m)} \in \overline{B}^*$, $m \to \infty$. Prove that the following statements are equivalent:

(a) $d(0, v^{(m)}) \to 0$; (b) $v_k^{(m)} \to 0$ for all $k \in \mathbb{N}$; (c) $\phi(v^{(m)})u \to 0$ for all $u \in \ell^1$.

(iv) Let $\ell^1 \ni u^{(n)} \rightharpoonup 0$ be a weakly convergent sequence. Given $\varepsilon > 0$, denote

$$F_N := \{v \in \overline{B}^* : |\phi(v)u^{(n)}| \le \varepsilon \text{ for all } n \ge N\}.$$

Check that F_N is a closed subset of the complete metric space $(\overline{B}^*, d)$. Prove that there exists $N \in \mathbb{N}$ and $\rho > 0$ such that $F_N \supset \{v \in \overline{B}^* : d(0,v) < \rho\}$.

(v) Conclude that the weak convergence $u^{(n)} \rightharpoonup 0$ implies that $\|u^{(n)}\|_1 \to 0$.

Remark. Note that though – <u>in the space ℓ^1</u> – the notions of convergent sequences $u^{(n)} \rightharpoonup u$ and $u^{(n)} \to u$ are equivalent, the two topologies are very different: e.g., as discussed above $\overline{S}^{\mathrm{w}} = \overline{B} \supsetneq S = \overline{S}$. In particular, the *weak topology on ℓ^1 is not metrizable* and, moreover, not first-countable: otherwise, we would have $\overline{S}^{\mathrm{w}} = \overline{S}$. One can also prove this directly relying upon the fact that ℓ^∞ is not separable.

13.2. **Stone–Čech compactification.** Let X be a *Tychonoff space,* this means that X satisfies the separation axioms (T_1) and

$$(T_{3\frac{1}{2}}): \quad \begin{array}{l} \text{for each } x \in X \text{ and a closed set } F \subset X \text{ s.t. } x \notin F \\ \text{there exists a continuous function } f : X \to [0,1] \\ \text{such that } x \in f^{-1}(\{0\}) \text{ and } F \subset f^{-1}(\{1\}). \end{array}$$

0. Recaps, basics.

(i) Quote (T_1), (T_2), (T_4), and Urysohn's lemma; argue that *normal Hausdorff spaces* (i.e., topological spaces satisfying (T_1) and (T_4)) are Tychonoff spaces.

(ii) Prove that *compact Hausdorff spaces* (='compacts en français' = compact topological spaces satisfying (T_2)) are normal Hausdorff (i.e., satisfy (T_4)).

Definition. A *Stone–Čech compactification* of a topological space X is a compact Hausdorff space βX and a continous mapping $\iota = \iota_X : X \to \beta X$ such that the following universal property holds:

> for each continuous mapping $f : X \to K$ of the space X to a compact Hausdorff space K there exists a unique continuous mapping $\beta f : \beta X \to K$ such that $\beta f \circ \iota_X = f$.

(iii) Prove that the space βX, if exists, is unique up to a homeomorphism.

(iv) Provided X is a Tychonoff space, prove that $\iota : X \to \iota(X)$ has to be a bijection and, moreover, a *homeomorphism* (the topology on $\iota(X)$ is induced from βX). [*Hint:* For the latter, prove that $\iota(F)$ is closed in $\iota(X)$ if F is closed in X.]

Informally speaking, we aim to homeomorphically embed a Tychonoff space X into a (huge) compact Hausdorff space such that all continuous functions $f : X \to K$ admit(!) a unique(!) continuation onto this bigger space.

Our first goal is to prove that, for Tychonoff spaces X, the Stone–Çech continuation βX *exists* by giving an explicit construction based upon Tychonoff's theorem.

1. Construction. Denote $C_X := C(X; [0,1])$, the space of all continuous functions from X to $[0,1]$, and consider a *Tychonoff cube*

$$K(C_X) := [0,1]^{C_X} = \prod_{g \in C_X} [0,1] = \{ \, \Phi : C_X \ni g \mapsto \Phi(g) \in [0,1] \, \}$$

(where one does *not* assume any property of Φ), equipped with the usual product topology. Recall that $K(C_X)$ is a compact space due to the Tychonoff theorem.

Now consider the mapping $I : X \to K(C_X)$, $x \mapsto I(x)$, where the *evaluation functional* $I(x) \in K(C_X)$ is defined as follows: $[I(x)](g) := g(x)$.

(i) Prove that $I : X \to I(X)$ is a bijection. Recall the definition of the topology on $I(X) \subset K(C_X)$ and prove that I is continuous.

(ii) Using the fact that X is a Tychonoff space, prove that $I : X \to I(X) \subset K(C_X)$ is a homeomorphism.

Let us *define* $\beta X := \overline{I(X)}^{K(C_X)}$, the closure of $I(X)$ in the topology of $K(C_X)$.

(iii) Argue that thus defined βX is a compact Hausdorff space.

2. Universal property. Assume now that $f : X \to K$ is a continuous function, where K is a compact Hausdorff space. We aim to prove that there exists a unique continuous function $\beta f : \beta X \to K$ s.t. $f = \beta f \circ I$, where βX is constructed above.

(i) Argue that such a function $\beta f : \beta X \to K$, if exists, is unique.

(ii) Consider first the case $K = [0,1]$. Given a continuous function $f : X \to [0,1]$, prove that the function $\beta f : K(C_X) \to [0,1]$, $\Phi \mapsto \Phi(f)$, is continuous and that its restriction onto $\beta X \subset K(C_X)$ satisfies the required property $f = \beta f \circ I$.

(iii) Now consider the case when $K = K(A) = [0,1]^A$ is also a Tychonoff cube (we do <u>not</u> make any assumption on A here). Again, prove that each continuous function $f : X \to K(A)$, $x \mapsto f(x) : A \to [0,1]$, admits a *continuous(!)* extension $\beta f : K(C_X) \to K(A)$ defined as follows: $[(\beta f)(\Phi)](\alpha) = \Phi(f(\cdot)(\alpha))$.

(iv) Finally, argue that each compact Hausdorff space K can be homeomorphically embedded into a Tychonoff cube $K(A)$ with $A = C_K$ and conclude the proof.

3*. Bonus: ultrafilters on $\mathbb{N}$. It is easy to see that $\beta X \cong X$ if X is a finite set. Let us consider the simplest nontrivial example: $X = \mathbb{N}$ (equipped with the discrete topology). Recall that $\mathcal{U} \subset 2^{\mathbb{N}} \smallsetminus \{\emptyset\}$ is called a (proper) *ultra-filter* if

- for each $Y \subset \mathbb{N}$ either $Y \in \mathcal{U}$ or $\mathbb{N} \smallsetminus Y \in \mathcal{U}$ (but not both);
- if $Y \in \mathcal{U}$, then $Y' \in \mathcal{U}$ for all larger sets $Y \subset Y' \subset \mathbb{N}$;
- if $Y_1, Y_2 \in \mathcal{U}$, then $Y_1 \cap Y_2 \in \mathcal{U}$.

Prove that the Stone–Čech compactification $\beta\mathbb{N}$ of $\mathbb{N}$ is homeomorphic to the set of all proper ultra-filters on $\mathbb{N}$ equipped with the *Stone topology*, i.e., the topology generated by the base sets $\mathcal{W}_Y := \{\mathcal{U} : Y \in \mathcal{U}\}$, where Y runs over all subsets of $\mathbb{N}$. The inclusion $\iota : \mathbb{N} \hookrightarrow \beta\mathbb{N}$ is defined as $n \mapsto \mathcal{U}_n := \{Y \subset \mathbb{N} : n \in Y\}$.

(i) Check that the Stone topology is correctly defined (i.e., that one can use the family $\{\mathcal{W}_Y\}_{Y \subset \mathbb{N}}$ as a base set to define a topology) and that it is Hausdorff.

(ii) Prove that thus defined topological space is compact. [*Hint:* This is equivalent to the following statement: given a family of sets Y_α such that no finite sub-family covers $\mathbb{N}$, one can define an ultrafilter on $\mathbb{N}$ that does not contain any of Y_α.]

(iii) *Ultra-limits:*

> let $\mathcal{U} \subset 2^{\mathbb{N}} \smallsetminus \{\emptyset\}$ be a proper ultra-filter and $(x_n)_{n \in \mathbb{N}}$ be a sequence of elements of K. We say that $\mathcal{U}\text{-}\lim x_n = x$ if for each open neighborhood $V_x \subset K$, the following holds: $\{n \in \mathbb{N} : x_n \in V_x\} \in \mathcal{U}$.

Prove that $\mathcal{U}\text{-}\lim x_n$ is unique provided that the topological space K is Hausdorff and that $\mathcal{U}\text{-}\lim x_n$ exists (for *all* sequences x_n) provided that K is compact.

(iv) Let K be a compact Hausdorff space, $f : \mathbb{N} \to K$, and $\mathcal{U}$ be a proper ultrafilter on $\mathbb{N}$. Define $(\beta f)(\mathcal{U}) := \mathcal{U}\text{-}\lim f(n)$ and prove that $\beta f : \beta\mathbb{N} \to K$ is a continuous extension of f from $\mathbb{N}$ to the topological space $\beta\mathbb{N}$ of all proper ultra-filters on $\mathbb{N}$.
