Topics in analysis

September 21, 2015

Remark. This is the live-texed notes of Topics in analysis course held by Tamás Tasnádi at CEU in the fall trimester of 2014. Every error and typo in the text is mine.

First Lecture, 25th of September

1 Metric spaces, topological properties

In the past, we have met the following spaces in analysis (listed in the order of generality): $\mathbb{R}$, $\mathbb{R}^n$, metric spaces (X, d), topological spaces (X, τ). Similarly, in functional analysis, we have met Hilbert spaces, Banach spaces and normed spaces. This course will deal with these notions as well. So first, we repeat the basics.

1.1 Metric spaces

Definition 1.1. A set X together with a function $d : X \times X \to \mathbb{R}$ is a *metric space*, if

1. $d(x, y) \geq 0$ and $d(x, y) = 0$ if and only $x = y$ for all $x, y \in X$
2. (Symmetry) $d(x, y) = d(y, x)$ for all $x, y \in X$
3. (Triangle inequality) $d(x, y) + d(y, z) \geq d(x, z)$ for all $x, y, z \in X$

Example 1.2. .

1. Euclidean distance: $X = \mathbb{R}^n$ where $n \in \mathbb{N}_+$. In this case the distance is

$$d(\underline{x}, \underline{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

2. Let $X = \mathbb{R}^n$ and $1 \leq p < \infty$. Then we get a metric by the formula

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

   For $p = 1$ it is called the Manhattan or (taxicab) distance. Similarly, for $X = \mathbb{R}^n$ and $p = \infty$ we can define $d_\infty(\underline{x}, \underline{y}) = \lim_{p\to\infty} d_p(\underline{x}, \underline{y}) = \max_{i=1,\dots,n}\{|x_i - y_i|\}$ (Note that all the above are defined by the aid of a norm.)

3. Let (M, g) be a Riemannian manifold and $x, y \in M$. The function

$$d(x, y) = \{\text{the length of the shortest path between } x \text{and } y\}$$

   is a distance function. This is an example for a metric, not given by a norm.

4. Let $X = \mathcal{C}[a,b] = \{f : [a,b] \to \mathbb{R} \mid f \text{ is continuous}\}$ and $||f|| = \sup_{x\in[a,b]}\{|f(x)|\}$. Then by this norm we can define the metric $d(f,g) = ||f-g||$.

Notation: Let (X,d) be a metric spaces. Then for the ε-radius balls we use the notation $B_\varepsilon(x) := \{y \in X \mid d(y,x) < \varepsilon\}$.

Reminder: (Even though the following notions were defined precisely on the lecture, I omit them since as I guessed it was not new for the audience) basic notions in the topic of metric spaces are listed, such as: limit, Cauchy-sequence, completeness, interior/exterior/boundary point of a subset, accumulation and isolated points, open and closed sets, closure, continuity (both by $\varepsilon - \delta$ and by open sets), compactness and sequential compactness.
In addition, some basic statements are collected, namely: union of open sets is open, the two notions of continuity are equivalent.

1.2 Topological spaces

Definition 1.3. A *topology* on the set X is $\tau \subseteq \mathcal{P}(X)$, if

1. $\emptyset, X \in \tau$
2. τ is closed under finite intersection
3. τ is closed under arbitrary union.

Definition 1.4. Let (X,τ_1) and (X,τ_2) be two topological spaces. If $\tau_1 \subseteq \tau_2$ then we say that τ_1 is *coarser* (or *weaker*) than τ_2 and τ_2 is *finer* (or *stronger*) than τ_1.

Remark 1.5. The weakest topology is the anti-discrete topology $\{\emptyset, X\}$ and the finest is the discrete topology: $\tau = \mathcal{P}(X)$.

Proposition 1.6. *If $\tau_1 \subseteq \tau_2$ and $A \subseteq X$ is compact in τ_2 then it is also compact in τ_1. (see Homework problem 1/2d.)*

Definition 1.7. (X,τ) is *Hausdorff*, if "any two points can be separated by open subsets" i.e., for all $x, y \in X$, $x \neq y$ there exists $U, V \in \tau$ such that $x \in U$, $y \in V$ and $U \cap V \neq \emptyset$.

Proposition 1.8. *If $\tau_1 \subseteq \tau_2$ and (X,τ_1) is Hausdorff then (X,τ_2) is Hausdorff. (see Homework problem 1/2e.)*

Definition 1.9. Let X be a topological space and $A \subseteq X$. Then the *relative topology* (or subspace topology) on A is $\tau_A = \{O \cap A \mid O \in \tau\}$

Remark 1.10. Similarly, we can define product- and factor-topology.

Definition 1.11. Let X be a set, (Y_α, τ_α) be topological spaces for all $\alpha \in I \neq \emptyset$ and let $f_\alpha : X \to Y_\alpha$ be maps for all $\alpha \in I$. Then we can define the *(weak) induced topology* on X: it is the weakest topology on X such that all the f_α's are continuous (for all $\alpha \in X$).

1.3 Compactness in metric spaces

Definition 1.12. A metric space (X,d) is *bounded*, if there exists an $x_0 \in X$ and an $R \in \mathbb{R}$ such that $B_R(x_0) = X$.

Definition 1.13. The metric space (X,d) is *totally bounded*, if for all $\varepsilon > 0$ there is a finite number of balls of radius ε covering on X.

Remark 1.14. These are notions that has no complete analog in the case of topological spaces.

Definition 1.15. Let (X,d) be a metric spaces and $\{U_i\}_{i\in I}$ be an open cover of X. We say that $\varepsilon > 0$ is a *Lebesgue number* of that cover, if for all $x_0 \in X$ there exists $i \in I$ such that $B_\varepsilon(x_0) \subseteq U_i$.

Theorem 1.16. *Let (X,d) be a metric space. Then X is compact (i.e. X has the Heine-Borel property) if and only if X is sequentially compact (i.e. X satisfies the Bolzano-Weierstrass property).*

Remark 1.17. For general topological spaces, this is not true in any direction, only with some countability assumptions.

Proof. $\Rightarrow$: Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence and assume indirectly that it has no convergent subsequence. It is equivalent to saying that it has no accumulation point in X. I.e. For all $y \in X$ there exists an $\varepsilon(y) > 0$ such that $B_{\varepsilon(y)}(y) \cap \{x_n\}_{n\in\mathbb{N}}$ is finite. We know that $\bigcup_{y\in X} B_{\varepsilon(y)}(y) = X$ is an open cover of X. By the compactness of X, we get that there exist $y_1,\dots,y_n$ such that $\bigcup_{i=1}^n B_{\varepsilon(y_i)}(y_i) = X$. However, the left hand side can contain only finitely many points and that is a contradiction.

To prove $\Leftarrow$ we need some lemmas:

Lemma 1.18. *If (X,d) is a sequentially compact metric space then it is totally bounded.*

Proof. Assume indirectly that X is not totally bounded, i.e. there exists an $\varepsilon > 0$ such that X cannot be covered by a finite number of epsilon-balls. We construct a sequence that has no convergent subsequence: choose $x_1 \in X$ arbitrarily. By induction, we can choose $x_{n+1} \in X \setminus \bigcup_{i=1}^n B_\varepsilon(x_i)$ where the right hand side is nonempty by the indirect assumption. This sequence cannot have a convergent subsequence because none of its subsequence can be Cauchy. ☐

Lemma 1.19. *If (X,d) is sequentially compact then any open cover of X has positive Lebesgue-number.*

Proof. Indirectly assume that $\{U_i\}_{i\in I}$ does not have a positive Lebesgue-number. For all $n \in \mathbb{N}_+$ there exists an x_n such that $B_{\frac{1}{n}}(x_n)$ is not contained in any U_i $(i \in I)$. X is sequentially compact so $\{x_n\}_{n\in\mathbb{N}}$ has a subsequence $\{y_n\}_{n\in\mathbb{N}}$ that converges to $y \in X$. By the assumptions we know hat $B_{\frac{1}{n}}(y_n)$ is not contained in any U_i. We know that U_i is an open cover of X so $y \in U_j$ for some j. Moreover, since U_j is open, there exists an $\varepsilon > 0$ such that $B_\varepsilon(y) \subseteq U_j$. So for $\frac{\varepsilon}{2}$ there is an $N \in \mathbb{N}$ such that for all $n > N$: $y_n \in B_{\frac{\varepsilon}{2}}(y)$. We can of course assume that N is bigger than $\frac{2}{\varepsilon}$, but that means that

$$B_{\frac{1}{n}}(y_n) \subseteq B_{\frac{1}{N}}(y_n) \subseteq B_{\frac{\varepsilon}{2}}(y_n) \subseteq B_\varepsilon(y) \subseteq U_j$$

and that is a contradiction. ☐

$\Leftarrow$: Let $\{U_i\}_{i\in I}$ be an open cover of X. Since X is sequentially compact, by Lemma 1.18 $\{U_i\}_{i\in I}$ has a Lebesgue number $\varepsilon > 0$ and by Lemma 1.19 X is totally bounded i.e. there exist $x_1,\dots,x_n$ such that $\bigcup_{k=1}^n B_\varepsilon(x_k) = X$ where ε is the previously chosen one. Now, for each $1 \le k \le n$ we can choose a U_{i_k} such that $U_{i_k} \supseteq B_\varepsilon(x_k)$. Then $\bigcup_{k=1}^n U_{i_k} \supseteq \bigcup_{k=1}^n B_\varepsilon(x_k) = X$ and that was the statement. ☐

Theorem 1.20. (Heine-Borel theorem for metric spaces) *Let (X,d) be a metric space. Then X is compact if and only if X is complete and totally bounded.*

Proof. $\Rightarrow$: Assume that X is compact but not complete. So there exists a Cauchy-sequence $\{x_n\}_{n\in\mathbb{N}}$ that has no limit point. By Theorem 1.16 we know X is also sequentially compact so $\{x_n\}_{n\in\mathbb{N}}$ has a convergent subsequence with limit point $x \in X$. However, if a Cauchy-sequence has a convergent subsequence then it is convergent so that is a contradiction. Totally boundedness follows from Lemma 1.18 and Theorem 1.16.

$\Leftarrow$: Assume that X is complete and totally bounded. We show that X is sequentially compact (that is enough by 1.16). Let $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ be a sequence. For each $i \in \mathbb{N}_+$ X has a finite cover of balls of radius $\frac{1}{i}$ namely $\bigcup_{j=1}^{n_i} B_{\frac{1}{i}}(x_j^{(i)}) = X$. Then at least one of the balls contain infinitely many points of $\{x_n\}_{n\in\mathbb{N}}$, so we can take one x_{n_1} arbitrarily from it. Similarly, we can define the whole subsequence by induction: if $x_{m_1},\dots,x_{m_k}$ are already given, then take a cover of the space of balls of radius $\frac{1}{k+1}$. One of the sets $B_{\frac{1}{k+1}}(x_j^{(k+1)}) \cap B_{\frac{1}{k}}(x_{m_k})$ (for all $j = 1,\dots,n_j$) contains infinitely many points of the sequence, so we can choose an $x_{m_{k+1}}$ in it. This way, we obviously get a Cauchy-sequence (we made it that way), which is convergent by the assumption of completeness. $\square$

Theorem 1.21. (Heine-Borel for Euclidean spaces) *Let $(\mathbb{R}^n, d)$ be a Euclidean space ($n \in \mathbb{N}$). Then $A \subseteq \mathbb{R}^n$ is compact if and only if A is closed and bounded.*

Proof. see Homework 1/3. $\square$

Theorem 1.22. Bolzano-Weierstrass theorem(s)*:*

1. *version: (on $\mathbb{R}$) Every bounded sequence in $\mathbb{R}$ has a convergent subsequence. (The proof by peaks was explained here.)*
2. *version: (in $\mathbb{R}^n$) Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*
3. *version: (in a metric space) Every sequence in a complete, totally bounded metric space has a convergent subsequence.*

Second Lecture, 2nd of October

1.4 Completion of a metric space

Motivation: the same procedure as in the case of $\mathbb{Q}$ and $\mathbb{R}$ can be carried over to the general metric case.

Definition 1.23. A *completion* of a metric space (X, d) is a pair $\big(\varphi, (X^*, d^*)\big)$ where $\varphi : X \to X^*$ is an isometry with a dense image, i.e. $\overline{\varphi(X)} = X^*$.

Theorem 1.24. *Every metric space (X, d) has a completion.*

Proof. (Constructive one) Let $C(X)$ denote the set of all Cauchy-sequences in X. We introduce an equivalence relation: $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ are equivalent if and only if $\lim_{n\to\infty} d(x_n, y_n) = 0$. We should check the followings: the equivalence is reflexive, symmetric, transitive. (Details omitted.) So we can define $X^* = C(X)/\sim$ and

$$d^*\big([(x_n)_{n\in\mathbb{N}}], [(y_n)_{n\in\mathbb{N}}]\big) = \lim_{n\to\infty} d(x_n, y_n)$$

To get a metric space this way, we have to check the followings: d^* is defined (i.e. the limit exists), d^* is well defined (i.e. it is independent of the choice of representative), d^* is a metric. (Details omitted.)

It is left to determine φ: the natural idea works, i.e.

$$\varphi : X \to X^*$$

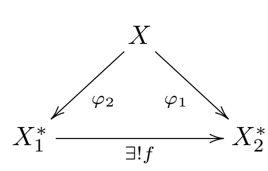
$$x \mapsto (x, x, x, \dots)$$

but we still have to check that φ is an isometry and $\operatorname{Im}\varphi$ is dense in X^*. The first is trivial by the definitions. The latter is immediately given by the construction: $\varphi\big((x_n)_{n\in\mathbb{N}}\big)$ converges to $[(x_n)_{n\in\mathbb{N}}]$. The only missing part now is that X^* is complete but the following lemma solves this problem as well.

Lemma 1.25. *Let A be a dense subset in a metric space (Y, ρ). Then (Y, ρ) is complete if and only if every Cauchy sequence in A converges in Y. (The proof is more or less straightforward.)*

Fixing A to be $\varphi(X)$ in $Y = X^*$ gives the theorem. $\square$

Theorem 1.26. *The completion of a metric space is unique up to metric isomorphism, i.e. if $\big(\varphi_1, (X_1^*, d_1^*)\big)$ and $\big(\varphi_2, (X_2^*, d_2^*)\big)$ are two completion of (X, d) then there is a unique surjective isometry of $f : X_1^* \to X_2^*$ such that*



is commutative.

1.5 Heine-Cantor theorem

Definition 1.27. Let (X, d) be a metric space, $f : X \to \mathbb{R}$ is a function f *uniformly continuous* on X, if

$$\forall \varepsilon > 0 \; \exists \delta(\varepsilon) \text{ such that } \forall x, y \in X \; \Big(d(x, y) < \delta(\varepsilon) \Rightarrow |f(x) - f(y)| < \varepsilon\Big)$$

Remark 1.28. In the case of continuity, the order of quantors are "totally different": in that case δ depends on x as well.

Theorem 1.29. (Heine-Cantor) *On a compact metric space, every real continuous function is uniformly continuous.*

Theorem 1.30. (Heine) *On closed, bounded interval, every continuous function is uniformly continuous. (well-known special case)*

2 Normed Linear spaces

Definition 2.1. Let X be a real or complex vector space (the field will be denoted by $\mathbb{K}$). $||.|| : X \to \mathbb{R}_{\geq 0}$ is a *norm* on X, if

1. $||x|| = 0 \iff x = 0 \in X$
2. (Positive homogenity) For all $\alpha \in \mathbb{K}$ and for all $x \in X$: $||\alpha x|| = |\alpha| \cdot ||x||$
3. (Triangle-inequality) For all $x, y \in X$: $||x + y|| \leq ||x|| + ||y||$

Remark 2.2. The induced metric of the norm is $d(x, y) = ||x - y||$.

Definition 2.3. The normed space $(X, ||.||)$ is a *Banach space* if it is complete with respect to the norm.

Example 2.4.

1. $(\mathbb{K}^d, ||.||_p)$ where $1 \leq p < \infty$ and $||x||_p = \left(\sum_{j=1}^{d} |x_j|^p\right)^{\frac{1}{p}}$
2. $(\mathbb{K}^d, ||.||_\infty)$ where $||x||_\infty = \max_j |x_j|$
3. Similarly, the spaces $\ell^p_{\mathbb{R}}$ and $\ell^p_{\mathbb{C}}$ of infinite sequences with the p-norm can be defined.
4. Let $(X, \mathcal{A}, \mu)$ be a measure space and $1 \leq p < \infty$. Then we get a normed space by

$$\mathcal{L}^p_{\mathbb{K}}(X, \mu) = \Big\{\big[f\big] \mid f : X \to \mathbb{K} \text{ measureable } \int_X |f|^p \mathrm{d}\mu < \infty\Big\}$$

$$\text{with the norm} \quad ||[f]|| = \left(\int_X |f|^p \mathrm{d}\mu\right)^{\frac{1}{p}}$$

   where $[f]$ means the equivalence class of functions with respect to the 'almost everywhere' equivalence.

5. Similarly, if $(X, \mathcal{A}, \mu)$ is a measure space. Then in the case of $p = \infty$ we can again get a normed space:

$$\mathcal{L}^p_{\mathbb{K}}(X, \mu) = \Big\{\big[f\big] \mid f : X \to \mathbb{K} \text{ measureable } \inf_{c>0}\{c \mid \mu(f > c) = 0\} < \infty\Big\}$$

$$\text{with the norm} \quad ||[f]|| = \inf_{c>0}\{c \mid \mu(f > c) = 0\}$$

Theorem 2.5. (Minkowski inequality – Triangle inequality for the p-norms) *Let* $1 \leq p \leq \infty$ *and* $x, y \in \mathcal{L}^p(X, \mu)$ *then*

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

Theorem 2.6. (Hölder inequality) *Let* $1 \leq p \leq \infty$, $x \in \mathcal{L}^p(X, \mu)$ *and* $y \in \mathcal{L}^p(X, \mu)$ *where* $\frac{1}{p} + \frac{1}{q} = 1$. *Then*

$$\int_X |xy| \mathrm{d}\mu \leq ||x||_p \cdot ||y||_q$$

Remark 2.7. The proof of Minkowski inequality uses the Hölder inequality.

Definition 2.8. The space of continuous functions on an interval is defined as

$$\mathcal{C}_{\mathbb{K}}[a, b] := \big\{ f : [a, b] \to \mathbb{K} \mid f \text{ is continuous} \big\}$$

With the *supremum norm* $||f|| = \sup\big\{ |f(x)| \mid x \in [a, b] \big\}$

These are, in some sense, the most important examples of Banach spaces.

2.1 Finite dimensional normed vector spaces

Proposition 2.9.

1. *Every finite dimensional normed spaces is complete.*
2. *On a finite dimensional normed space all norms are equivalent (i.e. they induce the same topology).*

Definition 2.10. Two norms $||.||_1$ and $||.||_2$ are *equivalent* if there exists $k, K \in \mathbb{R}_{>0}$ such that for all $x \in X$ $k||x||_2 \leq ||x||_1 \leq K||x||_2$. (It can be easily checked that it is an equivalence relation.)

Lemma 2.11. *If* $||.||_1$ *and* $||.||_2$ *are equivalent on* X *then they induce the same topology (in fact the converse is also true for normed spaces) and* $(X, ||.||_1)$ *is complete if and only is* $(X, ||.||_2)$ *is complete.*

Proof. By definition, every open ball of $||.||_1$ contains an open ball of $||.||_2$. The first statement follows. To see the second statement, first note that the Cauchy sequences are the same in the two norms. Therefore, because of the equivalence of topology, the limits are the same. □

Proposition 2.12. *Let* X *be a finite dimensional normed space and consider a basis* $\{b_1, b_2, \ldots, b_d\} \subseteq X$. *For arbitrary* $X \ni x = \sum_{i=1}^{d} \alpha_i b_i$ *where* $(\alpha_i \in \mathbb{K})$ *let the norm* $||.||_0$ *be defined as*

$$||x||_0 = \sqrt{\sum_{i=1}^{d} |\alpha_i|^2}$$

Then $(X, ||.||_0)$ *is complete.*

Proof. Let $x_n = \sum_{i=1}^{d} \alpha_i^{(n)} b_i$ be a Cauchy sequence in $(X, ||.||_0)$. Then the sequence of coordinates must be Cauchy sequences as well because $|\alpha_k^{(n)} - \alpha_k^{(m)}| \leq ||x_n - x_m||_0$. The normed space $\mathbb{K}$ is complete (known fact) so there exist the $\lim_{n\to\infty} \alpha_k^{(n)} = \beta_k$ for all $k = 1, \ldots, d$. Now, consider the candidate for being the limit $y = \sum_{i=1}^{d} \beta_i b_i$. It is indeed the limit in $||.||_0$ because

$$||x_n - y||_0 = \sqrt{\sum |\alpha_i^{(n)} - \beta_i|^2} \leq \sqrt{d\varepsilon^2} = \sqrt{d}\varepsilon$$

if $n > \max\{N(k, \varepsilon) \mid k = 1, \ldots, d\}$. □

Theorem 2.13. *Let X be a finite dimensional vector space over $\mathbb{K}$. Let $\{b_i\}_{i=1}^d \subseteq X$ be a basis and $||.||_0$ as before. Then for an arbitrary norm $||.||$, the previous $||.||_0$ is equivalent to $||.||$.*

Proof.

$$||x|| = ||\sum_{i=1}^{d} \alpha_i b_i|| \overset{\triangle}{\leq} \sum_{i=1}^{d} |\alpha_i| \cdot ||b_i|| \overset{\text{Hölder}}{\leq} ||\alpha||_2 \cdot \sqrt{\sum_{i=1}^{d} ||b_i||^2} = ||x||_0 \cdot K$$

with the definition $K = \sqrt{\sum_{i=1}^{d} ||b_i||^2}$.

Now, let $\Phi : \mathbb{K}^d \to \mathbb{R}$ be the usual Euclidean metric: $(\alpha_1, \dots, \alpha_2) \mapsto || \sum_{i=1}^{d} \alpha_i b_i||$. This obviously Φ is continuous. Denote by $\mathbb{S}^d = \{\lambda \in \mathbb{K}^d \mid \sum_{i=1}^{d} |\lambda_i|^2 = 1\}$ the unit sphere. By theorem 1.21 we know that $\mathbb{S}^d$ is compact because it is closed and bounded. Therefore, by Weierstrass's theorem there exists a minimum of Φ, $\beta \in \mathbb{K}^d$ i.e. for all $\alpha \in \mathbb{S}^d$ we have $\Phi(\beta) \leq \Phi(\alpha)$. Of course, $\Phi(\beta) > 0$ because if $\Phi(\beta) = 0$ then $\beta = 0 \notin \mathbb{S}^d$.

Therefore, if $||x||_0 = 1$ then $||x||_0 \geq \Phi(\beta)$. So we get the statement by

$$||x|| = \left|\left| ||x||_0 \cdot \frac{x}{||x||_0} \right|\right| = ||x||_0 \left|\left| \frac{x}{||x||_0} \right|\right| \geq ||x||_0 \cdot \Phi(\beta)$$

□

Corollary 2.14.

1. *Every finite dimensional normed spaces is complete.*
2. *All norms on a finite dimensional vector space over $\mathbb{K}$ are equivalent.*
3. *Any finite dimensional subspace of a normed vector space is closed.*

2.2 Linear subspaces of normed vector spaces

Example 2.15. (For a non-closed subspace in a necessarily infinite dimensional normed space) Take ℓ^∞ and the subspace $c_0 = \{x \in \ell^\infty \mid \exists N \; \forall n > N : \; x_n = 0\}$. Then c_0 is not closed because the sequence

$$x_k = \left(1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{k}, 0, 0, \dots\right)$$

is Cauchy in ℓ^∞ so it is also convergent, but the limit point is not in c_0.

Lemma 2.16. *If $U \subseteq V$ is a subspace in a normed vector space, then $\overline{U}$ is also a subspace. (Trivial by the linearity of the limit operator.)*

Theorem 2.17. (Riesz lemma) *Let $W = \overline{W} \subsetneq V$ be a closed, proper linear subspace of a normed vector space V. Then for all $\alpha \in (0,1)$ there exist $x_\alpha \in V$ such that $||x_\alpha|| = 1$ and $||x_\alpha - y|| > \alpha$ for all $y \in W$.*

Proof. Let $x \in V \backslash W$ be arbitrary. Then we can define $\rho(x, W) = \inf_{y \in W}\{\rho(x,y)\} = d > 0$. Then, by the definition, for all $\varepsilon > 0$ we can find a $y_\varepsilon \in W$ such that $||x - y|| < d + \varepsilon$. Then we define the "quasi-projection" of x to the "perpendicular complement" to W, i.e. let $x_\varepsilon = \frac{x - y_\varepsilon}{||x - y_\varepsilon||}$. Then it is obvious that it has length one, and for all $y \in W$:

$$||x_\varepsilon - y|| = \left|\left| \frac{x - y_\varepsilon}{||x - y_\varepsilon||} - y \right|\right| = \frac{1}{||x - y_\varepsilon||} \cdot \left|\left| x - (y_\varepsilon + ||x - y_\varepsilon|| \cdot y) \right|\right| \geq \frac{1}{d + \varepsilon} \cdot d \overset{\varepsilon \to 0}{\longrightarrow} 1$$

so by an appropriate definition of α, one can get the statement. □

Corollary 2.18. *Let V be a normed vector space. Then the unit sphere $S = \{x \in V \mid ||x|| = 1\}$ is compact if and only if* $\dim V < \infty$

Proof. We have already seen $\Leftarrow$ in Theorem 2.13 (it was an easy consequence of Theorem 1.21).

For the other direction, assume that $\dim V = \infty$. Let $x_1 \in S$ and define $V_1 = \mathrm{Span}(x_1)$. Then we can use the Riesz lemma (2.17), and we can choose inductively $x_n \in S \backslash V_{n-1}$ such that $\rho(x_n, V_{n-1}) \geq \frac{1}{2}$ and define $V_n = \mathrm{Span}(x_n, V_{n-1})$. The V_i's are finite dimensional so we can always take a vector from $S \backslash V_i$. This way we got a non-Cauchy sequence, so V is not sequentially compact therefore it cannot be compact (see Theorem 1.16). □

Third Lecture, 9th of October

2.3 Arselá-Ascoli theorem

Definition 2.19. Let (X, ρ) and (Y, d) be metric spaces and $\mathcal{F}$ be a family of $X \to Y$ functions. All $f \in \mathcal{F}$ are *continuous*, if

$$(\forall f \in \mathcal{F})(\forall x \in X)(\forall \varepsilon > 0)(\exists \delta(f, \varepsilon, x) > 0): \ \rho(x, y) < \delta(\varepsilon, f, x) \Rightarrow d\big(f(x), f(y)\big) < \varepsilon$$

All $f \in \mathcal{F}$ are *uniformly continuous* if

$$(\forall \varepsilon > 0)(\forall f \in \mathcal{F})(\exists \delta(\varepsilon, f) > 0)(\forall x \in X): \ \rho(x, y) < \delta(\varepsilon, f) \Rightarrow d\big(f(x), f(y)\big) < \varepsilon$$

The family $\mathcal{F}$ is *equicontinuous*, if

$$(\forall \varepsilon > 0)(\forall x \in X)(\exists \delta(\varepsilon, x) > 0)(\forall f \in \mathcal{F}): \ \rho(x, y) < \delta(\varepsilon, x) \Rightarrow d\big(f(x), f(y)\big) < \varepsilon$$

And the family $\mathcal{F}$ is *uniformly equicontinuous* on X, if

$$(\forall \varepsilon > 0)(\exists \delta(\varepsilon) > 0)(\forall x \in X)(\forall f \in \mathcal{F}): \ \rho(x, y) < \delta(\varepsilon) \Rightarrow d\big(f(x), f(y)\big) < \varepsilon$$

Remark 2.20. So "equi" refers to "independent of f" and "uniform" refers to "independent of x"

Definition 2.21. A family of $X \to Y$ functions $\mathcal{F}$ is *uniformly bounded*, if there exists $C \in \mathbb{R}$ such that $|f(x)| < C$ for all $f \in \mathcal{F}$ and for all $x \in X$.

Theorem 2.22. (Ascoli, 1883) *Let $\{f_n : [0,1] \to \mathbb{R}\}_{n \in \mathbb{N}}$ be a uniformly bounded, equicontinuous family of function. (By compactness, it also means uniformly equicontinuity.) Then there is a subsequence $\big\{f_{n_k}\big\}_{k \in \mathbb{N}}$ which converges uniformly on $[0,1]$ to a uniformly continuous function.*

Proof. Consider the rationals in $[0,1]$ as a well ordered set $\{q_i\}_{i \in \mathbb{N}}$. Since the family is uniformly bounded: $\big\{f_n(q_1)\big\}_{n \in \mathbb{N}}$ is bounded, so it has a converging subsequence $\big\{f_{n_k^{(1)}}(q_1)\big\}$. By the same reason, we can do the same for $\big\{f_{n_k^{(1)}}(q_2)\big\}$ and we get a subsequence $\big\{f_{n_k^{(2)}}\big\}$ of the previous subsequence. It can be iterated: $f_{n_k^{(i)}}$ is chosen as the subsequence of $f_{n_k^{(i-1)}}$ in a way that $\big\{f_{n_k^{(i)}}(q_i)\big\}$ is a converging subsequence of $\big\{f_{n_k^{i-1}}(q_i)\big\}$.

Now, we can name $\lim_{k \to \infty} f_{n_k^{(i)}}(q_i) =: f(q_i)$ for which $\lim_{k \to \infty} f_{n_k^{(i)}}(q_j) = f(q_j)$ also holds, if $j \leq i$, by the choice of $n_k^{(i)}$'s. To get the subsequence that converges at every rational, we can choose the "diagonal, i.e. let

$$g_k = f_{n_k^{(k)}}$$

By the construction $\lim_{k \to \infty} g_k(q_i) = f(q_i)$ for all $i \in \mathbb{N}$. Now, we got a sequence in $\mathcal{F}$ that is convergent in every rational point. Let us see what happens at the irrationals:

Lemma 2.23. *For all $x \in [0,1]$ there exists the $\lim_{k \to \infty} g_k(x)$.*

Proof. We check that $\{g_k\}$ is Cauchy for all $x \in [0,1]$. Indeed,

$$|g_k(x) - g_l(x)| \leq |g_k(x) - g_k(q)| + |g_k(q) - g_l(q)| + |g_l(q) - g_l(x)| < \varepsilon + \varepsilon + \varepsilon$$

if $d(q, x) < \delta(\varepsilon)$ (by the uniform continuity of $\mathcal{F}$) and if $k, l \geq N(\varepsilon)$ (by the Cauchy-ness of $\big\{g_k(q)\big\}_{k \in \mathbb{N}}$). □

Let us denote $\lim_{k\to\infty} g_k(x)$ by $f(x)$. We have to prove two things: first that f is continuous (because then it is uniform continuous by Heine's theorem 1.30) and then that g_k uniformly converges to f.

Lemma 2.24. *f is continuous.*

Proof.
$$|f(x) - f(y)| \le |f(x) - g_k(x)| + |g_k(x) - g_k(y)| + |g_k(y) - f(y)| < \varepsilon + \varepsilon + \varepsilon$$
if $k > N(x, \varepsilon)$, $|x - y| < \delta(\varepsilon)$ and $k > N(y, \varepsilon)$ what can be satisfied by an appropriate choice of ε, x, y and k (in this order). □

Lemma 2.25. *$g_k \to f$ uniformly.*

Proof. By the uniform equicontinuity of $\{f_n\}$ we can choose a good $\delta(\varepsilon)$. Let $\{y_i\}_{i=1,\dots,N} \subseteq [0, 1]$ be a finite sequence such that $[0, 1] \subseteq \cup_{i=1}^{N} B_{\delta(\varepsilon)}(y_i)$. (It exists by compactness.) Then

$$|g_n(x) - f(x)| \le |g_n(x) - g_n(y_i)| + |g_n(y_i) - f(y_i)| + |f(y_i) - f(x)| < \varepsilon + \varepsilon + \varepsilon$$

if $|x - y_i| < \delta(\varepsilon)$ (by the uniform equicontinuity of $\{f_n\}$), $n > \max_{i=1,\dots,N} N(\varepsilon, y_i)$ (by the convergence of $\{g_n(y_i)\}_{n\in\mathbb{N}}$) and $|x - y < \delta'(\varepsilon, y_i)|$ (by the continuity of f).

So for a fixed $\varepsilon > 0$, we can choose δ and δ' take their minimum δ^* and choose y_i for that δ^*. □

The statement of the theorem follows. □

Corollary 2.26. *If $\mathcal{F} \subseteq C[0, 1]$ is uniformly bounded and equicontinuous then $\mathcal{F}$ has a converging subsequence in $C[0, 1]$. I.e., in this case, $\overline{\mathcal{F}}$ is sequentially compact (therefore it is compact as well because $C[0, 1]$ is a metric space). In other words, equicontinuity gives a sufficient condition for (uniformly) bounded, closed sets in $C[0, 1]$ to be compact.*

Theorem 2.27. (Arzelá-Ascoli) *Let (X, ρ) be a compact metric space. Let $\mathcal{F} \subseteq C_{\mathbb{R}}(X) = \{f : X \to \mathbb{R} \mid f \text{ is continuous}\}$. Then $\overline{\mathcal{F}}$ is compact if and only if $\mathcal{F}$ is equicontinuous and point-wise bounded. Moreover, the latter pair of properties on a compact space is equivalent to saying uniformly equicontinuous and uniformly bounded.)*

Proof. It is easy to see that if $\mathcal{F}$ is equicontinuous on a compact space X and point-wise bounded then $\mathcal{F}$ is uniformly bounded.

The direction $\Leftarrow$ is similar to the proof of the previous Theorem. The main difference that instead of $\mathbb{Q} \cap [0, 1]$ we take a dense set of $\{x_i\}_{i\in\mathbb{N}} \subseteq X$ (which is possible because a compact metric space is separable). The remaining part of the proof carries over literally.

Direction $\Rightarrow$: Let $\overline{\mathcal{F}}$ is compact in $C_{\mathbb{R}}(X)$ then $\overline{\mathcal{F}}$ is totally bounded in the complete space $C_{\mathbb{R}}(X)$. The point-wise boundedness of $\mathcal{F}$ follows. Therefore, we only have to prove equicontinuity.

Given $\varepsilon > 0$, $\delta > 0$ and $x \in X$ let

$$N_x(\varepsilon, \delta) = \{f \in C_{\mathbb{R}}(X) \mid \sup_{y_1, y_2 \in B_\delta(x)} \{|f(y_1) - f(y_2)| < \varepsilon\}\}$$

This $N_x(\varepsilon, \delta)$ is an open set in $C_{\mathbb{R}}(X)$. For a fixed $\varepsilon > 0$ and $x \in X$ the set $\overline{\mathcal{F}}$ can be covered by $\cup_{\delta>0} N_x(\varepsilon, \delta) = C_{\mathbb{R}}(X)$ because every element is uniformly continuous (by compactness of X). Therefore, by compactness of $\overline{\mathcal{F}}$, this cover has a finite subcover. So for this fixed x we have $\delta_1, \dots, \delta_n$ such that $\cup_{i=1}^{n} N_x(\varepsilon, \delta_i) \supseteq \overline{\mathcal{F}}$. However, these are containing each other, so $\delta_x = \min_i \delta_i$ is a "good δ" in a sense that $N_x(\varepsilon, \delta_x) \supseteq \overline{\mathcal{F}}$. In other words, the elements of $\overline{\mathcal{F}}$ satisfy uniform continuity in $B_{\delta_x}(x)$ by δ_x. (So $\overline{\mathcal{F}}$ uniformly equicontinuous in $B_{\delta_x}(x)$)

To get that there exists a good δ for all x's, we have to use the compactness of the space as well. Consider the open cover $\cup_{x\in X} B_{\delta_x/2}(x) \supseteq X$ where δ_x is defined in the above way for all x's. Then – by compactness of X – it has a finite subcover

$$\bigcup_{i=1}^{n} B_{\delta_{x_i}/2}(x_i) \supseteq X$$

then $\delta = \min_i \delta_{x_i}/2$ is a globally good δ for the fixed ε because for all $y_1, y_2 \in X$ with the property $d(y_1, y_2) < \delta$ there exists an $i \le n$ such that $y_1 \in B_{\delta_{x_i}/2}(x_i)$ so $y_1, y_2 \in B_{\delta_{x_i}}(x_i)$. Therefore, $|f(y_1) - f(y_2)| < \varepsilon$ for all $f \in \overline{\mathcal{F}}$. That is exactly (uniform) equicontinuity. □

3 Existence of local solution of differential equations

Theorem 3.1. (Peano theorem, on the existence of local solutions) *Let $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ be a continuous function in the neighborhood of (t_0, y_0). Then the initial value problem*

$$y'(t) = f\big(t, y(t)\big) \qquad y(t_0) = y_0$$

has at least one solution defined in a neighborhood of t_0.

Proof. Let $a, b > 0$ such that f is continuous on $Q := [t_0, t_0 + a] \times \overline{B_b(y_0)}$ so there exists an $L > 0$ such that $||f(t, y)|| \le L$ for all $(t, y) \in Q$. Let $c := \min\left\{a, \frac{b}{L}\right\}$, $I = [t_0, t_0 + c]$ and

$$\mathcal{A} := \{g : I \to \mathbb{R} \mid g(t_0) = y_0,\ g \text{ is Lipschitz with Lipschitz constant } L\}$$

In this case, for $t \in I$ and $g \in \mathcal{A}$ we have $g(t) \in \overline{B_b(y_0)}$. Now, we can define the following (non-linear) functional $F : \mathcal{A} \to [0, \infty)$

$$g \mapsto F(g) = \max_{t \in I} \left\{ \left|\left| qg(t) - y_0 - \int_0^t f\big(s, g(s)\big)\, \mathrm{d}s \right|\right| \right\}$$

Using this function, we can characterize the solutions: g solves the equation if and only if $F(g) = 0$. However, F is continuous on $\mathcal{A}$ and $\mathcal{A}$ is compact by Arzelá-Ascoli, therefore F attends its minimum on $\mathcal{A}$. Therefore exists $y \in \mathcal{A}$ such that $F(y) = \inf_{x \in \mathcal{A}} F(x)$. "By hand", one can construct functions on which F is arbitrarily small (for example piecewise linear functions) so the previously found minimal y is in fact a solution. □

Remark 3.2. This theorem has loose requirements (only continuity) but the consequence is neither strong: only locally and not necessarily uniquely states the existence.

Theorem 3.3. (Banach fixed point theorem) *Let (X, d) be a complete metric space and $T : X \to X$ be a contraction (i.e. there exists a $q < 1$ such that $d(T(x), T(y) \le qd(x, y)$ for all $x, y \in X$). Then there exists a unique $x^* \in X$ such that $T(x^*) = x^*$. Moreover, for any $x_0 \in X$, $T^n(X) \to x_0$ as $n \to \infty$.*

Proof. We have to show that

1. $T^n(x_0)$ is Cauchy, where x_0 is arbitrary.
2. $y = \lim T^n(x_0)$ is a fixed point of T
3. if $Ty_1 = y_1$ and $Ty_2 = y_2$ then $y_1 = y_2$

The statement follows. □

Theorem 3.4. (Picard - Lindelöf, Existence and uniqueness criterion) *Given the initial value problem*

$$y'(t) = f\big(t, y(t)\big) \qquad y(t_0) = y_0$$

If f is Lipschitz-continuous in y and continuous in t then there exists an $\varepsilon > 0$ such that the problem has a unique solution in $[t_0 - \varepsilon, t_0 + \varepsilon]$

Remark 3.5. This theorem states stronger properties than Peano's theorem (i.e. unique existence) but also the assumptions are stronger.

Proof. The problem can be rewritten in integral form:

$$y(t) = y_0 + \int_{t_0}^{t} f\big(s, y(s)\big) \, \mathrm{d}s$$

This can give the idea for the Picard-iteration, i.e. let $\varphi_0(t) :\equiv y_0$ and let

$$\varphi_{n+1}(t) = y_0 + \int_{t_0}^{t} f\big(s, \varphi_n(s)\big) \, \mathrm{d}s = (T\varphi_n)(t)$$

then we can prove that φ converges to the (unique) solution. This is based on $T\varphi = \varphi$ if and only if φ is a solution.

Let $a, b > 0$, $M := \sup\{|f(t, y)| \mid (t, y) \in R\}$ and $R = \overline{B_b(y_0)} \times [t_0 - a, t_0 + a]$ where a should be chosen the following way:

1. for all $t \in [t_0 - a, t_0 + a] : \varphi_{n+1}(t) \in \overline{B_b(y_0)}$. This is possible because $||\varphi_{n+1}(t) - y_0|| = || \int_{t_0}^{t} f\big(s, \varphi_n(s)\big) \, \mathrm{d}s|| \leq a \cdot M < b$ if $a < \frac{b}{M}$

2. $||T\varphi_1 - T\varphi_2|| \leq q \cdot ||\varphi_1 - \varphi_2||$ for some $q < 1$. This is possible because $||T\varphi_1 - T\varphi_2|| = || \int_{t_0}^{t} f\big(s, \varphi_1(s)\big) - f\big(s, \varphi_2(s)\big) \, \mathrm{d}s|| \leq a \cdot L \cdot ||\varphi_1(s) - \varphi_2(s)||$ where $q = a \cdot L < 1$ if $a < \frac{1}{L}$.

Therefore, if we choose a as $a < \min\left\{\frac{b}{M}, \frac{1}{L}\right\}$ then T will be a contraction on $[t_0 - a, t_0 + a]$ so by the Banach fixed point theorem, it converges to the unique solution. □

Fourth Lecture, 16th of October

4 Linear operators in Banach spaces

Definition 4.1. Let $(X_i, || \cdot ||_i)$ be Banach-spaces $(i = 1, 2)$. $A : X_1 \to X_2$ is *linear*, if $A(\alpha x + \beta y) = \alpha A(x) + \beta A(y)$ for all $x, y \in X_1$ and $\forall \alpha, \beta \in \mathbb{K}$.

Definition 4.2. The linear operator $A : X_1 \to X_2$ is *bounded* (notation: $A \in B(X_1, X_2)$), if there is a $K > 0$ such that for all $x \in X_1$:

$$||Ax||_2 \leq K \cdot ||x||_1$$

Definition 4.3. If $A \in B(X_1, X_2)$ is bounded then its *norm* is

$$||A|| := \inf\{K \in [0, \infty) \mid \forall x \in X_1 \ ||Ax||_2 \leq K||x||_1\} =$$

$$= \sup\{||Ax||_2 \mid ||x||_1 = 1\} = \sup\left\{\frac{||Ax||_2}{||x||_1} \mid 0 \neq x \in X_1\right\}$$

Lemma 4.4. *For all $A, B \in B(X_1, X_2)$ then the above $||.||$ is indeed a norm. Moreover, if $A, B \in B(X_1, X_2)$ then $||A \cdot B|| \leq ||A|| \cdot ||B||$, i.e. it is a normed algebra. (The proof is straightforward.)*

Lemma 4.5. *Let $A : X_1 \to X_2$ be a linear operator. Then the followings are equivalent:*

1. *A is bounded.*

2. *A is continuous at $0 \in X_1$.*

3. *A is continuous.*

(The proof is straightforward if not trivial.)

Theorem 4.6. *If X_1 is normed, X_2 is Banach then $B(X_1, X_2)$ is a Banach space with the operator norm.*

Proof. Only the completeness of $B(X_1, X_2)$ is nontrivial. To prove that, let $\{A_n\}_{n\in\mathbb{N}} \subseteq B(X_1, X_2)$ be a Cauchy sequence, i.e. $||A_n - A_m|| \overset{n,m\to\infty}{\longrightarrow} 0$. for any $x \in X_1$, $\{A_n x\}_{n\in\mathbb{N}} \subseteq X_2$ is Cauchy since

$$||A_n x - A_m x|| = ||(A_n - A_m)x|| \leq ||A_n - A_m|| \cdot ||x|| \to 0 \cdot ||x|| = 0 \qquad \text{as } n, m \to \infty$$

X_2 is complete so there exists the pointwise limit $\lim_{n\to\infty} A_n x$. Let us denote it by $Ax = \lim_{n\to\infty} A_n x$. This operator A is linear (trivially) and is bounded (see below, basically follows from Banach Steinhaus) so we have to prove that A_n converges to A in the operator-norm.

Let $||x|| = 1$ and fix an $\varepsilon > 0$. Then

$$||Ax - A_n x|| \leq ||Ax - A_m x|| + ||A_m x - A_n x|| \leq$$

$$\leq ||Ax - A_m x|| + ||A_m - A_n|| \cdot ||x|| < ||Ax - A_m x|| + \varepsilon \cdot 1$$

if $m, n > N_1(\varepsilon)$. However, $||Ax - A_m x|| < \varepsilon$ if $m > N_2(\varepsilon, x)$. Therefore, for given x, we can choose a large enough m such that

$$||Ax - A_n x|| \leq 2\varepsilon$$

for all $n, m > N_1(\varepsilon)$. Note that the inequality is already independent of m so the statement becomes that $||Ax - A_n x|| \leq 2\varepsilon$ if $n > N_1(\varepsilon)$.

Now let us prove boundedness of A: let $||x|| = 1$

$$||Ax|| = ||Ax - A_m x + A_m x|| \leq ||(A - A_m)x|| + ||A_m x|| \leq$$

$$\leq ||(A_n - A)x|| + ||A_m|| \cdot ||x|| < \varepsilon + C \cdot 1$$

if $n > N(\varepsilon)$ where we used that if A_m is Cauchy then $||A_m||$ is Cauchy as well. So it is convergent, hence bounded. □

Definition 4.7. Let $(X, ||.||)$ be a normed space. Its *dual space* is defined as

$$X^* = B(X, \mathbb{K}) = \{\varphi : X \to \mathbb{K} \mid \varphi \text{ is bounded}\}$$

If $\varphi \in X^*$ then we can also define its functional norm: $||\varphi|| : \sup\{|\varphi(x)| \mid ||x|| = 1\}$. For this, we have $|\varphi(x)| \leq ||\varphi|| \cdot ||x||$.

Example 4.8. $\Phi : \mathcal{C}_\mathbb{R}[-1, +1] \to \mathbb{R}$, $\Phi(f) = \int_1^1 x f(x)\ \mathrm{d}x$. We are searching for the value of $||\Phi||$. It is clear that Φ is linear. Boundedness can be seen by

$$|\Phi(f)| = \left| \int_{-1}^{1} x f(x)\ \mathrm{d}x \right| \leq \int_{-1}^{1} |x| \cdot |f(x)|\ \mathrm{d}x \leq ||f||_\infty \int_{-1}^{1} |x|\ \mathrm{d}x = ||f||_\infty$$

therefore $||\Phi|| \leq 1$. Equality can be reached by approximating the sign-function with a sequence of continuous functions.

Remark 4.9. Let X be a Banach space. Then – without other structure – X and X^* cannot be canonically identified. They are not necessarily isomorphic generally, but even if if they are, there is no natural choice for a canonical isomorphism. However, in the case of second dual the following holds:

Theorem 4.10. *Let X be a Banach space. Then $\eta : X \hookrightarrow X^{**}$, $v \mapsto \big(\varphi \mapsto \varphi(v)\big)$ is a canonical linear isometric injection.*

Remark 4.11. The proof relies on the Hahn-Banach theorem.

Theorem 4.12. (Hahn-Banach, algebraic version) *Let $(X, ||.||)$ be a normed space and $X_0 \leq X$ be a linear subspace. Then a $\varphi : X_0 \to \mathbb{C}$ bounded linear functional can be norm-preservingly extended to X, i.e. there exists $\tilde{\varphi} \in X^*$ such that $\tilde{\varphi}|_{X_0} = \varphi$ and $||\tilde{\varphi}|| = ||\varphi||$. (See later the proof.)*

Corollary 4.13. *Let X be a normed space.*

1. *The elements of X^* separate the vectors in X in the sense that $x = y \iff \varphi(x) = \varphi(y)$ for all $\varphi \in X^*$. (If they differ then we can extend a non-zero functional from the 1-dimensional subspace $\langle x - y \rangle$)*
2. *For all $x \in X$ there exists $\varphi_x \in X^*$ such that $||\varphi_x|| = 1$ and $\varphi(x) = ||x||$. (Again, we can extend the 1-dimensional functional $\varphi : \mathbb{C}x \to \mathbb{C}$, $\lambda x \mapsto \lambda ||x||$.)*

Proof. (of Theorem 4.10) We have to prove three things: η is linear, isometric and injective. Linearity is straightforward to check. To get the isometry property we can first give an upper bound:

$$||\eta v|| = \sup\{|\varphi(v)| \mid \varphi \in X^*,\ ||\varphi|| = 1\} \leq 1 \cdot ||v||$$

The fact that it can be satisfied with equality follows from the second part of Consequence 4.13 (sometimes called the Little Hahn-Banach).

While the injectivity follows from the first part of Consequence 4.13: if $\tilde{x} = \tilde{y}$ then $\varphi(x) = \varphi(y)$ is true for all $\varphi \in X^*$ but that means that $x = y$. ☐

Definition 4.14. The Banach space X is *reflexive* if $\eta : X \hookrightarrow X^{**}$ is surjective.

Theorem 4.15. *Let X be a Banach space. Then*

1. *X is reflexive if and only if X^* is reflexive*
2. *If X is not reflexive then X^*, X^{**}, X^{***} are neither reflexive. (Consequence of the first.)*

4.1 Duality Theorems

Theorem 4.16. $(\mathbb{K}^n, ||\cdot||_p)^* \cong (\mathbb{K}^n, ||\cdot||_q)$ *for $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and $1 \leq p, q \leq \infty$ if $\frac{1}{p} + \frac{1}{q} = 1$. In the sense that if $\Phi \in (\mathbb{C}^n, ||\cdot||_p)^*$ then there exists a unique $f_\Phi \in (\mathbb{C}^n ||\cdot||_q)$ such that*

$$\Phi(x) = \sum_{k=1}^{n} \overline{f_{\Phi,k}} x_k$$

and $||\Phi|| = ||f_\Phi||_q$. Moreover, the map $\Phi \mapsto f_\Phi$ is an isometric conjugate-linear bijection.

Proof. Conjugate-linearity and bijectivity are straightforward. To get the isometry property, we have to use Hölder:

$$|\Phi(x)| = \left| \sum_{k=1}^{n} \overline{f_{\Phi,k}} x_k \right| \leq \sum_{k=1}^{n} |\overline{f_{\Phi,k}}| \cdot |x_k| \leq ||x||_p \cdot ||\overline{f_\Phi}||_q$$

i.e. $||\Phi|| \leq ||f_\Phi||_q$. On concrete vectors one can show that the inequality is strict. (The definition of $f_{\Phi,k}$ is $f_{\Phi,k} = \Phi(e_i)$.) ☐

Theorem 4.17. *If $1 < p, q < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$ then $\left(\ell^p\right)^* \cong \ell^q$ i.e. there is a conjugate-linear isometry $\eta : \ell^q \to (\ell^p)^*$ with $(f_k)_{k \in \mathbb{N}} \mapsto (x \mapsto \sum_{k \in \mathbb{N}} \overline{f_k} x_k)$.*

Proof. (Sketch) for finitely supported elements of ℓ^p we can use the construction seen in the previous theorem 4.16. This can be extended to the whole space by continuity because the set of finitely supported elements is dense. ☐

Remark 4.18. This method works in the case $p = 1$ as well, so $(\ell^1)^* \cong \ell^\infty$ in the same way. However, $(\ell^\infty)^* \not\cong \ell^1$.

Theorem 4.19. $(\ell^\infty)^* \supseteq (c_0)^* \cong \ell^1$ *with the same pairing as we used before.*

Theorem 4.20. *Let $c = \{(a_n)_{n\in\mathbb{N}} \mid \exists \lim_{n\to\infty} a_n \in \mathbb{C}\} < \ell^\infty$. Then $c^* \cong \ell^1$ with the pairing*

$$\ell^1 \ni x \mapsto \left(y \mapsto \overline{x_0} \lim y_n + \sum_{i=1}^{\infty} \overline{x_i} y_{i-1} \right) \in c^*$$

Remark 4.21. What is more c and c_0 are not isomorphic as Banach spaces. I.e. the predual is not a well-defined concept.

Theorem 4.22. (Duality theorem for measure spaces) *For the L^p-spaces: $\left(L^p(X,\mu)\right)^* \cong L^q(X,\mu)$ for $1 \le p, q < \infty$ where $\frac{1}{p} + \frac{1}{q} = 1$. What is more, $\left(L^1(\mathbb{R},\lambda)\right)^* \cong L^\infty(\mathbb{R},\lambda)$ (in general it is not true.) The pairing is*

$$L^q(X,\mu) \ni f \mapsto \left(g \mapsto \int_X \overline{f} g \,\mathrm{d}\mu \right) \in \left(L^p(X,\mu)\right)^*$$

Remark 4.23. It does not hold for $p = \infty$, $q = 1$.

Example 4.24. Let $\varphi \in (\ell^4)^*$ be defined by

$$\varphi(a) = \sum_{n=1}^{\infty} \frac{a_{n+2}}{\sqrt{n^3}}$$

The corresponding $\ell^{\frac{4}{3}}$-element is $f = (0, 0, 0, 1, \frac{1}{\sqrt{2^3}}, \frac{1}{\sqrt{3^3}}, \dots)$ so this way we can find

$$||\varphi|| = ||f|| = \left(\sum_{n=1}^{\infty} \left(\frac{1}{\sqrt{n^3}} \right)^{\frac{4}{3}} \right)^{\frac{3}{4}} = \left(\frac{\pi^2}{6} \right)^{\frac{3}{4}} < \infty$$

Example 4.25. Let $\varphi \in \left(L^{\frac{4}{3}}[0,1]\right)^*$ be defined by

$$\varphi(f) = \int_0^{\frac{1}{3}} f(x) x^3 \,\mathrm{d}x = \int_0^1 f(x) F_\varphi(x) \,\mathrm{d}x$$

where F_φ is defined as

$$F_\varphi(x) = \chi_{[0,\frac{1}{3}]}(x) \cdot x^3$$

so $||\varphi|| = ||F_\varphi||_4 = \left(\int_0^{\frac{1}{3}} (x^3)^4 \,\mathrm{d}x \right)^{\frac{1}{4}} = \dots$ can be computed a lot easier than by the definition.

Fifth Lecture, 30th of October

4.2 Weak topologies

Definition 4.26. Let X be a set and $(Y_i)_{i\in I}$ are topological spaces. Let $f_i : X \to Y_i$ be a function for all $i \in I$. Denote $\{f_i \mid i \in I\}$ by $\mathcal{F}$. The *$\mathcal{F}$-topology* on X is the coarsest topology for which all the maps f_i are continuous.

Definition 4.27. Let X be a Banach space and X^* be its dual. Then the *weak topology* $\sigma(X, X^*)$ on X is the X^*-topology on X i.e. it is the coarsest one for which all the mapping $\varphi \in X^*$ are continuous.

Theorem 4.28.

1. *The weak topology is weaker than the norm topology.*

2. *If $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ converges in the weak topology (i.e. there exists an $x \in X$ such that $\varphi(x_n) \to \varphi(x)$) then the sequence x_n is norm-bounded. (Proof goes with Banach-Steinhaus.)*

3. The weak topology is Hausdorff.

Remark 4.29. $\sigma(X^*, X^{**})$ is the weak topology on X^*.

Definition 4.30. Let X be a Banach space and X^* be its dual. The *weak-∗-topology* $\sigma(X^*, X)$ on X^* is the topology generated by the functionals $\eta X \subseteq X^{**}$ where $\eta : X \to X^{**}$ is the canonical embedding.

Remark 4.31. Since $X \subseteq X^{**}$, $\sigma(X^*, X)$ is coarser than $\sigma(X^*, X^{**})$.

Theorem 4.32. (Banach-Alaoglu) *If X is Banach then the unit ball of X^* is compact in the weak-∗-topology $\sigma(X^*, X)$.*

5 Hilbert spaces

Example 5.1. The following two reflexive Banach spaces $(\ell^2)^* \cong \ell^2$ and $(L^2(\mathbb{R}, \lambda))^* \cong L^2(\mathbb{R}, \lambda)$ are in fact Hilbert spaces as well.

Definition 5.2. $(\mathcal{H}, \langle ., . \rangle)$ is an *inner product space* (or *pre-Hilbert space*) if $\langle ., . \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is a $\frac{1}{2}-1$-linear, positive definite, Hermitian bilinear function.

Proposition 5.3. *If $(\mathcal{H}, \langle ., . \rangle)$ is a pre-Hilbert space then*

1. $||x|| = \sqrt{\langle x, x \rangle}$ *is a norm*
2. *the Cauchy-inequality* $|\langle x, y \rangle| \le ||x|| \cdot ||y||$ *holds for all* $x, y \in \mathcal{H}$
3. *the parallelogram law* $2||x||^2 + 2||y||^2 = ||x + y||^2 + ||x - y||^2$ *is satisfied for all* $x, y \in \mathcal{H}$
4. *polarization is true, i.e. the norms determine the inner product by the formulas:*
   (a) *if it is a real Hilbert space then* $||x + y||^2 - ||x - y||^2 = 4\langle x, y \rangle$
   (b) *if it is a complex Hilbert space then*

$$\sum_{i=0}^{3} i^{-n} ||x + i^n y||^2 = ||x + y||^2 - i||x + iy||^2 - ||x - y||^2 + i||x - iy||^2 = 4\langle x, y \rangle$$

Definition 5.4. A complete pre-Hilbert space is called a *Hilbert space*.

Theorem 5.5. *If $(X, ||.||)$ is a normed space where $||.||$ satisfies the parallelogram law then there exists an inner product $\langle ., . \rangle$ such that for all $x \in X$ $||x|| = \sqrt{\langle x, x \rangle}$.*

Proof. (Only in the real case.) By the polarization identity we can define the should-be inner product by the norms. We only have to prove that it is bilinear. This goes by first proving only additivity. That will give the $\mathbb{Q}$-bilinearity and by continuity we get the full linearity.

In details, consider the following two parallelogram laws,

$$2||x_1 + y||^2 + 2||x_2||^2 = ||x_1 + x_2 + y||^2 + ||x_1 - x_2 + y||^2$$
$$2||x_1 - y||^2 + 2||x_2||^2 = ||x_1 + x_2 - y||^2 + ||x_1 - x_2 - y||^2$$

and subtract them:

$$8\langle x_1, y \rangle = 2||x_1 + y||^2 - 2||x_1 - y||^2 = ||x_1 + x_2 + y||^2 - ||x_1 + x_2 - y||^2 + ||x_1 - x_2 + y||^2 - ||x_1 - x_2 - y||^2$$

Now, interchanging x_1 and x_2 we get a similar expression. Adding them up gives:

$$8\langle x_1, y \rangle + 8\langle x_2, y \rangle = 2||x_1 + x_2 + y||^2 - 2||x_1 + x_2 - y||^2 = 8\langle x_1 + x_2, y \rangle$$

I.e. we got the additivity in the first argument. By symmetry we get biaddivity. By the usual arithmetic argument, we get $\mathbb{Q}$-bilinearity. Finally, the defining formula for $\langle ., . \rangle$ was continuous so it will be $\mathbb{R}$-linear as well. Positive definiteness is now straightforward. ☐

6 Hahn-Banach theorem

Theorem 6.1. *(Hahn-Banach theorem, algebraic version) Let X be a normed space $X_0 \leq X$ be a subspace and $\varphi \in X_0^*$ arbitrary. Then there exists $\tilde{\varphi} \in X^*$ such that $\tilde{\varphi}|_{X_0} = \varphi$ and $||\tilde{\varphi}|| = ||\varphi||$.*

Proof. We will use Zorn's lemma:

Lemma 6.2. (Zorn's lemma) *If in non-empty partially ordered set P every chain (i.e. totally ordered subset) $\mathcal{L}$ has an upper bound in P then P has a maximal element.*

By continuity, φ can be uniquely extended to $\overline{X_0}$. If $\overline{X_0} \neq X$ then there exists $z \in X \backslash \overline{X_0}$, $||z|| = 1$. Now, let us define the extension to $\operatorname{Span}(\overline{X_0}, z)$ such that $\frac{|\varphi(z)|}{|z|} \leq ||\varphi||_{X_0}$. Let P b the set of all extensions of φ which do not increase the norm. The partial order on P is the "being extension of the other". Clearly, the chains have upper bounds by the union so we can take a maximal element by Zorn's lemma. This maximal element is defined in the whole space, else we could again extend it by a dimension. ☐

Definition 6.3. Let $(X_1, ||.||_1)$ and $(X_2, ||.||_2)$ are Banach spaces. Then their *direct sum* $X_1 \oplus X_2 = \{(x_1, x_2) \mid x_1 \in X_1, x_2 \in X_2\}$ with the norm $||(x_1, x_2)|| = ||x_1|| + ||x_2||$ is a Banach space.

Remark 6.4. The direct sum of Hilbert spaces is different than their direct sum as Banach spaces because in the first case we get $||(x_1, x_2)|| = ||x_1||^2 + ||x_2||^2$.

7 Baire Category theorem

Philosophy: Sets with nonempty interiors are important. See, for example, the following lemma:

Lemma 7.1. *Let X, Y be normed spaces, $T : X \to Y$ a linear map. Then T is bounded if and only if $T^{-1}\big(\overline{B_1}(0; Y)\big)$ has non-empty interior.*

Proof. Direction $\Leftarrow$: By the assumption, there exists $x \in X$ and $\varepsilon > 0$ such that $TB_\varepsilon(x; X) \subseteq \overline{B_1}(0; Y)$. Then if $\Delta \in X$ such that $||\Delta|| < \varepsilon$ then

$$||T(\Delta)|| = ||T(\Delta + x - x)|| \leq ||T(\Delta + x)|| + ||Tx|| \leq 1 + 1$$

Therefore $||T|| \leq \frac{2}{\varepsilon}$. Conversely, if $||x|| \leq \frac{1}{||T||}$ then $||Tx|| \leq 1$ so $B_{\frac{1}{||T||}}(0)$ is mapped into $\overline{B_1}(0; Y)$. ☐

Definition 7.2. A set S in a metric space M is *nowhere dense*, if the interior of $\overline{S}$ is non-empty. In general, a set $A \subseteq M$ is a set *of first category* (or is a *meager* set), if it is a countable union of nowhere dense sets. All other sets are *of second category.*

Theorem 7.3. (Baire category theorem) *A complete metric space is not a countable union of nowhere dense sets (i.e. it is of second category). In other words, if a complete metric space is a countable union of closed sets then at least one of the closed sets has an interior point.*

Proof. Indirectly assume that $M = \cup_{n\in\mathbb{N}} A_n$ where the A_n's are nowhere dense sets in M. We construct a Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$ such that its limit is not in M.

A_1 is nowhere dense so we can choose $x_1 \in M \backslash \overline{A_1}$. Then for some small enough radius r_1: $B_{r_1}(x_1) \cap \overline{A_2} = \emptyset$ so we can choose arbitrary $x_2 \in B_{r_1}(x_1) \backslash \overline{A_2}$. Inductively, if $x_1, \ldots, x_{n-1}$ is already chosen then we can choose $x_n \in B_{r_{n-1}}(x_{n-1}) \backslash \overline{A_n} \subseteq B_{r_k}(x_k) \backslash \overline{A_k}$ (for all $k < n-1$) for some radius r_{n-1}. We can also assume that $r_n \leq \frac{1}{2^n}$. Then it is obviously a Cauchy sequence, and for all $n \in \mathbb{N}$ there exists an $N \in \mathbb{N}$ such that for all $m > N$ we have $x_m \notin \overline{A_n}$. So the limit is not in A_n for any n which is a contradiction. ☐

Theorem 7.4. (Banach-Steinhaus, the principle of uniform boundedness) *Let X be a Banach space, Y a normed space and $\mathcal{F} = \{T_i\}_{i\in I}$ a set of bounded linear operator such that for all $x \in X$, $\{||T_i x|| \mid i \in I\}$ is bounded (or in short $\mathcal{F}$ is point-wise bounded). Then $\mathcal{F}$ is uniformly bounded.*

Proof. Let us define the following sets:

$$B_n := \{x \in X \mid ||T_i x|| \le n \ \forall i \in I\} = \bigcap_{i \in I} \{x \in X \mid ||T_i x|| \le n\}$$

these are closed by the boundedness of T_i's. By point-wise boundedness $X = \cup_{n\in\mathbb{N}} B_n$. So by Baire category theorem: one of B_n's have nonempty interior. This proves the uniform boundedness. □

Theorem 7.5. (Open mapping theorem, Banach-Schauder) *Let X and Y be two Banach spaces and $T : X \to Y$ be a bounded linear surjection. Then T is open.*

Sixth Lecture, 6th of November

Proof. By linearity, it is enough to see that there is an open ball in Y which contains the image of an open ball in X. Or equivalently, it is enough to prove that there exists an ε such that

$$TB_1(0; X) \supseteq B_\varepsilon(0; Y)$$

We prove this by Baire's category theorem. It is clear that $X = \cup B_n(0; X)$ so by the surjectivity of T we get $Y = T(\cup B_n(0; X)) = \cup T(B_n(0; X))$. By the category theorem (7.3, this is the point, where we use the completeness of Y) we know that there exists an $n \in \mathbb{N}$ such that $\overline{TB_n(0; X)}$ has nonzero interior, i.e. there exists an ε such that $B_\varepsilon(0; Y) \subseteq \overline{TB_1(0; X)}$. We would need that it holds without closure as well, i.e. $B_\varepsilon(0; Y) \subseteq TB_1(0; X)$. That is a bit too much, we instead prove that $B_{\frac{\varepsilon}{2}}(0; Y) \subseteq TB_1(0; X)$.

Take an arbitrary $y \in B_{\frac{\varepsilon}{2}}(0; Y)$. By $B_\varepsilon(0; Y) \subseteq \overline{TB_1(0; X)}$ it is clear that $y \in \overline{TB_{\frac{1}{2}}(0; X)}$ so there exists a $y_1 \in TB_{\frac{1}{2}}(0; X) : ||y - y_1|| < \frac{\varepsilon}{4}$. By definition there exists x_1 such that $y_1 = Tx_1$. Similarly, apply this procedure on $y - Tx_1 \in B_{\frac{\varepsilon}{4}}(0; Y)$. Then we get an $y_2 = Tx_2$ such that $||y - Tx_1 - Tx_2|| < \frac{\varepsilon}{8}$ holds. Therefore, $y - Tx_1 - Tx_2 \in \overline{TB_{\frac{1}{8}}(0; X)}$ by again $B_\varepsilon(0; Y) \subseteq \overline{TB_1(0; X)}$ and linearity.

Iterate this: we get an $x_n \in B_{\frac{1}{2^n}}(0; X)$ such that $||y - \sum_{j=1}^n Tx_j|| < \frac{\varepsilon}{2^{n+1}}$. This we can define the sequence $z_n = \sum_{j=1}^n x_n \in X$ what is a Cauchy-sequence since

$$||z_n - z_m|| \le \sum_{j=m+1}^{n} ||x_j|| \le \sum_{j=m+1}^{n} \frac{\varepsilon}{2^j} < \varepsilon$$

Therefore, by the completeness of X, we get that (z_n) is convergent. Its limit point $x \in B_1(0; X)$ is the one we are searching for, i.e. by the continuity of T we get $y = Tx$. Hence, $B_{\frac{\varepsilon}{2}}(0; Y) \subseteq TB_1(0; X)$ and that proves the theorem. □

Theorem 7.6. (Inverse mapping theorem) *Let $T : X \to Y$ be a bounded linear bijection between two Banach spaces. Then T^{-1} is also a bounded linear map.*

Proof. T is a bijection, in particular T is surjective. Therefore, by the open mapping theorem T is open. This is equivalent to the continuity of T^{-1} which is – assuming linearity – means boundedness. □

Definition 7.7. Let X and Y be normed spaces and $T : X \rightarrowtail Y$ linear (not necessarily defined on the whole X). Then the *graph* of T is

$$\Gamma(T) = \{(x, Tx) \in X \oplus Y \mid x \in \mathrm{Dom}(T)\}$$

An operator is called *closed* if its graph is a closed subset in $X \oplus Y$.

Remark 7.8. The difference between continuity and closedness can be understood by

- continuity means that if $x_n \to x$ then Tx_n is convergent and the limit is Tx
- closedness means that if $x_n \to x$ and (!) Tx_n is convergent then its limit is Tx

Therefore, it is clear that continuity implies closedness. The following theorem is about the converse in the case of Banach spaces and linear operators.

Theorem 7.9. *(Closed graph theorem) Let X, Y be Banach spaces and let $T : X \to Y$ be linear (not necessarily bounded). Then T is bounded if and only if $\Gamma(T)$ is closed in $X \oplus Y$.*

Proof. If T is bounded then $x_n \to x$ implies $Tx_n \to Tx$ so $(x_n, Tx_n) \to (x, Tx) \in \Gamma(T) \subseteq X \oplus Y$. Conversely, assume that $\Gamma(T)$ is closed. Define the projections $\pi_1 : X \oplus Y \to X$ and $\pi_2 : X \oplus Y \to Y$ on the two coordinates. Then $\pi_1|_{\Gamma(T)} : \Gamma(T) \to X$ is a bounded bijection so – by Inverse mapping theorem – $\pi_1|_{\Gamma(T)}^{-1}$ is bounded as well. (Here, we used the closedness of T because that means that $\Gamma(T)$ is a Banach space which is a requirement to apply Inverse mapping theorem.) Therefore, $T = \pi_2 \circ \pi_1|_{\Gamma(T)}^{-1}$ is bounded as well. □

Corollary 7.10. *When proving continuity, we take a sequence $x_n \to x$ and because of the above theorem, we do not have to care about the convergence of Tx_n. It is enough to prove that $\lim Tx_n = Tx$ where we can assume that Tx_n is convergent. It can help a lot in applications.*

8 Riesz's representation theorem

Lemma 8.1. (Riesz's lemma No. 253... No, I'm just joking.) *Let $\mathcal{H}$ be a Hilbert space and let K be a nonempty, convex, closed set in $\mathcal{H}$. Let $x \in \mathcal{H}$ then the distance between x and K is realized on a unique point of K. In other words, if we introduce the notion*

$$\operatorname{dist}(x, K) := \inf\{||x - k|| \mid k \in K\}$$

then there exists a unique $y \in K$ such that $||x - y|| = \operatorname{dist}(x, K)$.

Proof. If $x \in K$ then the statement is empty. If $x \notin K$ then there exists a sequence $\{y_n\}_{n\in\mathbb{N}} \subseteq K$ such that $||x - y_n|| \to \operatorname{dist}(x, K)$. We show that $\{y_n\}$ is Cauchy. Consider the normsquare instead of the norm:

$$||y_n - y_m||^2 = ||(y_n - x) - (y_m - x)||^2 = 2||y_n - x||^2 + 2||y_m - x||^2 - ||(y_n - x) + (y_m - x)||^2 =$$

where we used the parallelogram identity. This can be rearranged and estimated as:

$$= 2||y_n - x||^2 + 2||y_m - x||^2 - 4\left|\left|\frac{y_n + y_m}{2} - x\right|\right|^2 \leq 2||y_n - x||^2 + 2||y_m - x||^2 - 4\operatorname{dist}(x, K)^2$$

where the first two terms tend to $2\operatorname{dist}(x, K)^2$ so the whole expression tends to zero. Therefore, $\{y_n\}_{n\in\mathbb{N}}$ is in fact a Cauchy-sequence so – by completeness and closedness of K – its limit exists and is in K. Hence, by the continuity of the norm $||y - x|| = ||\lim y_n - x|| = \lim ||y_n - x|| = \operatorname{dist}(x, K)$.

To prove uniqueness, assume that y and y' are elements of K such that $||y - x|| = ||y' - x|| = \operatorname{dist}(x, K)$. Then we can prove $||\frac{y+y'}{2} - x|| < \operatorname{dist}(x, K)$ which contradicts the convexness of K. Indeed, apply the parallelogram identity:

$$4\operatorname{dist}(x, K) = 2||y - x||^2 + 2||y' - x||^2 = ||(y - x) - (y' - x)||^2 + ||(y - x) + (y' - x)||^2 = ||y - y'||^2 + 4\left|\left|\frac{y + y'}{2} - x\right|\right|^2$$

so we got a contradiction. □

Theorem 8.2. (Projection Theorem) *Let $\mathcal{H}$ be a Hilbert space, $\mathcal{M} \subseteq \mathcal{H}$ be a closed linear subspace of $\mathcal{H}$. Then for all $x \in \mathcal{H}$ there exists a unique $x_{||} \in \mathcal{M}$ and $x_\perp \in \mathcal{M}^\perp$ such that $x = x_{||} + x_\perp$.*

Proof. $\mathcal{M}$ is closed and convex, so we can apply the Riesz lemma: there exists an $x_{||} \in \mathcal{M}$ such that for all $y \in \mathcal{M}$, $||y - x|| \geq ||x_{||} - x||$. Obviously, let $x_\perp = x - x_{||}$. To prove existence, we have to show that $x_\perp \in \mathcal{M}^\perp$.

We know that for all $t \in \mathbb{R}$, and for all $y \in \mathcal{M}$

$$||x - (x_{||} + ty)||^2 \geq ||x - x_{||}||^2$$

where both sides can be expressed in terms of the scalar product:

$$||x - x_{||}|| + t^2||y||^2 - 2t\mathrm{Re}\langle x - x_{||}, y\rangle \geq ||x - x_{||}||$$

i.e. $t^2||y||^2 - 2t\mathrm{Re}\langle x, y\rangle \geq 0$ for all $t \in \mathbb{R}$ where we fixed y. But it can happen only if $\mathrm{Re}\langle x, y\rangle = 0$ for all $y \in \mathcal{M}$. This also implies that $\langle x, y\rangle = 0$ what can be seen by multiplying y by arbitrary complex unit. So $x_\perp \in \mathcal{M}^\perp$.

The uniqueness can be checked easily. □

Definition 8.3. Let $\mathcal{M}$ be a closed subspace of the Hilbert space $\mathcal{H}$. Then $P_\mathcal{M} : \mathcal{H} \to \mathcal{M}$, $x \mapsto x_{||}$ is a bounded linear mapping called the *orthogonal projection* to $\mathcal{M}$ where $\mathrm{Ker}P_\mathcal{M} = \mathcal{M}^\perp$ and $\mathrm{Ran}P_\mathcal{M} = \mathcal{M}$.

Lemma 8.4. *An equivalent (algebraic) definition of orthogonal projection $P \in \mathcal{B}(\mathcal{H})$ is requiring $P^2 = P = P^*$.*

Remark 8.5. Let $\mathcal{P}(\mathcal{H})$ be the set of closed subspaces (or, what is the same, the orthogonal projections) in $\mathcal{H}$. Then $\mathcal{P}(\mathcal{H})$ has a lattice structure: the $\vee$ is the generated subspace and $\wedge$ is the intersection.

This $\mathcal{P}(\mathcal{H})$ is the event lattice in Quantum Mechanics. And the reason why quantum mechanics behaves weird is that this lattice is not distributive. (e.g. three 1-dimensional linear subspaces in two dimension contradicts distributivity.)

Theorem 8.6. (Riesz representation theorem) *Let $\mathcal{H}$ be a Hilbert space and denote by $\mathcal{H}^*$ its dual. Then for all $\Phi \in \mathcal{H}^*$ there exists a unique $x_\Phi \in \mathcal{H}$ such that for all $y \in \mathcal{H}$: $\Phi(y) = \langle x_\Phi, y\rangle$.*

Moreover, $||\Phi|| = ||x_\Phi||$. The mapping $\mathcal{H}^* \ni \Phi \mapsto x_\Phi \in \mathcal{H}$ is a norm-preserving conjugate-linear bijection.

Proof. If $\mathrm{Ker}\Phi = \mathcal{H}$ then we choose $x_\Phi := 0$. Assume that $\mathrm{Ker}\Phi \neq \mathcal{H}$ and take a vector $y \in \mathcal{H} \setminus \mathrm{Ker}\Phi$. We know that $\mathrm{Ker}\Phi$ is a closed subspace so there exists a nonzero $\tilde{x} \in (\mathrm{Ker}\Phi)^\perp$ by the projection theorem. Then for all $y \in \mathcal{H}$ we have a decomposition

$$y = \frac{\Phi(y)}{\Phi(\tilde{x})}\tilde{x} + \left(y - \frac{\Phi(y)}{\Phi(\tilde{x})}\tilde{x}\right)$$

where $\frac{\Phi(y)}{\Phi(\tilde{x})}\tilde{x} \in (\mathrm{Ker}\Phi)^\perp$ and $y - \frac{\Phi(y)}{\Phi(\tilde{x})}\tilde{x} \in \mathrm{Ker}\Phi$. It is clear that $\dim(\mathrm{Ker}\Phi)^\perp = 1$ so the representing vector must be scalar times $\tilde{x}$. Let $x_\Phi = \frac{\overline{\Phi(\tilde{x})}}{||\tilde{x}||^2}\tilde{x}$ because in that case

$$\langle x_\Phi, \tilde{x}\rangle = \frac{\Phi(\tilde{x})}{||\tilde{x}||^2}\langle \tilde{x}, \tilde{x}\rangle = \Phi(\tilde{x})$$

so at least the required identity holds for $\tilde{x}$. In fact, it holds for all vectors by the decomposition:

$$\langle x_\Phi, y\rangle = \left\langle \frac{\overline{\Phi(\tilde{x})}}{||\tilde{x}||^2}\tilde{x}, \frac{\Phi(y)}{\Phi(\tilde{x})}\tilde{x} + \left(y - \frac{\Phi(y)}{\Phi(\tilde{x})}\tilde{x}\right)\right\rangle = \frac{\Phi(\tilde{x})}{||\tilde{x}||^2}\frac{\Phi(y)}{\Phi(\tilde{x})}\langle \tilde{x}, \tilde{x}\rangle = \Phi(y)$$

To finish the theorem we only have to compute the norm of x_Φ:

$$||x_\Phi|| = \left|\left|\frac{\overline{\Phi(\tilde{x})}}{||\tilde{x}||^2}\tilde{x}\right|\right| = \frac{|\Phi(\tilde{x})|}{||\tilde{x}||} \leq ||\Phi||$$

and conversely,

$$|\Phi(y)| = |\langle x_\Phi, y\rangle| \leq ||x_\Phi|| \cdot ||y|| \qquad \text{so} \qquad ||\Phi|| \leq ||x_\Phi||$$

□

9 Topologies on Hilbert spaces

Definition 9.1. Vector topologies:

1. Norm topology: $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathcal{H}$ converges to $x \in \mathcal{H}$ in the norm if $||x_n - x|| \to 0$. Notation: $x_n \to x$
2. *Weak topology:* $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathcal{H}$ converges weakly to $x \in \mathcal{H}$ in the norm if for all $y \in \mathcal{H}$: $\langle x_n, y\rangle \to \langle x, y\rangle$. Notation: $x_n \overset{w}{\to} x$ or $x_n \rightharpoonup x$

Lemma 9.2. *$x_n \to x$ implies $x_n \overset{w}{\to} x$. (The converse is not true.)*

Proof. Homework. ☐

Lemma 9.3. *$x_n \to x$ is equivalent to $x_n \overset{w}{\to} x$ and $||x_n|| \to ||x||$.*

Proof. $\Rightarrow$ is trivial. $\Leftarrow$ follows from $||x_n - x||^2 = ||x_n||^2 + ||x||^2 - 2\mathrm{Re}\langle x_n, x\rangle$ and $\mathrm{Re}||x||^2 = ||x||^2$. ☐

Example 9.4. If $\mathcal{H}$ is the (or 'a', if you like it that way) separable Hilbert space then a complete orthonormal system converges to zero weakly but not in norm.

Definition 9.5. Operator topologies:

1. *Operator norm topology:* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}(\mathcal{H})$ converges to $A \in \mathcal{H}$ in the norm if $||A_n - A|| \to 0$. Notation: $A_n \to A$
2. *Strong operator topology:* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}(\mathcal{H})$ converges to $A \in \mathcal{H}$ in strong operator topology if for all $x \in \mathcal{H}$ $A_n x \to Ax$, i.e. $||Ax_n - Ax|| \to 0$ for all $x \in \mathcal{H}$ Notation: $A_n \overset{s.o.}{\to} A$
3. *Weak operator topology:* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}(\mathcal{H})$ converges to $A \in \mathcal{H}$ in weak operator topology if for all $x \in \mathcal{H}$ $A_n x \overset{w}{\to} Ax$, i.e. for all $x, y \in \mathcal{H}$: $\langle y, A_n x\rangle \to \langle y, Ax\rangle$. Notation: $A_n \overset{w.o.}{\to} A$

Remark 9.6. $A_n \to A \Rightarrow A_n \overset{s.o.}{\to} A \Rightarrow A_n \overset{w.o.}{\to} A$ but none of the implications can be reversed.

Example 9.7. The powers of left and right shifts are general counterexamples for such statements.

Seventh Lecture, 13th of November

9.1 Adjoint of a bounded operator

Theorem 9.8. *Let $\mathcal{H}$ be a Hilbert-space and $A \in \mathcal{B}(\mathcal{H})$. Then there exists a unique $A^* \in \mathcal{B}(A)$ for which*

$$\langle A^* x, y\rangle = \langle x, Ay\rangle \qquad \forall x, y \in \mathcal{H}$$

Definition 9.9. A^* with the above property is the *adjoint* of $A \in \mathcal{B}(\mathcal{H})$.

Proof. (Relies on the Riesz representation theorem) Let $A \in \mathcal{B}(\mathcal{H})$. Then we can consider the map $\mathcal{H} \to \mathbb{C}$; $y \mapsto \langle x, Ay\rangle$ is a bounded linear operator such that $||\langle x, A\cdot\rangle|| \leq ||A|| \cdot ||x||$. So – by the Riesz representation theorem – there exists a unique $A^* x \in \mathcal{H}$ such that $\langle A^* x, \cdot\rangle = \langle x, A\cdot\rangle$. It can be shown that the map $x \mapsto A^* x$ is linear. ☐

Remark 9.10. Properties of the adjoint:

1. $A \to A^*$ is conjugate linear, i.e. $(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$
2. $||A^*|| = ||A||$

   Proof. It follows from the fact that for any $A \in \mathcal{B}(\mathcal{H})$ we have $||A|| = \sup\{|\langle x, Ay\rangle| \mid x, y \in \mathcal{H},\ ||x|| \leq 1,\ ||y|| \leq y\}$. ☐

3. $(AB)^* = B^*A^*$
4. If A is invertible then $(A^*)^{-1} = (A^{-1})^*$
5. $(A^*)^*$

Example 9.11. Let L and $R : \ell^2(\mathbb{N}) \to \ell^2(\mathbb{N})$ be the left and right shift. It is easy to see that $L^* = R$ and $R^* = L$.

10 Compact operators

Definition 10.1. Let $K \in \mathcal{B}(\mathcal{H})$ is *compact* if for all weakly converging sequence $(x_n) \subseteq \mathcal{H}$, (Kx_n) is a convergent sequence in the norm topology.

Lemma 10.2. *If K is compact then $Kx_n \to K \lim x_n$*

Proof. A bounded operator keeps weak convergence: For all $y \in \mathcal{H}$,

$$|\langle y, Ax_n\rangle - \langle y, Ax\rangle| = |\langle y, A(x_n - x)\rangle| = |\langle A^*y, x_n - x\rangle| \to 0$$

What is more, norm-convergence implies weak convergence so if (Kx_n) would converge to a different point in the norm then (Kx_n) would converge to it in the weak-topology. This contradicts the uniqueness of the weak-limit. □

Notation: Let us denote the set of all compact operators of $\mathcal{H}$ by $\mathcal{K}(\mathcal{H})$.

Theorem 10.3. *$\mathcal{K}(\mathcal{H})$ is a norm-closed two-sided $*$-ideal (i.e. it is an ideal closed under the taking-the-adjoint) in $\mathcal{B}(\mathcal{H})$.*

Proof. It is clear that $\mathcal{K}(\mathcal{H})$ is a subspace. To prove that it is also an ideal, we use only that an arbitrary bounded operator keeps weak- and norm-convergence. So let $K \in \mathcal{K}(\mathcal{H})$ and $A \in \mathcal{B}(\mathcal{H})$ and consider a weakly converging sequence (x_n). Then Ax_n is still weakly convergent so KAx_n is norm-convergent. Similarly, on the other side.

$\mathcal{K}(\mathcal{H})$ is closed in the norm in $\mathcal{B}(\mathcal{H})$: Let $K_n \in \mathcal{K}(\mathcal{H})$ and $A \in \mathcal{B}(\mathcal{H})$ such that $||K_n - A|| \to 0$ $(n \to \infty)$. Let (x_l) be a weakly convergent sequence in $\mathcal{H}$. We have to show that $||Ax_l - Ax|| \to 0$.

$$||A(x_l - x)|| = ||(A - K_n)(x_l - x) + K_n(x_l - x)|| \leq ||A - K_n|| \cdot ||x_l - x)|| + ||K_n(x_l - x)|| = ||A - K_n|| \cdot C + ||K_n(x_l - x)||$$

because a weakly convergent sequence is bounded (see Homework). What is more, for big enough n the first part is at most $\frac{\varepsilon}{2C}$ and for that n we can choose an l such that the second term is $< \frac{\varepsilon}{2}$ since $K_nx_l \to K_nx$ in the norm.

To get that $\mathcal{K}(\mathcal{H})$ is closed under taking the adjoint: consider a compact operator K and a weakly convergent sequence (x_n). Then

$$||K^*x_n - K^*x||^2 = |\langle K^*x_n - K^*x, K^*x_n - K^*x\rangle| = |\langle x_n - x, KK^*(x_n - x)\rangle| \leq ||x_n - x|| \cdot ||KK^*(x_n - x)|| \to 0$$

since $||x_n - x||$ is bounded (because of the weak convergence) and the second term tends to zero. □

Remark 10.4. Compact operators behave similarly to finite dimensional operators. In the following we will make it precise.

Theorem 10.5. $\mathcal{K}(\mathcal{H}) = \overline{\{A \in \mathcal{B}(\mathcal{H}) \mid \text{rank}\,(A) < \infty\}}^{||.||}$

Proof. (We only prove for separable Hilbert spaces.) Let us denote the set $\{A \in \mathcal{B}(\mathcal{H}) \mid \operatorname{rank}(A) < \infty\}$ by $\mathcal{F}(\mathcal{H})$. First, one has to note that every finite rank operator is compact because in a finite dimensional space weak convergence is equivalent to norm-convergence. After that it is clear that $\overline{\mathcal{F}(\mathcal{H})} \subseteq \mathcal{K}(\mathcal{H})$ by Theorem 10.3.

Conversely, we show that any compact operator can be approximated by finite rank operators in the norm topology. Let (e_n) be a complete orthonormal system in $\mathcal{H}$. Let $P_n \in \mathcal{B}(\mathcal{H})$ be the orthogonal projections to $\operatorname{Span}\{e_0, e_1, \ldots, e_n\}$. We claim that the finite rank operators KP_n tend to K.

Indirectly, assume that KP_n does not converge to K. Then we can choose a subsequence KP_{n_i} such that $||K - KP_{n_i}|| > \varepsilon$ for all $i \in \mathbb{N}$. I.e. there exists a sequence of unit length vectors $(x_i) \subseteq \mathcal{H}$ such that

$$||K(I - P_{n_i})x_i|| > \varepsilon$$

while $(I - P_{n_i})x_i \xrightarrow{w} 0$ since if we take its inner product with an arbitrary $z \in \mathcal{H}$:

$$|\langle z, (I - P_{n_i})x_i\rangle| = |\langle (I - P_{n_i})^* y, x_i\rangle| = |\langle (I - P_{n_i})y, x_i\rangle| \leq ||(I - P_{n_i})y|| \cdot ||x_i|| = ||(I - P_{n_i})y|| \to 0$$

so $(I - P_{n_i})x_i \xrightarrow{w} 0$ but $||K(I - P_{n_i})x_i|| > \varepsilon$ which is a contradiction. □

Theorem 10.6. *(Riesz-Schauder) Let $K \in \mathcal{K}(\mathcal{H})$ be a compact operator. Then, if $0 \neq \lambda \in \sigma(K)$ (spectrum) then it is an eigenvalue with finite multiplicity. What is more, the only possible accumulation point of $\sigma(K)$ is zero.*

11 Orthogonal polynomials

Definition 11.1. Let us consider the Hilbert space $\mathcal{H} = L^2(I, \rho)$ where I is a not necessarily finite interval, and $\rho : I \to \mathbb{C}$ is a weight function. As a set

$$L^2(I, \rho) := \{f : I \to \mathbb{C} \mid f \text{ is measurable}, ||f||_p^2 = \int_I |f(x)|^2 \rho(x) \mathrm{d}\lambda(x) < \infty\} / \sim$$

where λ stands for the Lebesgue measure and $\sim$ is the almost-everywhere equivalence. On this space, we have a natural inner product:

$$\langle f, g\rangle = \int_I \overline{f} g \rho \mathrm{d}\lambda$$

With that it will be a Hilbert space.

11.1 Legendre polynomials

Fact 11.2. *The polynomials constitute a dense set in $\mathcal{H}$ with respect to $||.||_2$, by the Weierstrass theorem.*

Definition 11.3. *Legendre polynomials:* Let $\mathcal{H} = L^2([-1, 1], \rho)$ with $\rho \equiv 1$. The Legendre polynomials $\{P_n\}$ are obtained by applying the Gram-Schmidt orthogonalization procedure to the power functions $1, x, \ldots, x^n, \ldots$. We denote by p_n the scalar-multiples of Legendre polynomials with leading coefficient 1.

Corollary 11.4. *(of the previous fact) $\{p_n\}$ is a complete orthogonal system in $L^2[-1, 1]$.*

Example 11.5. $p_0(x) = 1$, $p_1(x) = x$ (by $p_0 \perp p_1$), $p_2(x) = x^2 - \frac{1}{3}$.

Theorem 11.6. (Rodriquez-formula for Legendre polynomials) $P_n(x) = \frac{1}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2 - 1^n)$

Remark 11.7. $(x^2 - 1)^n$ has leading coefficient x^{2n} so after the n-th derivation it has leading coefficient $\frac{(2n)!}{n!}$. Therefore, by the theorem, the leading coefficient of P_n is $\alpha_n := \frac{1}{2^n}\binom{2n}{n}$

Proof. It is clear that the formula on the right hand side has degree n. So it is enough to prove that for any polynomial ν_k of degree $k < n$ we have $\nu_k \perp P_n$. And the latter can be computed:

$$\langle P_n, \nu_k \rangle = \frac{1}{2^n n!} \int_{-1}^{1} \left(\frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2 - 1^n) \right) \cdot \nu_k(x)\mathrm{d}x = \dots (n \text{ times partial derivation}) \dots =$$
$$= \frac{1}{2^n n!}(-1)^n \int_{-1}^{1} (x^2-1)^n \frac{\mathrm{d}^n}{\mathrm{d}x^n} \nu_k(x)\mathrm{d}x = 0$$

because all the other terms made by the partial derivation fall out. □

11.2 Hermite polynomials

Definition 11.8. *Hermite-polynomials:* $\mathcal{H} = L^2([-\infty, \infty], e^{-x^2})$ so we can again define the Hermite polynomials as the ones obtained by applying the Gram-Schmidt orthogonalization procedure to the power functions $1, x, \dots, x^n, \dots$. Notation: $H_n(x)$. The equality $\overline{\mathrm{Span}\{x^n \mid n \in \mathbb{N}\}} = \mathcal{H}$ implies that we will again get a complete system.by the Hermite polynomials.

Theorem 11.9. *(Rodriguez formula for Hermite polynomials)* $H_n(x) = (-1)^n e^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2}$

Remark 11.10. The Leading coefficient is $(-1)^n \cdot (-2)^n = 2^n$ by this formula.

Proof. First, it can be easily checked the the formula gives polynomials. After, the proof goes the same way as in the case of Legendre polynomials, i.e. it is enough to prove that for any polynomial ν_k of degree $k < n$ we have $\nu_k \perp H_n$. That can be calculated. □

Corollary 11.11. *Equations for Hermite polynomials: (can be proved by the above formula)*

1. $H_n'(x) = 2xH_n(x) - H_{n+1}(x)$
2. *(Recursion)* $H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x)$
3. *(Derived Recursion)* $H_n'(x) = 2nH_{n-1}(x)$
4. *(Differential equation)* $H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0$

Proposition 11.12. *Generating function for Hermite polynomials:*

$$G(x, z) = e^{z(2x-z)} = \sum_{n=0}^{\infty} H_n(x) z^n$$

Corollary 11.13. *We have seen that H_n's form a complete orthogonal system in $L^2(\mathbb{R}, e^{-x^2})$. By this one gets that $H_n(x)e^{-\frac{x^2}{2}}$ is a complete orthogonal system in $L^2(\mathbb{R})$.*

11.3 Laguerre polynomials

Definition 11.14. Let $\mathcal{H} = L^2([0, \infty], e^{-x})$ and by the same definition as before, one gets the Laguerre polynomials.

Theorem 11.15. *Rodriquez-formula:* $L_n(x) = \frac{e^x}{n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^n e^{-x})$

Remark 11.16. The leading coefficient is $\frac{(-1)^n}{n!}$.

Eighth lecture, 20th of November

12 Spectrum of operators in $\mathcal{B}(\mathcal{H})$

Definition 12.1. Let $T \in \mathcal{B}(\mathcal{H})$. Then the *resolvent set* of T is

$$\rho(T) := \{\lambda \in \mathbb{C} \mid \exists(\lambda I - T)^{-1} \in \mathcal{B}(\mathcal{H})\}$$

While its complement $\sigma(T) := \mathbb{C}\backslash\rho(T)$ is called the *spectrum* of T. A related notation is the *resolvent operator* of T for a given $\lambda \in \rho(T)$ is $R_\lambda(T) := (\lambda I - T)^{-1}$.

Example 12.2. If $A \in M_m(\mathbb{C})$ then $\lambda \in \sigma(A)$ if and only if λ is an eigenvalue. So we just got an infinite dimensional generalization of the notion of eigenvalue.

Remark 12.3. For a bounded operator $A \in \mathcal{B}(\mathcal{H})$ its invertibility is equivalent to $\operatorname{Ker} A = \{0\}$ and $\operatorname{Ran} A = \mathcal{H}$ by the Inverse Mapping Theorem 7.6.

Definition 12.4. Fix a bounded operator T. Then the parts of the spectrum are

1. Point spectrum: $\sigma_p(T) := \big\{\lambda \in \mathbb{C} \mid \operatorname{Ker}(\lambda I - T) \neq \{0\}\big\}$
2. Continuous spectrum: $\sigma_c(T) := \{\lambda \in \sigma(T) \mid \operatorname{Ker}(\lambda I - T) = \{0\}\ \overline{\operatorname{Ran}(\lambda I - T)} = \mathcal{H}\}$
3. Residual spectrum: $\sigma_r(T) := \{\lambda \in \sigma(T) \mid \operatorname{Ker}(\lambda I - T) = \{0\}\ \overline{\operatorname{Ran}(\lambda I - T)} \neq \mathcal{H}\}$

Corollary 12.5. *By the definitions we got a decomposition of the spectrum:* $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$ *where the unions are disjoint.*

Remark 12.6. Note that in the case of continuous spectrum the range is only dense but is proper all the time, because if $\lambda I - T$ is surjective as well then it is invertible.

Theorem 12.7. $\sigma(T)$ *is a non-empty, closed, bounded set in* $\overline{B_0(||T||)}$.

Definition 12.8. If $\lambda > ||A||$ then the corresponding Neumann series is

$$(\lambda I - A)^{-1} = \frac{1}{\lambda}\sum_{n=0}^{\infty}\left(\frac{A}{\lambda}\right)^n \in \mathcal{B}(\mathcal{H})$$

It is a short exercise to prove the equality and the norm-convergence of the right hand side.

Proof. (of theorem 12.7) If $|\lambda| > ||T||$ then the convergence of the Neumann series proves that the inverse exists and is bounded.

The closedness of $\sigma(T)$ is equivalent to that the resolvent set is open. So take a $\lambda_0 \in \rho(T)$ and let $\lambda \in \mathbb{C}$. Then

$$R_\lambda(T) = (\lambda I - T)^{-1} = \big((\lambda - \lambda_0)I + \lambda_0 I - T\big)^{-1} = \Big((\lambda_0 I - T)\big((\lambda - \lambda_0)(\lambda_0 I - T)^{-1} + I\big)\Big)^{-1} =$$

$$= \big((\lambda - \lambda_0)R_{\lambda_0}(T) + I\big)^{-1} \cdot R_{\lambda_0}(T) = \big(I - (\lambda_0 - \lambda)R_{\lambda_0}(T)\big)^{-1} \cdot R_{\lambda_0}(T) = \sum_{n=0}^{\infty}(\lambda_0 - \lambda)^n R_{\lambda_0}^{n+1}(T)$$

so we even got an explicit expression for $R_\lambda(T)$ if $|\lambda - \lambda_0| < \frac{1}{R_{\lambda_0}(T)}$ namely

$$R_\lambda(T) = \sum_{n=0}^{\infty}(\lambda_0 - \lambda)^n R_{\lambda_0}^{n+1}(T)$$

which proves that the resolvent set is open.

To ensure that $\sigma(T)$ is nonempty, assume on the contrary that $\rho(T) = \mathbb{C}$. Let $x, y \in \mathcal{H}$ and define

$$f_{x,y} : \mathbb{C} \to \mathbb{C}$$

$$\lambda \mapsto f_{x,y}(\lambda) = \langle x, R_\lambda(T)y\rangle$$

We will prove that this function is holomorphic and bounded and by the aid of Liouville's Theorem we conclude that it is constant which is a contradiction.

To prove holomorphicity at $\lambda_0 \in \mathbb{C}$ we can express $f_{x,y}$ in a neighborhood of λ by the previous expression of $R_\lambda(T)$: If $|\lambda - \lambda_0| < \frac{1}{R_{\lambda_0}(T)}$ then

$$f_{x,y}(\lambda) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n \langle x, R_{\lambda_0}^{n+1}(T)y\rangle$$

so it is in fact holomorphic because the sum is uniformly convergent in a small enough neighborhood by

$$|f_{x,y}(\lambda)| \leq \sum_{n=0}^{\infty} |\lambda_0 - \lambda|^n \cdot ||x|| \cdot ||y|| \cdot ||R_{\lambda_0}(T)||^{n+1} = ||x|| \cdot ||y|| \cdot ||R_{\lambda_0}(T)|| \cdot \frac{1}{1 - |\lambda_0 - \lambda| \cdot ||R_{\lambda_0}(T)||}$$

Now, $f_{x,y}$ is bounded because if $\lambda > ||T||$

$$|f_{x,y}(\lambda)| = |\langle x, (\lambda I - T)^{-1} y\rangle| = \left| \frac{1}{\lambda} \sum_{n=0}^{\infty} \left\langle x, \left(\frac{T}{\lambda}\right)^n y \right\rangle \right| \leq \frac{1}{|\lambda|} \sum ||x|| \cdot ||y|| \cdot \left(\frac{||T||}{||\lambda||}\right)^n =$$

$$\frac{1}{|\lambda|} \cdot ||x|| \cdot ||y|| \cdot \frac{1}{1 - \frac{||T||}{|\lambda|}} = ||x|| \cdot ||y|| \cdot \frac{1}{|\lambda| - ||T||} \overset{|\lambda| \to \infty}{\longrightarrow} 0$$

so if $|\lambda| > 2||T||$ then

$$|f_{x,y}(T)| \leq ||x|| \cdot ||y|| \cdot \frac{1}{||T||} < \infty$$

and in $|\lambda| \leq 2||T||$ $f_{x,y}$ is continuous hence bounded. So $f_{x,y}$ is bounded on the whole $\mathbb{C}$ as we stated. $\square$

Theorem 12.9. *If $U \in \mathcal{B}(\mathcal{H})$ is a unitary operator then $\sigma(U) \subseteq \mathbb{S}^1 = \{|z| = 1\}$. (Homework)*

Lemma 12.10. *The kernel $\operatorname{Ker} A$ of a bounded operator (which is a closed subspace) is the orthogonal complement of $\operatorname{Ran} A^*$. In short, $\operatorname{Ker} A = (\operatorname{Ran} A^*)^\perp$. (Homework.)*

Remark 12.11. Note that $\operatorname{Ran} A^*$ is a subspace which is not necessarily closed.

Theorem 12.12. *(The proof is homework.)*

1. *$\lambda \in \sigma(A)$ if and only if $\overline{\lambda} \in \sigma(A^*)$,*
2. *If $\lambda \in \sigma_p(A)$ then $\overline{\lambda} \in \sigma_p(A^*) \cup \sigma_r(A^*)$,*
3. *If $\lambda \in \sigma_r(A)$ then $\overline{\lambda} \in \sigma_p(A^*)$,*
4. *$\lambda \in \sigma_c(A)$ if and only if $\sigma_c(A^*)$.*

Example 12.13. The spectrum and its parts of the shift operators on ℓ^2:

Let $L : \ell^2 \to \ell^2$ be $L(x_0, x_1, \dots) = (x_1, x_2, \dots)$. If $\lambda \in \sigma_p(L)$ i.e. it is an eigenvalue then $x_i = \lambda x_{i-1}$ must hold for all $i > 0$. So if $x_0 = 0$ then $x = 0$ which cannot be an eigenvector and if $x_0 \neq 0$ then

$$v = (x_0, \lambda x_0, \lambda^2 x_0, \dots)$$

is an eigenvector of L. So $\sigma_p(L) = \{\lambda \in \mathbb{C} \mid |\lambda| < 1\}$. By $||L|| = 1$ we get that $\sigma(L) \subseteq \overline{B_1(0)}$ but it must be closed so in fact $\sigma(L) = \overline{B_1(0)}$. The spectrum of the right shift $R = L^*$ is also the closed unit ball by the above Theorem 12.12.

Claim 12.14. $\sigma_p(R) = \emptyset$

Proof. If $R(x_0, x_1, x_2, \dots) = (0, x_0, x_1, \dots) = (\lambda x_0, \lambda x_1, \lambda x_2, \dots)$ then in the case of $\lambda \neq 0$ we get that $x_i = 0$ for all $i \geq 0$. Hence, we cannot find a nonzero eigenvector for λ. If $\lambda = 0$ then it is clear that $x_i = 0$ for all $i \geq 0$ again. The claim follows. $\square$

From the claim we can conclude that $\sigma_p(R) = \emptyset$ so the residual spectrum of the Left shift $\sigma_r(L) = \emptyset$ by the above Theorem 12.12. So the result is that

$$\sigma_c(L) = \mathbb{S}^1 = \sigma_c(R) \qquad \text{and} \qquad \sigma_r(R) = \{z \mid |z| < 1\} = \sigma_p(L)$$

which completely describes the parts of the spectrum of the two operators. The example shows that considering the spectrum of the adjoint can help to compute the spectrum of the original operator.

Exercise 12.15. What is the spectrum and its parts for a multiplication operator

$$M_f : L^2[0,1] \to L^2[0,1]$$
$$g \mapsto M_f(g) = fg$$

where f is bounded.

Exercise 12.16. Let $f, g \in \mathcal{H}$. What is the spectrum of the rank one operator $|f\rangle\langle g| \in \mathcal{B}(\mathcal{H})$?

Definition 12.17. Let $A \in \mathcal{B}(\mathcal{H})$. The spectral radius of A is

$$r(A) := \sup \{|\lambda| \mid \lambda \in \sigma(A)\} \in \mathbb{R}_{\geq 0}$$

Remark 12.18. Obviously $r(A) \leq ||A||$.

Theorem 12.19. *Let $A \in \mathcal{B}(\mathcal{H})$. Then the spectral radius can be expressed as*

$$r(A) = \lim_{n \to \infty} \left(||A^n||\right)^{\frac{1}{n}}$$

Proof. Omitted. $\square$

Lemma 12.20. *If $A = A^* \in \mathcal{B}(\mathcal{H})$ then $r(A) = ||A||$.*

Proof. Homework. Hint: Use the previous theorem and the basic facts $||A|| = ||A^*||$ and $||AA^*|| = ||A||^2$. $\square$

Theorem 12.21. *Let $N \in \mathcal{B}(\mathcal{H})$ be a normal operator (i.e. $NN^* = N^*N$). Then $\lambda \in \rho(N)$ if and only if $||(\lambda I - N)x|| \geq c||x||$ for all $x \in \mathcal{H}$.*

Proof. To prove one implication, assume that $\lambda \in \rho(N)$. Then

$$||x|| = ||R_\lambda(N) \cdot (\lambda I - N)x|| \leq ||R_\lambda(N)|| \cdot ||(\lambda I - N)x||$$

so by the choice $c = \frac{1}{||R_\lambda(N)||}$ the statement is satisfied.

Conversely, if $||(\lambda I - N)x|| \geq c||x||$ where $c > 0$ then it is clear that $\mathrm{Ker}(\lambda I - N) = \{0\}$. We can consider the inverse of $\lambda I - N$ on $\mathrm{Ran}(\lambda I - N)$. We want to prove that it is defined on a dense subspace. For that, consider the following equation:

$$||(\lambda I - N)x||^2 = \langle(\lambda I - N)x, (\lambda I - N)x\rangle = \langle x, (\overline{\lambda} I - N^*)(\lambda I - N)x\rangle \overset{\text{normality}}{=}$$
$$= \langle(\overline{\lambda} I - N^*)x, (\overline{\lambda} I - N^*)x\rangle = ||(\overline{\lambda} I - N^*)x||^2$$

so by the assumption $||(\lambda I - N)x|| \geq c||x||$ we get $||(\overline{\lambda} I - N^*)x|| \geq c||x||$, in particular

$$\left(\mathrm{Ran}(\lambda I - N)\right)^\perp = \mathrm{Ker}(\overline{\lambda} I - N^*) = \{0\}$$

therefore $\overline{\mathrm{Ran}\,(\lambda I - N)} = \mathcal{H}$.

Now, let $(\lambda I - N)x = y$. We know that $||y|| \geq c||x||$ so $||x|| = ||(\lambda I - N)^{-1}y|| \leq \frac{1}{c}||y||$ what means $||(\lambda I - N)^{-1}|| \leq \frac{1}{c}$. Hence, we are in a situation that $(\lambda I - N)^{-1}$ is a bounded operator defined on a dense subspace so it extends uniquely to the whole space $\mathcal{H}$. Therefore $\lambda \in \rho(N)$. $\square$

Theorem 12.22. *Let $H = H^* \in \mathcal{B}(\mathcal{H})$. Then*

1. *Then $\sigma_r(H) = \emptyset$*
2. *$\sigma(H) \subseteq [-||H||, ||H||] \subseteq \mathbb{R}$*
3. *If λ and μ are distinct points of $\sigma_p(H)$ then the corresponding eigenvectors are orthogonal. In the language of spectral projections: $P_\lambda P_\mu = 0 = P_\mu P_\lambda$ for all $\lambda \neq \mu$.*

Proof. The first statement is clear by the self-adjointness since $\lambda \in \sigma_r(H)$ implies $\lambda \in \sigma_p(H^*) = \sigma_p(H)$ by Theorem 12.12, which contradicts the disjointness of $\sigma_r(H)$ and $\sigma_p(H)$. To verify the second statement, take a $\lambda = a + bi \in \mathbb{C}$ with nonzero b. For this,

$$||(\lambda I - H)x||^2 = \langle((a+bi)I - H)x, ((a+bi)I - H)x\rangle = ||(aI - H)x||^2 + b^2||x||^2 \geq b^2||x||^2$$

so $||(\lambda I - H)x||^2 \geq b^2||x||^2$ which – by the previous theorem 12.21 and the fact the self-adjointness implies normality – means that $a + bi \in \rho(H)$. (Self-adjointness is used in the calculation.)

To get the third statement let v_λ and v_μ be eigenvectors with eigenvalue λ and μ respectively. Then

$$(\lambda - \mu)\langle v_\mu, v_\lambda\rangle = \langle v_\mu, Hv_\lambda\rangle - \langle Hv_\mu, v_\lambda\rangle = \langle Hv_\mu, v_\lambda\rangle - \langle Hv_\mu, v_\lambda\rangle = 0$$

so either $\lambda = \mu$ or $\langle v_\mu, v_\lambda\rangle = 0$. □

13 Spectral decomposition of a normal operator

If $N \in M_n(\mathbb{C})$ is a normal matrix then its spectral decomposition is

$$N = \sum_{l=1}^{k} \lambda_l P_l$$

where k is the number of eigenvalues and P_l's are the orthogonal spectral projections. This helps to compute the functions of N, namely

$$f(N) = \sum_{l=1}^{k} f(\lambda_l) P_l$$

The decomposition will have a generalization to infinite dimensional normal operators:

$$N = \int_{\lambda \in \sigma(N)} \lambda \, \mathrm{d}P(\lambda)$$

where P is a projection-valued measure.

Ninth Lecture, 27th of November

13.1 Functions of self-adjoint operators

Question: For a self-adjoint operator $A \in \mathcal{B}(\mathcal{H})$ and a function $f : \mathbb{C} \to \mathbb{C}$ how can we give a meaning of $f(A)$?

If f is a polynomial $f(z) = \sum a_n z^n$ then there is no problem since we can take $f(A) = \sum a_n A^n$. Generally, if f is an analytic function $f(z) = \sum_{n=0}^{\infty} a_n z^n$ and this series is convergent for all z such that $|z| < R$ where $R > ||A||$ then $f(A) = \sum_{n=0}^{\infty} a_n A^n$ is still a valid definition since it gives an absolutely convergent series of operators.

Question: What if we consider continuous functions?

Let $\mathcal{C}(T) := \{f : T \to \mathbb{C} \mid f \text{ is continuous}\}$. Then $(\mathcal{C}(T), ||.||_\infty)$ becomes a commutative Banach algebra with pointwise multiplication. For $f \in \mathcal{C}(T)$ we can define $f^* := \overline{f}$. With this definition $(\mathcal{C}(T), ||.||_\infty)$ becomes a C^*-algebra.

Definition 13.1. $\mathcal{A}$ is a C^*-algebra, if it is a Banach-algebra and there is an operation $* : \mathcal{A} \to \mathcal{A}$ with the following: for all $A, B \in \mathcal{A}$:

1. $(A^*)^* = A$,
2. $(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$ for all $\alpha, \beta \in \mathbb{C}$,
3. $(AB)^* = B^* A^*$, and
4. $||AA^*|| = ||A||^2$ called the C^*-property

Theorem 13.2. (Gelfand, Naimark) *Every C^*-algebra is isometrically $*$-isomorphic to a (norm-)closed subalgebra of $\mathcal{B}(\mathcal{H})$ for a Hilbert space $\mathcal{H}$.*

Theorem 13.3. (Gelfand, Naimark) *Any commutative C^*-algebra is isometrically $*$-isomorphic to*

$$\mathcal{C}_0(T) = \{f \in \mathcal{C}(T) \mid \forall \varepsilon > 0 \; \exists K \subseteq T \text{ compact such that } f|_{T \setminus K} < \varepsilon\}$$

the rapidly decaying continuous functions on T where T is a locally compact Hausdorff space. Moreover, the C^*-algebra is unital if and only if the resulting T is compact (and in this case $\mathcal{C}_0(T) = \mathcal{C}(T)$).

Theorem 13.4. *(Continuous function calculus of self-adjoint operators) If $\mathcal{H}$ is a Hilbert space and $A = A^* \in \mathcal{B}(\mathcal{H})$ is a self-adjoint operator then there exists a unique*

$$\Phi_A : \mathcal{C}(\sigma(A)) \to \mathcal{B}(\mathcal{H}) \qquad f \mapsto f(A)$$

such that

1. *it is a $*$-algebra homomorphism, i.e. it is a unit-preserving algebra homomorphism with $\Phi_A(\overline{f}) = \Phi_A(f)^*$*
2. *it is norm-preserving, i.e. $||\Phi_A(f)||_{\mathcal{B}(\mathcal{H})} = ||f||_\infty$*
3. *$\Phi_A(\mathrm{id}_{\sigma(A)}) = A$*
4. (spectral mapping property) $\sigma(f(A)) = f(\sigma(A))$

Definition 13.5. Let us denote the set of all Borel sets of $\mathbb{R}$ by $\mathcal{B}orel(\mathbb{R})$. Then a function $E : \mathcal{B}orel(\mathbb{R}) \to \mathcal{B}(\mathcal{H})$ is called a *positive operator valued measure* (in short POVM) if

1. For all $B \in \mathcal{B}orel(\mathbb{R})$, $E(B)$ and $I - E(B)$ are positive operators on $\mathcal{H}$ and $E(\mathbb{R}) = Id_{\mathcal{H}}$
2. For an arbitrary set $\{B_i \mid i \in \mathbb{N}\}$ of Borel sets $B_i \in \mathcal{B}orel(\mathbb{R})$ such that $B_i \cap B_j = \emptyset$ for $i \neq j$ the following is true:
$$E\Big(\bigcup_{i \in \mathbb{N}} B_i\Big) = \sum_{i \in \mathbb{N}} E(B_i)$$
where the infinite sum is taken in the strong operator topology.

Definition 13.6. We denote the set of orthogonal projections of a Hilbert space by $\mathcal{P}(\mathcal{H})$ as before. A function $E : \mathcal{B}orel([a, b]) \to \mathcal{P}(\mathcal{H})$ is a *projection valued measure* (in short PVM), if

1. $E(\emptyset) = 0$, $E([a, b]) = I$

2. For an arbitrary set $\{B_i \mid i \in \mathbb{N}\}$ of Borel sets $B_i \in \mathcal{B}orel([a,b])$ such that $B_i \cap B_j = \emptyset$ for $i \neq j$ the following is true:

$$E\Big(\bigcup_{i\in\mathbb{N}} B_i\Big) = \sum_{i\in\mathbb{N}} E(B_i)$$

where the infinite sum is – again – taken in the strong operator topology.

Lemma 13.7. *Let P and Q be two orthogonal projections. Then $P+Q$ is an orthogonal projection if and only if $PQ = QP = 0$. (The proof is homework.)*

Corollary 13.8. *If $B_1 \cap B_2 = \emptyset$ then $E(B_1) \cap E(B_2)$*

Example 13.9.

1. Let $\mathcal{H} = \ell^2$ and $P_k = \mid \delta_k\rangle\langle\delta_k \mid$ and $\lambda_k \in [a,b]$. Then for $B \in \mathcal{B}orel[a,b]$ let $E(B) = \sum_{\lambda_k \in B} P_k$. This is a projection valued measure.
2. Let $\mathcal{H} = L^2[0,1]$. Then for $B \in \mathcal{B}orel([0,1])$ define $E(B) = M_{\chi(B)}$ the multiplication operator by $\chi(B)$ on $L^2[0,1]$. This is also a projection valued measure.

13.2 Integration with respect to a projection valued measure

Let $E : \mathcal{B}orel(\mathbb{R}) \to \mathcal{P}(\mathcal{H})$ be a projection valued measure and $x \in \mathcal{H}$. Then

$$\mu_x : \mathcal{B}orel(\mathbb{R}) \to [0,\infty) \qquad B \mapsto \mu_x(B) = \langle x, E(B)x\rangle$$

is a real valued measure. Moreover, the system of these measures $\{\mu_x \mid x \in \mathbb{R}\}$ fully describes $E(B)$.

Proposition 13.10. *(Without proof) Let f be a bounded, measurable function and E be a projective valued measure. Then there exists a unique operator $A \in \mathcal{B}(\mathcal{H})$ denoted by*

$$\int_{\mathbb{R}} f \,\mathrm{d}E \qquad \textit{or} \qquad \int_{\lambda\in\mathbb{R}} f(\lambda)\,\mathrm{d}E(\lambda)$$

called the integral of f with respect to E such that for all $x \in \mathcal{H}$:

$$\langle x, Ax\rangle = \int_{\mathbb{R}} f \,\mathrm{d}\mu_x$$

Example 13.11. Let $\mathcal{H} = L^2[0,1]$ and $E(B) = M_{\chi(B)}$. For a bounded measurable function $f : [0,1] \to \mathbb{C}$ we get

$$\int_{[0,1]} f \,\mathrm{d}E = M_f$$

Indeed, on the left hand side of the defining equation there is

$$\langle x, M_f x\rangle = \int_{[0,1]} \overline{x} \cdot f \cdot x \,\mathrm{d}\lambda$$

while on the right hand side the measure is

$$\mu_x(B) = \langle x, E(B)x\rangle = \int_{[0,1]} \overline{x} \cdot \chi_B \cdot x \,\mathrm{d}\lambda = \int_B |\overline{x}|^2 \,\mathrm{d}\lambda$$

therefore

$$\int_{[0,1]} f \,\mathrm{d}\mu_x = \int_{[0,1]} f \cdot |\overline{x}|^2 \,\mathrm{d}\lambda$$

which is the same as $\langle x, M_f x\rangle$.

Theorem 13.12. *(Spectral theorem for self-adjoint operators) Let $H = H^* \in \mathcal{B}(\mathcal{H})$ be an arbitrary self-adjoint operator. Then there exists a unique projective valued measure $E_H : \mathcal{B}orel(\sigma(H)) \to \mathcal{P}(\mathcal{H})$ called the spectral measure of H such that*

$$f(H) = \int_{\sigma(H)} f \, \mathrm{d}E_H$$

for all $f \in \mathcal{C}(\sigma(H))$.

In particular, $\int 1 \, \mathrm{d}E_H = \mathrm{Id}_{\mathcal{H}}$ and $\int \mathrm{id}_{\sigma(H)} \, \mathrm{d}E_H = H$ which is the infinite dimensional generalization of the projector decomposition of $H = \sum \lambda_k P_k$.

Remark 13.13. Interestingly, the following is true:

$$\int f \, \mathrm{d}E \cdot \int g \, \mathrm{d}E = \int fg \, \mathrm{d}E$$

since $f(H) \cdot g(H) = fg(H)$.

14 Distributions

In the topic of analysis there are three important, generally appearing spaces of functions: First,$\mathcal{D}$ the space of test functions, i.e. the space of compactly supported, infinitely many times differentiable functions. Second, the Schwartz space $\mathcal{S}$ consisting of those functions that are arbitrarily small in a neighborhood of infinity and their differentials, their products and their polynomial multiple still have the above property. And third, the L^2-space.

There is a natural containment among these three: $\mathcal{D} \subseteq \mathcal{S} \subseteq L^2$. Therefore, their (continuous) duals are in reversed order: $(L^2)' \subseteq \mathcal{S}' \subseteq \mathcal{D}'$. The topic of distributions is about these dual spaces. Note that we have not specified the the underlying topological space on which the functions are defined. So, precisely,

Definition 14.1. Let U be an arbitrary open subset of $\mathbb{R}^n$. The *space of test functions* is

$$\mathcal{D}(U) = \mathcal{C}_0^\infty(U, \mathbb{R})$$

where ∞ stands for the arbitrarily many differentiability and 0 means compactly supported. This space admits a (very strong) topology: Let $\varphi_k, \varphi \in \mathcal{D}(U)$. We say that $\varphi_k \to \varphi$ in $\mathcal{D}(U)$, if

1. there exists a compact subset $K \subseteq U$ such that $\cup_k \mathrm{supp}(\varphi_k) \subseteq K$
2. for all multiindeces $\alpha \in \mathbb{N}^n : \partial^\alpha \varphi_k \to \partial^\alpha \varphi$ uniformly on U.

Remark 14.2. The above convergence notion uniquely defined the topology on $\mathcal{D}(U)$: if we use the notation:

$$B_\varepsilon(\psi; \alpha, K) = \left\{ \varphi \in \mathcal{D}(U) \mid \left|\left| \partial^\alpha \varphi - \partial^\alpha \psi \right|\right|_\infty < \varepsilon, \ \mathrm{supp}(\varphi) \subseteq K \right\}$$

then we can give a topology on $\mathcal{D}_K(U) := \{ f \in \mathcal{D}(U) \mid \mathrm{supp}(\varphi) \subseteq K \}$ for any compact set K by

$$\mathcal{T} = \bigvee_{\alpha \in \mathbb{N}^n} \tau \left\{ B_\varepsilon(\psi; \alpha, K) \;\middle|\; \psi \in \mathcal{D}(U), \ \mathrm{supp}(\psi) \subseteq K, \ \varepsilon > 0 \right\}$$

where τ means generated topology by a sub-basis. (The reason behind using $\vee$ and not $\wedge$ is the philosophy that open sets represent obstructions for a sequence to be convergent hence if we require convergence for all α's then we have to include all the open balls for each α in the topology.) To extend this to the whole $\mathcal{D}(U)$ we have to use the notion of inductive limit: the topology on $\mathcal{D}(U)$ is the finest so that the natural inclusions $\mathcal{D}_K(U) \to \mathcal{D}(U)$ are all continuous. For details, see references in the topic.

Theorem 14.3. *$\mathcal{D}(U)$ is a complete, locally convex topological vector space with the Heine-Borel property (i.e. compactness is equivalent to bounded and closed).*

Definition 14.4. The *space of distributions* on U is the continuous dual space $\mathcal{D}'(U)$, i.e. $f \in \mathcal{D}'(U)$ if

1. $f : \mathcal{D}(U) \to \mathbb{R}$
2. f is linear
3. for all $\varphi_k \overset{\mathcal{D}}{\to} \varphi$ we have $\langle f, \varphi_k\rangle \to \langle f, \varphi\rangle$ as $k \to \infty$

The topology on $\mathcal{D}'(U)$ is the $\sigma(\mathcal{D}', \mathcal{D})$-topology (i.e. the weak-$*$ topology).

Goal: We will see in the next lecture that $\mathcal{D} \hookrightarrow \mathcal{D}'$ with the embedding $f \mapsto \big(\varphi \mapsto \int f\varphi \,\mathrm{d}\lambda\big)$

Tenth Lecture, 4th of December

Example 14.5. elements of $\mathcal{D}'(U)$:

1. If $f \in L^1_{loc}(U)$ then f defines a so called *regular distribution* T_f by the formula

   $$\varphi \mapsto \langle T_f, \varphi\rangle = \int_{\mathbb{R}^n} f(x)\varphi(x)\,\mathrm{d}x$$

   From now on, we will identify T_f with f. One can easily check that T_f is indeed continuous: assume that $\varphi_k \to \varphi$ in $\mathcal{D}(U)$. Then

   $$|\langle T_f, \varphi_k\rangle - \langle T_f, \varphi\rangle| = \Big|\int_{\mathbb{R}^n} f(x)\big(\varphi_k(x) - \varphi(x)\big)\,\mathrm{d}x\Big| \le \sup_U\{|\varphi_k - \varphi|\} \cdot \int_K |f(x)|\,\mathrm{d}x \to 0$$

   where K is a compact set that contains $\operatorname{supp}\varphi_k$ for all $k \in \mathbb{N}$.

2. Radon measures are also elements of $\mathcal{D}'(U)$:

   Definition 14.6. A *Radon measure* μ is a locally finite, inner regular measure on the Borel σ-algebra of a Hausdorff space X. Here, inner regularity means that for all Borel sets B of X

   $$\mu(B) = \sup\{\mu(K) \mid K \subseteq B,\ K \text{ is compact}\}$$

   (Similarly, a measure is outer regular if its value on a set can be approximated by the measures of containing open sets.) Besides, μ is locally finite if for all $x \in X$ has a neighborhood U_x such that $\mu(U_x) < \infty$. As a consequence, μ is then finite on compact sets.

   In our case, $X = \mathbb{R}^n$ or an open subset $U \subseteq \mathbb{R}^n$ which is a locally compact Hausdorff space. So a Radon measure $\mu : \mathcal{B}orel(U) \to \mathbb{R}$ then

   $$\langle \mu, \varphi\rangle = \int_U \varphi(x)\,\mathrm{d}\mu(x)$$

   defines an element of $\mathcal{D}'$. Continuity can be proved by similar estimation as in the firs example.

3. Define $\delta \in \mathcal{D}'$ as $\langle \delta, \varphi\rangle = \varphi(0)$. This is a (very important) special case of the previous example.

14.1 Operations on distributions

- Addition of distributions and multiplying by a constant are obviously (continuous) operations on $\mathcal{D}'(U)$.
- Differentiation:

  Let $f \in \mathcal{D}(\mathbb{R}) = \mathcal{C}_c^\infty(\mathbb{R}) \subseteq L^1_{loc}(\mathbb{R})$. Then the differentiation on the corresponding regular distribution is defined as $T_f' := T_{f'}$. Note that, by this definition we get

  $$\langle (T_f)', \varphi\rangle = \langle T_{f'}, \varphi\rangle = \int f'(x)\varphi(x)\,\mathrm{d}x = [f, \varphi]_{-\infty}^{+\infty} - \int_{\mathbb{R}} f(x)\varphi'(x)\,\mathrm{d}x = -\langle T_f, \varphi'\rangle$$

  so it is reasonable to define the differential of an arbitrary distribution as

Definition 14.7. If $f \in \mathcal{D}'$ then let $\langle f', \varphi \rangle := -\langle f, \varphi' \rangle$. This is indeed a continuous functional since if $\varphi_k \to \varphi$ in $\mathcal{D}$ then $\varphi_k' \to \varphi'$ in $\mathcal{D}$ by definition. Therefore, by the continuity of f we have

$$\langle f', \varphi_k \rangle = -\langle f, \varphi_k' \rangle \to \langle f, \varphi' \rangle = \langle f', \varphi \rangle$$

More generally, if $U \subseteq \mathbb{R}^n$ is an open subset then $f \in \mathcal{D}'(U)$, $\alpha \in \mathbb{N}^n$ then

$$\langle \partial^\alpha f, \varphi \rangle = (-1)^{|\alpha|} \cdot \langle f, \partial^\alpha \varphi \rangle$$

where $|\alpha| = |(\alpha_1, \alpha_2, \dots, \alpha_n)| = \alpha_1 + \alpha_2 + \cdots + \alpha_n$.

Example 14.8. The derivative of the Dirac delta is:

$$\langle \delta', \varphi \rangle = -\langle \delta, \varphi' \rangle = -\varphi'(0)$$

Example 14.9. $\ln|x| \in L^1_{loc}(\mathbb{R})$ since

$$\int_0^1 \ln x \, \mathrm{d}x = \lim_{\varepsilon \to 0+} \int_\varepsilon^1 \ln x \, \mathrm{d}x = \lim_{\varepsilon \to 0+} [x \ln x - x]_\varepsilon^1 = \lim_{\varepsilon \to 0+} (-1 - (\varepsilon \ln \varepsilon - \varepsilon)) = -1$$

but its usual derivative $\frac{1}{x}$ is not a locally integrable function. So the question arise: what is its distributional derivative? Let's compute:

$$\langle (\ln|x|)', \varphi \rangle = -\langle \ln|x|, \varphi' \rangle = -\int_{\mathbb{R}} \ln|x| \cdot \varphi'(x) \, \mathrm{d}x =$$

$$= \lim_{\varepsilon \to 0+} \left(\int_{-\infty}^{-\varepsilon} \ln|x| \cdot \varphi'(x) \, \mathrm{d}x + \int_\varepsilon^\infty \ln|x| \cdot \varphi'(x) \, \mathrm{d}x \right) =$$

$$= -\lim_{\varepsilon \to 0+} [\ln|x| \varphi(x)]_{-\infty}^{-\varepsilon} - \int_{-\infty}^{-\varepsilon} \frac{1}{x} \varphi(x) \, \mathrm{d}x + [\ln|x| \varphi(x)]_\varepsilon^\infty - \int_\varepsilon^\infty \frac{1}{x} \varphi(x) \, \mathrm{d}x =$$

$$= \lim_{\varepsilon \to 0+} (\ln \varepsilon (\varphi(\varepsilon) - \varphi(-\varepsilon))) + \lim_{\varepsilon \to 0+} \int_{|x| > \varepsilon} \frac{1}{x} \varphi(x) \, \mathrm{d}x = \left\langle P.V. \frac{1}{x}, \varphi(x) \right\rangle$$

Definition 14.10. More generally, if $A : \mathcal{D} \to \mathcal{D}$ and $A' : \mathcal{D} \to \mathcal{D}$ are continuous linear mappings such that for all $\varphi, \psi \in \mathcal{D}$ we have

$$\langle A\psi, \varphi \rangle = \langle \psi, A'\varphi \rangle$$

then for all $f \in \mathcal{D}'$ we can define $Af \in \mathcal{D}'$ as

$$\langle Af, \varphi \rangle = \langle f, A'\varphi \rangle$$

- Multiplication by a smooth function $g \in \mathcal{C}^\infty(U)$: let $f \in \mathcal{D}'$ and $\varphi, \psi \in \mathcal{D}$. Then $g\psi \in \mathcal{D}$ satisfies the assumptions of the previous definition since

$$\langle g\psi, \varphi \rangle = \int g\psi\varphi = \langle \psi, g\varphi \rangle$$

so with $A = A' = g\cdot$ we got that we can define the product $g \cdot f$ as

$$\langle gf, \varphi \rangle = \langle f, g\varphi \rangle$$

for all $f \in \mathcal{D}'$, $\varphi \in \mathcal{D}$ and $g \in \mathcal{C}^\infty(U)$

- Translation is also possible among distributions: let $y \in \mathbb{R}^n$ and define $\tau_y : \mathcal{D} \to \mathcal{D}$ as

$$\mathcal{D} \ni \big(x \mapsto \varphi(x)\big) \mapsto \big(x \mapsto \varphi(x-y)\big) \in \mathcal{D}$$

  Then one can observe that

$$\langle \tau_y \psi, \varphi \rangle = \int \psi(x-y)\varphi(x)\,\mathrm{d}x = \int \psi(z)\varphi(z+y)\,\mathrm{d}z = \langle \psi, \tau_{-y}\varphi \rangle$$

  therefore again $A = \tau_y$ and $A' = \tau_{-y}$ satisfies the above definition therefore for a general distribution $f \in \mathcal{D}'$ one can define

$$\langle \tau_y f, \varphi \rangle = \langle f, \tau_{-y}\varphi \rangle$$

  for all $\varphi \in \mathcal{D}$.

- Composition by a linear map is also possible: let $S : \mathbb{R}^n \to \mathbb{R}^n$ be a linear, invertible map. Then we define

$$\langle f \circ S, \varphi \rangle = \frac{1}{|\det(S)|} \langle f, \varphi \circ S^{-1} \rangle$$

  for all $f \in \mathcal{D}'$, $\varphi \in \mathcal{D}$.

14.2 Schwartz space

Definition 14.11. The elements of the dual space $\mathcal{S}'$ of $\mathcal{S}$ are called *tempered distributions*. The main property of these two spaces is being closed under Fourier-transform (while in the case of $\mathcal{D}$ and $\mathcal{D}'$ the analogous main property was that they are closed under differentiation.) Here, $\mathcal{S}$ stands for the rapidly decreasing functions, i.e.

$$\mathcal{S} := \{ f \in \mathcal{C}^\infty(\mathbb{R}^n) \mid \exists K \text{ compact } f|_{U\setminus K} < \varepsilon,\ \forall \alpha, \beta \in \mathbb{N}^n :\ ||x^\alpha \partial^\beta f||_\infty =: \mathfrak{p}_{\alpha,\beta} < \infty \}$$

with the topology: $\varphi_k \to \varphi$ in $\mathcal{S}$ if for all $\alpha, \beta \in \mathbb{N}^n$: $\mathfrak{p}_{\alpha,\beta}(\varphi_k - \varphi) \to 0$ as $k \to \infty$.

There is a unique topology defined by the above convergence. With this topology $\mathcal{S}$ is a locally compact, metrizable complete topological vector space. While, the topology on the continuous dual $\mathcal{S}'$ is the $\sigma(\mathcal{S}', \mathcal{S})$ weak topology. In the following, n is treated as a fixed number.

Definition 14.12. (*Fourier transform of distributions*) If $\varphi \in \mathcal{S}$ then

$$\mathcal{F}[\varphi](\xi) = \int_{\mathbb{R}^n} \varphi(x) e^{i\langle \xi, x\rangle}\,\mathrm{d}x$$

where $\langle .,. \rangle$ stands for the standard Euclidean scalar product of $\mathbb{R}^n$.

Proposition 14.13. *The Fourier-transform $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ on the Schwartz space is a continuous linear bijection.*

Remark 14.14. Note that $\mathcal{F}(\mathcal{D})$ is not contained in $\mathcal{D}$ however.

Proof. Linearity is obvious so to prove bijectivity, first note that

$$\partial^\alpha \mathcal{F}[\varphi](\xi) = \partial_\xi^\alpha \int \varphi(x) e^{i\langle \xi, x\rangle}\,\mathrm{d}x = \int \varphi(x)(ix)^\alpha e^{i\langle \xi, x\rangle}\,\mathrm{d}x = \mathcal{F}[x \mapsto (ix)^\alpha \varphi(x)](\xi)$$

and similarly for the derivation in the argument:

$$\mathcal{F}[\partial_x^\alpha \varphi](\xi) = \int \partial^\alpha \varphi(x) e^{i\langle \xi, x\rangle}\,\mathrm{d}x = (-1)^\alpha \int \varphi(x) \cdot \partial_x^\alpha e^{i\langle \xi, x\rangle}\,\mathrm{d}x = (-1)^\alpha \int \varphi(x)(i\xi)^\alpha e^{i\langle \xi, x\rangle} = (-i\xi)^\alpha \mathcal{F}[\varphi](\xi)$$

where we used the multiindex notation. Therefore,

$$\xi^\alpha \mathcal{F}[\varphi](\xi) = \mathcal{F}[(i\partial)^\alpha \varphi](\xi)$$

So with these identities, we can prove that $\mathcal{F}(\varphi) \in \mathcal{S}$ for $\varphi \in \mathcal{S}$:

$$\xi^\beta \partial_\xi^\alpha \mathcal{F}[\varphi](\xi) = \xi^\beta \mathcal{F}[(ix)^\alpha \varphi(x)](\xi) = \mathcal{F}[(i\partial)^\beta (ix)^\alpha \varphi(x)](\xi)$$

so we just have to estimate its absolute value:

$$|\xi^\beta \partial_\xi^\alpha \mathcal{F}[\varphi](\xi)| = \Big| \int (i\partial)^\beta (ix)^\alpha \varphi(x) \cdot e^{i\langle \xi, x\rangle} \, \mathrm{d}x \Big| \leq \int \left|\partial^\beta (x^\alpha \varphi(x))\right| \mathrm{d}x$$

where the right hand side is already finite since every element of $f \in \mathcal{S}$ (in our case $\partial^\beta(x^\alpha \varphi(x))$) has finite integral. Indeed, let $f \in \mathcal{S}$ be arbitrary. Then

$$\int |f(x)| \, \mathrm{d}x = \int |f(x)|(1+|x|^{n+1}) \frac{1}{1+|x|^{n+1}} \, \mathrm{d}x \leq ||f(x)(1+|x|^{n+1})||_\infty \cdot \int \frac{1}{1+|x|^{n+1}} \, \mathrm{d}x < \infty$$

where n is the dimension so $\frac{1}{1+|x|^{n+1}}$ is indeed finitely integrable. This proves that $\mathcal{F}$ is into the $\mathcal{S}$.

Now, we show that it is onto. For this purpose, let us define the inverse Fourier transform:

$$\mathcal{F}^{-1}[\psi](x) = \frac{1}{(2\pi)^n} \int \psi(\xi) e^{-i\langle \xi, x\rangle} \, \mathrm{d}x$$

Lemma 14.15. $\varphi(x) = \mathcal{F}^{-1}[\mathcal{F}[\varphi]](x) = \mathcal{F}[\mathcal{F}^{-1}[\varphi]](x)$ *(It is an exercise computation.)*

The bijectivity of $\mathcal{F}$ almost follows from the lemma since it would show both injectivity and surjectivity if we would know that $\mathcal{F}^{-1}[\psi] \in \mathcal{S}$. But that follows trivially from the previous computation since

$$\mathcal{F}^{-1}[\psi](x) = \frac{1}{(2\pi)^n} \mathcal{F}[\xi \mapsto \psi(-\xi)](x)$$

and multiplication by a constant and $\circ(-1)$ does not escape from $\mathcal{S}$. ☐

Lemma 14.16. *$\mathcal{F} : \mathcal{S} \to \mathcal{S}$ is continuous (and so $\mathcal{F}^{-1}$).*

Proof. Assume that $\varphi_k \to 0$ in $\mathcal{S}$. Let $\alpha, \beta \in \mathbb{N}^n$ be arbitrary. Then

$$|\xi^\beta \partial^\alpha \mathcal{F}[\varphi_k]| \leq \int |\partial^\beta x^\alpha \varphi_k| \, \mathrm{d}x = ||\partial^\beta x^\alpha \varphi(x) \cdot (1+|x|^{n+1})||_\infty \cdot \int \frac{1}{1+|x|^n} \, \mathrm{d}x \to 0$$

so $\mathcal{F}$ is indeed continuous. ☐

Definition 14.17. *Fourier Transform* of a distribution in $\mathcal{S}'$: Let $f \in L^1(\mathbb{R}^n)$. Then

$$|\mathcal{F}[f](\xi)| = \Big| \int f(x) e^{i\langle \xi, x\rangle} \, \mathrm{d}x \Big| \leq \int |f(x)| < \infty$$

so $\mathcal{F}[f]$ is a bounded function (in fact, it is also continuous). However, a bounded function is always in $\mathcal{S}'$ as a regular distribution. In formula, $T_{\mathcal{F}[f]} \in \mathcal{S}'$. Therefore, the Fourier transform of a regular distribution is

$$\langle \mathcal{F}[T_f], \varphi\rangle = \langle T_{\mathcal{F}[f]}, \varphi\rangle = \int \mathcal{F}[f](\xi)\varphi(x) \, \mathrm{d}\xi = \int\int f(x) e^{i\langle \xi, x\rangle} \varphi(\xi) \, \mathrm{d}x\mathrm{d}\xi = \int f(x) \mathcal{F}[\varphi](x) \, \mathrm{d}x = \langle T_f, \mathcal{F}[\varphi]\rangle$$

so it is reasonable to define the so called transpose of Fourier transform on general tempered distributions as

$$\langle \mathcal{F}[f], \varphi\rangle = \langle f, \mathcal{F}[\varphi]\rangle \qquad \forall f \in \mathcal{S}', \ \forall \varphi \in \mathcal{S}$$

Lemma 14.18. *With the above definition $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ becomes a continuous linear bijection with the continuous linear inverse $\mathcal{F}^{-1}$.*

Example 14.19. The Fourier transform of the Dirac delta can be computed as

$$\langle \mathcal{F}[\delta_{x_0}], \varphi \rangle = \langle \delta_{x_0}, \mathcal{F}[\varphi] \rangle = \mathcal{F}[\varphi](x_0) = \int \varphi(\xi) e^{i\langle \xi, x \rangle} \, \mathrm{d}\xi = \langle e^{i\langle \xi, x \rangle}, \varphi(\xi) \rangle$$

In particular, for $x_0 = 0$ one gets $\mathcal{F}[\delta] = 1$.

Eleventh Lecture, 9th of December

15 Sobolev spaces

Goal: To define a function space (hopefully a Banach space) which is similar to the L^p-spaces, but its elements are differentiable up to a given order k, where $1 \leq p \leq \infty$ and $k \in \mathbb{N}^+$. To make this possible (or to make it a complete space), instead of the usual differentiability we require weak differentiability. This notion is first explained in an example.

Example 15.1. Let $f(x) = |x|$ on the real line. In this case $f'(0)$ does not exist in the classical sense. But for all $g \in \mathcal{C}_c^1(\mathbb{R})$ we have

$$\int_{-\infty}^{\infty} f g' \, \mathrm{d}x = \int_{-\infty}^{0} f g' \, \mathrm{d}x + \int_{0}^{\infty} f g' \, \mathrm{d}x = \int_{-\infty}^{0} (-x) g' \, \mathrm{d}x + \int_{0}^{\infty} x g' \, \mathrm{d}x =$$

$$= [(-x)g]_{-\infty}^{0} - \int_{-\infty}^{0} (-1) g \, \mathrm{d}x + [xg]_0^{\infty} - \int_0^{\infty} 1 g \, \mathrm{d}x = \int_{-\infty}^{\infty} \operatorname{sgn}(x) \cdot g \, \mathrm{d}x$$

therefore it is acceptable to define the weak derivative of $|x|$ as $\operatorname{sgn}(x)$ as we did in the case of regular distributions.

Definition 15.2. For $\alpha \in \mathbb{N}^n$ and $\Omega \subseteq \mathbb{R}^n$ an open subset, the *weak $\underline{\alpha}$-th partial differential $\partial^\alpha u$* of $u : \Omega \to \mathbb{K}$ is $v \in L^1_{loc}(\Omega)$ if

$$\int_\Omega v \cdot \partial^\alpha \varphi = (-1)^{|\alpha|} \int_\Omega v \cdot \varphi \qquad \forall \varphi \in \mathcal{C}_c^\infty(\Omega)$$

In other words, the α-th weak differential of u is its α-th distributional derivative of the corresponding regular distribution if the result is the regular distribution of a locally integrable function (and not a general distribution).

Definition 15.3. Let $k \in \mathbb{N}_+$, $1 \leq p \leq \infty$ and $\Omega \subseteq \mathbb{R}^n$ an open subset. Then we can define the *Sobolev space* as

$$W^{k,p}(\Omega) := \{ u \in L^p(\Omega) \mid \partial^\alpha u \in L^p(\Omega) : \ \forall |\alpha| \leq k \}$$

We can also define two convenient, equivalent norms on $W^{k,p}(\Omega)$:

$$||u||_{W^{k,p}(\Omega)} = \begin{cases} \left(\sum_{|\alpha| \leq k} ||\partial^\alpha u||^p_{L^p(\Omega)} \right)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ \max_{|\alpha| \leq k} ||\partial^\alpha u||_{L^\infty(\Omega)} & \text{if } p = \infty \end{cases} \qquad \text{and} \qquad ||u||'_{W^{k,p}(\Omega)} = \sum_{|\alpha| \leq k} ||\partial^\alpha u||_{L^p(\Omega)}$$

where in the second case the definition is given by the same formula for all $1 \leq p \leq \infty$.

Theorem 15.4. *(Without proof) These two norms are equivalent and $W^{k,p}(\Omega)$ is a Banach space with these norms.*

Remark 15.5. $W^{k,p}(\Omega) = H^k(\Omega)$ is a Hilbert space with the inner product:

$$\langle u, v\rangle_{H^k(\Omega)} = \sum_{|\alpha|\le k} \langle \partial^\alpha u, \partial^\alpha v\rangle_{L^2(\Omega)} = \sum_{|\alpha|\le k} \int_\Omega \overline{\partial^\alpha u} \cdot \partial^\alpha v$$

Example 15.6. In the case of $k = 1$ the scalar product is the following: let $f, g \in H^1(\mathbb{R})$. Then

$$\langle f, g\rangle_{H^1(\mathbb{R})} = \langle f, g\rangle + \langle f', g'\rangle = \int_{\mathbb{R}} \overline{f} \cdot g \,\mathrm{d}x + \int_{\mathbb{R}} \overline{f'} \cdot g' \,\mathrm{d}x$$

15.1 Green's function method for solving linear differential equations

Example 15.7. (Physical problem) Given a charge distribution $\rho(\underline{r})$, find its electric potential $\Phi(\underline{r})$. Assume that $\mathrm{supp}(\rho)$ is bounded in $\mathbb{R}^3$. In particular, $\lim_{|\underline{r}|\to\infty} \Phi(\underline{r}) = 0$. So we are left with the differential equation:

$$\Delta\Phi = -\frac{\rho}{\varepsilon_0}$$

$$\lim_{|\underline{r}|\to\infty} \Phi(\underline{r}) = 0$$

where ε_0 is a constant and we are searching for Φ. The heuristic argument for a solution is the following: note that an electrostatic field $\underline{E}$ is conservative (what means its 'rot' is zero), i.e. there exists a Φ such that $\underline{E}(\underline{r}) = -\mathrm{grad}(\Phi)$. So we get

$$\Delta\Phi = \mathrm{div}(\mathrm{grad}(\Phi)) = -\mathrm{div}(\underline{E})$$

so we get the Second Maxwell Equation on ρ:

$$\mathrm{div}(\underline{E}) = \frac{\rho}{\varepsilon_0}$$

The generalizable idea in the above solution is that defining a potential can help.

The potential of a point charge Q at $\underline{r}_0$ is

$$\Phi_{\underline{r}_0}(\underline{r}) = \frac{Q}{4\pi\varepsilon_0} \cdot \frac{1}{|\underline{r} - \underline{r}_0|}$$

Mathematically: we got $\Delta\Phi_{\underline{r}_0} = -\frac{Q\cdot\delta_{\underline{r}_0}}{\varepsilon_0}$. By this the general solution of the original equation is

$$\Phi_\rho(\underline{r}) = \int_{\underline{r}'\in\mathbb{R}^3} \frac{\rho(\underline{r}')}{4\pi\varepsilon_0} \frac{1}{|\underline{r} - \underline{r}'|} \,\mathrm{d}^3\underline{r}'$$

where the motivation of it is that the equation is linear so the solution is the superposition of the solutions for infinitely many Dirac deltas.

Strategy for the solution of the linear differential equation $Lu = f$ with some boundary or initial conditions where L is a linear differential operator and f is a given function:

1. Solve the equation $LG_{x_0} = \delta_{x_0}$ for G_{x_0} with the same boundary or initial conditions for all $x_0 \in \Omega$. This G_{x_0} is called Green's function of L.
2. The general solution can be expressed as

$$u = G_{x_0} * f$$

15.2 Convolution

Definition 15.8. The *convolution* of $f, g : \mathbb{R} \to \mathbb{R}$ is

$$(f * g)(x) = \int_{\mathbb{R}} f(y)g(x-y)\,\mathrm{d}y$$

if it makes sense. For example, if one of the factors is compactly supported and the other is locally integrable then $f * g$ is defined (but it is not the only case). It "smooths out" the irregularities of the functions. It can also be defined for distributions.

Properties: It is a commutative, associative operation that is also linear. Moreover, it behaves well under differentiation, namely:

$$\partial^\alpha(f * g) = (\partial^\alpha f) * g = f * (\partial^\alpha g)$$

Besides convolution has the very useful property that $\delta * f = f * \delta = f$

Let $x \mapsto G_0(x)$ be the Green's function at $x_0 = 0$ of the linear differential operator L with certain boundary or initial conditions. Then the solution of $Lu = f$ with the above extra conditions can be expressed as$u = G_0 * f$ since

$$L(G_0 * f) = LG_0 * f = \delta_0 * f = f$$

Example 15.9. Consider the following problem:

$$y''(x) = f(x) \qquad y(x) \equiv 0 \quad \text{if } x \leq 0$$

The corresponding Green's equation is $G'' = \delta$ with the boundary condition $G(x) = 0$ for all $x \leq 0$. The solution of it is $x \mapsto \Theta(x)x$ where $\Theta(x)$ is the Heaviside function (i.e. the indicator of the positive numbers)

15.3 Finding the Green's function by Fourier-transform

Apply the Fourier transform on $LG = \delta$. Since $L = \sum_{n=0}^{k} a_n \frac{\mathrm{d}^n}{\mathrm{d}x^n}$ the Fourier transform of the left hand side is

$$\mathcal{F}[LG] = \sum_{n=0}^{k} a_n(-i\xi)^n \mathcal{F}[G](\xi)$$

and the Fourier transform $\mathcal{F}[\delta]$ is explicitly computable. So now one can solve this algebraic equation in the coefficients and the solution of the original equation is the inverse-Fourier transform of the solution of the transformed equation.

16 Summary on Measure Theory

Definition 16.1. The triple $(X, \mathcal{A}, \mu)$ is a *measure-space*, if $\mathcal{A} \subseteq \mathcal{P}(X)$ is a σ-algebra on X and $\mu : \mathcal{A} \to [0, \infty]$ is a σ-additive function. A measure-space is called finite if $\mu(X) < \infty$. Similarly, it is called σ-finite, if there exist A_i's such that $X = \bigcup_{i\in\mathbb{N}} A_i$ and $\mu(A_i) < \infty$ for all $i \in \mathbb{N}$.

Properties: Let $(X, \mathcal{A}, \mu)$ be a measure space. Then

1. μ is monotone
2. μ is subadditive i.e. $\mu(\bigcup_{n\in\mathbb{N}} A_i) \leq \sum_{i\in\mathbb{N}} \mu(A_i)$
3. μ is continuous: for measurable sets $A_1 \subseteq A_2 \subseteq \dots$ we have $\lim_{n\to\infty} \mu(A_n) = \mu(\cup_{n\in\mathbb{N}} A_n)$
4. for measurable sets $A_1 \supseteq A_2 \supseteq \dots$ we have $\lim_{n\to\infty} \mu(A_n) = \mu(\cap_{n\in\mathbb{N}} A_n)$ provided that (!) there exists a $k \in \mathbb{N}$ such that $\mu(A_k) < \infty$

Definition 16.2. A measure space $(X, \mathcal{A}, \mu)$ is *complete* if for all $A \in \mathcal{A}$ with $\mu(A) = 0$ all of its subsets $B \subseteq A$ are also measurable, i.e. $B \in \mathcal{A}$.

Remark 16.3. Every measure have a unique completion so we can always assume that the measure-space is complete.

Definition 16.4. A function $\mu : \mathcal{P}(X) \to [0, \infty]$ is an outer (or exterior) measure if X is any set and μ is σ-subadditive, i.e. for countably many $A_i \subseteq X$ $(i \in \mathbb{N})$ we have

$$A \subseteq \bigcup_{i \in \mathbb{N}} A_i \qquad \text{implies} \qquad \mu(A) \leq \sum_{i \in \mathbb{N}} \mu(A_i)$$

It follows that $\mu(\emptyset) = 0$ and μ is monotone in this case.

Definition 16.5. Let $\overline{\mu}$ be an outer measure. A set $A \subseteq X$ is $\overline{\mu}$-measurable if for all $B \subseteq X$:

$$\overline{\mu}(B) = \overline{\mu}(B \backslash A) + \overline{\mu}(B \cap A)$$

Notation: $\mathcal{A}_{\overline{\mu}} := \{\overline{\mu}\text{-measurable sets}\}$ sometimes called the "well-cutting" sets.

Theorem 16.6. *The set of $\overline{\mu}$-measurable subsets form a σ-algebra $\mathcal{A}_{\overline{\mu}} \subseteq \mathcal{P}(X)$ and $\overline{\mu}|_{\mathcal{A}_{\overline{\mu}}}$ is a complete measure.*

Theorem 16.7. *Let $\mathcal{H} \subseteq \mathcal{P}(X)$ and $\nu : \mathcal{H} \to [0, \infty]$ an arbitrary function. Then for any $A \subseteq X$:*

$$\overline{\mu}_\nu(A) := \inf \Big\{ \sum_{i \in I} \nu(A_i) \mid I \textit{ is countable}, A_i \in \mathcal{H}, A \subseteq \cup_{i \in I} A_i \Big\}$$

is an outer measure generated by ν.

Corollary 16.8. *Putting together the two theorem we get that for any $\nu : \mathcal{H} \to [0, \infty]$ function we can construct an associated outer measure that is in fact a complete measure on an appropriate σ-algebra.*

Remark 16.9. Given a measure, one can associate an outer measure to it in the above way. The restriction of this to the measurable sets is the completion of μ.

16.1 Lebesgue measure on $\mathbb{R}$

The above construction can be applied in the case of $X = \mathbb{R}$ and $\mathcal{H}$ being the set of intervals with the candidate measure:

$$\nu([a, b]) = \nu([a, b]) = \nu([a, b]) = \nu([a, b]) = b - a$$

if $b > a$. The resulting measure is denoted by λ and called the Lebesgue measure.

Lemma 16.10. *$\lambda(N) = 0$ if and only if for all $\varepsilon > 0$ there exist intervals I_n such that*

$$N \subseteq \bigcup_{n \in \mathbb{N}} I_n \qquad \textit{and} \qquad \sum_{n \in \mathbb{N}} \nu(I_n) < \varepsilon$$

Remark 16.11. We assume that the notion of measurable functions is known to the reader.

Lemma 16.12. *If $f, g : X \to \overline{\mathbb{R}}$ are measurable then $c \cdot f$, $|f|$, $\frac{1}{f}$, $\max(f, g)$, $f + g$, $f \cdot g$ are also measurable.*

Lemma 16.13. *If $f_i : X \to \overline{\mathbb{R}}$ are measurable then $\sup(f_n)$, $\inf(f_n)$, $\limsup(f_n)$, $\liminf(f_n)$ are also measurable.*

Definition 16.14. Let $(X, \mathcal{A}, \mu)$ be a measure space and $f : X \to [0, \infty]$ a measurable function. Then

$$\int_X f \,\mathrm{d}\mu := \sup \Big\{ \sum_{i=1}^{n} y_i \mu(A_i) \;\Big|\; n \in \mathbb{N}, A_i \in \mathcal{A}, A_i \cap A_j = \emptyset, x \in A_i \Rightarrow f(x) \geq y_i \Big\}$$

Properties: It is monotone and modifying the function on a zero set does not modify the value of the integral.

Theorem 16.15. (Fatou's lemma) *Let $(X, \mathcal{A}, \mu)$ be a measure-space, and $f_n : X \to [0, \infty]$ be measurable functions for all $n \in \mathbb{N}$. Then*

$$\liminf_{n\to\infty} \int_X f_n \,\mathrm{d}\mu \geq \int_X \liminf_{n\to\infty} f_n \,\mathrm{d}\mu$$

Theorem 16.16. *(Beppo-Levi or Lebesgue monotone convergence theorem) Let $(X, \mathcal{A}, \mu)$ be a measure-space, and $f_n : X \to [0, \infty]$ be measurable functions for all $n \in \mathbb{N}$. If $f_1 \leq f_2 \dots$ then*

$$\lim \int_X f_n \,\mathrm{d}\mu = \int_X \lim f_n \,\mathrm{d}\mu$$

Proof. $\leq$ is obvious by the monotonicity of the integral. $\geq$ is verified by Fatou's lemma. □

Definition 16.17. The integral of a function not necessarily non-negative function f is defined by its decomposition $f = f_+ - f$ to positive and negative parts. f is called integrable if and only if both f_+ and f_- has finite integral with respect to μ. Then the integral of f is defined as

$$\int f \,\mathrm{d}\mu = \int f_+ \,\mathrm{d}\mu - \int f_- \,\mathrm{d}\mu$$

Property: $|\int f \,\mathrm{d}\mu| \leq \int |f| \,\mathrm{d}\mu$

Theorem 16.18. *(Lebesgue dominated convergence theorem) Let $f_n : X \to \overline{\mathbb{R}}$ measurable and $g : X \to \overline{\mathbb{R}}$ is integrable such that $|f_n| \leq g$ almost everywhere. If $f_n \to f$ then*

$$\int |f_n - f| \,\mathrm{d}\mu \to 0 \qquad \textit{and} \qquad \int f_n \,\mathrm{d}\mu \to \int f \,\mathrm{d}\mu$$

Remark 16.19. The above defined integral is absolutely continuous.