Notes on

Functional Analysis and Partial Differential Equations

Herbert Koch
Universität Bonn
Winter term 2019

These are short incomplete notes. They do not substitute textbooks. The following textbooks are recommended.

- H. W. Alt, Linear functional analysis, Springer, 2016.
- E. H. Lieb, M. Loss: Analysis, AMS 2001.
- E. M. Stein, R. Shakarchi, Functional analysis. Princeton 2011.
- D. Werner, Funktionalanalysis, Springer, 2011.

Correction are welcome and should be sent to `koch@math.uni-bonn.de` or told me during office hours. The notes are only for participants of the course V3B1/F4B1 *PDG und Funktionalanalysis /PDE and Functional Analysis* at the University of Bonn, Winter term 2019/2020.

Examination: Saturday 15.02.2019 and Wednesday, 26.03.2019

Assistant: Dr. Franz Gmeineder

Contents

[9.10.2019]

1 Introduction

Functional analysis is the study of normed complete vector spaces (called Banach spaces) and linear operators between them. It is built on the structure of linear algebra and analysis. Functional analysis provides the natural frame work for vast areas of mathematics including probability, partial differential equations and numerical analysis. It expresses an important shift of viewpoint: Functions are now points in a function space.

Let $\Omega \subset \mathbb{R}^n$ be an open bounded set with smooth boundary. One of the deepest results in *Einführung in die PDG* was that the Green's function $g(x, y)$ provides a map

$$f \to u$$

given by

$$u(x) = \int g(x, y) f(y) dy =: Tf$$

so that

$$-\Delta u = f \qquad \text{in } \Omega$$

$$u = 0 \qquad \text{on } \partial\Omega$$

whenever f is sufficiently regular. It is not hard to see that

$$T : L^2(\Omega) \to L^2(\Omega)$$

and T is one of the most relevant operators in functional analysis.

Functional analysis provides the crucial language for many areas in mathematics.

The main abstract objects are topological vector spaces over $\mathbb{K}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We will focus on normed spaces, the most important class of topological vector spaces.

Definition 1.1. *Let X be a $\mathbb{K}$ vector space. A map $\|.\| : X \to [0, \infty)$ is called norm if*

$$\|x\| = 0 \qquad \Longleftrightarrow \qquad x = 0, \tag{1.1}$$

if for all $x, y \in X$

$$\|x + y\| \le \|x\| + \|y\| \qquad \textit{(triangle inequality)}\ , \tag{1.2}$$

and if for all $\lambda \in \mathbb{K}$ and $x \in X$

$$\|\lambda x\| = |\lambda| \|x\|. \tag{1.3}$$

It is called normed space and Banach space if it is complete as metric space.

Remark 1.2. *A norm defines a metric by* $d(x,y) = \|x-y\|$.

First examples.

1. $\mathbb{R}^n$ and $\mathbb{C}^n$ equipped with the Euclidean norm are real resp. complex Banach spaces.

2. Let X be a set. The space of bounded functions $\mathbb{B}(X)$ equipped with the supremum norm is a Banach space.

3. Let (X,d) be a metric spaces. The space of bounded continuous functions $C_b(X)$ equipped with the supremum norm is a Banach space, or more precisely a closed sub vector space of $\mathbb{B}(X)$.

4. Let $U \subset \mathbb{R}^d$ be open. $C_b^k(U)$ is the vector space of k times differentiable functions on U which are together with there derivatives bounded. The norm
$$\|u\|_{C^k(U)} = \max_{|\alpha|\le k} \|\partial^\alpha u\|_{\mathbb{B}(U)}$$
turns C_b^k into a Banach space. *Exercise*

5. Let $U \subset \mathbb{C}$ be open. The space of bounded holomorphic functions $H^\infty(U)$ is a Banach space when equipped with the supremum norm.

Lemma 1.3. *Suppose that X is a Banach space and $U \subset X$ is a vector space which is a closed subset of X. Then U is a Banach space.*

Definition 1.4. *Let X and Y be normed spaces. We define $L(X,Y)$ as the set of all continuous linear maps from X to Y.*

Theorem 1.5. *Let $T : X \to Y$ be in $L(X,Y)$. Then*
$$\|T\|_{X\to Y} := \sup_{\|x\|_X \le 1} \|T(x)\|_Y < \infty$$
and $\|.\|_{X\to Y}$ defines a norm on $L(X,Y)$. A linear operator $T : X \to Y$ is continuous if and only if its norm $\|T\|_{X\to Y}$ is finite. $L(X,Y)$ is a Banach space if Y is a Banach space.

Proof. Continuous linear maps from X to Y are a vector space with the obvious sum and multiplication. Let $T : X \to Y$ be a continous linear map. We choose $\varepsilon = 1$ and $x_0 = 0$. Then there exists $\delta > 0$ so that
$$\|Tx\|_Y \le 1 \qquad \text{if } \|x\|_X \le \delta,$$
and hence, if $x \in X$, $x \ne 0$, then $\left\|\frac{\delta x}{\|x\|_X}\right\|_X \le \delta$ and
$$\|Tx\|_Y = \frac{\|x\|_X}{\delta}\left\|T\frac{\delta x}{\|x\|_X}\right\|_Y \le \delta^{-1}\|x\|_X$$

and thus

$$\|T\|_{X\to Y} \le \delta^{-1}.$$

Vice versa: Let $T : X \to Y$ be linear so that $\|T\|_{X\to Y} < \infty$. For $\varepsilon > 0$ we choose $\delta = \varepsilon/\|T\|_{X\to Y}$. Then

$$\|Tx - Ty\|_Y = \|T(x-y)\|_Y \le \|T\|_{X\to Y}\|x-y\|_X \le \varepsilon$$

provided $\|x-y\|_X \le \delta$. In particular T is uniformly continuous.

Now assume that Y is a Banach space and let $T_n \in L(X,Y)$ be a Cauchy sequence. For all x, $T_n x$ is a Cauchy sequence in Y since

$$\|T_n x - T_m x\|_Y \le \|T_n - T_m\|_{X\to Y}\|x\|_X.$$

Let

$$Tx := \lim_{n\to\infty} T_n x.$$

The convergence is uniform on bounded sets, and hence the limit T is continuous and in $L(X,Y)$. Moreover

$$\begin{aligned}\|T - T_n\|_{X\to Y} = \sup_{\|x\|_X\le 1}\|(T-T_n)x\|_Y &= \sup_{\|x\|_X\le 1}\limsup_{m\to\infty}\|(T_m - T_n)x\|_Y \\ &\le \limsup_{m\to\infty}\|T_m - T_n\|_{X\to Y} \to 0\end{aligned}$$

as $n \to \infty$. Here we used continuity of addition and the map to the norm. □

Definition 1.6 (Dual space)**.** *Let X be a normed space. We define the dual (Banach) space as $X^* = L(X,\mathbb{K})$.*

Example: Let $X = \mathbb{R}^n$ with the Euclidean norm. The map

$$\mathbb{R}^n \ni y \to (x \to \sum_{j=1}^n x_j y_j) \in (\mathbb{R}^n)^*$$

is isometric and surjective. It allows to identify $\mathbb{R}^n$ and $(\mathbb{R}^n)^*$.

Some relations to other fields.

1. Calculus of variations. Let $\Omega \subset \mathbb{R}^d$ be open,

   $$E(u,U) = \int_U |\nabla u|^2 dx$$

   where $u \in C^1(\Omega)$ and $U \subset \Omega$ open. Suppose that for all $U \subset \Omega$ open and $\phi \in C^\infty(U)$ with compact support

   $$E(u+\phi, U) \ge E(u,U).$$

We expand

$$E(u + t\phi, U) = E(u, U) + 2t \int_U \nabla u \nabla \phi dx + t^2 E(\phi, U)$$

and deduce using the divergence theorem

$$0 = \int_U \nabla u \nabla \phi dx = -\int u \Delta \phi dx.$$

It follows that $u \in C^\infty(\Omega)$ is harmonic (EPDG). Functional analysis addresses the questions about the existence of such functions u.

2. Probability. Let Ω be a set, $\mathcal{A}$ a σ algebra and μ are measure defined on $\mathcal{A}$ with $\mu(\Omega) = 1$. Such measures are called probability measures. Often one has a family of measure spaces $(\Omega_n, \mathcal{A}_n, \mu_n)$ and wants to have notions of convergence. A good notion of convergence is based on the observation that a probability measure on the Borel σ algebra defines an element in $C_b(X)^*$. Suppose that $\Omega_n = \Omega$ is a metric space, $\mathcal{A}_n$ the Borel σ algebra. Then we can talk about so called weak* convergence

$$\mu_n \to^* \mu$$

which means that

$$\int f d\mu_n \to \int f d\mu$$

for every continuous function $f \in C(\Omega)$.

We will study such structures.

[9.10.2019]

[11.10.2016]

Lemma 1.7. *Let X be a normed space. Addition, scalar multiplication and the map to the norm are continuous.*

Lemma 1.8. *The closure of a subvector space of a normed spaces is a Banach space.*

Definition 1.9. *Two norms $\|.\|$ and $|.|$ on a normed spaces X are called equivalent, if there exits $C \geq 1$ so that for all $x \in X$*

$$C^{-1}\|x\| \leq |x| \leq C\|x\|.$$

Theorem 1.10. *All norms on finite dimensional spaces are equivalent. Finite dimensional normed spaces are Banach spaces.*

Proof. Let $|.|$ be the Euclidean norm on $\mathbb{K}^d$ and $\|.\|$ a second norm. Let $\{e_j\}_{j=1,\cdots,d}$ be the standard basis. Then

$$\Big\| \sum_{j=1}^d a_j e_j \Big\| \le \sum_{j=1}^d |a_j| \max_k \|e_k\| \le (\sqrt{d} \max_k \|e_k\|) \left| \sum_{j=1}^d a_j e_j \right| .$$

Thus $v \to \|.\|$ is continuous with respect to $|.|$. It attains the infimum on the Euclidean unit sphere (which is compact). This minimum has to be positive and we call it λ^{-1}. Then

$$|v| = |v||v/|v|| \le \lambda |v| \|v/|v|\| = \lambda \|v\|.$$

The two inequalities imply the equivalence of the norms $\|.\|$ and $|.|$ by choosing

$$C = \max\big\{\sqrt{d} \max \|e_k\|, \lambda, 1\big\}.$$

Thus every norm on $\mathbb{K}^d$ is equivalent to the Euclidean norm, and any two norms are equivalent.

A Cauchy sequence $v_m = (v_{m,j})$ with respect to $\|.\|$ is also a Cauchy sequence with respect to the Euclidean norm, hence it converges to a vector v with respect to the Euclidean norm, and hence also $\|v_m - v\| \to 0$.

This proves the claim for $\mathbb{K}^d$. Now let X be a vector spaces of dimension d. Then there is a basis of d vectors, and a bijective linear map ϕ from $\mathbb{K}^d$ to X. If $\|.\|_X$ is a norm on X then $x \to \|\phi(x)\|_X$ is a norm on $\mathbb{K}^d$. Thus the first part follows. Since $\phi(x_n)$ is a Cauchy sequence with respect to $\|.\|_X$ iff (x_n) is a Cauchy sequence in $\mathbb{K}^d$ with respect to the second metric, and one converges iff the second converges.This completes the proof.

□

Lemma 1.11. *Let X be a Banach space and U be a closed subvector space. Then X/U is a vector space,*

$$\|\tilde{x}\|_{X/U} = \inf_{y \in U} \|y - x\|$$

defines a norm (here $\tilde{x}$ is the equivalence class of x) and X/U is a Banach space.

Proof. Exercise □

Lemma 1.12. *Let X and Y be normed spaces. Their direct sum $X \oplus Y (= X \times Y)$ is a vector space. If $1 \le p \le \infty$ then*

$$\|(x, y)\|_p = \begin{cases} (\|x\|_X^p + \|y\|_Y^p)^{\frac{1}{p}} & \text{if } p < \infty \\ \max\{\|x\|_X, \|y\|_Y\} & \text{if } p = \infty \end{cases}$$

defines a norm with which $X \oplus Y$ becomes a Banach space if X and Y are Banach spaces.

Proof. Exercise, see also l^p below. □

1.1 Further examples of Banach spaces

Lemma 1.13. *The space $C_0(\mathbb{R}^d) \subset C_b(\mathbb{R}^d)$ of functions converging to 0 at ∞ is a Banach space. Similarly the space c_0 of sequences converging to 0 equipped with the sup norm is a Banach space.*

Further examples: Let $1 \le p \le q \le \infty$. We define the sequence spaces

Definition 1.14. *A $\mathbb{K}$ sequence $(x_j)_{j\in\mathbb{N}}$ is called p summable if*

$$\sum_{j=1}^{\infty} |x_j|^p < \infty \text{ if } p < \infty, \qquad \sup_{j\in\mathbb{N}} |x_j| < \infty \quad \text{ if } p = \infty.$$

The set of all p summable sequences is denoted by $l^p(\mathbb{N}) = l^p$.

Theorem 1.15. *The set of p summable sequences is a vector space. The expressions*

$$\|(x_j)\|_{l^p} = \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{1/p}, \quad p < \infty$$

resp.

$$\|(x_j)\|_{l^\infty} = \sup_{j\in\mathbb{N}} |x_j|$$

are norms on $l^p(\mathbb{N})$, which turn $l^p(\mathbb{N})$ into Banach spaces. If $\frac{1}{p} + \frac{1}{q} = 1$, $1 \le p, q \le \infty$ and $(x_j) \in l^p$, $(y_j) \in l^q$ then $(x_j y_j)$ is summable and Hölder's inequality holds:

$$\left|\sum_{j=1}^{\infty} x_j y_j\right| \le \sum_{j=1}^{\infty} |x_j y_j| \le \|(x_j)\|_{l^p} \|(y_j)\|_{l^q}.$$

Remark 1.16. *We may replace $\mathbb{N}$ by $\mathbb{Z}$, by a finite set, or even an arbitrary set. Then $l^\infty(X) = \mathbb{B}(X)$. The triangle inequality is called Minkowski inequality.*

We recall Young's inequality

$$|xy| \le \frac{1}{p}|x|^p + \frac{1}{q}|y|^q$$

for $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p, q < \infty$ and $x, y \in \mathbb{R}$. Without loss of generality we assume $x, y > 0$ and this can be proven by searching the maximum of

$$x \to xy - \frac{1}{p}x^p$$

for $y > 0$ which is attained at $x_0 = y^{1/(p-1)}$:

$$x_0 y - \frac{1}{p} x_0^p = \frac{p-1}{p} y^{\frac{p}{p-1}} = \frac{1}{q} y^q.$$

As a consequence

$$\sum_j |x_j y_j| \le \sum_j \frac{1}{p}|x_j|^p + \frac{1}{q}|y_j|^q = \frac{1}{p}\|(x_j)\|_{l^p}^p + \frac{1}{q}\|(y_j)\|_{l^q}^q$$

and we obtain Hölder's inequality

$$\begin{aligned}\sum_j |x_j y_j| =& \|(x_k)\|_{l^p}\|(y_k)\|_{l^q} \sum_j |x_j| \|(x_k)\|_{l^p}^{-1} |y_j| \|(y_k)\|_{l^q}^{-1} \\ =& \|(x_k)\|_{l^p}\|(y_k)\|_{l^q}(\frac{1}{p} + \frac{1}{q}) \\ =& \|(x_j)\|_{l^p}\|(y_j)\|_{l^q}.\end{aligned}$$

Proof. Since $l^\infty(\mathbb{N}) = \mathbb{B}(\mathbb{N})$ there is nothing to show if $p = \infty$. Moreover the triangle inequality is obvious if $p = 1$. Then, if $1 < p, q < \infty$

$$\begin{aligned}\sum_{j=1}^\infty |x_j + y_j|^p \le& \sum_{j=1}^\infty |x_j + y_j|^{p-1}|x_j + y_j| \\ \le& \sum_{j=1}^\infty |x_j + y_j|^{p-1}|x_j| + \sum_{j=1}^\infty |x_j + y_j|^{p-1}|y_j| \\ \le& \|(|x_j + y_j|^{p-1})\|_{l^q}\Big(\|(x_j)\|_{l^p} + \|(y_j)\|_{l^p}\Big) \\ =& \|(x_j + y_j)\|_{l^p}^{p-1}\Big(\|(x)_j\|_{l^p} + \|(y_j)\|_{l^p}\Big)\end{aligned}$$

and

$$\|(x_j + y_j)\|_{l^p} \le \|(x_j)\|_{l^p} + \|(y_j)\|_{l^p}.$$

provided we can divide by $\|(x_j + y_j)\|_{l^p}$. There is nothing to show if this quantity is 0 and it is finite whenever we sum over a finite number of indices. Then a limit argument gives the full statement.

In particular we obtain the triangle inequality. One easily sees that $\|(x_j)\|_{l^p} = 0$ implies $(x_j) = 0$ and

$$\|(\lambda x_j)\|_{l^p} = |\lambda| \|(x_j)\|_{l^p}.$$

Thus the spaces l^p are normed vector spaces. Now suppose that $x_n = (x_{n,j})$ is a Cauchy sequence in l^p. Then for every j, $n \to x_{n,j}$ is a Cauchy sequence in $\mathbb{K}$. Let $y_j = \lim_{n\to\infty} x_{n,j}$ and $y = (y_j)$. Then, for every $m > 1$ (assuming

$p < \infty$, since $l^\infty = \mathbb{B}(\mathbb{N})$)

$$\begin{aligned}\|y - x_m\|_{l^p}^p &= \lim_{N\to\infty} \sum_{j=1}^{N} |y_j - x_{m,j}|^p \\ &= \lim_{N\to\infty} \lim_{n\to\infty} \sum_{j=1}^{N} |x_{n,j} - x_{m,j}|^p \\ &\le \lim_{n\to\infty} \|x_n - x_m\|_{l^p}^p \\ &\to \quad 0 \qquad \text{as } m \to \infty.\end{aligned}$$

□

16.11.2019

2 Hilbert spaces

2.1 Definition and first properties

Definition 2.1. *Let X be a $\mathbb{K}$ vector space. A map $\langle .,. \rangle : X \times X \to \mathbb{K}$ is called inner product if*

$$\langle x_1 + x_2, y\rangle = \langle x_1, y\rangle + \langle x_2, y\rangle \qquad \text{for all } x_1, x_2, y \in X \tag{2.1}$$

$$\langle \lambda x, y\rangle = \lambda \langle x, y\rangle \qquad \text{for all } x, y \in X, \lambda \in \mathbb{K} \tag{2.2}$$

$$\langle x, y\rangle = \overline{\langle y, x\rangle} \qquad \text{for all } x, y \in X \tag{2.3}$$

In particular $\langle x, x\rangle \in \mathbb{R}$ for all $x \in X$ and

$$\langle x, x\rangle \ge 0 \qquad \text{for all } x \in X \tag{2.4}$$

$$\langle x, x\rangle = 0 \qquad \Longleftrightarrow \qquad x = 0 \tag{2.5}$$

Examples:

1. Euclidean vector spaces over $\mathbb{K}$, Euclidean inner product.
2. Real and complex square summable sequences space $l^2(\mathbb{N})$ with $\langle (x_j), (y_j)\rangle = \sum x_j \overline{y_j}$.
3. Let $U \subset \mathbb{R}^d$ be measurable, $X = C_b(U)$, $\langle f, g\rangle = \int_U f\bar{g} dm^n$ where m^n denotes the Lebesgue measure.
4. Let $(X, \mathcal{A}, \mu)$ be a measure space (X a set, $\mathcal{A}$ a σ algebra, $\mu : \mathcal{A} \to [0, \infty]$ a measure. Let $L^2(\mu)$ be the space of square integrable functions with the inner product

$$\langle f, g\rangle = \int f\bar{g} d\mu.$$

Lemma 2.2 (Cauchy-Schwarz)**.** *Let X be a vector space with inner product. Then*

$$|\langle x, y\rangle| \le (\langle x, x\rangle \langle y, y\rangle)^{\frac{1}{2}}$$

for all $x, y \in X$.

Proof. Let $x, y \in X$ and $\lambda \in \mathbb{K}$. Then

$$\begin{aligned} 0 \le & \langle x - \lambda y, x - \lambda y\rangle \\ = & \langle x, x\rangle - \lambda \langle y, x\rangle - \bar{\lambda}\langle x, y\rangle + |\lambda|^2 \langle y, y\rangle. \end{aligned}$$

If $y = 0$ there is nothing to show, so we assume $y \neq 0$ and define $\lambda = \frac{\langle x,y\rangle}{\langle y,y\rangle}$. Then

$$0 \le \langle x, x\rangle - \frac{|\langle x, y\rangle|^2}{\langle y, y\rangle}$$

which implies the Cauchy-Schwarz inequality. □

Lemma 2.3. *The map*

$$x \to \|x\| := \sqrt{\langle x, x\rangle}$$

defines a norm.

With this notation the Cauchy-Schwarz inequality becomes

$$|\langle x, y\rangle| \le \|x\| \|y\|.$$

Proof. Clearly $\|x\| \ge 0$, $\|x\| = 0$ iff $x = 0$. Moreover

$$\|\lambda x\|^2 = |\lambda|^2 \|x\|^2$$

and by the Cauchy-Schwarz inequality

$$\begin{aligned} \|x + y\|^2 = & \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 \\ \le & \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ = & (\|x\| + \|y\|)^2. \end{aligned}$$

□

Definition 2.4. *A vector space X with an inner product is called pre-Hilbert space. It is a Hilbert space if it is a Banach space.*

Lemma 2.5. *The inner product defines a continuous map from $X \times X$ to $\mathbb{K}$.*

Proof. Exercise □

It is not hard to verify the parallelogram identity

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{2.6}$$

for $x, y \in H$ some pre-Hilbert space.

Theorem 2.6 (Jordan von Neumann)**.** *Let X be a normed $\mathbb{K}$ vector space with norm $\|.\|$. We assume that the norm satisfies the parallelogram identity*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \tag{2.7}$$

Then

$$\langle x, y\rangle = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2\Big) \tag{2.8}$$

if $\mathbb{K} = \mathbb{R}$ and

$$\langle x, y\rangle = \frac{1}{4}\Big(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2\Big) \tag{2.9}$$

otherwise defines an inner product such that the norm is the norm of the preHilbert space. Vice versa: The norm of a prehilbert space defines the parallelogram identity.

As a consequence we could define a Hilbert spaces as a Banach space whose norm satisfies the paralellogram identity. By an abuse of notation we call a normed space pre-Hilbert space if it satisfies the parallelogram identity.

Proof. We begin with a real normed spaces whose norm satisfies the parallelogram identity. We define

$$\langle x, y\rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

Then

$$\langle x, y\rangle = \langle y, x\rangle.$$

Since by the parallelogram identity

$$2\|x+z\|^2 + 2\|y\|^2 = \|x+y+z\|^2 + \|x-y+z\|^2$$

hence

$$\begin{aligned} \|x+y+z\|^2 =& 2\|x+z\|^2 + 2\|y\|^2 - \|x-y+z\|^2 \\ =& 2\|y+z\|^2 + 2\|x\|^2 - \|y-x+z\|^2 \end{aligned}$$

and

$$\|x+y+z\|^2 = \|x\|^2 + \|y\|^2 + \|x+z\|^2 + \|y+z\|^2 - \frac{1}{2}\|x-y+z\|^2 - \frac{1}{2}\|y-x+z\|^2$$

$$\|x+y-z\|^2 = \|x\|^2+\|y\|^2+\|x-z\|^2+\|y-z\|^2-\frac{1}{2}\|x-y-z\|^2-\frac{1}{2}\|y-x-z\|^2$$

and we arrive at

$$\begin{aligned}\langle x+y,z\rangle =&\frac{1}{4}(\|x+y+z\|^2-\|x+y-z\|^2)\\ =&\frac{1}{4}(\|x+z\|^2-\|x-z\|^2)+\frac{1}{4}(\|y+z\|^2-\|y-z\|^2)\\ =&\langle x,z\rangle+\langle y,z\rangle.\end{aligned}$$

We claim

$$\langle \lambda x, y\rangle = \lambda\langle x,y\rangle$$

for all $x, y \in X$ and $\lambda \in \mathbb{R}$. It obviously holds for $\lambda = 1$ by checking the definition, and for all $\lambda \in \mathbb{N}$ be the previous step, hence for all $\lambda \in \mathbb{Z}$. But then it holds for all rational λ and by continuity for $\lambda \in \mathbb{R}$.

We complete the proof for complex Hilbert spaces: We define

$$\langle x,y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k\|x+i^k y\|^2$$

and observe that $\langle ix, y\rangle = i\langle x, y\rangle$, $\langle x, y\rangle = \overline{\langle y, x\rangle}$ by definition, $\mathrm{Re}\langle x, y\rangle$ is the previous real inner product and $\mathrm{Im}\langle x, y\rangle = \mathrm{Re}\langle x, iy\rangle$. □

Corollary 2.7. *A normed space is a pre-Hilbert space if and only if all two dimensional subspaces are pre-Hilbert spaces.*

Proof. It is a pre-Hilbert space if and only if its norm satisfies the parallelogram identity which holds if and only if the parallelogram identity holds for all two dimensional subspaces. □

This has geometric consequences.

Lemma 2.8. *Let H be a Hilbert space, $K \subset H$ compact, and $C \subset H$ closed and convex, C and K disjoint. Then there exist $x \in K$ and $y \in C$ so that*

$$\|x-y\| = d(C,K)$$

Proof. Let $x_j \in K$ and $y_j \in C$ be a minimizing sequence. Since K is compact there is a subsequence which we denote again by (x_j, y_j) and $x \in K$ so that $x_j \to x$. By the triangle inequality

$$\|x-y_j\| \to d(C,K).$$

Then

$$\begin{aligned}\|y_n-y_m\|^2 =&\|(x-y_n)-(x-y_m)\|^2\\ =&2\|x-y_n\|^2+2\|x-y_m\|^2-\|2x-(y_n+y_m)\|^2\\ \le&2(\|x-y_n\|^2+\|x-y_m\|^2)-4d^2(C,K)\\ &\to 0 \qquad \text{as } n,m\to\infty\end{aligned} \tag{2.10}$$

since by convexity $\frac{1}{2}(y_n + y_m) \in K$. Thus (y_n) is a Cauchy sequence with limit $y \in C$. Moreover

$$d(C, K) = \lim_{n\to\infty} \|x - y_n\| = \|x - y\|.$$

□

Definition 2.9. *We call two elements $x, y \in H$ orthogonal if $\langle x, y\rangle = 0$.*

Lemma 2.10. *Suppose that C is a closed and convex subset of a Hilbert space H and $x \in H$. Then the closest point in C to x is unique and we denote it by $p(x)$. Moreover,*

$$\operatorname{Re}\langle x - p(x), z - p(x)\rangle \le 0 \tag{2.11}$$

for all $z \in C$. If $y \in C$ satisfies

$$\operatorname{Re}\langle x - y, z - y\rangle \le 0$$

for all $z \in C$ then $y = p(x)$. If C is a closed subspace then for all $z \in C$

$$\langle x - p(x), z\rangle = 0. \tag{2.12}$$

The point $y = p(x) \in C$ is uniquely determined by this orthogonality condition. Moreover

$$\|x\|^2 = \|x - p(x)\|^2 + \|p(x)\|^2. \tag{2.13}$$

Proof. Uniqueness is a consequence of the proof of Lemma 2.8. Let $y \in C$. Then by the triangle inequality, if $z \in C$ then for $0 \le t \le 1$

$$y(t) = y + t(z - y) \in C$$

and hence if $y = p(x)$,

$$\|x-y\|^2 \le \|x-y-t(z-y)\|^2 = \|x-y\|^2 - 2t\operatorname{Re}\langle x-y, z-y\rangle + t^2\|z-y\|^2.$$

This implies (2.11) and also the converse. In the case that C is a closed subspace (2.11) implies that

$$\operatorname{Re}\langle x - p(x), z\rangle \le 0$$

for all $z \in C$, hence, combined with this inequality for $-z$

$$\operatorname{Re}\langle x - p(x), z\rangle = 0$$

for all $z \in C$ and

$$\operatorname{Im}\langle x - p(x), z)\rangle = -\operatorname{Re}\langle x - p(x), iz\rangle = 0$$

for $z \in C$. Now we expand

$$\begin{aligned} \|x\|^2 =& \|x - p(x) + p(x)\|^2 \\ =& \|x - p(x)\|^2 + \|p(x)\|^2 + 2\operatorname{Re}\langle x - p(x), p(x)\rangle \\ =& \|x - p(x)\|^2 + \|p(x)\|^2. \end{aligned}$$

□

2.2 The Riesz representation theorem

Theorem 2.11 (Riesz representation theorem)**.** *Let H be a Hilbert space. Then*

$$J : H \ni x \to (y \to \langle y, x\rangle) \in H^*$$

is an $\mathbb{R}$ linear isometric isomorphism. It is conjugate linear:

$$J(\lambda x) = \overline{\lambda} J(x).$$

Proof. By the Cauchy-Schwarz inequality

$$\|J(x)\|_{H^*} = \sup_{\|y\|\le 1} |\langle y, x\rangle| \le \|x\|_H$$

and the map is well defined and conjugate linear:

$$J(\lambda x)(y) = \langle y, \lambda x\rangle = \bar{\lambda}\langle y, x\rangle.$$

Since

$$\|x\|_H \|J(x)\|_{H^*} \ge \langle x, J(x)\rangle = \langle x, x\rangle = \|x\|_H^2$$

we see that

$$\|J(x)\|_{H^*} \ge \|x\|_H.$$

Thus J is an isometry: $\|J(x)\|_{H^*} = \|x\|_H$. In particular J is injective. To show that J is surjective we assume that $x^* \in H^*$ and try to find x so that $x^* = J(x)$. Let

$$N = \{y \in H : x^*(y) = 0\}.$$

Then N is a subvector space and it is closed. Let p be the orthogonal projection to N as above. We choose $y_0 \in H$ with $x^*(y_0) = 1$ and define

$$x_0 = y_0 - p(y_0).$$

Then $x^*(x_0) = 1$ (since $p(y_0) \in N$) and, since $\|x_0\| = \|y_0 - p(y_0) = \operatorname{dist}(y_0, N) = \operatorname{dist}(x_0, N)$ we have $p(x_0) = 0$ and for all $y \in N$ by (2.12) $\langle y, x_0\rangle = 0$. Moreover obviously $x^*(x - x^*(x)x_0) = 0$ hence $x - x^*(x)x_0 \in N$ and by (2.12)

$$\langle x - x^*(x)x_0, x_0\rangle = 0.$$

Since $x = [x - x^*(x)x_0] + x^*(x)x_0$, then

$$\langle x, x_0\rangle = \langle x^*(x)x_0, x_0\rangle = x^*(x)\|x_0\|_H^2$$

and, solving this identity for $x^*(x)$

$$x^*(x) = \left\langle x, \frac{x_0}{\|x_0\|^2}\right\rangle = J(\frac{x_0}{\|x_0\|^2})(x).$$

□

18.10.2019

Theorem 2.12 (Lax-Milgram)**.** *Let H be a Hilbert space and*

$$Q : H \times H \ni (x, y) \to Q(x, y) \in \mathbb{K}$$

be linear in x, antilinear in y, bounded in the sense that

$$|Q(x, y)| \le C\|x\|\|y\|$$

and coercive in the sense that there exists $\delta > 0$ so that

$$\operatorname{Re} Q(x, x) \ge \delta\|x\|^2.$$

Then there exists a unique continuous linear map $A : H \to H$ with continuous inverse A^{-1} so that

$$Q(x, y) = \langle Ax, y\rangle.$$

Moreover

$$\|A\|_{H\to H} \le C, \qquad \|A^{-1}\|_{H\to H} \le \delta^{-1}.$$

Remark 2.13. *A map Q with these properties is called sesquilinear.*

Proof. Let $x \in H$. Then

$$y \to \overline{Q(x, y)} \in H^*.$$

By the Riesz representation theorem there exists a unique $z(x) \in H$ so that

$$\overline{\langle z(x), y\rangle} = \overline{Q(x, y)}$$

for all $y \in H$. Then

$$\|z(x)\| = \sup_{\|y\|\le 1} |\langle z(x), y\rangle| = \left|\sup_{\|y\|\le 1} Q(x, y)\right|.$$

Clearly $z(x_1 + x_2) = z(x_1) + z(x_2)$ and $z(\lambda x) = \lambda z(x)$ and we define the continuous linear operator $Ax = z$. Since

$$\operatorname{Re}\langle Ax, x\rangle \ge \delta\|x\|^2$$

we obtain

$$\|Ax\| \ge \delta\|x\|.$$

It particular A is injective and the range is closed. If it is not surjective there exists z with $\|z\| = 1$ and z is orthogonal to the range, i.e.

$$\langle Ax, z\rangle = 0$$

for all $x \in X$. In particular we reach the contradiction

$$0 = \langle Az, z\rangle \ge \delta\|z\|^2.$$

□

Let $q, h \in C([0,1],\mathbb{R})$. We consider the boundary value problem

$$- u'' + qu = h \qquad \text{in } (0,1) \qquad u(0) = u(1) = 0 \tag{2.14}$$

Theorem 2.14. *Suppose that* $\inf q(x) > -2$. *Then* (2.14) *has exactly one solution.*

Proof. We consider $\mathbb{K} = \mathbb{R}$. In particular $\langle f, g\rangle = \langle g, f\rangle$.

Step 1. The Hilbert space 1: Let

$$\tilde{H} : \{U \in C^1([0,1]) : U(0) = U(1) = 0\}$$

which we equipp with the norm

$$\langle U, V\rangle_H = \int U'V'dx$$

Since for $0 \le x_1 \le x_2 \le 1$

$$|U(x_2) - U(x_1)| = \left| \int_{x_1}^{x_2} U'(t)dt \right| \le \|U'\|_{L^2} \|\chi_{[x_1,x_2]}\|_{L^2} \le |x_2 - x_1|^{\frac{1}{2}} \|U'\|_{L^2}$$

we see that $\|.\|_H$ defines a norm since the other conditions are immediate. The problem is that $\tilde{H}$ is not complete.

Measurable square integrable functions on $[0,1]$ are integrable. For $f \in L^2((0,1))$ we define

$$F(x) = \int_0^x f(y)dx.$$

Then, if $0 \le x_1 < x_2 \le 1$

$$|F(x_2) - F(x_1)| = \left| \int_{x_1}^{x_2} f(t)dt \right| \le \|f\|_{L^2} \|\chi_{[x_1,x_2]}\|_{L^2} \le (x_2 - x_1)^{\frac{1}{2}} \|f\|_{L^2}$$

and hence F is uniformly continuous and even Hölder continuous of exponent $\frac{1}{2}$. Moreover $F(0) = 0$ the map $H_0 \ni f \to F \in C_b([0,1])$ is linear, injective and continuous.

Step 2. We define $H \subset C_b([0,1])$ by

$$F \in H \iff \text{There exists } f \in L^2 \text{ with } \int_0^1 f dx = 0 \text{ and } F(x) = \int_0^x f(y)dy$$

and equipp it with the real inner product

$$\langle F_1, F_2\rangle_H = \int_0^1 f_1 f_2 dx.$$

Then H is isometric isomorphic to the closed subspace of L^2 of functions with integral 0, and hence a Hilbert space. In the sequel we use small resp. capital letter to indicate this relation. Moreover $F \in H$ implies $F(0) = F(1) = 0$ and

$$\|F\|_{sup} \le 2^{-1/2} \|F\|_H.$$

Step 3. Relaxation and the sesquilinear bilinear form. Suppose that

$$-U'' + qU = h$$

and let $\phi \in C^1([0,1])$ with $\phi(0) = \phi(1) = 0$. We multiple by ϕ and integrate. Then

$$\int_0^1 h\phi dx = \int -U''\phi dx + \int qU\phi dx = \int_0^1 U'\phi' + qU\phi dx$$

We define

$$Q(F,G) = \int fg dx + \int qFG dx.$$

Then U satisfies

$$Q(U,\phi) = \int_0^1 h\phi dx$$

for all ϕ as above, and with an approximation it also holds for $\phi \in H$. We relax our problem by first searching for $U \in H$ so that this identity holds for all $\phi \in H$.

We

$$|Q(F,G)| \le \left| \int fg dx \right| + \|q\|_{sup} \|F\|_{sup} \|G\|_{sup} \le (1 + \|q\|_{sup}) \|f\|_{L^2} \|g\|_{L^2},$$

$F \to Q(F,G)$ is linear for all G and $G \to Q(F,G)$ is linear.We check the coercivity:

$$\operatorname{Re} Q(F,F) \ge \|f\|_{L^2}^2 + \min\{\inf q, 0\} \|F\|_{sup}^2 \ge (1 + \frac{1}{2} \min \inf q, 0\}) \|F\|_H^2.$$

Let A be the invertible map defined in the Lemma of Lax Milgram 2.12.

Step 4. The linear form We define a map

$$C_b([0,1]) \ni h \to (F \to \int_0^1 Fh dx) \in H^*.$$

By the Riesz representation theorem there exists a unique $G \in H$ so that

$$\int_0^1 Fh dx = \langle F, G \rangle$$

for all $F \in H^*$. We define

$$u = A^{-1} G.$$

Then

$$Q(u,F) = \langle Au, F\rangle_H = \langle G, F\rangle_H = \int_0^1 hF dx$$

for all $F \in H$.

Step 5: Regularity. Let x_0 be a Lebesgue point of u and, for $r > 0$ small

$$\phi(x) = \begin{cases} x/(x_0 - r) & \text{if } x < x_0 - r \\ 1 - (x - x_0 - r)/(2r) & \text{if } x_0 - r \le x < x_0 + r \\ 0 & \text{if } x \ge x_0 + r. \end{cases}$$

Then $\phi \in H$ (with $\int_0^x \phi'(y)dy = \phi(x)$ by the fundamental theorem of calculus))

$$\frac{1}{2r}\int_{x_0-r}^{x_0+r} u dx = \frac{1}{x_0 - r}\int_0^{x_0-r} u dx + \int_0^{x_0} (qU + h)\phi dx$$

Now we let r tend to zero and we obtain (assuming that $u(x_0)$ is the limit of the averages)

$$x_0 u(x_0) = \int_0^{x_0} u dx + \int_0^{x_0} (qU + h) x dx.$$

By the convergence theorem of Lebesgue the right hand side is continuous in $x \in (0,1)$, hence u has a continuous representative (the set of Lebesgue points has full measure), and $U' = u$. Then the right hand side is continuously differentiable on $(0,1)$ and differentiation gives

$$u(x) + xu'(x) = u(x) + x(qU + h)$$

hence

$$U''(x) = q(x)U(x) + h(x).$$

Thus $U \in C^2([0,1])$ and it satisfies the differential equation. □

23.10.2019

There is an alternative approach. Let $b, c \in C_b([0,1];\mathbb{R})$. We consider the boundary value problem

$$-u'' + bu' + cu = h \text{ in } (0,1), u(0) = u(1) = 0$$

for $h \in C_b([0,1])$.

Theorem 2.15. *The boundary value problem is solvable for all h if the homogeneous equation with $h = 0$ has only the trivial solution.*

Proof. Let U be the unique solution to the differential equation with initial data $U(0) = 0$, $U'(0) = 1$. If $U(1) = 0$ then the homogeneous problem has a nontrivial solution U. Suppose that $U(1) \neq 0$. Let V be the solution with $V(1) = 1$, $V'(1) = 1$. Then $V(0) \neq 0$ (otherwise V would be a multiple of U, which contradicts $V(1) = 0$, U and V are linearly independent. Let

$$W(x) = \det \begin{pmatrix} V & U \\ V' & U' \end{pmatrix}$$

Then

$$W'(x) = b(x)W(x)$$

and $W(0) = V(0)U'(0) = V(0) \neq 0$, hence W never vanishes. Let

$$g(x, y) = \begin{cases} \frac{1}{W(y)} U(y)V(x) & \text{if } x > y \\ \frac{1}{W(y)} V(y)U(x) & \text{if } x < y \end{cases}$$

Then

$$g \in C([0, 1] \times [0, 1]), g(0, y) = g(1, y) = 0,$$

g is differentiable unless $x = y$ and

$$\partial_x g(x, y) = \begin{cases} \frac{1}{W(y)} U(y)V'(x) & \text{if } x > y \\ \frac{1}{W(y)} V(y)U'(x) & \text{if } x < y \end{cases}$$

and hence it has a jump of size -1 at the diagonal. We define

$$u(x) = \int_0^1 g(x, y)h(y)dy \tag{2.15}$$

Then

$$\begin{aligned} u'(x) =& \partial_x \int_0^x g(x, y)h(y)dy + \int_x^1 g(x, y)h(y)dy \\ =& g(x, x)h(x) + \int_0^x \partial_x g(x, y)h(y)dy - g(x, x)h(x) + \int_x^1 \partial_x g(x, y)h(y)dy \\ =& \int_0^1 \partial_x g(x, y)h(y)dy \end{aligned}$$

and, ∂^+ resp ∂^- the derivative from above resp. from below,

$$\begin{aligned}
-u''(x) =& - \partial_x \int_0^x g_x(x,y)h(y)dy + \partial_x \int_x^1 g_x(x,y)h(y)dy \\
=& - \int_0^x \partial_x^2 g(x,y)h(y)dy - \int_x^1 \partial_x^2 g(x,y)dy \\
& - \partial_x^- g(x,x)h(x) + \partial_x^+ g(x,x)h(x) \\
=& b(x)\partial_x \int_0^1 g(x,y)h(y)dy + c(x) \int_0^1 g(x,y)h(y)dy \\
& - \frac{1}{W(x)}(V'(x)U(x) - V(x)U'(x))h(x) \\
=& b(x)u'(x) + c(x)u(x) + h(x)
\end{aligned}$$

and $u \in C_b^2([0,1])$ is a solution. Here we used that if $x \neq y$ the map $x \to g(x,y)$ is a solution to the homogeneous problem, which allows to replace the second order derivatives by $b\partial_x g + cg$.

The solution is unique since the homogeneous equations has only the trivial solution.

□

Remark 2.16. *We obtain much more than stated: There is an integral formula* (2.15) *for the solution. g is called Green's function.*

2.3 Operators on Hilbert spaces

The purpose of this section is to introduce several definitions.

Definition 2.17.

- *The adjoint of $T \in L(H_1, H_2)$ is the unique operator $T^* \in L(H_2, H_1)$ which satisfies*

$$\langle Tx, y\rangle_{H_2} = \langle x, T^*y\rangle_{H_1}.$$

for all $x \in H_1$, $y \in H_2$.

- *Let H_1 and H_2 be $\mathbb{K}$ Hilbert spaces. Then $U \in L(H_1, H_2)$ is called unitary, if it is invertible (i.e. there exists an inverse U^{-1} so that $U^{-1}U = 1_{H_1}$, $UU^{-1} = 1_{H_2}$) and if*

$$\langle Ux, Uy\rangle_{H_2} = \langle x, y\|_{H_1}$$

for all $x, y \in H_1$.

- *$T \in L(H, H)$ is called self adjoint if $T^* = T$.*

Remark 2.18. *1. Existence and unqiueness in the second part has to be shown. Let $y \in H_2$. Then*

$$x \to \langle Tx, y\rangle_{H_2} \in H_1^*$$

By the Riesz representation theorem there exists a unique $z \in H_1$ so that

$$\langle Tx, y\rangle_{H_2} = \langle z, y\rangle_{H_1}.$$

We define $T^*y = z$. It is clearly linear and satisfies

$$\|T^*\|_{H_2\to H_1} = \|T\|_{H_1\to H_2}.$$

2. *The composition of unitary operators is again unitary. The adjoint of a unitary operator is unitary, and it is the inverse: Let $U : H_1 \to H_2$ be unitary then*

$$\langle x, y\rangle_{H_1} = \langle Ux, Uy\rangle_{H^2} = \langle U^*Ux, y\rangle_{H_1}$$

for x, y. Thus U^*U is the identity in H_1. Moreover

$$U = U(U^*U) = (UU^*)U$$

and hence

$$(UU^* - 1_{H_2})U = 0.$$

Since U is surjective this implies $UU^* = 1_{H_2}$ and hence U^* is the left and right inverse of U.

3. *In particular the unitary operators in $L(H) = L(H, H)$ are a group, called the unitary group $U(H)$.*

Example: Let $H = l^2(\mathbb{N})$ and e_n be the sequence with 1 at the position n and 0 otherwise.

$$(T(x_n))_m = \begin{cases} 0 & \text{if } m = 1 \\ x_{m-1} & \text{if } m < 1 \end{cases}$$

Then

$$\|T(x_n)\| = \|(x_n)\|,$$
$$(T^*(x_n))_m = x_{m+1}$$

Then

$$T^*e_0 = 0, T^*e_{j+1} = e_j, Te_j = e_{j+1}.$$

T^*T is the identity, but TT^* is the projection to the subspace defined by $\langle x, e_1\rangle = 0$.

We define the null space or kernel of an operator T by

$$N(T) = \{x : Tx = 0\}$$

and the range

$$R(T) = \{y : \text{ there exists } x \text{ with } y = Tx\}$$

2.4 Orthonormal systems

We recall that N elements x_j of a vector space X are called linearly independent if

$$\sum_{j=1}^{N} \lambda_j x_j = 0$$

implies $\lambda_j = 0$. Let x_j be N linearly independent vectors of a Hilbert space. By the Gram-Schmidt procedure we obtain an orthonormal system with

$$\begin{aligned} y_1 &= \frac{1}{\|x_1\|} x_1 \\ \tilde{y}_2 &= x_2 - \langle x_2, y_1 \rangle y_1 \\ y_2 &= \frac{1}{\|\tilde{y}_2\|} \tilde{y}_2 \end{aligned}$$

recursively. We can do this with $N = \infty$.

Definition 2.19. *An orthonormal system of H is given by a map $A \to H$ denoted by x_α for $\alpha \in A$, A a set wuch that*

$$\langle x_\alpha, x_\beta \rangle = \begin{cases} 1 & \text{if } \alpha = \beta \\ 0 & \text{otherwise} \end{cases}$$

Typically A is a subset of the natural numbers.

Lemma 2.20 (Bessel inequality)**.** *Let $(x_n)_{n \le N}$ be an orthonormal system. Then*

$$0 \le \|x\|^2 - \sum_{n=1}^{N} |\langle x, x_n \rangle|^2 = \|x - \sum_{n=1}^{N} \langle x, x_n \rangle x_n\|^2$$

Proof. Let $M \subset H$ be the N dimensional subspace spanned by the elements x_j and let p be the projection to the closest point. Then by Lemma 2.10

$$\|x\|^2 = \|x - p(x)\|^2 + \|p(x)\|^2.$$

Moreover

$$p(x) = \sum_{j=1}^{N} \lambda_j x_j$$

and

$$\langle x, x_n \rangle = \langle p(x), x_n \rangle = \sum_{m=1}^{N} \lambda_m \langle x_m, x_n \rangle = \lambda_n$$

and

$$\|p(x)\|^2 = \sum_{n=1}^{\infty} |\lambda_n|^2.$$

□

Definition 2.21. *A subset A of a metrix space X is called dense if its closure is X. A metric space X is called separable, if there is a countable dense subset.*

Examples:

1. $\mathbb{N}$ is countable.
2. X and Y countable implies $X \times Y$ is countable. In particular $\mathbb{Q}^d$ is countable.
3. If X_j are countable sets then there union is countable.
4. Subsets of countable sets are countable.
5. $\mathbb{Q}$ is countable.
6. $\mathbb{Q}^N$ is countable.
7. $\mathbb{R}^N$ is separable since $\mathbb{Q}^N$ is countable and dense.
8. $l^2(\mathbb{N})$ is separable. This requires a proof. Let $Q \subset l^2(\mathbb{N})$ be the set of all sequences with rational coefficients, for which there exists N so that larger components are zero. This set is countable, as countable union of set for which there is a bijective map to $\mathbb{Q}^N$. Now let $x \in l^2$ and $\varepsilon > 0$ and $x_N \in l^2$ the vector with the same first N components, and $(x_N)_n = 0$ for $n > N$. There exists N so that

$$\|x - x_N\|_{l^2} < \varepsilon/2$$

and then an element y of Q with

$$\|x_N - y\|_{l^2} < \varepsilon/2$$

Thus Q is dense.

25.10.2019

Definition 2.22. *A orthonormal system (x_n) of a Hilbert space is called orthonormal basis if*

$$\langle x, x_n \rangle = 0 \quad \textit{for all } n \in \mathbb{N}$$

implies $x = 0$.

Theorem 2.23. *The following properties are equivalent for a Hilbert space H which is not finite dimensional.*

- *The space H is separable.*
- *There exists an orthonormal basis and $\|x\|^2 = \sum |\langle x, x_j \rangle|^2$.*

- *The exists a unitary map $l^2 \to H$*

Proof. Suppose that H is separable. Let (y_n) be a dense sequence and X_N the span of the first N y_n. Its dimension is at most N. We use the Gram-Schmidt procedure to find an orthonormal basis (x_n) of X_N. We do this recursively by increasing n. This leads to a countable orthonormal sequence (x_n) so that its span is dense. Now let $x \in X$. By the Bessel inequality

$$\sum_{j=1}^{N} \langle x, x_j \rangle^2 + \|x - \sum_{j=1}^{N} \langle x, x_j \rangle x_j\|^2 = \|x\|^2.$$

Thus

$$N \to \|x - \sum_{n=1}^{N} \langle x, x_n \rangle x_n\|$$

is monotonically decreasing. Since (y_n) is dense and $\sum_{n=1}^{N} \langle x, x_n \rangle x_n = p_N x$ is the closed point in the span of $(x_n)_{n \le N}$ it converges to 0, which is equivalent to

$$\sum_{n=1}^{N} \langle x, x_n \rangle x_n \to x$$

in L^2, which in turn implies

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2.$$

Now suppose that (x_n) is an orthonormal basis. We want to define

$$l^2 \ni (a_n) \to \sum_{n=1}^{\infty} a_n x_n \in H.$$

We claim that for $M \ge N$

$$\sum_{n=1}^{N} a_n x_n - \sum_{n=1}^{M} a_n x_n = \sum_{n=N+1}^{M} a_n x_n$$

the norm of which is given by

$$\sqrt{\sum_{n=N+1}^{M} |a_j|^2}.$$

This implies that the partial sums are a Cauchy sequence and we define

$$x := \lim_{N \to \infty} \sum_{n=1}^{N} a_n x_n$$

Then

$$\|x\|^2 = \lim_{N\to\infty} \sum_{n=1}^{N} |a_n|^2 = \|(a_n)\|^2.$$

The map is clearly linear, surjective (by the previous part) and an isometry. To complete the proof we recall that $l^2(\mathbb{N})$ is separable.

□

In particular a Hilbert space is either isomorphic (there exists an isometric surjective linear map) to $\mathbb{R}^d$ resp. $\mathbb{C}^d$, to l^2, or it is not separable.

Example: The space $l^2(\mathbb{R})$ with inner product

$$\langle f, g\rangle = \sum_{x\in\mathbb{R}} f(x)\overline{g(x)}$$

is not separable since the vectors

$$e^x_y = \begin{cases} 0 & \text{if } y \neq x \\ 1 & \text{if } y = x \end{cases}$$

are an uncountable orthonormal system. In particular the pairwise distance is $\sqrt{2}$ and there cannot be a countable dense set.

2.5 Orthogonal polynomials

Let μ be a Borel measure on $\mathbb{R}$ so that all moments exists, i.e.

$$\int_{\mathbb{R}} |x|^N \mu dx < \infty$$

for all $N \geq 0$. Let $H = L^2(\mu)$ with inner product

$$X \times X \ni (f, g) \to \langle f, g\rangle := \int f\overline{g}\mu dx$$

We assume that the monomials

$$f_n = x^n$$

are linearly independent.

We consider the case

$$\mu(x) = \begin{cases} 1/2 & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

with $X = C([-1, 1])$. It leads to (multiples) of the Legendre polynomials.

Definition 2.24 (Legendre-polynomial)**.** *The Legendre polynomial P_n is the unique monic polynomial of degree n (i.e. with leading term x^n)*

$$\int_{-1}^{1} x^m P_n(x) dx = 0$$

for all $0 \le m < n$.

There is a very compact formula for them.

Lemma 2.25 (Rodrigues formula)**.**

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}[(x^2-1)^n]$$

Proof. The degree of P_n defined by the right hand side is obviously n and the leading term of $P_n(x)$ reads $\frac{(2n)!}{2^n (n!)^2} x^n$. If $0 \le m < n$ then after m integrations by parts

$$\int_{-1}^{1} x^m \frac{d^n}{dx^n}(x^2-1)^n dx = (-1)^m m! \int_{-1}^{1} \frac{d^{n-m}}{dx^{n-m}}(x^2-1)^n dx = 0.$$

Finally

$$\frac{d^n}{dx^n}(x^2-1)^n \bigg|_{x=1} = 2^n \frac{d^n}{dx^n}(x-1)^n \bigg|_{x=1} = 2^n n!.$$

□

A more difficult calculation gives

$$\begin{aligned}\int_{-1}^{1} (P_n(x))^2 \, dx =& \frac{1}{2^n n!} \frac{(2n)!}{2^n (n!)^2} \int_{-1}^{1} x^n \frac{d^n}{dx^n}(x^2-1)^n dx \\ =& \frac{(2n)!}{2^{2n}(n!)^2} (-1)^n \int_{-1}^{1} (x^2-1)^n dx \\ =& \frac{(2n)!}{2^{2n}(n!)^2} 2 \cdot 2^{2n} \int_0^1 s^n (1-s)^n ds = \frac{2}{2n+1}.\end{aligned}$$

and

$$\sqrt{\frac{2n+1}{2}} \sqrt{\frac{2m+1}{2}} \int_{-1}^{1} P_n(x) P_m(x) \, dx = \delta_{n,m}$$

and the functions

$$\sqrt{\frac{2n+1}{2}} P_n(x)$$

are an orthonormal system in $L^2(\mu)$.

Theorem 2.26. *These functions are a basis of $L^2(\mu)$.*

Proof. See introduction to PDEs. □

We consider the complex Hilbert space $L^2([0,2\pi])$ with inner product

$$\langle f, g\rangle = \frac{1}{2\pi}\int f\overline{g}dx.$$

Lemma 2.27. *The functions e^{inx} are a basis.*

Proof. We compute

$$\frac{1}{2\pi}\int_0^{2\pi} e^{inx}\overline{e^{-imx}}dx = \frac{1}{2\pi}\int_0^{2\pi} e^{i(n-m)x}dx = 0$$

if $n \neq m$ and $= 1$ if $n = m$. See Analysis 3 for a proof of the basis property. □

2.6 Sturm-Liouville problems

Let $q \in C_b([0,1])$. For $\lambda \in \mathbb{C}$ we consider the eigenfunction problem

$$-u'' + qu = \lambda u \qquad \text{in } (0,1)$$

$$u(0) = u(1) = 0.$$

This is a special case of Theorem 2.15, with complex function $q - \lambda$. It is not hard to see that there is no nontrivial solution unless $\lambda \in \mathbb{R}$: Suppose that $\lambda \notin \mathbb{R}$ and $u \in C^2([0,1];\mathbb{C})$ satisfies the boundary value problem. Then

$$0 = \int_0^1 -u''\bar{u} + qu\bar{u} - \lambda u\bar{u}dx = \int_0^1 |u'|^2 + q|u|^2 - \operatorname{Re}\lambda |u|^2 dx - i\operatorname{Im}\lambda \int_0^1 |u|^2 dx$$

and thus $\operatorname{Im}\lambda = 0$ or $\int_0^1 |u|^2 dx = 0$. Hence there is no nontrivial solution unless $\lambda \in \mathbb{R}$.

If $u(x) = 0$ then $u'(x) \neq 0$, otherwise u vanishes identically. As a consequence zeros are isolated.

Theorem 2.28. *Given $\lambda \in C$ the space of solutions to (2.14) is vector space of dimension 0 or 1. There exists a monotone sequence $\lambda_n \to \infty$ and a sequence of nontrivial real valued functions $u_n \in C^2([0,1])$ which satisfy*

$$-u_n'' + qu_n = \lambda_n u_n$$

$$\int u_n u_m dx = \delta_{nm}.$$

The functions u_n have exactly $n-1$ zeroes in $(0,1)$. The function u_{n+1} has one zero between two zeros of u_n. Moreover

$$\pi^2 n^2 + \inf q \leq \lambda_n \leq \pi^2 n^2 + \sup q. \tag{2.16}$$

If $\lambda \neq \lambda_n$ for some j then (2.14) has only trivial solutions.

We consider a lemma before we turn to the proof.

Lemma 2.29. *Let $I=[a,b]$. Suppose that $q_1, q_2 \in C_b([a,b];\mathbb{R})$, $q_2 < q_1$, suppose that $u \in C^2([a,b])$ is positive on (a,b), $u(a)=u(b)=0$ and*

$$-u'' + q_1 u = 0.$$

Suppose that $v \in C^2([a,b])$ satisfies

$$-v'' + q_2 v = 0$$

Then v has a zero in (a,b).

Proof. Suppose that v has no zero in $[a,b]$. Then we may assume that v is postive in this interval. Let

$$w = v/u \in C^2((a,b))$$

Then $w(x) \to \infty$ as $x \to a, b$ and hence w assume the positive minimum at a point x_0. On the other hand

$$w'(x) = v'/u - wu'/u,$$

$$w'' = v''/u - \frac{v'u'}{u^2} - \frac{u'}{u}w' - w\frac{u''}{u} = -(q_2 - q_1)w - 2\frac{u'}{u}w' - \left(\frac{u'}{u}\right)^2 w$$

and hence w does not assume an interior minimum. Now suppose that $v(a)=0$. Since $u'(a) \neq 0$ and $v'(a) \neq 0$ we have by the rule of de l'Hospital

$$w(a) = v'(a)/u'(a), \qquad w'(a) = v''(a)/u'(a) - w(a)u''(a)/u'(a) = 0$$

and hence $w'' < 0$ in (a, x_1) for some $x_1 > a$, hence it cannot assume the minimum at a. The same argument applies at b. Hence v has a zero in (a,b). □

Proof of Theorem 2.28 : We apply this with $q_1 = q - \lambda_1$ and $q_2 = q - \lambda_2$ with $\lambda_2 > \lambda_1$. If u_1 , u_2 are corresponding solutions then the zeros are interlaced. In particular, if λ_j are eigen values and u_j eigenfunctions then u_2 has a zero between two zeros of u_1 and the number of zeros is monotonically increasing with λ. In particular there are at most countably many eigenfunctions, and for each n there exists at most one eigenfunction with n zeros in $(0,1)$.

Since with

$$U_\mu(x) = \sin(\pi\mu x)$$
$$-U''_\mu(x) = \pi^2\mu^2 U_\mu(x)$$

we see that when

$$\mu^2 \leq \lambda - \sup q$$

then zeros have at least the distance $1/\mu$ and if

$$\mu^2 \geq \lambda - \inf q$$

then zeros have at least the distance $1/\mu$. In particular, if u is an eigenfunction with $n-1$ zeroes, then the inequalities for λ are true.

We define

$$\Phi(\lambda) = U(1)$$

where

$$-U'' + qU = \lambda U.$$

Then Φ is continuous as a function of λ. The zeros of Φ are the eigenvalues. Let $N(\lambda)$ be the number of zeros of U in $(a, b]$. The map

$$\lambda \to N(\lambda)$$

is monontoically increasing and there exists a minimal λ_n so that

$$N(\lambda) < n \qquad \text{if } \lambda < \lambda_n$$

(we choose $\mu < 1$ in the comparison argument. If

$$\lambda < \mu^2 + \inf q$$

then zeroes have a distance larger than 1),

We consider U as a function of λ and x. It is continuous as a function of both variables. Now assume that $\Phi(\lambda) \neq 0$. Then there exists δ so that two zeros of eigenfunctions to an eigenvalue $\lambda' < |\lambda| + 1$ have distance at least δ. Let

$$A = [0,1] \backslash \bigcup (x_j - \delta/2, x_j + \delta/2)$$

where x_j are the zeroes of U. Then there exists ε so that for $|\lambda' - \lambda| < \varepsilon$ there is no zero in A. Checking the signs and using the intermediate value theorem we see that each of the intervals $(x_j - \varepsilon, x_j + \varepsilon)$ contains one zero. By the choice of δ there is at most one zero there, hence $N(\lambda)$ is constant near λ. The same argument shows that $N(\lambda)$ jumps by 1 if $\Phi(\lambda) = 0$.

Orthogonality is an exercise.

□

There are natural questions:

- Which of the concrete orthonormal systems constructed above are a basis? We will see that the answer is all of them, but we need more tools to prove this.
- Is there a good theory of not necessarely orthonormal basis? This is more tricky.

30.10.2019

3 Lebesgue spaces

3.1 Review of measure spaces

Reference:

1. Alt: Linear functional analysis, Springer.
2. Lieb and Loss: Analysis, AMS 2001.
3. Sharkarchi and Stein: Real Analysis: Measure theory, Integration and Hilbert spaces. Princeton University Press. 2009.

Theorem 3.1 (Banach-Tarski)**.** *There exists finitely many pairwise disjoints sets A_n, B_m of $\mathbb{R}^3$ and isometric maps $\phi_j, \psi_j \colon \mathbb{R}^3 \to \mathbb{R}^3$ so that*

$$B_1(0) = \bigcup_{n=1}^{N} \phi_n(B_n) = \bigcup_{m=1}^{M} \psi_m(A_m) = \bigcup_{n=1}^{N} B_n \cup \bigcup_{m=1}^{M} A_m$$

Remark: Makes use of the axiom of choice.

Definition 3.2. *Let X be a set. A family of subset $\mathcal{A}$ is called a σ algebra if*

1. $\{\} \in \mathcal{A}$
2. $A \in \mathcal{A}$ *implies* $X \backslash A \in \mathcal{A}$
3. $A_n \in \mathcal{A}$ *implies* $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

A map $\mu : \mathcal{A} \to [0, \infty]$ is called a measure if whenever $A_n \in \mathcal{A}$ are pairwise disjoint then

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n \Big) = \sum_{n=1}^{\infty} \mu(A_n).$$

The triple $(X, \mathcal{A}, \mu)$ is called a measure space.

Examples:

1. X a set, $\mathcal{A} = 2^X$ the set of all subsets, and $\mu(A)$ the number of elements.
2. If (X, d) is a metric space then there is a smallest σ algebra containing all open sets. It is called the Borel σ algebra of X.
3. In probability theory the σ algebra encodes the available information on a system.

4. $X = \mathbb{R}^n$, $\mathcal{A}$ the Borel sets, μ the Lebesgue measure restricted to the Borel sets.

5. $X = \mathbb{R}^n$, $\mathcal{A}$ the Lebesgue sets, μ the Lebesgue measure.

6. $X = \mathbb{R}^n$, $0 \le s \le n$, $\mathcal{H}^s$, $\mathcal{A}$ the Borel sets, $\mathcal{H}^s$ the Hausdorf measure.

Definition 3.3. *Let X be a set and $\mathcal{A}$ a σ algebra. A map $f : X \to \mathbb{R} \cup \{-\infty, \infty\}$ is called measurable if*

$$f^{-1}((t, \infty]) \in \mathcal{A}$$

for all $t \in \mathbb{R}$. If $(X, \mathcal{A}, \mu)$ is a measure space and $f : X \to [0, \infty]$ then we define the Lebesgue integral by the Riemann integral

$$\int f d\mu = \int_0^\infty \mu(f^{-1}((t, \infty]))dt \in [0, \infty]$$

We call a measurable function f integrable if $|f|$ is integrable. Let $1 \le p < \infty$. We call a measurable function f p integrable if $|f|^p$ is integrable and denote

$$\|f\|_{L^p} = \left(\int |f|^p d\mu \right)^{1/p}.$$

We call a measurable function ∞ integrable or essentially bounded if there is a constant C so that

$$\mu(\{x : |f(x)| > C\}) = 0.$$

The best constant is denoted by $\|f\|_{L^\infty}$.

There are convergence theorems about the relation between the limit of integrals, and the integral over limits: The *theorem of Lebesgue* on dominated convergence, the *Lemma of Fatou* and the *theorem of Beppo Levi on monoton convergence.*

Definition 3.4. *A measure space $(X, \mathcal{A}, \mu)$ is called sigma finite if there exists a sequence of measurable sets A_j of finite measure so that $X = \bigcup_{n=1}^{\infty} A_n$.*

3.2 Construction of measures

Measures are often constructed by first constructing *outer measures.*

Definition 3.5. *Let X be a set. An outer measure μ maps subsets of X to $[0, \infty]$ so that*

1. $\mu(\{\}) = 0$.

2. $A \subset B$ *implies* $\mu(A) \le \mu(B)$.

3. $\mu\Big(\bigcup\limits_{j=1}^{\infty} A_j\Big) \le \sum_{j=1}^{\infty} \mu(A_j)$.

Examples:

- Let $X = \mathbb{R}^d$. We define the measure of a coordinate rectangle as the product of the sidelengths and the measure of a countable disjoint union of coordinate rectangles as the sum over the measures of the rectangles. Finally we define the *outer* measure of a general set as the infimum of all measures of coverings by unions of coordinate rectancles.

- If $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ are σ finite measure spaces we define rectangles in the cartesian product as the cartesian product of measurable sets and their measure as the product of the measures. Then we proceed in the same way as above to obtain an outer measure on $X \times Y$.

- The Hausdorff measure: Let X be a metric space and $s \ge 0$. We define the premeasure of a set A of diameter r

$$\phi(A) = 2^{-s} \frac{\pi^{s/2}}{\Gamma(\frac{s}{2}+1)} r^s$$

and the Hausdorff measure

$$\mathcal{A} = \inf\Big\{\sum_{n=1}^{\infty} \phi(A_n) : A \subset \bigcup_{n=1}^{\infty} A_n\Big\}.$$

Definition 3.6. *Let X be a metric space. We call μ an outer metric measure if it is an outer measure which satisfies*

$$\mu(A \cup B) = \mu(A) + \mu(B)$$

for all $A, B \subset X$ with $\operatorname{dist}(A, B) > 0$.

Definition 3.7. *Let μ be an outer measure on X. We call a subset $A \subset X$ Caratheodory measurable if for all $B \subset X$*

$$\mu(B) = \mu(B \cap A) + \mu(B \cap (X \backslash A)).$$

Theorem 3.8 (Caratheodory)**.** *Let μ be an outer measure on the set X. Then the Caratheodory measurable sets $\mathcal{C}$ are a σ algebra and $(X, \mathcal{C}, \mu|_{\mathcal{C}})$ is a measure space. Moreover $\mathcal{C}$ contains all sets of exterior measure 0. If X is a metric space and μ is a metric outer measure than $\mathcal{C}$ contains all open sets. In the case of the Cartesian product $\mathcal{C}$ contains all Cartesian products of measurable sets.*

Theorem 3.9 (Fubini-Tonelli)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ-finite measure spaces, $\mathcal{A} \times \mathcal{B}$ the product σ algebra and $\mu \times \nu$ the product measure. Let f be $\mu \times \nu$ integrable. Then for almost of $x \in X$ $y \to f(x, y)$ is ν integrable, $x \to \int_Y f(x, y) d\nu(y)$ is μ integrable and*

$$\int_{X \times Y} f(x, y) d\mu \times \nu = \int_X \int_Y f(x, y) d\nu(y) d\mu(x).$$

3.3 Jensen's and Hölder's inequalities

Lemma 3.10. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is convex. Then both one sided derivatives exist and if $x < y$ then*

$$\frac{df}{dx}^{+}(x) \le \frac{df}{dy}^{-}(y) \le \frac{df}{dy}^{+}(y)$$

and for all z

$$f(z) \ge \max\{f(x) + \frac{df}{dx}^{+}(x)(z-x), f(x) + \frac{df}{dx}^{-}(x)(z-x)\}.$$

Proof. If $x_0 < x_1 < x_2$ then

$$\frac{f(x_1)-f(x_0)}{x_1-x_0} \le \frac{f(x_2)-f(x_0)}{x_2-x_0} \le \frac{f(x_2)-f(x_1)}{x_2-x_1}$$

and

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

is monotonically increasing. This implies the differentiability from the right, and similarly from the left and the relation between the derivatives. □

Lemma 3.11. *Let $(X, \mathcal{A}, \mu)$ be a measure space, $\mu(X) = 1$, $F : \mathbb{R} \to \mathbb{R}$ convex and f real valued and integrable. Then $F \circ f$ is measurable, $(F \circ f)_-$ is integrable and*

$$F \circ \int_X f d\mu \le \int_X F \circ f d\mu$$

Proof. Let $t_0 = \int_X f d\mu$. Since $F : \mathbb{R} \to \mathbb{R}$ is convex, we have for any t

$$F(t) \ge F(t_0) + \frac{dF}{dt}^{+}(t_0)(t-t_0).$$

Thus

$$\mu(\{F \circ f \le s\}) \le \mu(\{F(t_0) + \frac{dF}{dt}^{+}(t_0)(f-t_0) \le s\})$$

and $\min\{F \circ f, 0\}$ is integrable since $x \to F(t_0) + \frac{dF}{dt}^{+}(t_0)(f-t_0)$ (which is affine in f) is integrable. Then

$$\begin{aligned}\int_X F \circ f d\mu \ge & \int_X F(t_0) + \frac{dF}{dt}^{+}(t_0)(f-t_0) d\mu \\ =& F(t_0) + \frac{dF}{dt}^{+}(t_0)\left(\int_X f d\mu - t_0\right) = F(t_0).\end{aligned}$$

□

We call a function $f : X \to \mathbb{C}$ integrable if the real and imaginary parts are both integrable. We say a property holds almost everywhere, if it holds outside a set of measure 0.

Lemma 3.12. *Let $1 \le p, q \le \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. If $f \in L^p$ and $g \in L^q$ then fg is integrable and*

$$\left| \int fg d\mu \right| \le \|f\|_{L^p} \|g\|_{L^q}.$$

If $1 < p < \infty$, then equality implies that $g = \lambda |f|^{p-2} \bar{f}$ almost everywhere for some $\lambda \in \mathbb{K}$ with $|\lambda| = 1$.

Proof. We copy the proof basically from the one for the sequence space. As there it suffices to consider f and g with $\|f\|_{L^p} = 1$ and $\|g\|_{L^q} = 1$ and prove

$$\int |f||g| d\mu \le \int \frac{1}{p}|f|^p + \frac{1}{q}|g|^q d\mu = 1.$$

The inequality is strict unless

$$|fg(x)| = \frac{1}{p}|f(x)|^p + \frac{1}{q}|g(x)|^q$$

almost everywhere, which implies $|g| = |f|^{p-1}$. Now

$$\left| \int fg d\mu \right| \le \int |fg| d\mu \le \|f\|_{L^p} \|g\|_{L^q}$$

and in the case of equality all inequalities must be equalities. Hence $|g| = |f|^{p-1}$. Now suppose for some integrable function h

$$\left| \int h d\mu \right| = \int |h| d\mu.$$

Then there exits $\lambda \in \mathbb{C}$ with $|\lambda| = 1$ so that $\int h d\mu \in [0, \infty)$ and

$$\int \lambda^{-1} h d\mu = \int \operatorname{Re} \lambda^{-1} h d\mu = \int |h| d\mu$$

and hence

$$h = \lambda |h|$$

almost everywhere. Back to our situation above this implies

$$g = \lambda |f|^{p-2} \bar{f}$$

almost everywhere for some complex number λ of modulus 1. □

3.4 Minkowski's inequality

Theorem 3.13. *Let $1 \le p < \infty$ and let X and Y be spaces with σ finite measures μ and ν respectively. Let f be $\mu \times \nu$ measurable. Then*

$$\left(\int_X \left|\int_Y |f(x,y)| d\nu(y)\right|^p d\mu(x)\right)^{1/p} \le \int_Y \left(\int_X |f(x,y)|^p d\mu(x)\right)^{1/p} d\nu(y).$$

If $1 < p$, if the integrals above are finite, and if

$$\left(\int_X \left|\int_Y f(x,y) d\nu(y)\right|^p d\mu(x)\right)^{1/p} = \int_Y \left(\int_X |f(x,y)|^p d\mu(x)\right)^{1/p} d\nu(y).$$

then there exist a μ-measurable function α and a ν-measurable function β so that

$$f(x,y) = \alpha(x)\beta(y)$$

almost everywhere. A special case is the triangle inequality (which holds without assuming σ finiteness)

$$\|f+g\|_{L^p(\mu)} \le \|f\|_{L^p(\mu)} + \|g\|_{L^p(\mu)},$$

whenever f and g are p-integrable, with equality for $p > 1$ iff f and g are linearly dependent.

Proof. We assume first that f is nonnegative and omit the absolute value. We claim that

$$y \to \int_X f^p(x,y) d\mu(x) \qquad \text{and } H(x) := \int_Y f(x,y) d\nu(y)$$

are measurable functions. This follows from the Theorem of Fubini if f resp f^p are $\mu \times \nu$ integrable, and by an approximation argument in the general case. Then

$$\begin{aligned}
\int_X H^p(x) d\mu(x) &= \int_X \int_Y f(x,y) d\nu(y) H(x)^{p-1} d\mu(x) \\
&= \int_Y \int_X f(x,y) H^{p-1}(x) d\mu(x) d\nu(y) \\
&\le \int_Y \left(\int_X f^p(x,y) d\mu(x)\right)^{1/p} \left(\int H^p d\mu(x)\right)^{\frac{p-1}{p}} d\nu(y) \\
&= \int_Y \left(\int_X f^p(x,y) d\mu(x)\right)^{1/p} d\nu(y) \left(\int H^p d\mu(x)\right)^{\frac{p-1}{p}}
\end{aligned}$$

where we used Hölder's inequality with $q = \frac{p}{p-1}$.

We want to divide by the right hand. We can do that whenever the left hand side is neither 0 nor ∞, and we can achieve that in the same fashion as for sequences.

Now assume that $p > 1$, f is complex valued and integrable. Then, with

$$\int_X \left| \int_Y f(x,y) d\nu(y) \right|^p d\mu(x) \le \int_X \left(\int_Y |f(x,y)| d\nu(y) \right)^p d\mu(x)$$

and we continue as in the previous step, assuming equality. Then we have equality in the application of Hölder's inequality

$$|f(x,y)| = \alpha_0(x) \int_Y |f(x,y')| d\nu(y')$$

for almost all x and y. Since we must have also equality in the equality above we must have

$$f(x,y) = \alpha(x)\beta(y)$$

for some measurable function α and β.

For the last part we apply the first part with the counting measure on $Y = \{0,1\}$. The product measure is defined in the obvious fashion, even without assuming that μ is σ finite. If f is p integrable then by the definition of the integral

$$\mu(\{x : |f(x)| > t\}) \le t^{-p} \|f\|_{L^p}^p$$

Let

$$A = \bigcup_{j=1}^{\infty} \{x : |f(x)| + |g(x)| > \frac{1}{j}\}$$

which is a countable union of sets of finite measure. We replace X by A, take as σ algebra the sets in $\mathcal{A}$ which are contained in A, and μ restricted to this σ algebra as measure. This is σ additive. $\square$

Definition 3.14. *Let $(X, \mathcal{A}, \mu)$ as above, $1 \le p < \infty$. We define the space $L^p(\mu)$ as equivalence classes of measurable p integrable functions and equipp it with the norm $\|.\|_{L^p}$. If $p = 0$ we define $L^\infty(\mu)$ as the space of equivalence classes of measurable almost everywhere bounded functions equipped with*

$$\|f\|_{L^\infty} = \inf\{C : \text{ there exists a set } N \text{ of measure } 0 \text{ so that } |f(x)| \le C \text{ for } x \in X \backslash \{N\}\}.$$

The Minkoswki inequality implies the triangle inequality. Here $f \sim g$ if $f(x) = g(x)$ almost everywhere, i.e. if there exists a set of measure 0 such that $f(x) = g(x)$ for $x \in X \backslash N$.

06.11.2019

3.5 Hanner's inquality

There is an improvement of the triangle inequality.

Theorem 3.15 (Hanner's inequality)**.** *Let $(X, \mathcal{A}, \mu)$ be a measure space and f, g be p-integrable functions, $1 < p < \infty$. If $1 \le p \le 2$ then*

$$\|f+g\|_{L^p}^p + \|f-g\|_{L^p}^p \ge (\|f\|_{L^p} + \|g\|_{L^p})^p + \big|\|f\|_{L^p} - \|g\|_{L^p}\big|^p, \tag{3.1}$$

$$(\|f+g\|_{L^p}+\|f-g\|_{L^p})^p+\big|\|f+g\|_{L^p}-\|f-g\|_{L^p}\big|^p \le 2^p(\|f\|_{L^p}^p+\|g\|_{L^p}^p). \tag{3.2}$$

If $2 \le p < \infty$ all inequalities are reversed.

The inequalities reduce to the parallelogram identity if $p = 2$. Both are equivalent: The second is obtained from the first by replacing f by $f + g$ and g by $f - g$. It suffices to prove the first inequality.

Proof. We may assume that $\|g\|_{L^p} \le \|f\|_{L^p}$ (otherwise we exchange the two) and $\|f\|_{L^p} = 1$ (otherwise we multiply f and g by the inverse of the norm).

The first inequality follows from the following pointwise inequality: Let

$$\alpha(r) = (1+r)^{p-1} + (1-r)^{p-1}, \quad \beta(r) = [(1+r)^{p-1} - (1-r)^{p-1}]r^{1-p}.$$

We claim that

$$\alpha(r)|f|^p + \beta(r)|g|^p \le |f+g|^p + |f-g|^p \tag{3.3}$$

for $1 \le p \le 2$, $0 \le r \le 1$ and complex numbers f and g (and the reverse inequality for $2 \le p < \infty$). Indeed, (3.3) implies

$$\alpha(r)|f(x)|^p + \beta(r)|g(x)|^p \le |f(x)+g(x)|^p + |f(x)-g(x)|^p$$

and by integration

$$\alpha(r)\|f\|_{L^p}^p + \beta(r)\|g\|_{L^p}^p \le \|f+g\|_{L^p}^p + \|f-g\|_{L^p}^p.$$

We apply the inequality with $r = \|g|_{L^p}$ and recall that $\|f\|_{L^p} = 1$. The left hand side becomes

$$\begin{aligned}\Big[(\|f\|_{L^p} + \|g\|_{L^p})^{p-1} + (\|f\|_{L^p} - \|g\|_{L^p})^{p-1}\Big]\|f\|_{L^p}&\\ + \Big[(\|f\|_{L^p} + \|g\|_{L^p})^{p-1} - (\|f\|_{L^p} - \|g\|_{L^p})^{p-1})\Big]\|g\|_{L^p}&\\ =(\|f\|_{L^p} + \|g\|_{L^p})^p + (\|f\|_{L^p} - \|g\|_{L^p})^p.&\end{aligned}$$

It remains to prove (3.3).

Let for $0 \le R \le 1$,

$$F_R(r) = \alpha(r) + \beta(r)R^p.$$

We claim that it attains its maximum at $r = R$ if $1 \le p < 2$ and resp. its minimum at $r = R$ if $p > 2$. We compute

$$F_R' = \alpha' + \beta' R = (p-1)[(1+r)^{p-2} - (1-r)^{p-2}](1 - (R/r)^p)$$

and the derivative vanishes only at $r = R$ and changes sign there. Thus

$$\alpha(r) + \beta(r)R^p \le (1 + R)^p + (1 - R)^p$$

if $0 \le R \le 1$ and $p \le 2$ with the opposite inequality if $p \ge 2$. Now let $R \ge 1$. Since $\beta \le \alpha$ if $p \le 2$ we obtain

$$\alpha(r)+\beta(r)R^p \le R^p\alpha(r)+\beta(r) \le R^p[(1+R^{-1})^p+(1-R^{-1})^p] = (1+R)^p+(R-1)^p$$

and the reverse inequality if $p > 2$. This implies (3.3) for real f and g. We claim that (3.3) holds for complex f and g. It suffices to consider $f = a > 0$ and $g = be^{i\theta}$. Since

$$(a^2 + b^2 + 2ab\cos\theta)^{p/2} + (a^2 + b^2 - 2ab\cos\theta)^{p/2}$$

has its minimum at $\theta = 0$ (resp. its maximum if $p \ge 2$) since $x \to x^{p/2}$ is concave if $p \le 2$ (convex if $p \ge 2$). □

3.6 The Lebesgue spaces $L^p(\mu)$

Lemma 3.16. *The set of p-integrable functions is a vector space. The Minkowski inequality*

$$\|f + g\|_{L^p} \le \|f\|_{L^p} + \|g\|_{L^p}$$

holds. Moreover

$$\|\lambda f\|_{L^p} = |\lambda|\|f\|_{L^p}$$

and

$$\|f\|_{L^p} = 0$$

if and only if f vanishes outside a set of zero

$$\mu(\{f \ne 0\}) = 0.$$

These functions are p integrable for all p. They are a subvector space.

Proof. The vector space property follows from the Minkowski inequality. The other statements are obvious. □

Definition 3.17. *We call two measurable functions equivalent $f \sim g$ if $\mu(\{f \ne g\}) = 0$. We define $L^p(\mu)$ as the space of equivalence classes of p integrable functions.*

If $f \sim g$ then $\|f - g\|_{L^p} = 0$. The equivalence relation is compatible with the vector space structure.

Theorem 3.18 (Fischer-Riesz)**.** *The space $L^p(\mu)$ is Banach space.*

Proof. It is straight forward to verify that $L^p(\mu)$ is a vector space (using Minkowksi's inequality), and that $\|.\|_{L^p}$ is a norm. Completeness is more involved. We only consider $p < \infty$ and leave the case $p = \infty$ to the reader.

Let f_n be representatives of a Cauchy sequence. By taking subsequences if necessary we may assume

$$\|f_n - f_m\|_{L^p} \le 2^{-\min\{m,n\}}.$$

We define the monotone sequence of functions

$$F_n(x) = |f_1(x)| + \sum_{m=1}^{n-1} |f_{m+1}(x) - f_m(x)|$$

and $F = \lim_{n\to\infty} F_n(x)$. F is measurable and by monotone convergence

$$\int |F|^p d\mu = \lim_{n\to\infty} \int |F_n|^p d\mu \le \|f_1\|_{L^p}^p + 1$$

and in particular it is finite almost everywhere. Thus

$$f_n = f_1 + \sum_{m=1}^{n-1} (f_{m+1} - f_m)$$

converges if $F(x) < \infty$. Let f be the limit if $F(x) < \infty$, and 0 otherwise. It is measurable. Since $\max\{f, f_n\} \le F$ we obtain by dominated convergence

$$\|f - f_n\|_{L^p}^p = \int |f - f_n|^p d\mu \to 0.$$

□

3.7 Projections and the dual of $L^p(\mu)$

Lemma 3.19. *Let $1 < p < \infty$ and let K be a closed convex set in $L^p(\mu)$. Let $f \in L^p(\mu)$. Then there exists a unique $g \in K$ with*

$$\|f - g\|_{L^p(\mu)} = \operatorname{dist}(f, K).$$

Moreover

$$\operatorname{Re} \int_X (h - g)(\bar f - \bar g)|f - g|^{p-2} d\mu \le 0, \quad \forall h \in K.$$

Proof. Let h_n be a minizing sequence. Since $\frac{1}{2}(h_n + h_m) \in K$ and $\|h_n - f + h_m - f\|_{L^p} \le \|h_n - f\|_{L^p} + \|h_m - f\|_{L^p}$, we see that

$$\|h_n - f + h_m - f\|_{L^p} \to 2\operatorname{dist}(f, K).$$

Now let $p \le 2$, from the second Hanner's inequality we obtain

$$\begin{aligned}\big(\|h_n - f + h_m - f\|_{L^p} + \|h_n - h_m\|_{L^p}\big)^p \\ + \big|\|h_n - f + h_m - f\|_{L^p} - \|h_n - h_m\|_{L^p}\big|^p \\ \le 2^p\big(\|h_n - f\|^p_{L^p} + \|h_m - f\|^p_{L^p}\big).\end{aligned}$$

Let $A = \limsup_{n,m\to\infty} \|h_n - h_m\|_{L^p}$. This limsup is obtained along two subsequences $n, m \to \infty$. Let $D = \operatorname{dist}(f, K)$. Then

$$(2D + A)^p + |2D - A|^p \le 2^{p+1} D^p$$

which implies $A = 0$ by the strict convexity of $A \to |2D + A|^p$.

If $p > 2$ we argue similarly with the first inequality.

Now let $g \in K$ be the point of minimal distance and let $h \in K$. Let

$$N(t) = \int |f - (g + t(h - g))|^p d\mu.$$

Then $N(t)$ attains its minimum at $t = 0$ on the interval $[0, 1]$. We claim that its derivative at $t = 0$ is

$$\frac{d}{dt} N|_{t=0} = p \operatorname{Re} \int |f(x) - g(x)|^{p-2}(f(x) - g(x))(\bar g(x) - \bar h(x)) d\mu.$$

This implies the assertion.

To calculate the derivative we assume that $f, g \in L^p(\mu)$ and define

$$N(t) = \|f + tg\|^p_{L^p}.$$

Since almost everywhere

$$\frac{d}{dt}|_{t=0}|f + tg|^p = p|f|^{p-2} \operatorname{Re} f\bar g$$

and since the pth power is convex also $t \to |f + tg|^p$ is convex and

$$|f|^p - |f - g|^p \le \frac{1}{t}(|f + tg|^p - |f|^p) \le |f + g|^p - |f|^p,$$

by convexity the formula follows by dominated convergence. □

Theorem 3.20. *Let $(X, \mathcal{A}, \mu)$ be a measure space, $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$j : L^q \ni g \to (f \to \int fg d\mu) \in (L^p(\mu))^*$$

is a linear isometric isomorphism.

08.11.2019

The proof is the same as for Hilbert spaces: By Hölder's inequality the map is well defined and

$$\|j(f)\|_{(L^p)^*} \le \|f\|_{L^q}.$$

Since

$$j(f)(|f|^{q-2}\bar f) = \int |f|^q d\mu$$

we conclude as for Hilbert spaces that

$$\|j(f)\|_{(L^p)^*} \ge \|f\|_{L^q}.$$

Surjectivity is proven exactly as for Hilbert spaces.

Corollary 3.21. *Suppose that μ is σ finite. Then*

$$L^\infty \ni g \to (f \to \int fg d\mu) \in (L^1(\mu))^*$$

is an isometric isomorphism.

The proof is an exercise on sheet 5.

3.8 Borel and Radon measures

Let (X,d) be a metric space. We recall that the Borel sets $\mathcal{B}(X)$ are the smallest σ algebra containing all open sets.

Definition 3.22. *Let (X,d) be a metric space. A Borel measure is a measure on the Borel sets. A Radon measure is a Borel measure, such that for every $x \in X$ there exists an open environment $U \ni x$ so that $\mu(U) < \infty$ and such that for every Borel set A*

$$\mu(A) = \sup\{\mu(K) : K \subset A, K \text{ compact }\}.$$

This is called inner regularity.

Definition 3.23. *We call a measure complete, if the σ algebra contains every subset of a set of measure zero.*

The theorem of Fubini in the form stated above holds for $\mu \times \nu$ with the smallest σ algebra containing all cartesian products of measurable sets. The Lebesgue measure restricted to the Borel sets is not complete. We can easily complete σ algebras.

Lemma 3.24. *Let μ be a Radon measure. Then the measure of compact sets is finite and for $\varepsilon > 0$ and K compact there exists an open set $U \supset K$ of finite measure with $\mu(U) \le \mu(K) + \varepsilon$. This is called outer regularity.*

Proof. Let K be compact. For every $x \in K$ exists an open set U_x containing x with $\mu(U_x) < \infty$. Since K is compact and

$$K \subset \bigcup U_x$$

there exists a finite subcovering

$$K \subset \bigcup_{j=1}^{N} U_{x_j} =: U$$

and

$$\mu(K) \le \mu(U) \le \sum_{j=1}^{N} \mu(U_{x_j}).$$

We define $U_j = U \cap \{x : d(x,K) < \frac{1}{j}\}$. By the theorem of Lebesgue

$$\mu(U_j) \to \mu(K).$$

□

Definition 3.25. *We call X locally compact if for every point x there is a neighborhood whose closure is compact. We call (X,d) σ compact if it is locally compact and if it is a countable union of compact sets.*

Lemma 3.26. *Let μ be a Borel measure on a σ compact space (X,d) and let B be a Borel set with $\mu(B) < \infty$ and $\varepsilon > 0$. Then there exists a closed set $C \subset B$ with $\mu(B \backslash C) < \varepsilon$. If μ is in addition Radon then there exists an open set U containing B with $\mu(U \backslash B) < \varepsilon$.*

Proof. For the first part we may assume $\mu(X \backslash B) = 0$ - otherwise we define $\nu(A) = \mu(A \cap B)$. we define

$$\mathcal{F} = \left\{ \begin{array}{ll} A \subset \mathbb{R}^d : & A \text{ is Borel and for every } \varepsilon > 0 \text{ there exists a closed set } C \\ & \text{with } \mu(A \backslash C) < \varepsilon. \end{array} \right\}$$

It contains all closed sets. We claim:

1. If $A_j \in \mathcal{F}$ then $\bigcap A_j \in \mathcal{F}$.
2. If $A_j \in \mathcal{F}$ then $\bigcup A_j \in \mathcal{F}$.
3. Since open sets are countable unions of closed sets every open set is in $\mathcal{F}$.

We define

$$\mathcal{G} = \{A : X \backslash A, A \in \mathcal{F}\}$$

Then $\mathcal{G}$ contains complements of elements and countable unions of elements of $\mathcal{G}$. Hence it is a σ algebra containing all open sets, and thus it is the Borel σ algebra. This implies the first claim.

Let now μ be Radon and K_j as above. Then $K_j \backslash B$ is Borel with $\mu(\mathring{K}_j \backslash B) < \infty$. Then there exists a closed set $C_j \subset \mathring{K}_j \backslash B$ with $\mu((\mathring{K}_j \backslash C_j) \backslash B) < \varepsilon 2^{-j}$. Let

$$U = \bigcup_{j=1}^{\infty} (\mathring{K}_j \backslash C_j).$$

It is open and

$$B = \bigcup_{j=1}^{\infty} (\mathring{K}_j \cap B) \subset \bigcup \mathring{K}_j \backslash C_j = U$$

Moreover

$$\mu(U \backslash B) = \mu(\bigcup (\mathring{K}_j \backslash C_j) \backslash B) < \varepsilon.$$

□

Lemma 3.27. *Let (X, d) be σ compact, assume that every compact set has an open neighborhood whose closure is compact, and let μ be a Borel measure such that any compact set is of finite measure. Then μ is Radon and it is outer regular.*

Proof. Only inner regularity has to be proven since outer regularity follows then by Lemma 3.26. Let A be Borel with finite measure (why does this suffice?). By Lemma 3.26 there exists a closed set $C \subset A$ such that $\mu(A \backslash C) < \varepsilon$. Let K_j be compact subsets with $X = \bigcup K_j$ and K_j contained in the interior of K_{j+1}. Then

$$\mu(C \cap K_j) \to \mu(C)$$

and $C \cap K_j$ is compact. □

The most important example is the Lebesgue measure. A Radon measure on a compact metric space is finite. If (X, d) is a countable union of compact sets and μ is a Radon measure then μ is σ finite.

The counting measure on $\mathbb{R}$ is not a Radon measure.

Remark 3.28. *Continuous functions on compact metric spaces are integrable with respect to Radon measures.*

Lemma 3.29. *Let (X, d) be a σ compact metric space and assume that every compact set has an open neighborhood whose closure is compact. Let μ be a Radon measure on X and $1 \leq p < \infty$. Then continuous functions with compact support are dense in $L^p(\mu)$.*

Proof. Let f be integrable. We decompose it into real and the imaginary part and it suffices to prove the assertion for real functions. Similar we decompose a real valued function into positive and negative part, and it suffices to approximate a nonnegative integrable function f.

Since

$$\int f d\mu = \int_0^\infty \mu(\{f > t\}) dt$$

given $\varepsilon > 0$ there exists $0 = t_0 < t_1 < \dots t_j < t_{j+1} < t_N < \infty$ so that (with $t_0 = 0$)

$$0 < \int_0^\infty \mu(\{f > t\}) dt - \sum_{j=1}^{N} (t_j - t_{j-1}) \mu(\{f > t_j\}) < \varepsilon$$

Let

$$A_j = \{x : f(t) > t_j\}.$$

Then

$$\|f - \sum (t_j - t_{j-1}) \chi_{A_j}\|_{L^1} < \varepsilon$$

and it suffices to approximate a characteristic function of a measurable set A of finite measure by a continuous function. Let $\varepsilon > 0$. By inner and outer regularity there exists a compact set K and an open set U so that

$$K \subset A \subset U \qquad \mu(U) \le \mu(K) + \varepsilon.$$

Then $d(K, X \backslash U) := d_0 > 0$ and we define

$$f_L(x) = \max\{1 - L d(x, K), 0\} \in C(X)$$

Then if $d_0 L \ge 1$

$$\|f_L - \chi_A\|_{L^p} < \varepsilon^{\frac{1}{p}}.$$

If L is sufficiently large then $\operatorname{supp} f$ is compact. Thus continuous functions with compact support are dense. □

3.9 Compact sets

Lemma 3.30. *If (X, d) is σ compact and μ is Radon measure, then Lipschitz continuous functions with compact support are dense in $L^p(\mu)$ for $1 \le p < \infty$.*

Proof. We prove that for every $\varepsilon > 0$ and $f \in C(X)$ with compact support, there exists f_ε Lipschitz continuous with

$$\operatorname{supp} f_\varepsilon \subset \operatorname{supp} f$$

and

$$\sup |f_\varepsilon - f| < \varepsilon.$$

It suffices to do this for $f \geq 0$. Since $\operatorname{supp} f$ is compact it is uniformly continuous: There exists $\delta > 0$ so that $|f(x) - f(y)| < \varepsilon$ if $d(x,y) < \delta$. With

$$L = \sup \left\{ \frac{|f(x) - f(y)|}{d(x,y)} : d(x,y) \geq \delta \right\}$$

which is finite since it is the supremum of a continuous function on a compact set, we obtain the inequality

$$|f(x) - f(y)| \leq \varepsilon + Ld(x,y), \quad \forall x, y \in X.$$

We define

$$g(x) = \min_y \{f(y) + 2Ld(x,y)\}.$$

One easily checks that g has Lipschitz constant $2L$, and the mimimum is attained in $B_\delta(x)$ and

$$\max\{0, f(x) - \varepsilon\} \leq g(x) \leq f(x).$$

□

13.11.2019

Theorem 3.31 (Heine-Borel). *Let X be a normed space. The close unit ball is compact if and only if X is finite dimensional.*

Proof. Let X be finite dimensional and let (x_n) be a basis. Then

$$\mathbb{K}^N \ni (a) \to \sum a_n x_n$$

is an invertible linear map. It is easy to see that it together with its inverse is continuous. The preimage of the closed unit ball is a closed bounded set in $\mathbb{R}^N$, hence compact.

Now let the dimension of X be infinite. We construct a sequence $(x_n)_{n\in\mathbb{N}}$ if unit vectors with distance at least $\frac{1}{2}$. It has no convergent subsequence, and hence the closed unit ball is not compact.

Suppose we have found $x_1 \dots x_N$. Let X_N be the subspace spanned by these vectors. We claim that there exists x_{N+1} of length 1 with $\operatorname{dist}(x_{N+1}, X_N) \geq \frac{1}{2}$. Then we construct the sequence recursively.

Let $x \notin X_N$ with $\operatorname{dist}(x, X_N) = 1$. Then there exists $y \in X_N$ by definition so that $1 \leq \|x - y\| \leq \frac{5}{4}$. We define

$$x_N = \frac{x - y}{\|x - y\|}.$$

□

Theorem 3.32 (Arzela-Ascoli). *Let (X,d) be a compact metric space. Then a closed set $A \subset C_b(X)$ is compact if and only if*

1. *A is bounded.*

2. *A is equicontinuous, i.e. for $\varepsilon > 0$ there exists $\delta > 0$ so that*

$$|f(x) - f(y)| < \varepsilon \qquad \text{if } f \in A \text{ and } d(x,y) < \delta.$$

Proof. Let A be compact. Since

$$C_b(X) \ni f \to \|f\|_{C_b(X)}$$

is continuous and hence attains its maximum in A, we deduce that A is bounded. Let $\varepsilon > 0$. For every f there exists $\delta_f > 0$ and an open neighborhood $U_f \subset C_b(X)$ so that

$$|g(x) - g(y)| < \varepsilon \qquad \text{if } g \in U_f \text{ and } d(x,y) < \delta_f.$$

Then $A \subset \bigcup_{f \in A} U_f$, and since A is compact there is a finite subcovering, $A \subset \bigcup_{j=1}^N U_{f_j}$. We define $\delta = \min \delta_{f_j}$.

Vice versa, assume that A is closed, bounded and equicontinuous. Let $f_j \in A$ be a sequence. Given $\varepsilon > 0$ we claim that there exists g such that $B_{3\varepsilon}(g) \subset C_b(X)$ contains infinitely many f_j. By a recursive argument this gives a convergent subsequence, and hence comapctness. Let $\varepsilon > 0$ and $\delta > 0$ as in the second condition. Then there exist a finite number N of points x_k so that $B_{\delta/2}(x_k)$ cover X since X is compact. There exists a subsequence so that $f_{j_l}(x_k)$ converges for all x_k. In particular, after relabeling, there are infinitely many $\{f_{j_l}\}_{l \in \mathbb{N}}$ so that $|f_{j_l}(x_k) - f_{j_m}(x_k)| < \varepsilon$. Then

$$f_{j_l} \subset B_{3\varepsilon}(f_{j_1}).$$

□

Lemma 3.33. *Let (X,d) be a compact set. Then it is separable.*

Proof. Given ε there exists a finite number of points x_n^ε, $1 \le n \le N(\varepsilon)$ so that the union of the balls $B_\varepsilon(x_n^\varepsilon)$ cover X. Take a sequence $\varepsilon = 2^{-j}$ we obtain the countably many points $(x_n^{2^{-j}})_{n,j}$ which are dense. □

Corollary 3.34. *Let (X,d) be compact. Then $C_b(X)$ is separable.*

Proof. By the proof of Lemma 3.30 the Lipschitz continuous functions are dense. The countable union of separable sets is separable and its closure is separable. Hence it suffices to prove that

$$K = \{f \in C_b(X) : \|f(x)\|_{C_b(X)} \le n, |f(x) - f(y)| \le n d(x,y)\}$$

is separable. This set is compact by Theorem 3.32 and hence separable. □

Corollary 3.35. *Let (X,d) be σ compact and μ a Radon measure. If $1 \le p < \infty$ then $L^p(\mu)$ is separable.*

Proof. Since Lipschitz continuous functions with compact support are dense we argue as for $C_b(X)$ if X is compact. If X is σ compact there is a sequence of compact sets $(K_n)_n$ with $\bigcup K_n = X$, We may assume that $K_n \subset K_{n+1}$. If $f \in L^2(\mu)$ then $\chi_{K_n} f \to f$ in $L^p(\mu)$ and the claim follows. □

Corollary 3.36. *Suppose that $1 \le p < \infty$, $f \in L^p(\mathbb{R}^d)$, $\varepsilon > 0$. Then there exist $\delta > 0$ and $R > 0$ so that for all $|h| < \delta$*

$$\|f(.+h) - f(.)\|_{L^p} < \varepsilon, \quad \|\chi_{\mathbb{R}^d \setminus B_R(0)} f\|_{L^p} < \varepsilon.$$

Proof. The second claim is a consequence of monotone convergence. For the first we approximate f by a Lipschitz continuous function g with compact support and Lipschitz constant $\|g\|_{Lip}$, , $\|g - f\|_{L^p} < \varepsilon/4$ and estimate

$$\begin{aligned}\|f(.+h) - f\|_{L^p} \le& \|f(.+h) - g(.+h)\|_{L^p} + \|f - g\|_{L^p} + \|g(\cdot + h) - g\|_{L^p} \\ \le& \varepsilon/4 + \varepsilon/4 + |h| \|g\|_{Lip} \Big(2m^d(\operatorname{supp} g) \Big)^{1/p} \\ \le& \varepsilon\end{aligned}$$

by choosing $|h| \le r$ for some small r. □

We want to characterize compact subsets of L^p spaces.

Theorem 3.37 (Kolmogorov)**.** *Let $1 \le p < \infty$. A closed subset $C \subset L^p(\mathbb{R}^n)$ is compact iff*

1. *C is bounded.*
2. *For every $\varepsilon > 0$ there exists δ so that for all $|h| < \delta$ and all $f \in C$*

$$\|f(.+h) - f\|_{L^p(\mathbb{R}^d)} < \varepsilon$$

3. *For every $\varepsilon > 0$ there exists R so that for all $f \in C$*

$$\|\chi_{\mathbb{R}^d \setminus B_R(0)} f\|_{L^p} < \varepsilon.$$

Proof. Let C be compact. Since $f \to \|f\|_{L^p}$ is continuous it attains its maximum and hence C is bounded. Suppose there exists $\varepsilon > 0$ and $h_j \to 0$ and $f_j \in C$ so that

$$\|f_j(.+h_j) - f_j\|_{L^p} \ge \varepsilon.$$

Since C is compact we may assume that f_j is a Cauchy sequence with limit f. Then there exists $\delta > 0$ so that

$$\|f(.+h) - f\|_{L^p} < \varepsilon/2$$

for $|h| < \delta$. This contradicts the previous inequality. Similarly we deduce the third part.

Vice versa: Suppose that $C \subset L^p(\mathbb{R}^d)$ is closed, bounded, and satisfies the three claims. We choose a smooth function η supported in the unit ball with values between 0 and 1 and $\int \eta = 1$, define $\eta_r(x) = r^{-d}\eta(x/r)$ and we fix $\varepsilon > 0$. Then there exists δ so that by Minkowski's inequality and the second assumption

$$\|f_r - f\|_{L^p} \le \sup_{|h|\le r} \|f(.+h) - f\|_{L^p} < \varepsilon/4, \quad f_r = \eta_r * f,$$

for all $f \in C$ and $r \le \delta$. Moreover f_r is Lipschitz continuous with Lipschitz constant depending on δ. We choose R large so that

$$\|f - \chi_{B_R(0)} f\|_{L^p} < \varepsilon/4$$

for all $f \in C$. Then also

$$\|f_r - \chi_{B_R(0)} f_r\|_{L^p} < \varepsilon/2.$$

By Theorem 3.32 the set

$$\{f_r|_{\overline{B_R(0)}} : f \in C\}$$

is compact in $C_b(\overline{B_R(0)})$, and hence in $L^p(B_R(0))$, and we can cover it by a finite number of balls of radius $\varepsilon/4$. But then the balls with radius ε cover C. Thus C is precompact, and compact since it is closed. □

15.11.2019

3.10 The Riesz representation theorem for $C_b(X)$

Definition 3.38. *Let (X,d) be a σ compact metric space and $K_j \subset X$ compact with K_j in the interior of K_{j+1} and $X = \bigcup K_j$. We denote by $C_0(X) \subset C_b(X)$ the continuous functions f with limit 0 at ∞, i.e. for all $\varepsilon > 0$ there exists j so that f is at most of size ε outside K_j. We define $C_c(X)$ as the subspace of continuous functions with compact support.*

Let K be compact. Under these assumptions there exists n so that $K \subset K_n$. Indeed, since $X = \bigcup_n \operatorname{int} K_n$ also $K \subset \bigcup_n \operatorname{int} K_n$. Since K is compact there exists a finite subcover.

Thus $C_c(X) \subset C_0(X)$ and it is not hard to see that the latter is the closure of the first.

Definition 3.39. *Let (X,d) be a metric space. We call $L \in (C_b(X))^*$ nonnegative if*

$$L(f) \ge 0 \qquad \textit{whenever } f \ge 0.$$

Theorem 3.40. *Let (X,d) and K_n be a above and let $L : C_0(X) \to \mathbb{K}$ be linear and assume that it satisfies*

$$|L(f)| \leq C_K \|f\|_{C_0(X)}$$

for $f \in C_0(X)$ Then there exists a Radon measure μ and a measureable function $\sigma : X \to \{\pm 1\}$ so that

$$L(f) = \int_X f \sigma d\mu.$$

Definition 3.41. *Let L be as above. We define the variation measure of L by*

$$\mu^*(U) = \sup\{|L(f)| : f \in C_c(X), \operatorname{supp} f \subset U, |f| \leq 1\}$$

for open sets U and for general sets

$$\mu^*(A) = \inf\{\mu(U) : A \subset U, U \text{ open }\}.$$

Proof. We prove the theorem by several steps as follows:

1. μ^* is an outer metric measure, which defines a Radon measure on the Borel sets.
2. For $f \in C_0(X)$ nonnegative we define

   $$\lambda(f) = \sup\{|L(g)| : |g| \leq f\}$$

   and prove for f nonnegative

   $$\lambda(f) = \int f d\mu$$

3. As a consequence

   $$|L(f)| \leq \int |f| d\mu$$

   and we can extend L to $L^1(\mu)$, hence $L \in (L^1(\mu))^*$ and there exists $\sigma \in L^\infty(\mu)$ so that

   $$L(f) = \int f \sigma d\mu$$

   for all $f \in C_c(X)$ with $\|\sigma\|_{L^\infty(\mu)} \leq 1$.
4. We complete the proof by $|\sigma(x)| = 1$ for almost all x. Since we may change on a set of μ measure 0, we obtain $|\sigma| = 1$.

Step 1: We claim that μ^* is an outer metric measure. Clearly

$$\mu^*(X) = \|L\|_{C_0^*(X)}, \qquad \mu^*(\{\}) = 0$$

and $A \subset B$ implies $\mu^*(A) \le \mu^*(B)$.

To show that it is an outer measure, we prove first σ subadditivity for open sets. let U_j be open sets, $U = \cup U_j$ and we have to show that

$$\mu^*(U) \le \sum \mu^*(U_j).$$

Let $0 \le f \le 1$ with $\operatorname{supp} f \subset U$. We have to show that

$$L(f) \le \sum \mu^*(U_j).$$

Let $K = \operatorname{supp} f$ which is compact. Thus K is covered by finitely many $\cup_{j=1}^N U_j$ for some $N < \infty$. Moreover we may assume that the U_j's are contained in a fixed compact set, or even replacing X by this compact set, that X is compact. We claim that there exist g_j, $0 \le g_j \le 1$, $\operatorname{supp} g_j \subset U_j$ and $\sum g_j = 1$ on K. We define $f_j = g_j f$ for $1 \le j \le N$. Then

$$L(f) = \sum_j L(f_j) \le \sum_{j=1}^N \mu^*(U_j).$$

To see the existence of the g_j, take $U_0 = X \backslash K$ (and recall that we may assume that X is compact). Then $X = \cup_{j=0}^N U_j$ and we take a subordinate partition of unity, i.e. functions $\eta_j \in C_b(X)$ with $0 \le \eta_j$ and $\operatorname{supp} \eta_j \in U_j$ so that

$$1 = \sum_{j=0}^N \eta_j.$$

The functions $g_j = \eta_j$ for $1 \le j \le N$ have this property. More precisely, let $A_0 = X \backslash \bigcup_{j=1}^N U_j$. It is compact and satisfies $A_0 \subset U_0$. There exists $\tilde\eta_0 \in C_b(X)$ supported in U_0, identically 1 on A_0. Let $A_1 = X \backslash (\{x : \tilde\eta_0(x) > \frac12\} \cup \bigcup_{j=2}^N U_j) \subset U_1$ and we repeat the contruction. Recursively we obtain $\tilde\eta_j$ with

$$\rho = \sum_{j=0}^N \tilde\eta_j \ge \frac12$$

in X. We define

$$\eta_j = \frac{\tilde\eta_j}{\rho}.$$

A standard argument in measure theory implies subadditivity for all sets.

Finally, if A, B are Borel sets with positive distance there exist disjoint open sets V and W containing A resp. B. Then

$$\mu^*(A \cup B) = \inf \mu^*(U) = \inf \mu^*(U \cap V) + \mu^*(U \cap W) = \mu^*(A) + \mu^*(B).$$

and hence μ^* is a metric outer measure.

Let μ be the measure defined by the Caratheodory construction, hence by restricting μ^* to measurable sets. Since μ^* is a metric outer measure all Borel sets are measurable, and we consider μ as restricted to Borel sets. In particular open sets are Borel sets and

$$\mu(U) = \mu^*(U)$$

for open sets. By construction μ is bounded hence bounded on compact sets and thus its restriction to Borel sets is a Radon measure.

Step 2: Let $f \in C_c(X)$ be non negative. We define

$$\lambda(f) = \sup\{|L(g)| : g \in C_c(X), |g| \le f\}.$$

Clearly $0 \le f_1 \le f_2$ implies $\lambda(f_1) \le \lambda(f_2)$ and for $c > 0$, $\lambda(cf) = c\lambda(f)$. We claim that

$$\lambda(f_1 + f_2) = \lambda(f_1) + \lambda(f_2)$$

for $f_1, f_2 \in C_c(X)$ nonnegative. Indeed, if $|g_1| \le f_1$ and $|g_2| \le f_2$ then $|g_1 + g_2| \le f_1 + f_2$, and, if in addition $L(g_1), L(g_2) \in [0, \infty)$,

$$|L(g_1) + L(g_2)| \le \lambda(f_1 + f_2).$$

This gives

$$\lambda(f_1) + \lambda(f_2) \le \lambda(f_1 + f_2).$$

Now let $|g| \le f_1 + f_2$. We define

$$g_1 = \begin{cases} \frac{f_1 g}{f_1 + f_2} & \text{if } f_1 + f_2 > 0 \\ 0 & \text{otherwise} \end{cases}$$

and similarly g_2. Then $|g_i| \le f_i$ and hence

$$|L(g)| \le \lambda(f_1) + \lambda(f_2)$$

which gives

$$\lambda(f) = \lambda(f_1) + \lambda(f_2).$$

We claim that

$$\lambda(f) = \int f d\mu.$$

It suffices to consider $0 \le f \le 1$. We approximate f by step function so that

$$\|f - \frac{1}{N} \sum_{j=1}^{N-1} \chi_{U_j}\|_{sup} < \frac{1}{N} \tag{3.4}$$

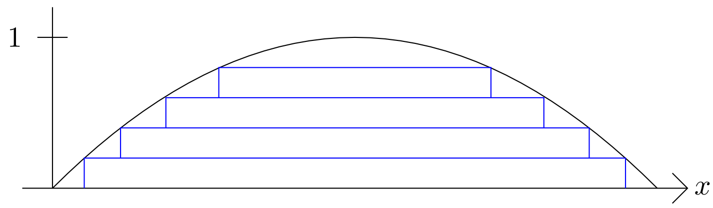


Figure 1: Formula (3.4)

with

$$U_j = \{x : f(x) > \frac{j}{N}\}.$$

By continuity $U_{j+1} \subset U_j$. We approximate the characteristic function by continuous functions so that $\operatorname{supp} \eta_j \subset U_{j-1}$, $\eta_j = 1$ on U_j and $\mu(\operatorname{supp} \eta_j \backslash U_j) < 1/j$. We define

$$g = \frac{1}{N} \sum_{j=2}^{N} \eta_j$$

so that

$$0 \le g \le f \le g + \frac{2}{N}.$$

and $\operatorname{supp} g$ is compact.

Thus

$$\lambda(g) \le \lambda(f) \le \lambda(g) + \frac{2}{N} \|L\|_{C_0^*}$$

and

$$\int g d\mu \le \int |f| d\mu \le \int g d\mu + \frac{2}{N} \|L\|_{C_0^*}.$$

By definition

$$\mu(U_j) \le \lambda(\eta_j) \le \mu(U_{j-1})$$

and hence

$$\frac{1}{N} \sum_{j=2}^{N} \mu(U_j) \le \lambda(g) \le \frac{1}{N} \sum_{j=1}^{N-1} \mu(U_j)$$

and

$$\frac{1}{N} \sum_{j=1}^{N-1} \mu(U_j) - \frac{1}{N} \sum_{j=2}^{N} \mu(U_j) = \frac{1}{N} \mu(U_1) \le \frac{1}{N} \|L\|_{C_0^*}.$$

We obtain

$$\left| \lambda(f) - \int_X |f| d\mu \right| \le \frac{5}{N} \|L\|_{C_0^*}$$

and the claim follows.

Step 3: Now

$$|L(f)| \leq \lambda(|f|) = \int |f| d\mu.$$

We extend L to an element in $(L^1(\mu))^*$, which is represented by an infinite integrable function σ by Corollary 3.21 . Moreover

$$\|\sigma\|_{L^\infty(\mu)} \leq \|L\|_{(L^1(\mu))^*} = 1.$$

Step 4: We claim that $|\sigma| = 1$ almost everywhere. By definition

$$\mu(U) = \sup\{\int f\sigma d\mu = L(f) : f \in C_c(X), |f| \leq 1, \operatorname{supp} f \subset U\}$$

We choose a sequence of functions with

$$\int f_j \sigma d\mu = L(f_j) \to \mu(U).$$

Since $\int f_j \sigma d\mu \leq \int |\sigma| d\mu$ and $|\sigma| \leq 1$ we deduce $|\sigma| = 1$ almost everywhere. □

20.11.2019

3.11 Covering lemmas and Radon measures on $\mathbb{R}^d$

The space $\mathbb{R}^d$ is σ compact and the Lebesgue measure is σ finite Radon measure.

Theorem 3.42 (Covering theorem of Besicovitch)**.** *There exists M_d depending only on d so that every family $\mathcal{F}$ of closed balls with bounded radii contains M_d subfamilies G_m, $1 \leq m \leq M_d$ so that each G_m consists disjoint balls and if A is the set of the centers then*

$$A \subset \bigcup_m \bigcup_{B \in G_m} B.$$

The same statement with the same proof holds for open balls.

Proof. We assume first that A is bounded and define D as the supremum of the radii. There exists a ball $B_1 = \overline{B_{r_1}(x_1)}$ with $r_1 \geq \frac{3D}{4}$. We choose recursively $B_n = \overline{B_{r_n}(x_n)}$ with x_n in

$$A_n = A \backslash \bigcup_{j=1}^{n-1} \overline{B_{r_j}(x_j)}$$

so that

$$r_n \geq \frac{3}{4} \sup\{r : \overline{B_r(x)} \in \mathcal{F}, x \in A_n\}.$$

We stop if $A_n = \{\ \}$. For simplicity we consider the case when the procedure does not stop. Then whenever $j \geq n$, we have $r_j \leq \frac{4}{3} r_n$ (otherwise we would not have chosen $\overline{B_{r_n}(x_n)}$) and

$$|x_j - x_n| \geq r_n \geq \frac{r_n + r_j}{3}$$

and the balls $B_{r_j/3}(x_j)$ are all disjoint. Thus $r_n \to 0$ (otherwise there would be infinitely many disjoint balls in a bounded set, which is impossible) and

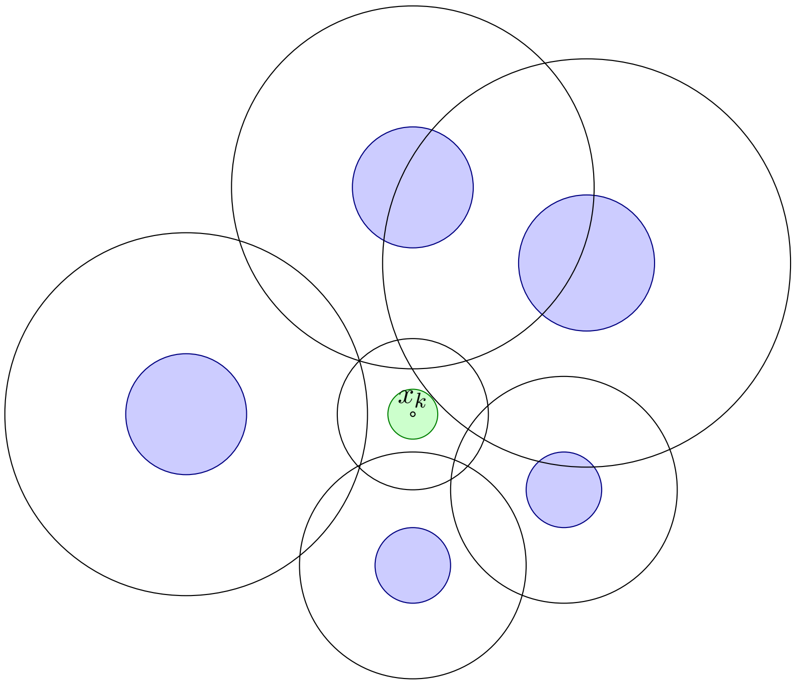
$$A \subset \bigcup_n B_n.$$

We fix $k > 1$ and define

$$I = \{j : 1 \leq j < k, B_j \cap B_k \neq \{\ \}\}.$$

We claim that there is a bound for the number of balls in I: $\#I \leq M_d$ with M_d depending only on d.

We first bound the number of small balls. Let $K := I \cap \{j : r_j \leq 3r_k\}$. Then $\#K \leq 20^d$.



To see that we consider $j \in K$ and choose $x \in B_{r_j/3}(x_j) \subset B_{5r_k}(x_k)$. The $\#K$ balls $B_{r_j/3}(x_j)$ are all disjoint and hence

$$(5r_k)^d \geq \sum_{j \in K} (r_j/3)^d \geq (r_k/4)^d \#K.$$

This implies the desired bound.

Next we bound the number of large balls, i.e $\#(I\backslash K)$. Let now $i, j \in I\backslash K$, $i \neq j$. We will give a *upper bound* on

$$\cos(\angle(x_k x_i, x_k x_j)) = \frac{\langle x_i - x_k, x_j - x_k\rangle}{|x_i - x_k||x_j - x_k|}.$$

This gives a lower bound on the distance of the points $\frac{x_n - x_k}{|x_n - x_k|}$ for $n < k$, $n \in I\backslash K$, and hence a upper bound on their numbers L_d depending only on the dimension since the unit sphere is compact. Therefore we can take $M_d = 20^d + L_d + 1$.

To simplify the notation we assume that $x_k = 0$ and $r_k = 1$. Let θ be the angle between the centers $\angle x_i, x_j$. Since $B_i \cap B_k \neq \{\ \}$ and $B_j \cap B_k \neq \{\ \}$, we have without loss of generality

$$r_i \leq |x_i| \leq |x_j|, |x_i| \leq r_i + r_k, |x_j| \leq r_j + r_k.$$

We claim that $\cos(\theta) > \frac{5}{6}$ implies $x_i \in B_j$. Firstly we notice that if $|x_i - x_j| \geq |x_j|$, then

$$\cos\theta = \frac{|x_i|^2 + |x_j|^2 - |x_i - x_j|^2}{2|x_i||x_j|} \leq \frac{|x_i|^2}{2|x_i||x_j|} = \frac{|x_i|}{2|x_j|} \leq \frac{1}{2} \leq \frac{5}{6}.$$

Hence $\cos\theta \geq \frac{5}{6}$ implies

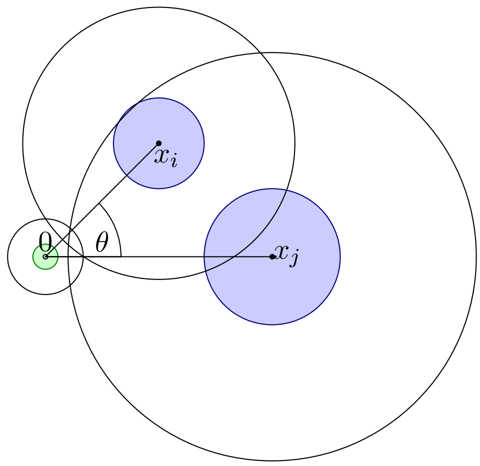
$$|x_i - x_j| \leq |x_j|.$$

We suppose by contradiction that $x_i \notin B_j$. Then $r_j \leq |x_i - x_j|$ and

$$\begin{aligned}\cos\theta =& \frac{|x_i|^2 + |x_j|^2 - |x_i - x_j|^2}{2|x_i||x_j|}\\ =& \frac{|x_i|}{2|x_j|} + \frac{(|x_j| - |x_i - x_j|)(|x_j| + |x_i - x_j|)}{2|x_i||x_j|}\\ \leq& \frac{1}{2} + \frac{|x_j| - |x_i - x_j|}{|x_i|}\\ \leq& \frac{1}{2} + \frac{r_j + r_k - r_j}{r_i}\\ =& \frac{1}{2} + \frac{r_k}{r_i} \leq \frac{5}{6}.\end{aligned}$$

Now it suffices to derive the upper bound for $\cos\theta$ when $x_i \in B_j$, since otherwise $\cos\theta \leq \frac{5}{6}$ has already a upper bound. So we assume $x_i \in B_j$ from now on. Then $i < j$, since otherwise B_i would not have been chosen, and thus $x_j \notin B_i$, and

$$3 \leq r_i < |x_i - x_j| < r_j \leq \frac{4}{3} r_i, \quad r_i < |x_i| \leq 1 + r_i, \quad r_j < |x_j| < 1 + r_j.$$



The proof becomes now an exercise in planar geometry. We have

$$\begin{aligned}
\frac{3}{16} \leq & \frac{1}{4}\frac{r_j}{|x_j|} \leq \frac{1}{3}\frac{r_i}{|x_j|} \leq \frac{\frac{2}{3}r_i - 1}{|x_j|} \leq \frac{r_i + r_i - r_j - 1}{|x_j|} \\
\leq & \frac{|x_i - x_j| + |x_i| - |x_j|}{|x_j|} \\
\leq & \frac{|x_i - x_j| + |x_i| - |x_j|}{|x_j|} \frac{|x_i - x_j| - |x_i| + |x_j|}{|x_i - x_j|} \\
= & \frac{|x_i - x_j|^2 - ||x_i| - |x_j||^2}{|x_j||x_i - x_j|} \\
= & 2(1 - \cos\theta)\frac{|x_i||x_j|}{|x_j||x_i - x_j|} \\
= & 2(1 - \cos\theta)\frac{|x_i|}{|x_i - x_j|} \\
\leq & 2(1 - \cos\theta)\frac{r_i + 1}{r_i} \\
\leq & \frac{8}{3}(1 - \cos\theta)
\end{aligned}$$

and hence $\cos\theta \leq \frac{119}{128}$.

It remain to define the sets G_m. We do this by defining a map

$$\sigma : \mathbb{N} \to \{1, \ldots M_d\}.$$

We choose it to be the identity for $j \leq M_d$. After that we proceed recursively, which we can do since

$$\#\Big\{j \leq k : B_j \cap B_{k+1} \neq \{\ \}\Big\} < M_d.$$

It remains to extend the result to unbounded sets. We do this by applying the first part in the annuli $6(m-1)D \leq |x| < 6mD$ and increasing M_d to $2M_d$.

□

Theorem 3.43. *Let μ be a Radon measure on $\mathbb{R}^d$ and let $\mathcal{F}$ be a family of closed balls and let A be a Borel set which is the union of the centers. We assume $\mu(A) < \infty$ and $\inf\{r : \overline{B_r(x)} \in \mathcal{F}\} = 0$ for $x \in A$. Let $U \subset \mathbb{R}^d$ be open. Then there exists a countable collection of disjoint closed balls $G \subset \mathcal{F}$ so that $B \subset U$ if $B \in G$ and*

$$\mu\Big((A \cap U) \backslash \bigcup_{B \in G} B\Big) = 0.$$

Proof. We fix θ so that $1 - \frac{1}{M_d} < \theta < 1$ and claim that there is a finite collection of disjoint balls B_j, $1 \le j \le M$ in $\mathcal{F}$ so that

$$\mu\left((A \cap U) \backslash \bigcup_{j=1}^{M} B_j\right) \le \theta \mu(A \cap U).$$

Suppose this is true. Then we define

$$A_1 = A \backslash \bigcup_{j=1}^{M} B_j, \quad U_1 = U \backslash \bigcup_{j=1}^{M} B_j$$

and repeat the argument with $\mathcal{F}_1$ the subset of balls with center in A_1. After the kth step the complement has a measure at most $\theta^j \mu(U \cap A)$. So it remains to prove the claim. Let $\mathcal{F}_0$ be the subset of balls with radii at most 1. Then we apply the Besicovitch covering theorem and obtain G_m. Then

$$A \cap U \subset \bigcup_{j=1}^{M_d} \bigcup_{B \subset G_j} B$$

and

$$\mu(A \cap U) \le \sum_{j=1}^{M_d} \mu\Big(A \cap U \cap \bigcup_{B \in G_j} B\Big).$$

There exists J so that

$$\frac{1}{M_d} \mu(A \cap U) \le \mu\Big(A \cap U \cap \bigcup_{B \in G_J} B\Big).$$

By monotone convergence there exist finitely many balls in G_J so that the claim holds. □

22.11.2019

We turn to derivatives of Radon measures.

Definition 3.44. *Let μ and ν be Radon measures on $\mathbb{R}^d$. For $x \in \mathbb{R}^d$ we define*

$$\overline{D_\mu}\nu(x) = \begin{cases} \limsup_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} & \text{if } \mu(B_r(x)) > 0 \text{ for all } r > 0 \\ \infty & \text{if for some } r > 0, \mu(B_r(x)) = 0 \end{cases}$$

$$\underline{D_\mu}\nu(x) = \begin{cases} \liminf_{r\to 0} \frac{\nu(B_r(x))}{\mu(B_r(x))} & \text{if } \mu(B_r(x)) > 0 \text{ for all } r > 0 \\ \infty & \text{if for some } r > 0, \mu(B_r(x)) = 0. \end{cases}$$

We say that ν is differentiable with respect to μ at x if $\overline{D_\mu}\nu(x) = \underline{D_\mu}\nu(x) < \infty$. Then we write $D_\mu\nu(x)$ when $\overline{D_\mu}\nu(x) = \underline{D_\mu}\nu(x)$ and call this quantity the density of ν with respect to μ.

Remark 3.45. *Let $f \in C_c(\mathbb{R}^d)$ and μ Radon measure. Then*

$$x \to \int f(y-x)d\mu(y)$$

is continuous and hence Borel measurable.

Since the characteristic function of open and closed balls can be obtained as pointwise limit of continuous functions with compact support, the map

$$x \to \nu(B_r(x)), \quad x \to \mu(B_r(x))$$

are measurable. Thus

$$x \to \begin{cases} \frac{\nu(B_r(x))}{\mu(B_r(x))} & \text{if } \mu(B_r(x)) > 0 \\ \infty & \text{if } \mu(B_r(x)) = 0 \end{cases}$$

is Borel measurable. The maps

$$r \to \mu(B_r(x)), \qquad r \to \nu(B_r(x))$$

is monoton and continuous from the left. There are at most countably many points of discontinuity. Thus also $\overline{D_\mu}\nu(x)$ and $\underline{D_\mu}\nu(x)$ are Borel measurable since we can write them as inf's and sup's over rational radii. Moreover by inner and outer regularity we obtain the same $\overline{D_\mu}\nu(x)$ and $\underline{D_\mu}\nu(x)$ if we use closed balls.

Theorem 3.46. *Let μ and ν be Radon measures on $\mathbb{R}^d$. Then*

1. *$D_\mu\nu(x)$ exists and is finite μ almost everywhere.*
2. *$D_\mu\nu$ is Borel measurable.*

Proof. We may assume that $\mu(\mathbb{R}^d) < \infty$ and $\nu(\mathbb{R}^d) < \infty$.

Step 1: We claim that for all Borel sets B and all $t > 0$

$$\nu(B \cap \{x : \underline{D_\mu}\nu(x) < t\}) \le t\mu(B \cap \{x : \underline{D_\mu}\nu(x) < t\}),$$

$$\nu(B \cap \{x : \underline{D_\mu}\nu(x) > t\}) \ge t\mu(B \cap \{x : \underline{D_\mu}\nu(x) > t\}),$$

$$\nu(B \cap \{x : \overline{D_\mu \nu}(x) < t\}) \le t\mu(B \cap \{x : \overline{D_\mu \nu}(x) < t\})$$

and

$$\nu(B \cap \{x : \overline{D_\mu \nu}(x) > t\}) \ge t\mu(B \cap \{x : \overline{D_\mu \nu}(x) > t\}).$$

The proofs are the same for $\overline{D_\mu \nu}(x)$ and $\underline{D_\mu}\nu(x)$, and we restrict ourselves to $\underline{D_\mu}\nu(x)$

By outer regularity $\mu(B) = \inf\{\mu(U) : B \subset U\}$ and it suffices to prove the claim for $B = U$ open. Let $A = \{x \in U : \underline{D_\mu}\nu(x) < t\}$ (opposite inequality for the second inequality). . Let (again with the reverse inequality for the second inequality)

$$\mathcal{F} = \{\overline{B_r(a)} : a \in A, \overline{B_r(a)} \subset U, \nu(\overline{B_r(a)}) < t\mu(\overline{B_r(a)})\}.$$

For every $x \in A$, $\mathcal{F}$ contains arbitrarily small balls and we apply Theorem 3.43 to obtain a sequence of disjoint closed balls B_j in $\mathcal{F}$ so that

$$\nu(A \backslash \bigcup_n B_j n) = 0$$

$(\mu(A \backslash \bigcup_n B_n) = 0$ for the second inequality). Then

$$\nu(A) = \sum \nu(B_j) \le t \sum \mu(B_j) \le t\mu(U)$$

resp. for the second inequality

$$\nu(U) \ge \sum \nu(B_j) \ge t \sum \mu(B_j) = t\mu(A).$$

Since

$$\mu(A) = \inf\{\mu(U) : A \subset U\}, \qquad \nu(A) = \inf\{\nu(U) : A \subset U\}$$

we obtain the first two inequalities.

Step 2: We claim that $\overline{D_\mu}\nu(x) < \infty$ outside a set of μ measure 0. Let $A = \{x : \overline{D_\mu}\nu(x) = \infty\}$. Then

$$\nu(A) \ge t\mu(A)$$

for all t, hence $\mu(A) = 0$.

Step 3: For $s < t$ we define

$$R(s,t) = \{x : \underline{D_\mu}\nu(x) < s < t < \overline{D_\mu}\nu(x)\}$$

Then

$$t\mu(R(s,t)) \leq \nu(R(s,t)) \leq s\mu(R(s,t))$$

which implies $\mu(R(s,t)) = 0$. Since

$$\{x : \underline{D_\mu}\nu(x) < \overline{D_\mu}\nu(x)\} = \bigcup_{s<t, s,t\in\mathbb{Q}} R(s,t)$$

we see that $\underline{D_\mu}\nu(x) = \overline{D_\mu}\nu(x)$ for μ almost all x. □

Definition 3.47. *Let μ and ν be Borel measures on $\mathbb{R}^d$. We say the measure ν is absolutely continuous with respect to μ, $\nu << \mu$, if $\mu(A) = 0$ implies $\nu(A) = 0$. The mearures ν and μ are mutally singular if there exists a Borel set B such that $\mu(X\backslash B) = \nu(B) = 0$. We write then $\nu \perp \mu$.*

Theorem 3.48 (Radon-Nikodym)**.** *Let ν and μ be Radon measures on $\mathbb{R}^d$ with $\nu << \mu$. Then*

$$\nu(A) = \int_A D_\mu \nu d\mu$$

for all Borel sets.

Proof. It suffices to consider the case $\mu(\mathbb{R}^d) < \infty$ and $\nu(\mathbb{R}^d) < \infty$. We have seen that

$$\mu(\{D_\mu\nu(x) = \infty\}) = 0$$

and hence, since $\nu << \mu$, $\nu((\{D_\mu\nu(x) = \infty\}) = 0$. In the same fashion

$$\nu(\{D_\mu\nu(x) = 0\}) = 0.$$

Let A be a Borel set. For $t > 1$ we define

$$A_n = A \cap \{t^n \leq D_\mu\nu < t^{n+1}\}.$$

Then

$$\nu(A) = \sum_{n=-\infty}^{\infty} \nu(A_n) \leq \sum_{n=-\infty}^{\infty} t^{n+1}\mu(A_n) \leq t\int_0^\infty \mu(\{D_\mu\nu > s\})ds = t\int_A D_\mu\nu d\mu$$

and

$$\nu(A) = \sum_{n=-\infty}^{\infty} \nu(A_n) \geq \sum_{n=-\infty}^{\infty} t^n\mu(A_n) \geq t^{-1}\int_0^\infty \mu(\{D_\mu\nu > s\})ds = t^{-1}\int_A D_\mu\nu d\mu.$$

We let now $t \to 1$. □

Theorem 3.49 (Lebesgue points)**.** *Let μ be a Radon measure on $\mathbb{R}^d$ and $\tilde{f} \in L^1_{loc}(\mu)$. Then*

$$f(x) := \lim_{r\to 0} \mu(B_r(x))^{-1} \int_{B_r(x)} \tilde{f} d\mu$$

exists almost everywhere and we define $f(x) = 0$ if it does not exist. Then f is in the equivalence class of $\tilde{f}$. If $\tilde{f} \in L^p_{loc}(\mu)$ then $f \in L^p_{loc}$ and

$$\lim_{r\to 0} \mu(B_r(x))^{-1} \int_{B_r(x)} |f(y) - f(x)|^p d\mu(y) = 0$$

almost everywhere.

Proof. It suffices to consider nonnegative $\tilde{f}$ and $\mu(\mathbb{R}^d) < \infty$. We define

$$\nu(A) = \int_A \tilde{f} \, d\mu.$$

This is a Radon measure by Lemma 3.27 which is absolutely continuous with respect to μ. Thus

$$\nu(A) = \int D_\mu \nu d\mu = \int \tilde{f} d\mu$$

and $D_\mu \nu$ lies in the equivalence class. Now the first claim follows from Theorem 3.46.

For every t, $|f(x) - t|^p$ is integrable. From the first part

$$\lim_{r\to 0} \mu(B_r(x))^{-1} \int_{B_r(x)} |f(y) - t|^p d\mu(y) = |f(x) - t|^p$$

almost every where. There is even a set N of μ measure zero so that this is true for all $t \in \mathbb{Q}$ outside the same set of measure zero. Let $\varepsilon > 0$. Thus the set of all x such that

$$\limsup_{r\to 0} \mu(B_r(x))^{-1} \int_{B_r(x)} |f(y) - f(x)|^p d\mu(y) > \varepsilon$$

is contained in N. To see this, chose $t \in \mathbb{Q}$ so that $|f(x) - t|^p < \varepsilon$. This completes the proof. □

Corollary 3.50. *Let μ be a Radon measure and $f \in L^p(\mu)$. Then there exists a cannonical representative of the equivalence class.*

27.11.2019

3.12 Young's inequality and Schur's lemma

Let $(X, \mathcal{A}, \mu)$ be a measure space and suppose that $1 \le p, q, r \le \infty$ and $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$. If $f \in L^p(\mu)$, $g \in L^q(\mu)$ and $h \in L^r(\mu)$, then fgh is integrable and

$$\left| \int fgh d\mu \right| \le \|f\|_{L^p(\mu)} \|g\|_{L^q(\mu)} \|h\|_{L^r(\mu)}.$$

This is a consequence of a multiple application of Hölder's inequality:

$$\left| \int fgh d\mu \right| \le \|f\|_{L^p} \|gh\|_{L^{\frac{p}{p-1}}}$$

and

$$\int |g|^{\frac{p}{p-1}} |h|^{\frac{p}{p-1}} d\mu \le \| |g|^{\frac{p}{p-1}} \|_{L^{\frac{q(p-1)}{p}}} \|h\|_{L^{\frac{r(p-1)}{p}}} = \|g\|_{L^q}^{\frac{p}{p-1}} \|h\|_{L^r}^{\frac{p}{p-1}}$$

since

$$\frac{p}{p-1} \left(\frac{1}{q} + \frac{1}{r} \right) = \frac{p}{p-1} (1 - \frac{1}{p}) = 1.$$

We denote $L^p(\mathbb{R}^d)$ (or even L^p) for $L^p(m^d)$ where m^d is the Lebesgue measure.

Lemma 3.51. *Suppose that* $1 \le p, q, r \le \infty$ *satisfy*

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$$

and that $f \in L^p(\mathbb{R}^d)$, $g \in L^q(\mathbb{R}^d)$ *and* $h \in L^r(\mathbb{R}^d)$. *Then*

$$\mathbb{R}^d \times \mathbb{R}^d \ni (x, y) \to f(-x) g(x - y) h(y)$$

is integrable and

$$I(f, g, h) := \int_{\mathbb{R}^d \times \mathbb{R}^d} f(-x) g(x - y) h(y) dm^{2d}(x, y)$$

satisfies

$$|I(f, g, h)| \le \|f\|_{L^p} \|g\|_{L^q} \|h\|_{L^r}$$

and

$$I(f, g, h) = I(g, f, h) = I(f, h, g) = I(h, g, f).$$

Proof. We assume $1 < p, q, r < \infty$ since the limit cases are simpler, and follow by obvious modifications. Measurability is a consequence of the theorem of Fubini. It suffices to prove the statement for nonnegative functions since

$$\left| \int fgh dm^{2d} \right| \le \int |fgh| dm^{2d}$$

and the integrability of fgh follows from the integrability of $|fgh|$. We define p', q' and r' by $\frac{1}{p}+\frac{1}{p'}=1$, i.e. $p'=\frac{p}{p-1}$ etc. Let

$$\begin{aligned}\alpha(x,y)=&|f(-x)|^{p/r'}|g(x-y)|^{q/r'},\\ \beta(x,y)=&|f(-x)|^{p/q'}|h(y)|^{r/q'},\\ \gamma(x,y)=&|g(x-y)|^{q/p'}|h(y)|^{r/p'}.\end{aligned}$$

Then $\frac{1}{p'}+\frac{1}{q'}+\frac{1}{r'}=1$ and

$$\begin{aligned}I=&\int \alpha(x,y)\beta(x,y)\gamma(x,y)dm^{2d}\\ \le&\|\alpha\|_{L^{r'}}\|\beta\|_{L^{q'}}\|\gamma\|_{L^{p'}}\\ =&\|f\|_{L^p}^{\frac{p}{r'}}\|g\|_{L^q}^{\frac{q}{r'}}\|f\|_{L^p}^{\frac{p}{q'}}\|h\|_{L^r}^{\frac{r}{q'}}\|g\|_{L^q}^{\frac{q}{p'}}\|h\|_{L^r}^{\frac{r}{p'}}\\ =&\|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r}.\end{aligned}$$

The second last equality is a consequence of the theorem of Fubini. □

Theorem 3.52 (Young's inequality)**.** *Suppose that* $1\le p,q,r'\le\infty$ *and*

$$\frac{1}{p}+\frac{1}{q}=1+\frac{1}{r'}.$$

If $f\in L^p$ *and* $g\in L^q$*, then for almost all* x

$$f(x-y)g(y)$$

is integrable and

$$f*g(x):=\begin{cases}\int f(x-y)g(y)dm^d(y) & \textit{if integrable}\\ 0 & \textit{otherwise}\end{cases}$$

defines a unique element in $L^{r'}(\mathbb{R}^d)$ *and*

$$\|f*g\|_{L^{r'}(\mathbb{R}^d)}\le\|f\|_{L^p}\|g\|_{L^q}.$$

Proof. We have $e^{-|x|^2}\in L^r$ for all $1\le r\le\infty$. Then

$$e^{-|x|^2}f(x-y)g(y)$$

is m^{2d} integrable by Lemma 3.51. We apply Fubini to see that $\int f(x-y)g(y)dm^d(y)$ exists for almost all x. By Theorem 3.20 the estimate follows once we prove

$$\left|\int f*ghdm^d\right|\le\|f\|_{L^p}\|g\|_{L^q}\|h\|_{L^r}$$

for

$$\frac{1}{r}+\frac{1}{r'}=1$$

and all $h \in L^r$. Since then

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2$$

and, by Fubini and Lemma 3.51

$$\begin{aligned}\left| \int f * gh(x)dx \right| \le & \int |f| * |g||h| dm^d \\ = & \int |f(x-y)||g(y)||h(x)| dm^{2d}(x,y) \\ \le & \|f\|_{L^p} \|g\|_{L^q} \|h\|_{L^r}.\end{aligned}$$

□

There is a particular case: if $q = 1$ and $p = r'$:

$$\|f * g\|_{L^p} \le \|f\|_{L^p} \|g\|_{L^1}.$$

There is an important variant. Let $1 \le p < \infty$, and $f \in L^p(\mu)$. Then

$$\begin{aligned}\|f\|_{L^p}^p &= \int |f|^p d\mu = \int_0^\infty \mu(\{|f|^p > \lambda\}) d\lambda = p \int_0^\infty t^{p-1} \mu(\{|f| > t\}) dt \\ &\ge \sup_{t>0} t^p \mu(\{|f| > t\}) =: \|f\|_{L^p_w}^p.\end{aligned}$$

This is an abuse of notation, since the right hand is not a norm. We write $f \in L^p_w$ if f is measurable and $\|f\|_{L^p_w} < \infty$.

The interesting example is $f(x) = |x|^{-d/p} \in L^p_w$.

Theorem 3.53. *Let $1 < p, q, r < \infty$ satisfy*

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 2,$$

$f \in L^p(\mathbb{R}^d)$, $g \in L^q_w(\mathbb{R}^d)$, $h \in L^r(\mathbb{R}^d)$. Then

$$\left| \int f(x) g(x-y) h(y) dm^d(x) dm^d(y) \right| \le c(d,p,q) \|f\|_{L^p} \|g\|_{L^q_w} \|h\|_{L^r}$$

and, with $\frac{1}{r} + \frac{1}{r'} = 1$

$$\|f * g\|_{L^r} \le c(d,p,q) \|f\|_{L^p} \|g\|_{L^q_w}.$$

Proof. We replace f, g and h by their absolute value. The left hand side does not decrease, and the right hand side does not change. Multiplying each function by a constant we assume

$$\|f\|_{L^p(\mathbb{R}^d)} = 1, \|g\|_{L^q_w(\mathbb{R}^d)} = 1, \|h\|_{L^r(\mathbb{R}^d)} = 1.$$

We denote the left hand side by $I(f,g,h)$. Then, with the fundamental theorem of calculus and Fubini

$$\begin{aligned} I(f,g,h) &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \int_0^{f(x)} dt_1 \int_0^{g(x-y)} dt_2 \int_0^{h(y)} dt_3 dm^{2d}(x,y) \\ &= \int_0^\infty \int_0^\infty \int_0^\infty I(\chi_{\{f>t_1\})}\chi_{\{g>t_2\}}\chi_{\{h>t_3\}})dt_1 dt_2 dt_3 \end{aligned}$$

Replacing one of the characteristic functions by 1, using for example

$$I(f,g,1) = \int f dm^d \int g dm^s$$

(and doing the optimal choice), and using $m^d(\{x : g(x) > t\}) \le t^{-q}$

$$\begin{aligned} I(\chi_{\{f>t_1\}},\chi_{\{g>t_2\}},\chi_{\{h>t_3\}}) &\le \min\{F(t_1)G(t_2), F(t_1)H(t_3), G(t_2)H(t_3)\} \\ &\le \min\{F(t_1)t_2^{-q}, F(t_1)H(t_3), t_2^{-q}H(t_3)\} \end{aligned}$$

where

$$F(t) = m^d(\{f > t\}), \qquad H(t) = m^d(\{h > t\}).$$

We decompose the outer integral with

$$1 = \chi_{\{F(t_1)<H(t_3)\}} + \chi_{\{F(t_1)\ge H(t_3)\}},$$

and write accordingly $I(f,g,h) = I_1(f,g,h)+I_2(f,g,h)$, we consider $I_1(f,g,h)$. We carry out the integration with respect to $t_2 \ge H^{-\frac{1}{q}}$ first and split the integral into $t_2 < H(t_3)^{-1/q}$ and $t_2 \ge H(t_3)^{-1/q}$:

$$\begin{aligned} I_1 &\le \int_0^\infty \int_0^\infty \chi_{F\le H} F(t_1)H^{\frac{q-1}{q}}(t_3)(\frac{1}{q-1}+1)dt_1dt_3 \\ &= \frac{q}{q-1}\int_0^\infty \int_0^\infty \chi_{F\le H}F(t_1)H^{\frac{q-1}{q}}(t_3)dt_1dt_3. \end{aligned}$$

Since $F \le H$ implies

$$F(t_1)H(t_3)^{\frac{q-1}{q}} \le F(t_1)^{\frac{q-1}{q}}H(t_3)$$

we obtain (splitting the t_3 integral into two integrations)

$$\begin{aligned} I_1 &\le \frac{q}{q-1}\int_{\{(t_1,t_3):F(t_1)<H(t_3)\}} \chi_{F\le H}F(t_1)(H(t_3))^{\frac{q-1}{q}}dt_1dt_3 \\ &\le \frac{q}{q-1}\left(\int_0^\infty H(t_3)\int_0^{t_3^{r/p}} F^{\frac{q-1}{q}}(t_1)dt_1dt_3 + \int_0^\infty F(t_1)\int_0^{t_1^{p/r}} H^{\frac{q-1}{q}}(t_3)dt_3dt_1\right) \end{aligned}$$

The roles of F and H are symmetric and we bound only the second integral.

We use Hölder's inequality with $m = (r-1)\frac{q-1}{q}$

$$\int_0^{t_1^{p/r}} H(t_3)^{\frac{q-1}{q}} t_3^m t_3^{-m} dt_3 \le \left(\int_0^{t_1^{p/r}} H(t_3) t_3^{r-1} dt_3 \right)^{\frac{q-1}{q}} \left(\int_0^{t_1^{p/r}} t_3^{-(r-1)(q-1)} dt_3 \right)^{\frac{1}{q}}$$

and

$$\left(\int_0^{t_1^{p/r}} t_3^{-(r-1)(q-1)} dt_3 \right)^{1/q} = c t_1^{(1-(r-1)(q-1))\frac{p}{qr}} = c t_1^{p-1}$$

hence

$$I_1 \le 2c \int_0^\infty t_1^{p-1} F(t_1) dt_1 \left(\int_0^\infty t_3^{r-1} H(t_3) dt_3 \right)^{\frac{q}{q-1}} \le c(d,p,q)$$

as claimed, due to the normalization $\|f\|_{L^p} = \|h\|_{L^r} = 1$.

□

Example: $u \in C_c^2(\mathbb{R}^d)$, $f = \Delta u$, $d \ge 3$. Then

$$u(x) = c_d f * |.|^{2-d}$$

and if

$$1 < p, r < \infty, \quad \frac{1}{r} = \frac{1}{p} - \frac{2}{d}$$

Then $|x|^{2-d} \in L^{\frac{d}{d-2}}$,

$$\frac{1}{p} + \frac{d-2}{d} = 1 + \frac{1}{p} - \frac{2}{d} = \frac{1}{r}.$$

Thus

$$\|u\|_{L^r} \le c\|f\|_{L^q(\mathbb{R}^d)}.$$

Schur's lemma gives a criterium for an integral kernel to define a linear map from $L^p(\nu)$ to $L^p(\mu)$ for $1 \le p \le \infty$.

Theorem 3.54 (Schur's lemma)**.** *Let $(X, \mathcal{A}, \mu)$ and $(Y, \mathcal{B}, \nu)$ be σ finite measure spaces and $k : X \times Y \to \mathbb{R}$ be $\mu \times \nu$ measurable. Suppose that $C_1, C_2 \in [0, \infty)$ and*

$$\sup_x \int |k(x,y)| d\nu(y) \le C_1, \qquad \sup_y \int |k(x,y)| d\mu(x) \le C_2.$$

If $1 \le p \le \infty$ and $f \in L^p(\nu)$, then

$$\int k(x,y) f(y) d\nu(y)$$

exists for almost all x and

$$\Big\| \int k(x,y)f(y)d\nu(y)\Big\|_{L^p(\mu)} \le C_1^{1-\frac{1}{p}} C_2^{\frac{1}{p}} \|f\|_{L^p(\nu)}.$$

The map

$$L^p(\nu) \ni f \to Tf := \int k(x,y)f(y)d\nu(y) \in L^p(\mu)$$

is a continuous linear map which satisfies

$$\|T\|_{L^p(\nu)\to L^p(\mu)} \le C_1^{1-\frac{1}{p}} C_2^{\frac{1}{p}}.$$

Proof. This is an immediate estimate if $p=1$ or $p=\infty$. The other cases follow from the next theorem. □

Theorem 3.55. *Suppose that $\mathbb{K}=\mathbb{C}$, $1 \le p_0, p_1 \le \infty$, $1 \le q_0, q_1 \le \infty$, $T: L^{p_0}(\nu) \cap L^{p_1}(\nu) \to L^{q_0}(\mu) \cap L^{q_1}(\mu)$ and assume*

$$\|Tf\|_{L^{q_0}(\mu)} \le C_0 \|f\|_{L^{p_0}(\nu)},$$

and

$$\|Tf\|_{L^{q_1}(\mu)} \le C_0 \|f\|_{L^{p_1}(\nu)}$$

for all $f \in L^{p_0}(\nu) \cap L^{p_1}(\nu)$. Let $0 \le \lambda \le 1$ and

$$\frac{1}{p} = \frac{1-\lambda}{p_0} + \frac{\lambda}{p_1} \qquad \frac{1}{q} = \frac{1-\lambda}{q_0} + \frac{\lambda}{q_1}.$$

Then T define a unique continuous linear map from $L^p(\nu) \to L^q(\mu)$ with

$$\|T\|_{L^p(\nu)\to L^q(\mu)} \le C_0^{1-\lambda} C_1^{\lambda}.$$

Clearly Theorem 3.55 implies Theorem 3.54 with $p_0 = q_0 = 1$, $p_1 = q_1 = \infty$, $\lambda = 1 - \frac{1}{p}$.

Proof of Theorem 3.55. To keep the notation simple we only consider $p_0 = q_0 = 1$, $p_1 = q_1 = \infty$. The argument immediately generalizes. Repeating the argument of Young's inequality we have to prove that

$$\int g(x)Tf(x)d\mu(x) \le C_0^{1-\lambda} C_1^{\lambda} \tag{3.5}$$

where

$$\frac{1}{p} + \frac{1}{q} = 1,$$

if $\|f\|_{L^p} = \|g\|_{L^q} = 1$. For $z \in \mathbb{C}$ with $0 \le \operatorname{Re} z \le 1$, we define

$$f_z = \begin{cases} |f|^{pz-1}f & \text{if } f \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

and

$$g_z = \begin{cases} |g|^{\frac{p}{p-1}(1-z)-1} g & \text{if } g \neq 0 \\ 0 & \text{otherwise .} \end{cases}$$

Then for $\sigma \in \mathbb{R}$,

$$\|f_{i\sigma}\|_{L^\infty} = \|g_{1+i\sigma}\|_{L^\infty} = 1,$$

and

$$\|f_{1+i\sigma}\|_{L^1} = \|f\|_{L^p}^p = \|g_{i\sigma}\|_{L^1} = \|g\|_{L^q}^q = 1,$$

and hence

$$\int |g_{i\sigma}(x)||Tf_{i\sigma}(x)|d\mu \le C_0$$

$$\int |g_{1+i\sigma}(x)|Tf_{1+i\sigma}(x)|d\mu \le C_1.$$

Moreover

$$f_{\frac{1}{p}} = f, \qquad g_{\frac{1}{p}} = g.$$

Notice that f_z and g_z are bounded and zero outside a set of finite measure. By dominated convergence

$$z \to H(z) = \int_X g_z(x) Tf_z(x) d\mu(x)$$

is continuous in the strip $\mathcal{C} = \{z : 0 \le \operatorname{Re} z \le 1\}$, differentiable and satisfies the Cauchy-Riemann differential equations. The claim follows from the three lines inequality below.

Lemma 3.56 (Three lines inequality)**.** *Suppose that $u \in C(\mathcal{C})$ is bounded and holomorphic in the interior. Then*

$$\sup_{\mathcal{C}} |u| = \sup_{\partial \mathcal{C}} |u|$$

We apply the lemma to

$$u(z) = C_1^{z-1} C_2^{-z} H(z).$$

□

Proof of Lemma 3.56: 1) Let $U \subset \mathbb{C}$ be a bounded open connected set and $u \in C(\overline{U}; \mathbb{C})$ be a holomorphic function in the interior. We claim that then

$$\sup_{x \in \overline{U}} |u(x)| = \sup_{x \in \partial U} |u(x)|.$$

We prove this by contradiction. Suppose that $|u|$ attains its maximum M at some interior point z_0 and suppose that this is larger than $\sup_{\partial U} |u(z)|$. Then

$$f(z) = \operatorname{Re} u(z)/u(z_0)$$

satisfies $0 \le f \le M$ and $f(z_0) = M$. Moreover f is harmonic. Let

$$f_\varepsilon(x+iy) = f(x+iy) + \varepsilon |x - \operatorname{Re} z_0|^2$$

where ε is so small that $f_\varepsilon(z) < M$ for $z \in \partial U$. Then f_ε has a maximum in an interior point z_1. At this point the Hessian is not negtative semidefinite by its trace $\Delta f_\varepsilon(z_1) = 4\varepsilon$. This is a contradiction.

2) Let u be as in the lemma amd let

$$u_\varepsilon(z) = e^{\varepsilon z^2} u(z).$$

Since $u_\varepsilon(z) \to 0$ as $|\operatorname{Im} z| \to \infty$

$$\sup_{z \in \mathcal{C}} |u_\varepsilon(z)| = \sup_{z \in \partial \mathcal{C}} |u_\varepsilon| \le e^\varepsilon \sup_{z \in \partial \mathcal{C}} |u(z)|.$$

Now we let ε tend to zero. ∎

29.11.2019

4 Distributions and Sobolev spaces

4.1 Baire category theorem and consequences

Lemma 4.1 (Baire category theorem)**.** *A countable intersection of dense open subsets of a complete metric space is dense.*

Proof. Let (X,d) be a complete metric space and A_j open dense sets. Let $x \in X$ and $\varepsilon > 0$. Let $x_1 \in A_1$ so that $d(x,x_1) < \varepsilon/3$ and $0 < \delta_1 < \varepsilon/3$ so that $B_{2\delta_1}(x_1) \subset A_1$. We pick recursively x_n, δ_n so that $d(x_{n-1}, x_n) < \delta_n$, $2\delta_n < \varepsilon/3^n$ and $B_{2\delta_n}(x_n) \in A_n \cap B_{\delta_{n-1}}(x_{n-1})$.

By construction, $d(x_{n-1},x_n) < \varepsilon/(2\cdot 3^{n-1})$ and, if $n < m$, $d(x_n, x_m) < \frac{\varepsilon}{2\cdot 3^n} \sum_{j=0}^{m-n} 3^{-j} \le \frac{3\varepsilon}{4\cdot 3^n}$ and (x_n) is a Cauchy sequence with limit y. Since $x_m \in \overline{B_{\delta_n}(x_n)}$ for $m \ge n$ the same is true for y, and $y \in A_n$ for all n. ∎

Theorem 4.2 (Banach-Steinhaus)**.** *Let X and Y be Banach spaces, $\mathcal{F} \subset L(X,Y)$. Suppose for each $x \in X$*

$$\sup\{\|Tx\|_Y : T \in \mathcal{F}\} < \infty.$$

Then

$$\sup\{\|T\|_{X\to Y} : T \in \mathcal{F}\} < \infty.$$

Proof. Let

$$C_n = \{x \in X : \sup_{T \in \mathcal{F}} \|Tx\|_Y \le n\}.$$

This set is closed since both map and norm are continuous, and C_n is an intersection of closed sets. By assumption $\bigcup C_n = X$. We claim that some

C_n has nonempty open interior. If not then the sets $U_n = X \backslash C_n$ are open and dense, with nonempty intersection, a contradiction to $\bigcup C_n = X$. Let $U \subset C_{n_0}$ be nonempty and open. It contains a ball $B_r(x_0)$. If $\|x\| < r$ then

$$\|Tx\|_Y \le \|T(x-x_0)\|_Y + \|T(x_0)\|_Y \le n_0 + \sup_{T \in \mathcal{F}} \|T(x_0)\|_Y =: R.$$

Then

$$\|T\|_{X \to Y} \le R/r$$

for all $T \in \mathcal{F}$. □

The Baire category theorem has interesting further consequences.

Theorem 4.3. *[Open mapping theorem] Let X and Y be Banach spaces and $T \in L(X,Y)$. T is surjective if and only if it is open, i.e. if the image of open sets is open.*

Proof. Let T be open. Then $T(B_1(0))$ is open. In particular it contains a ball $B_r(0)$. Then $Y = \bigcup T(B_n(0))$ and T is surjective. Now suppose that T is surjective. It suffices to show that $T(B_1(0))$ contains a ball around 0. (Why?). Let

$$Y_n = \overline{T(B_n(0))} = \overline{\{Tx | \|x\|_X < n\}}.$$

By surjectivity $Y = \bigcup Y_n$. As above we conclude that one (and hence all) of the Y_n contains an open ball. Hence there exists a ball $B_r(y_0) \subset Y_1$. Then $B_r(0) \subset Y_2$. □

Corollary 4.4. *Suppose that $T \in L(X,Y)$ is injective and surjective. Then $T^{-1} \in L(Y,X)$.*

Thus continuous linear maps which are invertible as maps between sets are invertible as continuous linear maps.

Proof. Linearity of the inverse map is immediate. By Theorem 4.3 T is open. So $T(B_1^X(0))$ contains a ball $B_r^Y(0)$ and hence

$$\|Tx\|_Y \ge r^{-1} \|x\|_X.$$

□

Let X and Y be Banach spaces. Then $X \times Y$ is a Banach space with norm

$$\|(x,y)\|_{X \times Y} = \max\{\|x\|_X, \|y\|_Y\}$$

If $T : X \to Y$ is a linear map not necessarely constinuous - then the graph is

$$\Gamma(T) = \{(x, Tx) : x \in X\}.$$

Theorem 4.5 (The closed graph theorem)**.** *The linear map T is continuous if and only if $\Gamma(T)$ is closed.*

Proof. Suppose that T is continuous and that $(x_n, Tx_n) \in \Gamma(T))$ is a Cauchy sequence in $X \times Y$. Then (x_n) is a Cauchy sequence in X with limit x , and

$$\lim_{n\to\infty} Tx_n = Tx =: y.$$

Thus $(x_n, Tx_n) \to (x, Tx)$ in $X \times Y$. Now assume that Γ is closed. It is a closed linear subspace, hence a Banach space. The map

$$\Gamma \ni (x, Tx) \to x \in X$$

in injective and surjective, hence, by the inverse mapping theorem, its inverse is continous. But then the composition to the second factor is also continuous, which is the map $x \to Tx$. □

Theorem 4.6. *The set of nowhere differentiable functions in $C_b([0,1])$ is dense.*

Proof. We define

$$A_L = \{f \in C_b([0,1]) : \text{ there exists } x \text{ such that } |f(x)-f(y)| \le L|x-y| \text{ for all } y\}$$

Then

1. A_L is closed. Let f_n be a Cauchy sequence converging to f, and x_n the sequence of points. Without loss assume that $x_n \to x$. Then

   $$|f(x)-f(y)| = \lim_{n\to\infty} |f(x_n)-f(y)| = \lim_{n\to\infty} \lim_{m\to\infty} |f_m(x_n)-f(y)| \le \lim_{n\to\infty} L|x_n-y| = L|x-y|.$$

2. If f is differentiable at x then there exists L so that

   $$|f(y) - f(x)| \le L|x - y|.$$

   By the definition of differentiability there exists $\delta > 0$ so that

   $$|f(y) - f(x)| \le (|f'(x)| + 1)|x - y|$$

   for $|x - y| < \delta$. Since continuous functions on compact set attain the supremum the difference quotient is bounded.

3.

4. A_L is nowhere dense. Let $f \in A_L$ and $\varepsilon > 0$. Since f is uniformly continuous there exists $\delta > 0$ so that $|f(x) - f(y)| < \varepsilon/3$ if $|x - y| \le \delta$. We choose n so that

   $$\delta n > 4$$

and define

$$g(x) = f(x) + \varepsilon \sin(2\pi n x).$$

In any interval of length $\delta/2$ there exist two points x_1, x_2 with

$$\sin(2\pi n x_1) = 1, \sin(2\pi n x_2) = -1.$$

We fix $x \in I$ and choose two such points in a $\delta/4$ neighborhood. Then

$$|g(x_1) - g(x_1)| \geq 2\varepsilon - 2\varepsilon/3 = \frac{4}{3}\varepsilon$$

hence either

$$|g(x_1) - g(x)| \geq \frac{2}{3}\varepsilon \qquad \text{of } |g(x_2) - g(x)| \geq \frac{2}{3}\varepsilon$$

and hence

$$\max\{\frac{|g(x_1) - g(x)|}{|x - x_1|}, \frac{|g(x_2) - g(x)|}{|x_2 - x|}\} \geq \frac{varepsilon}{\delta}$$

Decreasing δ if necessary we can ensure $\varepsilon/\delta > L$.

Let

$$B = \bigcap_L C_b([0,1] \backslash A_L$$

It is dense. Its complement consist of all functions there exists $x \in [0,1]$ so that

$$\limsup_{y \to x} \frac{|f(y) - f(x)|}{|y - x|} = \infty.$$

□

4.2 Distributions: Definition

We need a preliminary result.

Lemma 4.7. *Let $U \subset \mathbb{R}^d$ be open and $k \in \mathbb{N}$. Then for every $f \in C^k_c(U)$ there exists a compact set $K \subset U$ and a sequence $f_n \in C^\infty(U)$ supported in K so that $\partial^\alpha f_n \to \partial^\alpha f$ in $C_b(U)$ for $n \to \infty$ and $|\alpha| \leq k$.*

Proof. Let $K' = \operatorname{supp} f$ and $K = \{x \in U : d(x, K') < r\}$ for some small number r so that $K \subset U$. We define $f_n = \eta_{2^{-n}} * f$ with $\eta_r(x) = r^{-d}\eta(x/r)$, $\eta \in C^\infty_c(B_1(0))$ with $\int \eta dx = 1$. □

Recall that $C^k_b(U)$ is a Banach space with the norm

$$\|f\|_{C^k_b(U)} = \sup_{|\alpha| \leq k} \|\partial^\alpha f\|_{sup}$$

Definition 4.8. *Let $U \subset \mathbb{R}^d$ be open, and $\mathcal{D}(U) = C_c^\infty(U)$ be the vector space of infinitely differentiable functions with compact support called test functions. We say $f_j \to f$ in $C_c^\infty(U) = \mathcal{D}(U)$ if there is a compact set $K \subset U$ and $\operatorname{supp} f_j \subset K$ for all j and for all multiindices α*

$$\partial^\alpha f_j \to \partial^\alpha f \qquad \text{in } C_b(U).$$

A distribution T on U is a continuous linear map from $C_0^\infty(U) \to \mathbb{K}$. We denote the space of distributions by $\mathcal{D}'(U)$.

By continuous we mean that

$$Tf_j \to Tf$$

if $f_j \to f$ in the sense of test functions. It is immediate that the distributions define a $\mathbb{K}$ vector space.

Lemma 4.9. *Let $T \in \mathcal{D}'(U)$. For every compact set $K \subset U$ there exists k and $C > 0$ so that, if $f \in \mathcal{D}(U)$ with $\operatorname{supp} f \subset K$ then*

$$|T(f)| \le C\|f\|_{C_b^k(U)}.$$

Proof. We define for $K \subset U$ compact

$$X_K = \{f \in \mathcal{D}(U) : \operatorname{supp} f \subset K\}.$$

We define a metric on X_K

$$d(f,g) = \sup_{k\ge 0} 2^{-k} \min\{1, \|f-g\|_{C_b^k(U)}\}.$$

With this metric X_K is a complete metric space:

$$d(f_n, f) \to 0 \qquad \text{iff } f_n \to f \qquad \text{in } C_b^k(U) \qquad \text{for all } k \ge 0.$$

Moreover, if f_n is a Cauchy sequence in (X_k, d) then it is a Cauchy sequence in $C_b^k(U)$ for all k, and hence it has a smooth limit $f \in \mathcal{D}(U)$.

Then $T \in \mathcal{D}'(U)$ define a continuous linear map from $X_K \to \mathbb{K}$. Moreover

$$|Tf| < \infty$$

for all $f \in X_K$. Now we argue as for the uniform boundedness principle of Banach-Steinhaus: There exists m so that the set

$$\{f \in X_K : |Tf| < m\}$$

contains an open ball. Then as above there exists $r > 0$ so that

$$|Tf| < m \qquad \text{for all } f \in X_K \quad \text{with } d(f,0) < r.$$

Let k be so that $2^{1-k} < r$. Then $f \in X_K$ with

$$\|f\|_{C_b^k(U)} < 2^{-k}$$

implies $d(f,0) < r$. Thus, if $f \in X_K$,

$$|Tf| \le 2^k m \|f\|_{C_b^k(U)}.$$

□

06.12.2019

Definition 4.10. *We say that*

$$T_n \to T \qquad \text{in } \mathcal{D}'(U)$$

if

$$T_n(f) \to T(f) \qquad \text{for all } f \in \mathcal{D}(U).$$

If $f \in \mathcal{D}(u)$ and $g \in C^\infty(U)$ then $fg \in \mathcal{D}(U)$.

Definition 4.11. *Let $\phi \in C^\infty(U)$ and $T \in \mathcal{D}'(U)$. We define their product by*

$$(\phi T)(f) = T(\phi f)$$

and the derivative

$$(\partial_{x_j} T)(f) = T(-\partial_{x_j} f).$$

It is easy to see that the right hand side of the formulas defines a distribution. We can easily calculate Leibniz' formula in the form

$$\begin{aligned} \partial_{x_j}(\phi T)(f) = -T(\phi \partial_{x_j} f) &= T((\partial_{x_j}\phi) f) - T(\partial_{x_j}(\phi f)) \\ &= [(\partial_{x_j}\phi)T](f) + (\phi \partial_{x_j} T)(f) \end{aligned}$$

and the associative and distributive law:

$$\phi(\psi T) = \psi(\phi T).$$

Similarly the theorem of Schwarz holds

$$\partial_{x_j}\partial_{x_k} T = \partial_{x_k}\partial_{x_j} T.$$

Let $L^1_{loc}(U)$ be the set of measurable functions on U which are integrable on compact subsets. We say f_j converges to f in L^1_{loc} if $f_j|_K \to f|_K$ in $L^1(K)$ for all compact subsets K.

Definition 4.12. *We define $L^1_{loc}(U) \ni f \to T_f \in \mathcal{D}'(U)$ by*

$$T_f(\phi) = \int_U f \phi \, dm^d$$

for $\phi \in \mathcal{D}(U)$.

Lemma 4.13. *The map $L^1_{loc} \to \mathcal{D}'$ is linear, continuous and injective.*

Proof. Only injectivity has to be proven. After multiplying by a characteristic function of a ball we consider $f \in L^1(B)$. Suppose that

$$\int f\phi dx = 0$$

for all $\phi \in C_c^\infty$ supported in $B_1(0)$. Then

$$f * \phi(x) = 0$$

for all ϕ supported in $B_r(0)$ and $|x| < 1 - r$. But we have seen that there is such a sequence ϕ_j so that $f * \phi_j \to f$ in L^1. Then $f|_{B_{1-r}(0)} = 0$. This implies the full statement. $\square$

Similarly, any Radon measure μ on $U \subset \mathbb{R}^d$ defines a linear map from the continuous functions with compact support to $\mathbb{K}$, and we identify it with the restriction to $\mathcal{D}$.

Lemma 4.14. *The following identities hold*

$$T_{\phi\psi} = \phi T_\psi \text{ for } \phi,\, \psi \in C(U),$$
$$T_{\partial_{x_j}\phi} = \partial_{x_j} T_\phi \text{ for } \phi \in C^1(U).$$

Proof. We use Fubini, integration by parts and the fact that $f \in \mathcal{D}(U)$ has compact support, to get

$$T_{\phi\psi}(f) = \int_U \phi\psi f\, dm^d = (\phi T_\psi)(f),$$
$$T_{\partial_{x_j}\phi} f = \int f \partial_{x_j}\phi\, dm^d = \int (-\partial_{x_j} f)\phi\, dm^d = (\partial_{x_j} T_\phi)(f).$$

$\square$

Examples:

1. The Dirac measure δ_0. It is the unique distribution so that

$$T * \phi = \phi$$

for all test functions. Let T be such a distribution. Then

$$(T - \delta) * \phi = 0$$

for all $\phi \in \mathcal{D}(\mathbb{R}^d)$ and hence

$$T(\phi) = \phi(0)$$

for all $\phi \in \mathcal{D}(\mathbb{R}^d)$.

2. The Heaviside function $H(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0 \end{cases}$ satisfies

$$\partial_x H = \delta_0.$$

Firstly $H \in L^1_{loc}$. If $\phi \in \mathcal{D}(U)$

$$(\partial_x H)(\phi) = (\partial_x T_H)(\phi) = -T_H(\phi') = -\int_0^\infty H(x)\phi'(x)dx = \phi(0) = \delta_0(\phi).$$

3. Let

$$f = \begin{cases} 1/2 & \text{if } |x| < t \\ 0 & \text{if } |x| \geq t \end{cases}$$

Then

$$\begin{aligned}
(\partial_t^2 - \partial_x^2)T_f(\phi) &= \frac{1}{2}\int_{\{|x|<t\}} (\partial_t^2 - \partial_x^2)\phi(t,x)dxdt \\
&= -\frac{1}{2}\int_0^\infty \int_{-t}^t \partial_x^2 \phi dxdt + \frac{1}{2}\int_{\mathbb{R}} \int_{|x|}^\infty \phi_{tt} \\
&= \frac{1}{2}\int_0^\infty \phi_x(t,-t) - \phi_x(t,t)dt - \frac{1}{2}\int_{\mathbb{R}} \phi_t(|x|,x)dx \\
&= \frac{1}{2}\int_0^\infty \phi_x(t,-t) - \phi_t(t,-t) - (\phi_t(t,t) + \phi_x(t,t))dt \\
&= -\frac{1}{2}\int_0^\infty \frac{d}{dt}(\phi(t,-t) + \phi(t,t))dt \\
&= \phi(0)
\end{aligned}$$

hence

$$(\partial_t^2 - \partial_x^2)\frac{1}{2}\chi_{|x|<t} = \delta_0$$

4. Let $d > 2$ and

$$g(x) = \frac{2}{d-2}\frac{\Gamma(\frac{d}{2})}{\pi^{d/2}}|x|^{2-d}$$

In *Einführung in die PDG* we have seen that

$$-\Delta g * \phi = \int g(x-y)(-\Delta\phi(y))dy = \phi(x).$$

Thus

$$-\Delta T_g = \delta_0.$$

5. Let $\eta \in C^\infty(\mathbb{R}^d)$ with $\int \eta dx = 1$ and $\operatorname{supp} \eta \in B_1(0)$, $\eta_r(x) = r^{-d}\eta(x/r)$. Then

$$\eta_r \to \delta_0 \qquad \text{as } r \to 0$$

if $x \in U$. This holds already in the dual space of $C_b(U)$.

6. Now let $\eta_r : \mathbb{R}^m \to \mathbb{R}$ as above, $\phi \in C^1(U;\mathbb{R}^m))$ with $\nabla\phi \neq 0$. Then, using the coarea formula

$$\begin{aligned} T_{\eta_r\circ\phi}(\psi) &= \int_U \eta_r(\phi)\psi dx \\ &= \int_{\mathbb{R}^m} \eta_r(y) \int_{\phi^{-1}(y)} \det(D\phi D\phi^T)^{-1/2}\psi(x) d\mathcal{H}^{d-m}(x) dy \\ &\to \int_{\phi^{-1}(0)} \det(D\phi D\phi^T)^{-1/2}\psi(x) d\mathcal{H}^{d-m}(x) \\ &=: \delta_\phi(\psi) \end{aligned}$$

Of particular importance is the case $m = d-1$. Then

$$\det(D\phi D\phi^T) = |\nabla\Phi|^2.$$

7. Now recall Kirchhoff's formula for the wave equation for $d = 3$: Consider

$$u_{tt} - \Delta u = f, \qquad u(0,x) = u_t(0,x)) = 0$$

and let $t > 0$. Then, using the area formula

$$\begin{aligned} u(t,x) =& \frac{1}{4\pi} \int_{B_t(0)} \frac{1}{|y|} f(t-|y|, x-y) dm^3(y) \\ =& \frac{1}{4\pi\sqrt{2}} \int_{\{|t-s|^2=|x-y|^2, 0<s<t\}} (t-s)^{-1} f(s,y) d\mathcal{H}^3(s,y) \\ =& \frac{1}{2\pi} \delta_{(t-s)^2-|x-y|^2}(f) \end{aligned}$$

as first for f compactly supported $(0,t)\times\mathbb{R}^3$, but, using the formulas, for all $\psi \in C_c^\infty(\mathbb{R}^d)$ (but multiplying by $\chi_{t>0}$). Then

$$u = \frac{1}{2\pi}\chi_{t>0}\delta_{s^2-|y|^2} * f.$$

and hence

$$(\partial_{tt}^2 - \Delta)[\chi_{t>0}\delta_{s^2-|y|^2}] = 2\pi\delta_0.$$

This we want to verify by a direct calculation. We recall that $\eta_R(S) = r^{-1}\eta(s/r)$ and compute with $s = t^2 - |x|^2$

$$\begin{aligned} (\partial_t^2 - \Delta)\eta_r(t^2-|x|^2) &= 2\partial_t(\eta_r'(t^2-|x|^2)t) + 2\sum_j \partial_{x_j}[\eta_r'(t^2-|x|^2)x_j] \\ &= 4(t^2-|x|^2)\eta_r^{(2)}(t^2-|x|^2) + 8\eta_r'(t^2-|x|^2) \\ &= r^{-2}(4(s/r)\eta''(s/r) - 8\eta'(s/r)) \end{aligned}$$

with $s = t^2 - |x|^2$. Now, with a smooth function h

$$\begin{aligned}\int r^{-2}(4(s/r)\eta''(s/r) - 8\eta'(s/r))h(s)ds &= 4r^{-1}\int ((s\eta'(s))' - 4\eta'(s))h(rs)ds \\ &= -4\int (s\eta'(s) - \eta(s))h'(rs)ds \\ &= 4r\int s\eta(s)h''(rs)ds \\ &\to 0\end{aligned}$$

as $r \to 0$. We set for $\phi \in \mathcal{D}(\mathbb{R}^d)$ with support in $t > 0$

$$h(s) = \int_{(t-\tau)^2 - |x-y|^2 = s} |t-\tau|^{-1}\phi d\mathcal{H}^3(\tau, y)$$

and apply the coarea formula. Thus

$$\operatorname{supp}(\partial_t^2 - \Delta)\chi_{t>0}\delta_{t^2-|x|^2} \subset \{0\}.$$

in the sense that the distribution applied to ϕ vanishes unless 0 is in the suppport of ϕ.

Let $\lambda > 0$. We trace the effect of the scaling

$$(\partial_t^2 - \Delta)\delta_{t^2-|x|^2}(\phi(\lambda.)) = (\partial_t^2 - \Delta)\delta_{t^2-|x|^2}(\phi(.)).$$

Together with the support property (EPDE, exercises) this implies that there exists c so that

$$(\partial_t^2 - \Delta)\chi_{t>0}\delta_{t^2-|x|^2} = c\delta_0$$

and we have to determine δ. By a small abuse of notation (h does not have compact support in x, and we should multiply by a function of x with compact support, identically 1 on a large ball) we get

$$\begin{aligned}(\partial_t^2 - \Delta u)\delta_{t^2-|x|^2}(h(t)) =& \delta_{t^2-|x|^2}(h''(t)) \\ =& 2\pi\int_0^\infty th''(t)dt \\ =& -2\pi\int_0^\infty h'(t)dt \\ =& 2\pi h(0).\end{aligned}$$

and hence $c = 2\pi$.

11.12.2019

Definition 4.15. *Let $T \in \mathcal{D}'(U)$. We say that T vanishes near $x \in U$ if there exists $r > 0$ so that $T(f) = 0$ for all $f \in \mathcal{D}(U)$ with support in $B_r(x)$. We define the support of T as the complement of the points near which T vanishes.*

Lemma 4.16. *Let $\phi \in \mathcal{D}(U)$ and $T \in \mathcal{D}'(U)$ with disjoint supports. Then*

$$T(\phi) = 0.$$

Proof. Let K be the support of ϕ. Given $x \in K$ there exists r so that $T\psi = 0$ for every $\psi \in \mathcal{D}(U)$ supported in $B_r(x)$. Since K is compact there is a finite covering of such balls $B_{r_j}(x_j)$ with $1 \le j \le N$. We choose a partition of unity $\eta_j \in C^\infty(U)$ supported in $B_{r_j}(x_j)$ so that

$$\sum_{j=1}^N \eta_j(x) = 1$$

for $x \in K$. Then

$$T\phi = \sum_{j=1}^N T(\eta_j \phi) = 0$$

□

Definition 4.17. *Let $T \in \mathcal{D}'(\mathbb{R}^d)$ and $\phi \in \mathcal{D}(\mathbb{R}^d)$. We define their convolution by*

$$(\phi * T)(x) = T(\phi(x - \cdot)), \quad \forall x \in \mathbb{R}^d.$$

The righthand side denotes T acting on $\phi(x-y)$ as a function of y.

Lemma 4.18. *With the notation above, $\phi * T \in C^\infty(\mathbb{R}^d)$ and*

$$\partial_{x_j}(\phi * T) = (\partial_{x_j}\phi) * T = \phi * (\partial_{x_j} T).$$

If $\psi \in L^1(\mathbb{R}^d)$ then

$$\phi * T_\psi(x) = \phi * \psi(x).$$

Moreover, if $\operatorname{supp} \phi = K_1$ and $\operatorname{supp} T = K_2$ then

$$\operatorname{supp} \phi * T \subset K_1 + K_2 = \{x + y : x \in K_1 \text{ and } y \in K_2\}.$$

Proof. For $v \in \mathbb{R}^d$, we have to prove that

$$\begin{aligned}
\frac{(\phi * T)(x+tv) - (\phi * T)(x)}{t} &= T(\frac{1}{t}(\phi(x+tv-\cdot) - \phi(x-\cdot))) \\
\to T\left(\sum_{j=1}^d v_j (\partial_j \phi)(x - \cdot)\right) &= \Big(\sum_{j=1}^d v_j \partial_j \phi\Big) * T(x) \\
\equiv -T\left(\sum_{j=1}^d v_j \partial_j (\phi(x-\cdot))\right) &= \phi * \Big(\sum_{j=1}^d v_j \partial_j T\Big)(x).
\end{aligned}$$

This is a consequence of Lemma 4.9 and that the difference quotient

$$\frac{(\partial^\alpha \phi)(x+tv) - (\partial^\alpha \phi)(x)}{t} \to \sum_{j=1}^d v_j \partial_j (\partial^\alpha \phi)(x) \text{ as } t \to 0$$

uniformly in x since the support is compact.

The remaining properties are not hard to verify. (Exercise) □

Remark 4.19. *We can equivalently define the convolution of* $\phi \in \mathcal{D}(\mathbb{R}^d)$ *and* $T \in \mathcal{D}'(\mathbb{R}^d)$ *as the distribution:*

$$(\phi * T)(f) = T(\tilde{\phi} * f), \quad \forall f \in \mathcal{D}(\mathbb{R}^d),$$

where $\tilde{\phi}(x) = \phi(-x)$ *since*

$$T_{\phi * T}(\psi) = \int T(\phi(x - .))\psi(x)dx = T(\int \phi(x-.)\psi(x))dx = T(\tilde{\phi} * \psi)$$

by linearity, a Riemann type approximation of the integral and a limit.

In the same way we can easily define the convolution $\phi * T \in \mathcal{D}'(\mathbb{R}^d)$ *when* $\phi \in C^k(\mathbb{R}^d)$ *has compact support.*

By an abuse of notation we write the convolution evaluated at x *whenever it is defined, even if it is only defined on a subset of* $\mathbb{R}^d$.

Definition 4.20. *Let* $T \in \mathcal{D}'(\mathbb{R}^d)$ *and let* $S \in \mathcal{D}'(\mathbb{R}^d)$ *with compact support. We define their convolution by*

$$(S * T)(\phi) = T(\tilde{S} * \phi)$$

for $\phi \in \mathcal{D}(\mathbb{R}^d)$. *Here* $\tilde{S}(\psi) = S(\tilde{\psi}), \quad \psi \in \mathcal{D}(\mathbb{R}^d)$.

It is an exercise to formulate and prove reasonable properties of the convolution of distributions.

Lemma 4.21. *Suppose that* U *is connected and* $\partial_j T = 0$, $j = 1, \cdots, d$. *Then there exists a constant* c *so that* $T = T_c$.

Proof. The statement is clear if $T = T_f$ with $f \in C^1$. If $\partial_j T = 0$ then

$$\partial_j (T * \phi) = (\partial_j T) * \phi = 0,$$

whenever this is defined. So $T * \phi$ is constant. Since $\eta_r * T \to T$ in $\mathcal{D}'$ and since $\eta_r * T$ is constant, the same is true for T. □

Lemma 4.22. *Suppose that* $T_n \to T$ *in* $\mathcal{D}'(U)$ *and that* $K \subset U$ *is compact. Then there exists* k *and* C *so that*

$$\sup_n |T_n(f)| \le C \|f\|_{C^k_b(U)}$$

for all $f \in X_K = \{f \in \mathcal{D}(U) : suppf \subset K\}$ *and*

$$\sup\{|T_n(f) - T(f)| : f \in X_K, \|f\|_{C^k_b} \le 1\} \to 0$$

as $n \to \infty$.

Proof. The proof of the first part is the same is for Lemma 4.9. So, given K, there exist $k \geq 1$ and C so that for any n

$$|T_n(f)| \leq C\|f\|_{C_b^{k-1}}.$$

We define the closed subspace

$$X_k = \{f \in C_b^k(U) : \operatorname{supp} f \subset K\}.$$

Then

$$X_k \ni f \to (\partial^\alpha f)_{|\alpha| \leq k-1} \in C_b^{0,1}(K, \mathbb{K}^{\#\Sigma_{k-1}})$$

has norm at most 1. By a multiple application of Theorem 3.32 the closed unit ball in $C_b^{0,1}(K), \mathbb{K}^{\#\Sigma_{k-1}})$ is compact in $C_b(K), \mathbb{K}^{\#\Sigma_{k-1}})$. Given $\varepsilon > 0$ there exist finitely many functions $f_m \in C_c^\infty(U)$ with the support in K and $\|f_m\|_{C_b^k(U)} \leq 1$ so that the ε balls in C_b^{k-1} centered at f_m cover the closure of such functions in $C_b^{k-1}(U)$. By the assumption there exists n_0 so that

$$|(T_n - T)(f_m)| < \varepsilon$$

if $n \geq n_0$. Then, for any $f \in \bar{B}_1(0)$ there exists f_m such that $\|f - f_m\|_{C_b^{k-1}} < \varepsilon$ and hence for $n \geq n_0$ we have

$$|(T_n - T)f| \leq |(T_n - T)(f_m)| + |T_n(f - f_m)| + |T(f - f_m)| \leq \varepsilon + 2C\varepsilon.$$

□

Theorem 4.23. *Let $U \subset \mathbb{R}^d$ be open. Then $\mathcal{D}(U) \subset \mathcal{D}'(U)$ is dense.*

Proof. There are several steps.
Step 1: Distributions with compact support are dense. We choose a sequence $\phi_j \in \mathcal{D}(U)$ so that

$$\partial^\alpha \phi_j \to \partial^\alpha 1$$

on compact subsets. Then $\phi_j T$ has compact support and

$$(\phi_j T)(g) = T(\phi_j g) \to Tg$$

for all $g \in \mathcal{D}(U)$.

Step 2: Construction of the ϕ_j. Let K_j be a monotone sequence of compact sets so that K_j is contained in the interior of K_{j+1} and $U = \bigcup K_j$. Then for any j there exists $r_j > 0$ so that

$$\min\{\operatorname{dist}(K_{j-1}, \mathbb{R}^d \backslash K_j), \operatorname{dist}(K_j, \mathbb{R}^d \backslash K_{j+1})\} \geq r_j > 0.$$

Let $\eta \in C_c^\infty(B_1(0))$ be radial with $\int \eta dx = 1$ and let $\eta_r(x) = r^{-d}\eta(x/r)$. Then

$$\phi_j = \eta_{r_j} * \chi_{K_j} \in C^\infty,$$

$$\operatorname{supp}\phi_j \subset K_{j+1}, \quad \phi_j = 1 \text{ on } K_{j-1}.$$

Step 3 Let $T \in \mathcal{D}'(U)$ have compact support. Then, for r small and $g \in \mathcal{D}(U)$

$$\eta_r * T(g) = T(\tilde{\eta_r} * g) \to T(g), \quad r \to 0$$

by an abuse of notation. □

The same argument gives Lemma 4.7.

4.3 Schwartz functions and tempered distributions

We briefly cover the definition of Schwartz functions and tempered distributions, which are the proper frame work for the Fourier transform.

Definition 4.24. *The Schwartz space $\mathcal{S}(\mathbb{R}^d)$ consists of all Schwartz functions, which are functions f so that for all multiindices α, β*

$$\|x^\alpha \partial^\beta f\|_{sup} < \infty.$$

We say $f_j \to f$ as Schwartz functions if for all multiindices α and β $x^\alpha \partial^\beta f_j \to x^\alpha \partial^\beta f$ uniformly.

Remark 4.25. *It is easy to see that if $f \in \mathcal{S}(\mathbb{R}^d)$, then for any N there exists C_N such that $|f(x)| \le C_N(1+|x|)^{-N}$, and hence $f \in L^p(\mathbb{R}^d)$, for any $p \in [1,\infty]$. So is $\partial^\alpha f$ for any multiindex α.*

Roughly speaking, the Schwartz functions have two properties: they have infinite bounded derivatives and they decay fast at infinity. Recalling (with $\hat{}$ denoting the Fourier transform) $\widehat{\partial_{x_j} f}(\xi) = 2\pi i \xi_j \hat{f}(\xi)$ and $\widehat{x_j f}(\xi) = \frac{i}{2\pi}\partial_{\xi_j}\hat{f}(\xi)$, the Fourier transform works well in the framework of Schwartz space and tempered distributions (see below).

Since $\eta_n f \to f$ in $\mathcal{S}(\mathbb{R}^d)$, $\eta_n(x) = \eta(n^{-1}x)$ where $\eta \in C_c^\infty(B_2(0))$ takes value 1 on $B_1(0)$, the inclusion $\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$ is dense.

Lemma 4.26. *Let f be a Schwartz function.*

1. *If α is a multiindex then $\partial^\alpha f \in \mathcal{S}(\mathbb{R}^d)$.*
2. *If $g \in C^\infty$ and for any multiindex α there exist $c_{|\alpha|}$ and $\kappa_{|\alpha|}$ so that*

   $$|\partial^\alpha g| \le c_{|\alpha|}(1+|x|)^{\kappa_{|\alpha|}}$$

   then $gf \in \mathcal{S}(\mathbb{R}^d)$.
3. *If $g \in C(\mathbb{R}^d)$ satisfies for any multiindex α*

   $$\|x^\alpha g(x)\|_{sup} < \infty,$$

   then $g * f \in \mathcal{S}(\mathbb{R}^d)$.

Proof. The first property follows from the definition. By the first property, in order to prove the second property it suffices to show

$$\sup |x^\alpha (\partial^\beta g) f| < \infty,$$

which follows from the definition.

Since

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

and since $\partial^\alpha f$ is Schwartz, by the first property the proof of the third property is reduced to bounding

$$\|x^\alpha (g * f)\|_{sup}.$$

We observe that

$$x_j (g * f) = (x_j g) * f + g * (x_j f)$$

and the claim follows by induction on the length of α. □

Definition 4.27. *We define $d : \mathcal{S} \times \mathcal{S} \to [0, \infty)$ by*

$$d(f, g) = \sup_k 2^{-k} \min\{1, \sup_{|\alpha|+|\beta|=k} \|x^\alpha \partial^\beta (f - g)\|_{sup}\}.$$

Lemma 4.28. *The expression $d(f, g)$ defines a metric on $\mathcal{S}$ which turns it into a complete metric space.*

Proof. Easy exercise. □

Definition 4.29. *A tempered distribution is a continuous linear map from $\mathcal{S}(\mathbb{R}^d)$ to $\mathbb{K}$. We denote the space of tempered distributions by $\mathcal{S}'(\mathbb{R}^d)$. We say that $T_j \to T$ in $\mathcal{S}'(\mathbb{R}^d)$ if $T_j(\phi) \to T(\phi)$ for all $\phi \in \mathcal{S}(\mathbb{R}^d)$.*

Lemma 4.30. *Let $T \in \mathcal{S}'(\mathbb{R}^d)$. Then there exist k and c so that*

$$|T\phi| \le c \sup_{|\alpha|+|\beta| \le k} \|x^\alpha \partial^\beta \phi\|_{sup}.$$

If $T_n \to T$ in $\mathcal{S}'(\mathbb{R}^d)$ then there exist C and k so that

$$\sup_n |T_n \phi| \le C \sup_{|\alpha|+|\beta| \le k} \|x^\alpha \partial^\beta \phi\|_{sup}$$

and

$$\frac{|T_n(\phi) - T\phi|}{\sup_{|\alpha|+|\beta| \le k} \|x^\alpha \partial^\beta \phi\|_{sup}} \to 0.$$

Proof. The proof is similar as that of Lemma 4.22. The existence of C and k follows from the idea of the proof of Banach-Steinhaus theorem. The convergence result follows from the compactness of the ball $\{\phi \in \mathcal{S}(\mathbb{R}^d) \mid \sup_{|\alpha|+|\beta|\le k} \|x^\alpha \partial^\beta \phi\|_{sup} \le 1\}$ in the space $\{\phi \in \mathcal{S}(\mathbb{R}^d) \mid \sup_{|\alpha|+|\beta|\le k-1} \|x^\alpha \partial^\beta \phi\|_{sup} < +\infty\}$, which is easy to see if we notice that

$$\sup_{|\alpha|+|\beta|\le k-1} \|x^\alpha \partial^\beta \phi\|_{sup} \le R^{-1} \sup_{|\alpha|+|\beta|\le k} \|x^\alpha \partial^\beta \phi\|_{sup((B_R(0))^C)} + R^{k-1} \|\phi\|_{C_b^{k-1}(B_R(0))}$$

and we can choose R big enough. □

Remark 4.31. *We define the derivative and the multiplication by a smooth function with controlled derivatives for a tempered distribution as we did it for distributions. Similarly, since compactly supported distributions S can act on Schwartz functions f, we can define the convolution $(S * f)(x) \in \mathcal{S}(\mathbb{R}^d)$. We then can define the convolution of a tempered distribution with Schwartz functions and with compactly supported distributions.*

Let $1 \le p \le \infty$. There are the embeddings

$$\mathcal{D}(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d) \subset L^p(\mathbb{R}^d) \subset \mathcal{S}'(\mathbb{R}^d) \subset \mathcal{D}'(\mathbb{R}^d).$$

The embeddings are dense if $p < \infty$.

4.4 Sobolev spaces: Definition

Definition 4.32. *Let $U \subset \mathbb{R}^d$ be open, $k \in \mathbb{N}$ and $1 \le p \le \infty$. The Sobolev space $W^{k,p}(U) \subset L^p(U)$ is the set of all $L^p(U)$ functions, so that for all multiindices α of length at most k there exists $f_\alpha \in L^p(U)$ so that*

$$\partial^\alpha T_f = T_{f_\alpha}.$$

We define (identifying T_f and f and $\partial^\alpha T_f$ resp $\partial^\alpha f$ with f_α by an abuse of notation)

$$\|f\|_{W^{k,p}} = \left(\sum_{|\alpha|\le k} \|\partial^\alpha f\|^p_{L^p(U)} \right)^{\frac{1}{p}}$$

with the usual modification if $p = \infty$.

We have

$$g = \partial_{x_j} f, \quad f, g \in L^p(U)$$

if and only if

$$\int g\phi dm^d = -\int f \partial_{x_j} \phi dm^d$$

for all $\phi \in \mathcal{D}(U)$.

Lemma 4.33. *Let $g \in C_b^k(U)$ and $f \in W^{k,p}(U)$. Then $gf \in W^{k,p}(U)$.*

Proof: Easy exercise.

Definition 4.34. *Let $k \in \mathbb{N}$, $1 \le p < \infty$. We define $W_0^{k,p}(U)$ as the closure of $C_c^\infty(U)$ with respect to the norm $\|\cdot\|_{W^{k,p}(U)}$. If $V \subset U$ the extension defines a canonical (nonsurjective) isometry from $W_0^{k,p}(V)$ to $W_0^{k,p}(U)$.*

Theorem 4.35. *Let $k \in \mathbb{N}$ and $1 \le p \le \infty$. The Sobolev space $W^{k,p}(U)$ is a Banach space. If $V \subset U$ then the restriction defines a map of norm 1 from $W^{k,p}(U)$ to $W^{k,p}(V)$. Moreover $W_0^{k,p}(\mathbb{R}^d) = W^{k,p}(\mathbb{R}^d)$ if $1 \le p < \infty$.*

Proof. Let $\Sigma_k = \{\alpha : |\alpha| \le k\}$. There is an obvious isometry

$$W^{k,p}(U) \ni f \to (f_\alpha)_{|\alpha|\le k} \in L^p(U \times \Sigma_k).$$

Let f_j be a Cauchy sequence in $W^{k,p}(U)$ with limit $f \in L^p(U)$. Then $\partial^\alpha f_j \to f_\alpha$ in $L^p(U)$ and in $\mathcal{D}'(U)$. It is easy to check that $f_\alpha = \partial^\alpha f$ in $\mathcal{D}'(U)$: for any $\phi \in \mathcal{D}(U)$,

$$T_{f_\alpha}(\phi) = \lim_{j\to\infty} (-1)^{|\alpha|} \int_U f_j \partial^\alpha \phi \, dm^d = (-1)^{|\alpha|} \int_U f \partial^\alpha \phi \, dm^d = T_{\partial^\alpha f}(\phi).$$

Thus $f \in W^{k,p}(U)$ and $W^{k,p}(U)$ is complete and we can identify $W^{k,p}(U)$ with a closed subspace of $L^p(U \times \Sigma_k)$.

The restriction map with norm ≤ 1 follows from the definition, and the norm is indeed 1 since

$$\|f\|_{W^{k,p}(V)} = \|\tilde{f}\|_{W^{k,p}(U)}, \quad f \in W_0^{k,p}(V),$$

and $\tilde{f} \in W_0^{k,p}(U)$ is the trivial extension of f by 0.

Now let $1 \le p < \infty$. Density of $\mathcal{D}(\mathbb{R}^d) \subset W^{k,p}(\mathbb{R}^d)$ follows by the same argument as in the proof of Theorem 4.23. □

13.12.2019

Definition 4.36. *Let $k \in \mathbb{N}$ and $1 < p < \infty$. We define $W^{-k,p}(U) = (W_0^{k,p'}(U))^*$.*

We consider $W_0^{k,p'}(U)$ as a subset of the space of distributions $\mathcal{D}'(U)$

Lemma 4.37. *The map J: $L^p(U \times \Sigma_k) \to W^{-k,p}(U)$ defined by*

$$J((f_\alpha))(u) = \sum_{|\alpha|\le k} \int_U f_\alpha \partial^\alpha u \, dm^d, \quad \forall u \in W_0^{k,p'}(U)$$

has norm ≤ 1.

This follows from Hölder's inequality.

Proof. Exercise. □

Lemma 4.38. *Let $U \subset \mathbb{R}^d$ be bounded and open and $k \in \mathbb{N}$. We assume that $f \in C_b^k(U)$ and its derivatives of order up to $k-1$ extend to continous functions in $\bar{U}$ which vanish at ∂U. Then $f \in W_0^{k,p}(U)$ for $1 \le p < \infty$.*

Proof. We proceed as in Theorem 4.23. Let $K_j = \{x \in U : \operatorname{dist}(x, \partial U) \ge 2^{-j}\}$. We extend χ_{K_j} by 0 to $\mathbb{R}^d$ and convolve it with a smooth function (say $\varphi_{2^{-j-1}}$) of integral 1 supported in $B_{2^{-j-1}}(0)$, to obtain $\eta_j \in C_c^\infty(U)$. Then $\operatorname{supp} \eta_j \subset \{x \in U : \operatorname{dist}(x, \partial U)\} \le 2^{-j-1}$ and $\eta_j(x) = 1$ for $\operatorname{dist}(x, \partial U) \ge 2^{1-j}$ and

$$|\partial^\alpha \eta_j| \le c(|\alpha|) 2^{|\alpha| j}.$$

Since for $|\alpha| \le k$, by the Taylor formula,

$$|\partial^\alpha f(x)| \le c \operatorname{dist}(x, \partial U)^{k-|\alpha|}.$$

Thus the sequence $\eta_j f$ is uniformly bounded in $C_b^k(U)$, and hence in $W^{k,p}(U)$. Moreover

$$\partial^\alpha(\eta_j f) \to \partial^\alpha f$$

for every $x \in U$ and by dominated convergence

$$\eta_j f \to f \qquad \text{in } W^{k,p}(U).$$

We complete the proof by regularizing $\eta_j f$ as in Lemma 4.7. □

The cofactor matrix $\operatorname{cof} A$ of an $n \times n$ matrix has as (i,j) entry $(-1)^{i-j}$ times the determinant of the $(n-1) \times (n-1)$ matrix obtained from A by removing the ith row and the jth column. It is the same as the partial derivative of $\det(A)$ with respect to the (i,j)th entry. Then (linear algebra)

$$A^T \operatorname{cof} A = \det A 1_{n \times n} \tag{4.1}$$

Lemma 4.39. *Let $U \subset \mathbb{R}^d$ be open and $\phi \in C^2(U; \mathbb{R}^d)$. Then*

$$\sum_{j=1}^d \partial_{x_j} \operatorname{cof}(D\phi)_{ij} = 0.$$

Proof. There are two very different proofs. Suppose that $\phi \in C_b^2(W)$ for some $\overline{U} \subset W$ is a C^2 diffeomorphism to its image with $\det D\phi > 0$. Let $V = \phi(U)$. By the transformation formula

$$m^d(V) = \int_U \det D\phi dx.$$

The left hand side depends only on ∂V. Let $\psi \in C_c^\infty(U)$, t small and

$$\phi_t(x) = \phi(x) + t\psi(x) e_j$$

If t is small this is a diffeomorphism and

$$m^d(V) = \int_U \det D\phi_t dx.$$

Hence

$$0 = \frac{d}{dt}\int_U \det D\phi_t dx|_{t=0} = \int_U \sum_{k=1}^d (\operatorname{cof} D\phi)_{jk}\partial_k \psi dx = -\int_U \sum_{k=1}^d \partial_k (\operatorname{cof} D\phi)_{jk} dx$$

which holds for all test functions. This implies the lemma. □

Lemma 4.40. *If $\phi : V \to U$ is a C^k_b diffeomorphism (bounded derivatives of ϕ and ϕ^{-1}) then there exists $C > 1$ so that*

$$\|f \circ \phi\|_{W^{k,p}(V)} \le C\|f\|_{W^{k,p}(U)}.$$

Moreover the chain rule holds

$$\partial_{y_j}(f \circ \phi) = \sum_{k=1}^d (\partial_{x_k} f \circ \phi)\partial_{y_j}\phi_k.$$

Proof. The first claim follows from the chain rule and the transformation formula. We prove the chain rule for a smooth diffeomorphism. The general case follows by approximating the diffeomorphism and taking limits.

There are two very different proofs of the chain rule. The chain rule holds for C^1 functions. Taking limits it holds for functions in $W^{1,p}_0(U)$.

Given $x \in U$ is suffices to verify it in a small ball around x. If we multiply f by a mooth function $\psi \in \mathcal{D}(U)$ we get as above $f\psi f \in W^{k,p}_0(\mathbb{R}^d)$ (extending by 0, if $p < \infty$). By the previous argument the chain rule holds for this product, and hence for f.

The other possibility is to prove the chain rule for distributions. Again we could argue by approximation, but it is more elegant to do a direct argument. We want to prove for $f \in W^{1,p}(U)$. It is easier to give a formal proof first

$$\begin{aligned}
&\int_V \partial_{x_i}(f\circ\Phi)\psi\circ\Phi - \sum_{k=1}^d (\partial_k f)\circ\Phi \frac{\partial\Phi^k}{\partial x_i} dx \\
&\qquad = -\int_V f\circ\Phi\partial_{x_i}\psi\circ\Phi dx + \int_U (\det D\Phi)^{-1}\sum_{k=1}^d \frac{\partial\Phi^k}{\partial x_i})\partial_{y_k} f\, dy \\
&\qquad = \int_U (\det D\Phi)^{-1}\sum_{k=1}^d \partial_k\psi\Big(-\frac{\partial\Phi^k}{\partial x_i} + \frac{\partial\Phi^k}{\partial x_i}\Big) dy \\
&\qquad = 0
\end{aligned}$$

The left hand side first has to be understand as application of the distribution, the first identity as the definition of the distribution, the second as an integration by parts and the transformation formula. Here Ψ has compact support, and hence there are no boundary terms. The final identity then implies the chain rule. □

4.5 (Whitney) extension and traces

Definition 4.41. *Let $U \subset \mathbb{R}^d$ be open, bounded and connected. We say that U is a Lipschitz domain if there exist a continuous vector field ν on ∂U and a Lipschitz continuous function ρ and $c > 0$ so that $\partial U = \rho^{-1}(\{0\})$ and*

$$\rho(x + t\nu) - \rho(x + s\nu) \geq t - s$$

for $x \in \partial U$ and $-c < s < t < c$.

Examples:

1. Bounded connected open sets with C^1 boundary.
2. Let $h : \mathbb{R}^{d-1} \to \mathbb{R}$ be Lipschitz continuous with Lipschitz constant L. The set below the graph is not compact, but the other conditions are satisfied with $\rho(x) = x_d - h(x_1, \cdots, x_{d-1})$ and $\nu = e_d$.

Theorem 4.42 (Whitney)**.** *Let $1 \leq p < \infty$ and $U \subset \mathbb{R}^d$ be a Lipschitz domain. Then there exists a continuous linear extension map*

$$E : W^{k,p}(U) \to W^{k,p}(\mathbb{R}^d)$$

with $Ef|_U = f$ for all $f \in W^{k,p}(U)$.

The extension from f to g is characterized by $f = g|_U$.

18.12.2019

Proof. We prove the theorem under the stronger assumption that ∂U is a C^k manifold. By use of rotation, compactness and partition of unity, it suffices to consider the extension problem for $U = \{x : x_d < \psi(x_1, \dots, x_{d-1})\}$ where ψ is a function in C_b^k. By Lemma 4.40 we can choose $\phi(x) = (x_1, \dots, x_{d-1}, x_d - \psi(x_1, x_2, \dots x_{d-1}))$ to reduce the problem to extending Sobolev functions on the lower half space $V = \{x | x_d \leq 0\}$. Let f be defined on V.

We prove the theorem first for $d = 1$ and make the Ansatz

$$F(x) = \begin{cases} f(x) & \text{if } x_d \leq 0 \\ \sum_{j=1}^{k+1} a_j f(-jx_d) & \text{if } x > 0. \end{cases}$$

If $f \in C^k(-\infty, 0])$ we want to choose the a_j so that $F^{(l)}$ is continuous for $j \leq k$. We evaluate the left and the right limit of $F^{(l)}$ at $x = 0$: From the left we obtain $f^{(l)}(0)$ and from the right

$$(-1)^l \sum_{j=1}^{k+1} a_j j^l f^{(l)}(0)$$

which leads to the linear system with Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & 2 & 3 & \dots & k+1 \\ 1^2 & 2^2 & 3^2 & \dots & (k+1)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1^k & 2^k & 3^k & \dots & (k+1)^k \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_{k+1} \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \\ \dots \\ (-1)^k \end{pmatrix}.$$

The Vandermonde matrix is invertible and we can solve this system. The coefficients a_j hence exist and depend only on k. Then

$$\|F\|_{W^{k,p}(\mathbb{R})} \leq C\|f\|_{W^{k,p}((-\infty,0))}.$$

Now we would like to pass from the assumption $f \in C^k(U)$ to $f \in W^{k,p}(U)$. We define the extension F in the same way. Then we have to prove that for $l \leq k$ the distributional derivative is given by the distribution defined by the distributional derivatives on both sides. Let $1 \leq l \leq k$ and $\phi \in \mathcal{D}(\mathbb{R})$. Then, with the same Ansatz

$$\partial^l F = \begin{cases} \partial^l f & \text{if } x < 0 \\ \sum_{j=1}^{k+1} (-1)^l a_j j^l (f^{(l)}(-jx)) & \text{if } x > 0 \end{cases}$$

By an abuse of notation we denote by $F^{(l)}$ the function in L^p defined by these formulas when $x \neq 0$. This suffices since $x = 0$ has measure 0.

We obtain

$$\begin{aligned} (-1)^\kappa \int_{\mathbb{R}} F \frac{d^l}{dx^l} \phi dx &= + (-1)^l \sum_{j=1}^{k+1} a_j \int_0^\infty f(-jx) \frac{d^l}{dx^l} \phi(x) dx \\ &= \int_{-\infty}^0 f \frac{d^l}{dx^l} \Big[(-1)^l \phi - \sum_{j=1}^{k+1} j^{l-1} a_j \phi(-x/j) \Big] dx \\ &= \int_{-\infty}^0 \frac{d^l}{dx^l} f(x) [\phi(x) - (-1)^l \sum_{j=1}^{k+1} a_j j^{l-1} \phi(-x/j)] dx \\ &= \int_{\mathbb{R}} (\frac{d^l}{dx^l} F) \phi(x) dx. \end{aligned}$$

where the second equality follows from substitution and the third since

$$\frac{d^l}{dx^l}((-1)^l\phi(x) - \sum_{j=1}^{k+1} a_j j^{l-1}\phi(-x/j))\Big|_{x=0} = 0$$

for $x=0$ and $l \le k$ by the definition of the a_j: Indeed, we observe that

$$\Big(1 + (-1)^{l-1}\sum_{j=1}^{k+1} a_j(-j)^{l-1}\Big) = 0.$$

Hence it is in $W_0^{\kappa,p}((-\infty,0))$ by Lemma 4.38 and we can approximate it by functions in $\mathcal{D}((-\infty,0))$ if $p<\infty$. For those we can move the derivatives to f. The last equality is again a consequence of the transformation formula.

This also suffices for $p=\infty$, and for $d>1$ since we may treat the other independent variables as parameters. □

Corollary 4.43. *Suppose that U is an open, bounded with Lipschitz boundary, $k \in \mathbb{N}$ and $1 \le p < \infty$. Then the restrictions of $C_c^\infty(\mathbb{R}^d)$ functions is dense in $W^{k,p}(U)$.*

Theorem 4.44 (Traces)**.** *Let U be a bounded domain with C^1 boundary and let $f \in W^{1,p}(U)$, $1 \le p \le \infty$. Then there is a unique trace $Tf \in L^p(\partial U)$ so that*

$$\int_{\partial U} \sum_{j=1}^d F^j \nu^j Tf d\mathcal{H}^{d-1} = \int_U f \sum_{j=1}^d \partial_{x_j} F^j dm^d + \int_U \sum_{j=1}^d \partial_{x_j} f F^j dm^d, \quad (4.2)$$

for all $F^j \in C^1(\overline{U})$ and ν denotes the outer normal vector of ∂U. It satisfies

$$\|Tf\|_{L^p(\partial U)} \le c\|f\|_{L^p(U)}^{\frac{p-1}{p}} \|Df\|_{L^p}^{\frac{1}{p}}, \quad \|Df\|_{L^p} := \|\, |(\partial_{x_j} f)|\, \|_{L^p}.$$

We write by an abuse of notation $Tf = f|_{\partial U}$.

Proof. If $f \in C^1(\overline{U})$ then $g = f|_{\partial U}$ is obviously the trace. We fix F with $F\cdot\nu = 1$ and apply the divergence formula Then, for if $1<p<\infty$,

$$\begin{aligned} \|f\|_{L^p(\partial U)}^p = & \int |f|^p d\mathcal{H}^{d-1} = \int_{\mathbb{R}^{d-1}} |f|^p F\cdot \nu d\mathcal{H}^{d-1} \\ = & \int_U |f|^p \nabla\cdot F dx + p\int_U |f|^{p-1}\bar{f}\nabla f \cdot F dx \\ \le & \sup|\nabla\cdot F| \|f\|_{L^p(U)}^p + \sup|F| \int |f|^{p-1}|\nabla f| dx \end{aligned}$$

and by Hölder

$$\int |f|^{p-1}|\nabla f| dx \le \|f\|_{L^p}^{p-1} \||\nabla f|\|_{L^p} \le 2^d \|f\|_{W^{1,p}}^p$$

and, for g smooth,

$$\int fg d\mathcal{H}^{d-1} = \int fg\nabla c \cdot F + f\nabla g \cdot F + g\nabla f \cdot F.$$

The right hand side determined the left hand side. Now we approximate $f \in W^{1,p}(U)$ by smooth function and obtain the estimate and uniqueness.

The case $p = \infty$ is simpler and left to the reader.

□

4.6 Finite differences

Here we want to relate the analogue of finite differences to Sobolev functions.

Theorem 4.45. *a) Let $1 < p \le \infty$. If $f \in L^p$ and*

$$\sup_{h\neq 0} \frac{\|f(.+h) - f\|_{L^p(\mathbb{R}^d)}}{|h|} \le C \tag{4.3}$$

then $f \in W^{1,p}$ and

$$\|\partial_{x_j} f\|_{L^p} = \sup_{t\neq 0} \left\| \frac{f(.+te_j) - f}{t} \right\|_{L^p(\mathbb{R}^d)} \le C.$$

b) Now let $1 \le p \le \infty$ and $f \in W^{1,p}$. Then

$$\frac{f(.+te_j) - f(.)}{t} \to \partial_j f \qquad \text{in} \left\{ \begin{array}{ll} L^p & \text{if } p < \infty \\ \mathcal{D}' & \text{if } p = \infty \end{array} \right.$$

Proof. Suppose that (4.3) holds. Then

$$\lim_{t\to 0} \frac{f(x+te_1) - f(x)}{t} \to \partial_{x_1} f$$

as distribution:

$$T_{f_t}(\phi) = \frac{1}{t} \int (f(x+te_1) - f(x))\phi(x) dm^d = T_f\left(\frac{1}{t}(\phi(x-te_1) - \phi(x))\right)$$

and in $\mathcal{D}(\mathbb{R}^d)$

$$\frac{1}{t}(\phi(x-te_1) - \phi(x)) \to -\partial_{x_1}\phi(x).$$

The difference quotient defines an element in $L^{\frac{p}{p-1}}(\mathbb{R}^d)^*$. It is bounded by C uniformly with respect to t, since its L^p norm is bounded by C. For all $\phi \in \mathcal{D}(\mathbb{R}^d)$ we have

$$\left| \int f \partial_i \phi dx \right| \le C \|\Phi\|_{L^{\frac{p}{p-1}}}$$

hence the distribution $\partial_i T_f$ defines an element in $L^{\frac{p}{p-1}(\mathbb{R}^d)}$ of norm at most C. But then by duality $\partial_i f \in L^p(\mathbb{R}^d)$. This proves the first direction and

$$\|\partial_{x_j} f\|_{L^p(\mathbb{R}^d)} \le \liminf_{t\to 0} \Big\| \frac{f(te_j + .) - f}{t} \Big\|_{L^p}.$$

Consider now $f \in W^{1,p}(U)$, $1 \le p \le \infty$ and we would like to show (4.3). By the fundamental theorem of calculus and Minkowski's inequality

$$\begin{aligned} \|f(.+h) - f(.)\|_{L^p(\mathbb{R}^d)} &= \| \int_0^1 \sum_{j=1}^d h_j \partial_j f(.+sh) ds \|_{L^p(\mathbb{R}^d)} \\ &\le |h| \int_0^1 \| |\nabla f(.+sh)| \|_{L^p} = |h| \| |Df| \|_{L^p} \end{aligned}$$

for $C^1 \cap W^{1,p}$ functions. Density completes the argument for $p < \infty$. For $p = \infty$ we use that by the previous argument

$$\begin{aligned} \Big| \int (f(x+h) - f(x))\phi(x) dm^d(x) \Big| =& \Big| \int_0^1 \int_{\mathbb{R}^d} \sum_{j=1}^d \partial_{x_j} f(x+th) h_j \phi(x) dm^d(x) dt \Big| \\ \le& |h| \| |\nabla f| \|_{L^\infty} \|\phi\|_{L^1} \end{aligned}$$

and hence using a Dirac sequence

$$\| \frac{f(x+h) - f(x)}{|h|} \|_{L^\infty} \le \| |\nabla f| \|_{L^\infty}.$$

□

Corollary 4.46 (Poincaré inequality)**.** *Let* $U \subset \{x \in \mathbb{R}^d : a < x_1 < b\}$ *and* $f \in W_0^{1,p}(U)$*. Then*

$$\|f\|_{L^p(U)} \le |b-a| \| |\nabla f| \|_{L^p(U)}.$$

Proof. Extend f by 0 to $\mathbb{R}^d$ and apply the previous theorem. □

Corollary 4.47 (Compact embedding)**.** *Let* U *be a bounded Lipschitz domain and* $1 \le p < \infty$*. Then the closure in* $L^p(U)$ *of*

$$\{f \in W^{1,p}(U) : \|f\|_{W^{1,p}(U)} \le 1\}$$

is compact in $L^p(U)$*.*

Proof. By extension it suffice to prove that for K compact the closure in $L^p(\mathbb{R}^d)$ of

$$\{f \in W^{1,p}(U) : \|f\|_{W^{1,p}(U)} \le 1, \operatorname{supp} f \subset K\}$$

is compact in $L^p(\mathbb{R}^d)$. We obtain this from an application of Theorem 3.37 since this set is closed, bounded, uniformly small at infinity due to the uniform compact support, and 'equicontinuous' due to Theorem 4.45. □

20.12.2019

We define

$$f_B = \fint_B f dm^d = (m^d(B))^{-1} \int_B f dm^d.$$

Lemma 4.48 (Poincaré inequality on ball)**.** *If* $1 \leq p \leq \infty$ *and* $f \in W^{1,p}(B_R(0))$ *then*

$$\|f - f_{B_R(0)}\|_{L^p(B_R(0))} \leq 2^{\frac{d}{p}} R \||Df|\|_{L^p(B_R(0))}.$$

Proof. It suffices to consider $R = 1$ by replacing f by $f(x/R)$. We calculate again by Minkowski's and Jensen's inequality

$$\begin{aligned}
&\fint_{B_1(0)} |f - f_{B_1(0)}|^p dm^d(x) \\
=&\fint_{B_1(0)} \left| \fint_{B_1(0)} (f(x) - f(y)) dm^d(y) \right|^p dm^d(x) \\
\leq&\fint \fint_{B_1(0)\times B_1(0)} |f(x) - f(y)|^p dm^d(x) dm^d(y) \\
=&\fint_{B_1(0)\times B_1(0)} \left| \int_0^1 \nabla f(x + t(y - x)) dt \right|^p |x - y|^p dm^d(x) dm^d(y) \\
\leq&2^p \fint_0^1 \fint_{B_1(0)\times B_1(0)} |\nabla f(x + t(y - x))|^p dm^d(x) dm^d(y) dt \\
=&2^p \cdot 2 \fint_0^{\frac{1}{2}} \fint_{B_1(0)\times B_1(0)} |\nabla f(x + t(y - x))|^p dm^d(x) dm^d(y) dt
\end{aligned}$$

where we used symmetry in x and y in the last equality. However, if $y \in B_1(0)$ and $0 \leq t \leq \frac{1}{2}$ then

$$\fint_{B_1(0)} |\nabla f(x + t(y - x))|^p dm^d(x) \leq 2^d \fint_{B_1(0)} |\nabla f(x)|^p dm^d(x)$$

by the transformation formula. □

4.7 Sobolev inequalities and Morrey's inequality

Lemma 4.49. *Let* $f \in C_0(\mathbb{R})$*,* $f' \in L^1$*. Then the Sobolev inequality*

$$\|f\|_{sup} \leq \frac{1}{2} \|f'\|_{L^1} \tag{4.4}$$

holds. If $f' \in L^p$*,* $1 \leq p \leq \infty$ *then every point is a Lebesgue point and*

$$\sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|^{1 - \frac{1}{p}}} \leq \|f'\|_{L^p}. \tag{4.5}$$

Proof. It suffices to prove the estimate for smooth functions with compact support. The inequality (4.4) is a consequence of the fundamental theorem of calculus:

$$f(x) = \int_{-\infty}^{x} f'(y)dy = -\int_{x}^{\infty} f'(y)dy.$$

and we derive (4.5) by Hölder's inequality

$$|f(x) - f(y)| \le \int_x^y |f'(z)|dz \le \|f'\|_{L^p} \|\chi_{[x,y]}\|_{L^{\frac{p}{p-1}}} = |x-y|^{1-\frac{1}{p}} \|f'\|_{L^p}.$$

□

The Sobolev inequality and Morrey's inequality are the versions of these inequalities (4.4), (4.5) in higher space dimension.

Theorem 4.50 (Morrey). *Let U be open. Suppose that $p > d$ and $\tilde f \in W^{1,p}(U)$. Every point is a Lebesgue point and the canonnical representative f is continuous. There exists c depending on p and d so that the following is true: Let $x, y \in U$ with*

$$|x - y| < \operatorname{dist}(x, \mathbb{R}^d \backslash U).$$

Then

$$|f(x) - f(y)| \le c|x-y|^{1-\frac{d}{p}} \| |\nabla f| \|_{L^p(B_{|x-y|}(x))}.$$

Proof. The inequality follows from

$$\left| f(y) - f_{B_R(x_0)} \right| \le C R^{1-\frac{d}{p}} \| |\nabla f| \|_{L^p(B_R(x_0))} \tag{4.6}$$

for $|y - x_0| < R$ (and Lebesgue points y) with a constant $C(p, d)$ which is bounded as $p \to \infty$ by

$$|f(y) - f(x_0)| \le \left| f(y) - f_{B_R(x_0)} \right| + \left| f(x_0) - f_{B_R(x_0)} \right|$$

and two applications of (4.6) if y and x_0 are Lebesgue points. Thus the restriction to Lebesgue points is Hölder continuous, and hence there is a unique Hölder continuous function in the equivalence class. As a consequence there are no non-Lebesgue points.

It remains to prove (4.6). We have for $B_{R/2}(y) \subset B_R(x_0)$

$$\begin{aligned}
&\left| ⨍_{B_R(0)} f(x_0 + z) - f(y + z/2) dm^d(z) \right| \\
&= \left| ⨍_{B_R(0)} \int_0^1 \sum_{j=1}^d (x_0 + z/2 - y)_j \partial_{x_j} f(x_0 + (1 - t/2)z + t(y - x_0)) dt dm^d(z) \right| \\
&\le c(d) R ⨍_{B_R(x_0)} |\nabla f| dm^d(x) \\
&\le c(d) m^d(B_1(0)))^{-\frac{1}{p}} R R^{-d} R^{\frac{pd}{p-1}} \|\nabla f\|_{L^p(B_R(x_0))}
\end{aligned}$$

first for smooth functions, and then by approximation for Sobolev functions. Now $RR^{-d}R^{\frac{dp}{p-1}} = R^{1-\frac{d}{p}}$.

By a geometric series and an iterative application of the above inequality with $(x_0, y) = (x_{j-1}, x_j)$:

$$
\begin{aligned}
|f(y) - f_{B_R(x_0)}| =& \left| \sum_{j=1}^{\infty} f_{B_{2^{-j}R}(x_j)} - f_{B_{2^{1-j}R}(x_{j-1})} \right| \\
\leq& \sum_{j=1}^{\infty} |f_{B_{2^{-j}R}(x_j)} - f_{B_{2^{1-j}R}(x_{j-1})}| \\
\leq& cR^{1-\frac{d}{p}} \sum_{j=1}^{\infty} 2^{-j(1-\frac{d}{p})} \| |\nabla f| \|_{L^p(B_R(x_0))}
\end{aligned}
$$

provided y is a Lebesgue point and for $x_j = y - 2^{-j}(x-y)$. Here we use convergence of the means at Lebesgue points.

□

Theorem 4.51 (Rademacher)**.** *I) Functions in $W^{1,p}(U)$ with $p > d$ are almost everywhere differentiable. The derivative almost everywhere is the same as the weak derivative.*

II) Lipschitz continuous functions are almost everywhere differentiable.

Proof. Let $f \in W^{1,p}(U)$. By Morrey's theorem 4.50 we know that every point is a Lebesgue point and there is a uniformly continuous representative. By Theorem 3.49 there exists a set A whose complement has zero measure so that every $x \in A$ is a Lebesgue point for all partial derivatives, moreover

$$
\lim_{r \to 0} \fint_{B_r(x)} |\nabla f(y) - \nabla f(x)|^p dm^d(y) = 0.
$$

We apply the Morrey's inequality Theorem 4.50 to

$$
v(y) = f(x+y) - f(x) - \sum_{j=1}^{d} \partial_j f(x) y_j
$$

on $B_r(0)$ where again $x \in A$. Then

$$
\begin{aligned}
|v(y)| \leq& cr^{1-\frac{d}{p}} \left(\int_{B_r(0)} |\nabla f(x+z) - \nabla f(x)|^p dm^d(z) \right)^{\frac{1}{p}} \\
=& cr \left(m^d(B_r(x))^{-1} \int_{B_r(x)} |\nabla f(z) - \nabla f(x)|^p dm^d(z) \right)^{\frac{1}{p}} \\
=& o(r).
\end{aligned}
$$

This implies that f is differentiable at $x \in A$.

By localization and extension it suffices to prove the second part of $U = \mathbb{R}^d$. Morrey's theorem and the difference characterization imply that $W^{1,\infty}$ is the space of bounded Lipschitz continuous functions. Now the second part follows from the first. $\square$

Theorem 4.52 (Sobolev)**.** *Suppose that* $1 \le p < d$ *and*

$$\frac{1}{q} + \frac{1}{d} = \frac{1}{p}.$$

Then

$$\|f\|_{L^q(\mathbb{R}^d)} \le c \||Df|\|_{L^p(\mathbb{R}^d)}$$

whenever $f \in L^q(\mathbb{R}^d)$ *and* $|Df| \in L^p(\mathbb{R}^d)$.

Proof. We prove the estimate first for $p = 1$ and $q = \frac{d}{d-1}$. More precisely we prove the estimate

$$\|f\|^d_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} \le 2^{-d} \prod_{j=1}^{d} \|\partial_j f\|_{L^1(\mathbb{R}^d)} \tag{4.7}$$

by induction on the dimension. The case $d = 1$ has been contained in Lemma 4.49. Suppose we have proven the estimate for $d \le k-1$. Then by Fubini and Hölder's inequality (since $\frac{1}{k-1} + \frac{k-2}{k-1} = 1$)

$$\begin{aligned}
\|f\|^{\frac{k}{k-1}}_{L^{\frac{k}{k-1}}(\mathbb{R}^k)} &= \int_{\mathbb{R}} \int_{\mathbb{R}^{k-1}} |f|^{\frac{1}{k-1}} |f| dm^{k-1} dm^1(x_1) \\
&\le \int_{\mathbb{R}} \|f(x_1, \dots)\|^{\frac{1}{k-1}}_{L^1(\mathbb{R}^{k-1})} \|f(x_1, \dots)\|_{L^{\frac{k-1}{k-2}}(\mathbb{R}^{k-1})} dm^1(x_1) \\
&\le \sup_{x_1} \|f(x_1, .)\|^{\frac{1}{k-1}}_{L^1(\mathbb{R}^{k-1})} \int_{\mathbb{R}} \|f(x_1, .)\|_{L^{\frac{k-1}{k-2}}(\mathbb{R}^{k-1})} dm^1(x_1) \\
&\le 2^{-\frac{1}{k-1}} \|\partial_{x_1} f\|_{L^1(\mathbb{R}^k)} \int_{\mathbb{R}} \left(2^{-(k-1)} \prod_{j=2}^{k} \|\partial_{x_j} f(x_1, .)\|_{L^1(\mathbb{R}^{k-1})} \right)^{\frac{1}{k-1}} dm^1(x_1)
\end{aligned}$$

We take the inequality to the power $k-1$, and apply Hölder's inequality in the form

$$\left(\int \prod_{j=2}^{k} |g_j|^{\frac{1}{k-1}} dm^1 \right)^{k-1} \le \prod_{j=2}^{k} \int |g_j| dm^1$$

to arrive at

$$\|f\|^k_{L^{\frac{k}{k-1}}(\mathbb{R}^k)} \le 2^{-k} \prod_{j=1}^{k} \|\partial_j f\|_{L^1(\mathbb{R}^k)}.$$

Now let $1 < p < d$. We apply the above inequality (4.7) to $|f|^{\frac{(d-1)p}{d-p}}$. Then

$$\begin{aligned}\|f\|_{L^q(\mathbb{R}^d)}^{\frac{(d-1)p}{d-p}} =& \||f|^{\frac{(d-1)p}{d-p}}\|_{L^{\frac{d}{d-1}}(\mathbb{R}^d)} \\ \le& \|D|f|^{\frac{(d-1)p}{d-p}}\|_{L^1(\mathbb{R}^d)} \le \frac{(d-1)p}{d-p}\int |f|^{\frac{d(p-1)}{d-p}}|Df|dm^d \\ \le& \frac{(d-1)p}{d-p}\|f\|_{L^q(\mathbb{R}^d)}^{\frac{d(p-1)}{d-p}}\|Df\|_{L^p(\mathbb{R}^d)}\end{aligned}$$

where we first argue for smooth functions, and where we used Hölder's inequality in the last step. ☐

Corollary 4.53. *Let U be a bounded Lipschitz domain. $1 \le p, q < \infty$ and*

$$\frac{1}{q} + \frac{1}{d} \ge \frac{1}{p}. \tag{4.8}$$

Then there exists c such that

$$\|f\|_{L^q(U)} \le c\|f\|_{W^{1,p}(U)}$$

for all $f \in W^{1,p}(U)$. The map

$$W^{1,p}(U) \ni f \to f \in L^q(U)$$

is compact when the strict inequality holds in (4.8). This means that $\overline{B_1^{W^{1,p}}(0)}$ is a compact subset of $L^q(U)$.

Proof. The first statement follows from Theorem 4.42. Lemma 4.47 implies the second statement if $p = q$. Hölder's inequality gives the full result by

$$\|f_n - f_m\|_{L^q} \le \|f_n - f_m\|_{L^p}^{\lambda}\|f_n - f_m\|_{L^r}^{1-\lambda}$$

if $p < d$, $\frac{1}{r} + \frac{1}{d} = \frac{1}{p}$ and

$$\frac{1}{q} = \frac{\lambda}{p} + \frac{1-\lambda}{r}.$$

Thus convergence in L^p and boundedness in $L^{\frac{1}{q'}}$, $\frac{1}{q'} + \frac{1}{d} = \frac{1}{p}$ obtain the convergence in L^q if $p < d$. To conclude we observe that on bounded sets $W^{1,p} \subset W^{1,p'}$ whenever $p' \le p$ and we choose a suitable $p' < d$. ☐

08.01.2020

5 Linear Functionals

In this section we will study the dual space X^* of Banach spaces X.

5.1 The Theorem of Hahn-Banach

Definition 5.1. *Let X be a $\mathbb{K}$ vector space. A map $p : X \to \mathbb{R}$ is called sublinear if*

1. *$p(\lambda x) = \lambda p(x)$ for $x \in X$ and $\lambda \geq 0$,*
2. *$p(x + y) \leq p(x) + p(y)$ for $x, y \in X$.*

Examples:

1. The norm of a normed space is sublinear.
2. If $\mathbb{K} = \mathbb{R}$, any element of X^* is sublinear.
3. The Minkowski functional of a convex set. Let $K \subset X$ be convex such that for every $x \in X$ there exists $\lambda > 0$ so that $\lambda x \in X$. we define

   $$p_K(x) = \inf\{\lambda > 0 : \frac{1}{\lambda} x \in K\} \in [0, \infty).$$

   It is not difficult to verify that p_K is sublinear. A norm is the Minkowski functional of the unit ball.

Theorem 5.2 (Hahn-Banach, real case)**.** *Let X be a real vector space, $Y \subset X$ a subvector space, $p : X \to \mathbb{R}$ sublinear and $l : Y \to \mathbb{R}$ linear such that*

$$l(y) \leq p(y) \qquad \text{for all } y \in Y.$$

Then there exists $L : X \to \mathbb{R}$ linear so that

1. *$l(y) = L(y)$ for all $y \in Y$*
2. *$l(x) \leq p(x)$ for all $x \in X$.*

Proof. There are two very different steps.

Suppose that $Y \neq X$. Then there exists $x_0 \in X \backslash Y$. Let Y_1 be the space spanned by Y and x_0. Every element of Y_1 can uniquely be written as

$$y + r x_0, \quad y \in Y, r \in \mathbb{R}.$$

We want to find a linear map $l_1 : Y_1 \to \mathbb{R}$ such that

1. $l_1(y) = l(y)$ for $y \in Y$
2. $l_1(y + s x_0) \leq p(y + s x_0)$ for $s \in \mathbb{R}$ and $y \in Y$.

By the first condition and linearity we have to find $t = l_1(x_0)$ so that

$$l(y) + st \le p(y + sx_0)$$

for all $y \in Y$ and $s \in \mathbb{R}$. We consider $s > 0$. Then this inequality is equivalent to

$$st \le s(p(y/s + x_0) - l(y/s))$$

for all $s > 0$ and $y \in Y$

$$\iff t \le \inf_y p(y + x_0) - l(y).$$

Similarly the inequality holds for $s < 0$ if and only if

$$t \ge \sup_y l(y) - p(y - x_0).$$

We can find t if and only if

$$l(y) - p(y - x_0) \le p(\tilde{y} + x_0) - l(\tilde{y}) \qquad \text{for all } y, \tilde{y} \in Y$$

which follows from the inequality on Y and sublinearity of p:

$$l(y) + l(\tilde{y}) = l(y + \tilde{y}) \le p(y + \tilde{y}) \le p(y + x_0) + p(\tilde{y} - x_0).$$

This completes the first step.

For the second step we need the axiom of choice in the form of Zorn's lemma.

Let Z be a partially ordered set which contains an upper bound for every chain. Then there is a maximal element.

A chain is a totally ordered subset, i.e. a subset A so that always either $a \le b$ or $b \le a$. An element b is an upper bound for the chain A, if $a \le b$ for all $a \in A$. An element $a \in Z$ is maximal if $b \in Z$, $b \ge a$ implies $b = a$.

We define

$$Z = \{(W, l_W) : Y \subset W \subset X,\ l_W|_Y = l,\ l_W(w) \le p(w) \text{ for } w \in W\},$$

with the ordering

$$(W, l_W) \le (V, l_V) \qquad \text{if } W \subset V \quad \text{and } l_V|_W = l_W.$$

This is a partial order. If $\tilde{Z}$ is a chain then

$$V = \bigcup_{(W, l_W) \in \tilde{Z}} W$$

with the obvious l_V being an upper bound for the chain. Now let (V, l_V) be a maximal element. If $V = X$ we are done. Otherwise we obtain a contradiction by the first step. □

There is a complex version. It relies on the observation

Lemma 5.3. *Let X be a complex vector space and $l : X \to \mathbb{R}$ be $\mathbb{R}$ linear. Then it is the real part of the linear map*

$$l_{\mathbb{C}}(x) = l(x) - il(ix).$$

The real part determines $l_{\mathbb{C}}$.

Proof. We have to show complex linearity. Real linearity is obvious. We compute

$$l_{\mathbb{C}}(ix) = l(ix) - il(iix) = i(l(x) - il(ix)) = il_{\mathbb{C}}(x).$$

□

Theorem 5.4 (Complex Hahn Banach)**.** *Let X be a complex vector space, Y a subvector space, p sublinear and $l : Y \to \mathbb{C}$ linear so that*

$$\operatorname{Re} l(y) \le p(y) \text{ for } y \in Y.$$

Then there exists $L : X \to \mathbb{C}$ linear so that

1. $L|_Y = l$
2. $\operatorname{Re} L(x) \le p(x)$.

Proof. We apply the real theorem of Hahn Banach to the real part, and extend it to a complex linear map by Lemma 5.3. To complete the proof we observe that L and $\tilde{L}$ are the same iff the real parts are the same. □

We formulate the consequences for normed vector spaces, making use of the fact that norms are sublinear.

Theorem 5.5. *Let X be a normed $\mathbb{K}$ vector space, Y a subspace and $l : Y \to \mathbb{K}$ a continuous linear map. Then there exists $L : X \to \mathbb{K}$ linear and continuous so that*

1. $L|_Y = l$
2. $\|L\|_{X^*} = \|l\|_{Y^*}$.

Proof. We define

$$p(x) = \|l\|_{Y^*} \|x\|_X.$$

Then

$$\operatorname{Re} l(y) \le p(y)$$

for all $y \in Y$. We apply the theorems of Hahn-Banach to obtain $L \in X^*$ so that $L|_Y = l$ and

$$\operatorname{Re} L(x) \le p(x).$$

This implies

$$|L(x)| = \operatorname{Re} \alpha L(x) = \operatorname{Re} L(\alpha x) \le p(\alpha x) = p(x) = \|l\|_{Y^*} \|x\|_X$$

for some $\alpha \in \mathbb{C}$ with $|\alpha| = 1$. Thus

$$\|L\|_{X^*} = \|l\|_{Y^*}.$$

□

Example 5.6. *Let $0 < p < 1$ and $L^p(\mathbb{R}^d)$ the set of measurable p integrable functions. It is atraight forward to verify that*

$$d(f,g) = \int |f-g|^p dm^d$$

defines a norm and L^p becomes a closed metric space with this metric. Suppose that $T : L^p(\mathbb{R}^d) \to \mathbb{K}$ is a continous linear map. Then $T(f) = 0$. To see this is suffices to consider $\mathbb{K} = \mathbb{R}$. Any $f \in L^1([0,1]^d)$ is also (by a slight abuse of notation) in $L^p(\mathbb{R}^d)$. Since T is continous at 0 there exists $\varepsilon > 0$ so that

$$|T(f)| < 1 \qquad \text{if} \int |f|^p dx < \varepsilon$$

and hence $T|_{L^1([0,1]^d)} \in (L^1([0,1]^d))^*$. By the Riesz representation theorem there exists a unique $g \in L^\infty$ with

$$Tf = \int_{[0,1]^d} fg dm^d$$

Now suppose that there is $\varepsilon > 0$ and a subset $Q \subset [0,1]^d$ with positive measure so that

$$\varepsilon \le g(x) \le \|g\|_{L^\infty}$$

Let x_0 be a Lebesgue point of $g\chi_Q$ where $(g\chi_Q)(x_0) > \varepsilon$. Then

$$\begin{aligned}\int_Q |x-x_0|^{-r} g dx &\ge (1-2^{-r}) \sum_{s\in\mathbb{N}} 2^{rs} \int_{Q\cap B_{2^{-s}}(0)} g dm^d \\ &= (1-2^{-r}) m^d(B_1(0)) \sum_{s\in\mathbb{N}} 2^{(r-d)s} \fint_{Q\cap B_{2^{-s}}(0)} g dm^d \infty\end{aligned}$$

if $r > d$, but $\chi_Q |x-x_0|^{-r} \in L^p(\mathbb{R}^d)$ if $pr < d$. This is a contradiction and thus $g \le 0$. In the same fashion we see that $g = 0$. There is a small gap: We have seen the representation through g only for integrable functions. To complete the argument we have to truncate and take a limit.

Lemma 5.7. *$(l^\infty(\mathbb{N}))^* \neq l^1(\mathbb{N})$. More precisely the map*

$$l^1(\mathbb{N}) \ni (x_n) \to ((y_n) \to \sum_n x_n y_n) \in l^\infty(\mathbb{N})$$

is not surjective.

Proof. The space of converging sequences c is a closed subspace of l^∞. Let $l : c \to \mathbb{K}$ be defined by

$$l((x_j)) = \lim_{j\to\infty} x_j.$$

Then

$$|l((x_j))| \le \|(x_j)\|_{l^\infty}$$

for every converging sequence. Checking constant sequences we see that $\|l\|_{(l^\infty)^*} = 1$. By Theorem 5.5 it has an extension L to l^∞. Clearly $L(e_j) = l(e_j) = 0$. We claim that it cannot be represented in the form

$$L((x_j)) = \sum_{j=1}^{\infty} y_j x_j$$

for $(y_j) \in l^1$ - if it were represented in this fashion then all y_j would have to vanish. □

5.2 Consequences of the theorems of Hahn-Banach

Lemma 5.8. *Let X be a normed space and $x \in X$. There exists $x^* \in X^*$ with $\|x^*\|_{X^*} = 1$ and $x^*(x) = \|x\|_X$.*

Proof. Let $x \in X\backslash\{0\}$ and let Y be the span of x. Y is one dimensional and we define $y^* \in Y^*$ by $y^*(rx) = r\|x\|_X$. Then $\|y^*\|_{Y^*} = 1$. We apply Theorem 5.5 to obtain x^*. If $x = 0$ we choose $x_0 \ne 0$ and find x^* for x_0. □

Corollary 5.9. *Let X be a normed space. If $x \in X$ then*

$$\|x\|_X = \sup\{\operatorname{Re} x^*(x) : x^* \in X^*, \|x^*\|_{X^*} = 1\}.$$

If $x^* \in X^*$ then

$$\|x^*\|_{X^*} = \sup\{\operatorname{Re} x^*(x) : x \in X, \|x\|_X = 1\}.$$

Proof. The first claim is a consequence of Lemma 5.8. The second statement is an immediate consequence of the definition. □

10.01.2020

Lemma 5.10. *Let $(X, \|\cdot\|)$ be a normed vector space and $Y \subset X$ a closed vector subspace with $\overline{Y} \ne X$. Then there exists $x^* \in X^* \setminus \{0\}$ such that $x^*|_Y = 0$.*

Proof. We treat the real case exclusively. Let $x_0 \in Y^c$ and define on $\mathbb{M} := Y + \mathbb{R}\{x_0\}$ a functional $\ell\colon \mathbb{M} \to \mathbb{R}$ via

$$\ell(y + tx_0) := td,$$

where $d := \operatorname{dist}(x_0, Y)$; since Y is closed and $x_0 \notin Y$, $d > 0$. Moreover, note that ℓ is linear (by construction) and *bounded*. To see the latter, if $t = 0$, then we have $|\ell(y + tx_0)| = |\ell(y)| = 0 \le \|y\|$, and if $t \neq 0$, then

$$|\ell(y + tx_0)| = |t|d = |t|d\frac{\|y + tx_0\|}{\|y + tx_0\|} = d\frac{\|y + tx_0\|}{\|\frac{y}{t} + x_0\|} \le \|y + tx_0\|.$$

Indeed, since $(-\frac{y}{t}) \in Y$,

$$d = \operatorname{dist}(x_0, Y) \le \|x_0 - (-\tfrac{y}{t})\|.$$

In consequence, $\ell \colon Y \to \mathbb{R}$ is bouded and so we may extend ℓ to some $x^* \in X^*$ while preserving its norm: $\|x^*\|_{X^*} = \|\ell\|_{Y^*}$. Obviously, $x^*(y) = 0$ for all $y \in Y$ and $x^*(x_0) = \ell(0 + 1 \cdot x_0) = 1 \cdot d = d > 0$. The proof is complete. □

Lemma 5.11. *If X is a normed space and X^* is separable then X is separable.*

Proof. Since X^* is separable, and a subset of a separable set is separable also the unit sphere is separable. Let x_j^* be a dense sequence of unit vectors. We choose a sequence of unit vectors x_j with $x_j^*(x_j) \ge \frac{1}{2}$. We claim that the span U of the x_j is dense. Otherwise, by Lemma 5.10 there exists $x^* \in X^*$ of norm 1 which vanishes on the closure of U, and in particular $x^*(x_j) = 0$ for all j. By density there exists j such that $\|x^* - x_j^*\|_{X^*} < \frac{1}{2}$ and hence

$$\frac{1}{2} \le \operatorname{Re} x_j^*(x_j) = \operatorname{Re}(x^*(x_j) + (x_j^* - x^*)(x_j)) < \frac{1}{2}$$

This is a contradiction. □

Introduction to reflexivity

As one of the key results of the lecture, we shall now single out a class of Banach spaces which allow for *some* compactness results. Such compactness results are more than desirable since many problems from PDEs or the Calculus of Variations precisely require compactness argument to establish existence of solutions.

Recall that if any bounded sequence in a Banach space $(X, \|\cdot\|)$ possesses a convergent subsequence, then $\dim(X) < \infty$. So, in general, we cannot hope for compactness results for the norm topology in the infinite dimensional situation. However, weakening the underlying notion of convergence, we may indeed come up with a suitable concept of convergence: As we shall see later, this so-called *weak convergence* is particularly useful in reflexive spaces, a notion that we discuss now.

Setup. Let $(X, \|\cdot\|)$ be a normed space. We denote

$$X^{**} := (X^*)^* \qquad \text{(the } \textit{double dual} \text{ of } X)$$
$$X^{***} := (X^{**})^* \qquad \text{(the } \textit{triple dual} \text{ of } X)$$

Now consider the map $J\colon X \to X^{**}$ given by

$$J(x)(x^*) := \langle x^*, x\rangle := x^*(x), \qquad x^* \in X^* \tag{5.1}$$

for $x \in X$. Note carefully that, for any $x \in X$, $J(x)\colon X^* \to \mathbb{R}$; we call J the *evaluation functional* simply as $J(x)$ evaluates a given $x^* \in X^*$ at x. We then have the following

Corollary 5.12. *Let $(X, \|\cdot\|_X)$ be a normed space and define J as above. Then J is a linear isometry*

$$J\colon X \hookrightarrow X^{**}.$$

In particular, J is linear and injective, *and so any normed linear space is isometrically isomorphic to a dense subspace of a Banach space.*

Proof. Let $x \in X$. Then we have

$$\begin{aligned}\|J(x)\|_{X^{**}} &= \sup\{\langle J(x), x^*\rangle_{X^{**}\times X^*}\colon\ \|x^*\|_{X^*} \leq 1\}\\ &= \sup\{\langle x^*, x\rangle_{X^*\times X}\colon\ \|x^*\|_{X^*} \leq 1\}\\ &= \|x\|_X\end{aligned}$$

by the dual description of $\|x\|_X$ (in turn being a consequence of Hahn-Banach). For the rest of the statement, it now suffices to realie that $J(X)$ is clearly dense in $\overline{J(X)}$. Note that $\overline{J(X)}$ is a Banach space, and now the claim follows by the isometry and linearity of J. The proof is complete. □

As an upshot, J *embeds* X *into* X^{**}, and this embedding is linear and isometric. We shall often refer to J as the *canonical embedding* of X into X^{**}. The pivotal question leading to reflexivity then is as follows:

Q.: *Is the evaluation map $J\colon X \to X^{**}$ surjective*?

Put differently, does any element of X^{**} arise as an evaluation functional? This motivates the next

Definition 5.13 (Reflexive spaces)**.** *A normed linear space $(X, \|\cdot\|)$ is called* reflexive *if and only if the evaluation map $J\colon X \hookrightarrow X^{**}$ is surjective.*

Reflexive spaces are automatically Banach (why?). In general, the evaluation map J is not surjective as will become clear from the following examples:

- Every finite dimensional vector space is reflexive.
- If $1 < p < \infty$, then $\ell^p(\mathbb{N})$ is reflexive.
- The space $c_0(\mathbb{N})$ (sequences converging to zero equipped with the supremum norm) is not reflexive. Indeed: Reflexivity would mean that any $f \in c_0^{**}(\mathbb{N})$ arises as

$$f((x_j)) = \sum_j x_j y_j, \qquad (x_j) \in c_0^* \cong \ell^1,$$

  for some $(y_j) \in c_0(\mathbb{N})$. However, for given $(y_j) \in \ell^\infty \setminus c_0$ the functional

$$f((x_j)) = \sum_j x_j y_j, \qquad (x_j) \in c_0^* \cong \ell^1,$$

  certainly belongs to c_0^{**} – hence c_0 is not reflexive.

Remark 5.14. *If $X = L^p(\mu)$, $1 < p < \infty$ then X^* is isomorphic to $L^{\frac{p}{p-1}}(\mu)$ and J is surjective.*

Remark 5.15. *By Lemma 5.7 $J : l^1 \to (l^1)^{**}$ is not surjective - we can identify the second space with $(l^\infty)^*$.*

For reflexivity it is not enough that there is *some* isometric isomorphism between X and X^{**} – it needs to be the evalution map:

Remark 5.16 (James space)**.** *Denote $\mathcal{P}$ the family of all finite increasing sequences of integers of odd length. For a real sequence $x = (x_j)$ and $p = (p_1, p_2, ..., p_{2N+1}) \in \mathcal{P}$ put*

$$\|x\|_p := \Big(x_{p_{2N+1}}^2 + \sum_{j=1}^N |x_{p_{2j-1}} - x_{p_{2j}}|^2\Big)^{\frac{1}{2}}.$$

The James space *is defined as*

$$\mathbb{J} := \{x \in c_0 \colon \ \|x\|_{\mathbb{J}} := \sup_{p \in \mathcal{P}} \|x\|_p < \infty\}.$$

We list some of its properties without proof:

- *$\mathbb{J}$ is Banach.*
- *$\mathbb{J}$ is isometrically isomorphic to its double dual – yet it* fails to be reflexive*!*
- *$\dim(\mathbb{J}^{**}/J(\mathbb{J})) = 1$.*

Theorem 5.17. *Let X be reflexive and $Y \subset X$ a closed subspace. Then Y is reflexive. Moreover, a Banach space X is reflexive if and only if X^* is reflexive.*

Proof. Let $y^{**} \in Y^{**}$. For every $x^* \in X^*$ the restriction to Y is in Y^* with

$$\|x^*|_Y\|_{Y^*} \le \|x^*\|_{X^*}$$

We define $x^{**} \in X^{**}$ by

$$x^{**}(x^*) = y^{**}(x^*|_Y).$$

Since X is reflexive there exists $y \in X$ so that

$$x^{**}(x^*) = x^*(y) \qquad \text{for } x^* \in X^*.$$

If y were not in Y there would be $\tilde{x}^* \in X^*$ with $\tilde{x}^*|_Y = 0$ and $x^*(y) = 1$. This is impossible and hence $y \in Y$. Thus J_Y is surjective. The second statement follows by a standard application of Hahn-Banach. □

15.01.2020

Corollary 5.18. *Let $U \subset \mathbb{R}^d$ be open, $k \in \mathbb{N}$ and $1 < p < \infty$. Then $W^{k,p}(U)$ and $W_0^{k,p}(U)$ are reflexive.*

Proof. The Sobolev space $W^{k,p}(U)$ is isometrically isomorphic to a closed subspace of $L^p(U \times \Sigma_k)$ and thus reflexive. $W_0^{k,p}(U) \subset W^{k,p}(U)$ is a closed subspace of a reflexive space and hence reflexive. □

Lemma 5.19. *Let $U \subset \mathbb{R}^n$ be open, $1 \le p < \infty$ and $k \in \mathbb{N}$. The map*

$$J : L^{\frac{p}{p-1}}(U \times \Sigma_k) \ni (g_\alpha) \to (f \to \sum_{|\alpha| \le k} g^\alpha \partial^\alpha f) \in (W^{k,p})^*$$

is bounded and surjective.

Proof. The map

$$W^{k,p}(U) \ni f \to (\partial^\alpha f)_{|\alpha| \le k} \in L^p(U \times \Sigma_k)$$

map $W^{k,p}$ isometrically to a closed subspace. Any $y^* \in (W^{k,p})^*$ defines a linear functional on this closed subspace. By Theorem 5.5 we can extend it to the whole of $L^p(U \times \Sigma_k)$. This can be represented by a function in $L^{\frac{p}{p-1}}(U \times \Sigma_k)$. We have seen that J has norm ≤ 1 as a consequence of Hölder's inequality. □

5.3 Separation theorems and weak convergence

Lemma 5.20. *Let X be a normed space and $K \subset X$ convex. If 0 is in the interior of K then for every $x \in X$ there exists $\lambda > 0$ so that $\lambda x \in K$. Moreover there exists $C > 0$ so that*

$$p_K(x) \le C\|x\|_X.$$

If X is a Banach space, and for every x there exists $\lambda > 0$ so that $\lambda x \in K$ then 0 is in the interior of K.

Proof. If 0 is in the interior of K there exists $\varepsilon > 0$ so that $B_\varepsilon(0) \subset K$. An easy calculation shows that then $p_K(x) \le \varepsilon^{-1}\|x\|_X$.

Suppose that X is Banach and that for every x there exists $\lambda > 0$ so that $\lambda x \in K$. In particular $0 \in K$. Let

$$A_n = \left\{ x \in X : \frac{1}{n} x \in \overline{K} \right\}$$

The set A_n are closed, convex, $0 \in A_n$ and $X = \bigcup A_n$. By the Baire category theorem one and hence all of the A_n have nonempty interior. In particular there exist x and $\varepsilon > 0$ so that $B_\varepsilon(x) \subset A_1$. There exists A_n so that $-x \in A_n$ hence $-x/n \in A_1$. The convex hull of $B_\varepsilon(x)$ and $-x/n$ contains a ball $B_\varepsilon(0)$. Let $y \in \partial K = \partial \overline{K}$. Then there exists $x^* \in X^*$ with norm 1 and

$$x^*(y) \ge x^*(z)$$

for $z \in \overline{K} \backslash \{\}$. Since $K \cap B_\varepsilon(0)$ is dense in $B_\varepsilon(0)$ we have $|y| \ge \varepsilon$ and hence $B_\varepsilon \subset K$.

Hence 0 is in the interior of $A_{\frac{1}{2}} \subset \overline{K}$. □

Lemma 5.21. *Let X be a normed vector space and V convex, open with $0 \notin V$. Then there exists $x^* \in X^*$ with*

$$\operatorname{Re} x^*(x) < 0 \qquad \text{if } x \in V$$

Proof. It suffices to consider $\mathbb{K} = \mathbb{R}$. The complex case is a consequence of Lemma 5.3. Let $x_0 \in V$ and define the translate $U = V - x_0$. Let p_U be the Minkoski functional of U. It is sublinear. Let $y_0 = -x_0 \notin U$ and Y the span of y_0. We define

$$l(ty_0) = t p_U(y_0) \qquad t \in \mathbb{R}.$$

Then $l(y) \le p(y)$ for all $y \in Y$. By Theorem 5.2 there exists $L \in X^*$ with $L|_Y = l$, $l(x) \le p(x)$ for $x \in X$ (here we use Lemma 5.20). In particular $L(y_0) \ge 1$ and for $x \in V$ and $u = x + y_0$

$$L(x) = L(u) - L(y_0) \le p_U(u) - 1 < 0.$$

The strict inequality on the right holds since V is open, and hence for every $x \in V$ there is a ball centered at x in which $\le$ holds, which implies the strict inequality for x. □

Theorem 5.22 (Separation theorem 1)**.** *Let X be a normed space, V and W disjoint convex sets with V open. Then there exists $x^* \in X^*$ such that*

$$\operatorname{Re} x^*(v) < \operatorname{Re} x^*(w) \qquad \textit{for every } v \in V, w \in W$$

Proof. Let $\tilde{V} = V - W = \{v - w : v \in V, w \in W\}$. It is convex and open. Since V and W are disjoint $0 \notin \tilde{V}$. By Lemma 5.21 there exist $x^* \in X^*$ so that $\operatorname{Re} x^*(x) < 0$ for $x \in \tilde{V}$. This implies the desired inequality. □

Theorem 5.23 (Separation theorem 2)**.** *Let X be a normed space, V convex and closed, $x \notin V$. Then there exists $x^* \in X^*$ such that*

$$x^*(x) < \inf_{v \in V} x^*(v). \tag{5.2}$$

Proof. We may assume $x = 0$. Since V is closed there exists $\varepsilon > 0$ so that $B_\varepsilon(0) \cap V = \{\}$. We apply Theorem 5.22 to see that there is $x^* \in X^*$ with

$$\operatorname{Re} x^*(u) < \operatorname{Re} x^*(v) \qquad \text{for } u \in B_\varepsilon(0), v \in V$$

There exists $x \in B_\varepsilon(0)$ so that

$$\operatorname{Re} x^*(x) \geq \varepsilon \|x^*\|_{X^*}/2 > 0$$

which implies (5.2) □

Corollary 5.24. *Let X be a normed vector space, $K \subset X$ open and convex and $x \in \partial K$. Then there exists a half space containing K with x a boundary point.*

Proof. Apply Lemma 5.21 to $\{K + x_0\}$. □

Definition 5.25. *We call a sequence $x_n \in X$ weakly convergent against $x \in X$ if for all $x^* \in X^*$*

$$x^*(x_n) \to x^*(x)$$

Lemma 5.26. *Norm convergence implies weak convergence.*

Proof. This follows from the continuity of x^*. □

Lemma 5.27. *Weakly convergent sequences are bounded.*

Proof. This is an immediate consequence of uniform boundedness principle Theorem 4.2. Let $J : X \to X^{**}$ be the canonical map. We apply Theorem 4.2 to $(J(x_n))_n$. By convergence for all x^*

$$x^*(x) = J(x)(x^*)$$

is bounded for every x. Thus $(\|x_n\|)_n = (\|J(x_n)\|_{X^{**}})_n$ is bounded. □

17.10.2020

Lemma 5.28. *Let C be a closed convex set in X, $x_n \in C$ a sequence which converges weakly to $x \in X$. Then $x \in C$.*

Proof. It suffices to consider $\mathbb{K} = \mathbb{R}$. Suppose that $x \notin C$. By Theorem 5.23 there exists $x^* \in X^*$ with

$$\inf_{y \in C} x^*(y) > x^*(x).$$

This contradicts $x^*(x_n) \to x^*(x)$. □

Definition 5.29. *Let $A \subset X$ be a subset. The convex hull $C(A)$ consists of all convex linear combinations*

$$C(A) = \Big\{ \sum_{k=n}^{\infty} \mu_k x_k : 0 \le \mu_k, \sum_{k=n}^{\infty} \mu_k = 1, \text{ only finitely many } \mu_k \text{ are nonzero.} \Big\}$$

Lemma 5.30. *$C(A)$ is convex. If K is convex and $A \subset K$ then $C(A) \subset K$.*

Proof. Convexity is immediate. If $A \subset K$ and $x \in C(A)$ then clearly $x \in K$. □

Theorem 5.31 (Mazur). *Let X be a normed space and $(x_n)_n$ a sequence which converges weakly to x for $n \to \infty$. Then there exist real numbers $\lambda_{k,n}$ with $0 \le \lambda_{k,n} \le 1$, $\sum_{k=n}^{N(n)} \lambda_{k,n} = 1$ and*

$$\sum_{k=n}^{N(n)} \lambda_{k,n} x_k \to x$$

Proof. Let C_n be the closure of the convex hull of $(x_k)_{k \ge n}$. By Lemma 5.28 $x \in C_n$ for all n. Thus it is the limit of linear combinations of $\{x_k : k \ge n\}$. A diagonal sequence argument implies the statement. □

Lemma 5.32. *Let X be a Banach space and suppose that the unit ball in X is uniformly convex, x_n converges weakly to x and $\|x_n\|_X \to \|x\|_X$. Then $\|x_n - x\|_X \to 0$.*

Proof. Without loss of generality assume that $\|x\|_X = 1$. By Corollary 5.24 there exists $x^* \in X^*$ with $1 = \operatorname{Re} x^*(x) > \operatorname{Re} x^*(y)$ for all $y \in B_1(0)$ and hence $\|x^*\|_{X^*} = 1$. Replacing x_n by $x_n/\|x_n\|_X$ we may assume that $\|x_n\|_X = 1$. By the weak convergence $x^*(x_n) \to 1$. Uniform convexity (Problem Set 4, Exercise 1) says that with

$$\delta_x(\varepsilon) := \Big\{ 1 - \|\frac{1}{2}(x+y)\|_X : \|x\|, \|y\| \le 1, \|x - y\| > \varepsilon \Big\}$$

$\delta_X(\varepsilon) > 0$. We argue by contraction and suppose that (x_n) is not a Cauchy sequence. Hence, there exists $\varepsilon > 0$ so that for $N > 0$ there exists $n, m \geq N$ with $\|x_n - x_m\| > \varepsilon$. But then

$$\frac{1}{2}\|x_n + x_m\| < 1 - \delta(\varepsilon)$$

hence

$$1 - \delta(\varepsilon) > x^*(\frac{1}{2}(x_n - x_m)) = \frac{1}{2}(x^*(x_n) + x^*(x_m)) \to 1 \qquad \text{as } N \to \infty,$$

a contradiction. □

In particular Lemma 5.32 holds for all Sobolev spaces with $1 < p < \infty$ by Hanner's inequality.

5.4 Weak* topology and the theorem of Banach-Alaoglu

Let X be a normed space. The dual space X^* is a Banach space, and hence a metric space. The metric defines open sets, and hence a topology which we call norm topology.

Definition 5.33. *Let A be a set. A family τ of sets is called topology if*

1. $\{\}, A \in \tau$.
2. $B, C \in \tau$ *implies* $B \cap C \in \tau$.
3. *If* Λ *is a set, and for every* $\lambda \in \Lambda$ *there is a set* $B_\lambda \in \tau$ *then* $\bigcup_{\lambda\in\Lambda} B_\lambda \in \tau$.

We call the elements of τ open. A map is called continuous if the preimage of open sets is open. A set is called compact, if it is Borel and every open covering contains a finite subcover.

We want to define a topology on X^*. Desired properties are

1. For every x the map $X^* \ni x^* \to x^*(x) \in \mathbb{K}$ is continuous. Equivalently, for every open set $U \in \mathbb{K}$ and $x \in X$
$$U_x^* = \{x^* : x^*(x) \in U\}$$
is open.
2. The weak$*$ topology is the weakest topology with this property. This means that the open sets are the smallest subset of the power set, such that all sets above are contained in it, and arbitrary unions and finite intersections are contained in it.

Finite intersections of sets of the first type are sets

$$\bigcap_{j=1}^{N} (U_j)^*_{x_j} \tag{5.3}$$

and open sets are arbitrary unions of such sets. This follows from a multiple application of the distributive law of union and intersection.

Definition 5.34. *A local base of a topology is a family of open sets so that for every x and every open set U there exists $V \in \mathcal{S}$ so that $x \in V \subset U$. A subbase is a collection of open sets so that finite intersections form a local base.*

Examples.

1. In metric space the balls $\{B_{1/n}(x) : x \in X, n \in \mathbb{N}\}$ are a base.
2. The sets (5.3) form a base.
3. The sets $\{U_x^* : x \in X, U \subset \mathbb{K} \text{ open } \}$ are a subbase of the weak topology of X^*.

Let τ be a topology on X. Then X is compact if every open cover has a finite subcover.

Lemma 5.35 (Alexander)**.** *If $\mathcal{S}$ is a subbase of a topology then X is compact if every $\mathcal{S}$ cover has a finite subcover.*

Proof. We argue by contradiction and assume that X is not compact but that every $\mathcal{S}$ cover has a finite subcover. Let $\mathcal{P}$ be the collection of all covers without finite subcover. By assumption $\mathcal{P}$ is not empty. We take the partial order by inclusion. The union of every element of a chain is an upper bound. By Zorn's lemma there is a maximal cover Γ without finite subcover.

Now let $\tilde\Gamma = \Gamma \cap \mathcal{S}$. It has no finite subcover either since it is a subset of Γ. We show that $\tilde\Gamma$ covers X, which completes the proof.

Arguing by contradiction assume that $x \in X$ is in none of the elements of $\tilde\Gamma$. Γ covers X hence there is $W \in \Gamma$ so that $x \in W$. Since $\mathcal{S}$ is a subbase there are $V_j \in \mathcal{S}$ so that $\cap_{j=1}^N V_j \subset W$. Since x is not covered by $\tilde\Gamma$ $V_j \notin \Gamma$.By maximality for each j $\Gamma \cup \{V_j\}$ has a finite subcover

$$X \subset \bigcup_{k=1}^{M_j} Y_{jk} \cup V_j$$

Hence

$$X \subset W \cup \bigcup_{j=1}^{N} \bigcup_{k=1}^{M_j} Y_{jk}$$

is a finite subcover of Γ which contradicts the definition. □

Let Λ be a set, suppose that for every $\lambda \in D$ there is a set X_λ. The cartesian product

$$X = \prod_{\lambda \in D} X_\lambda$$

is the set of all 'maps' which assign to each λ and element of X_λ. There are the obvious projections

$$\pi_\lambda : X \to X_\lambda$$

Suppose that all spaces X_λ are topological spaces. Let τ be the smallest topology (subset of the power set) containing all preimages of open sets in X_λ under π_λ.

Lemma 5.36. *The preimages of open sets in X_λ under π_λ define a subbase.*

Proof. We define the collection of arbitrary unions of finite intersections of such sets. Then arbitrary unions and finite intersections have this form. Thus every open set is a union of finite intersections of such sets. Thus these sets are a subbase. □

Theorem 5.37 (Tychonoff)**.** *Any cartesian product X of compact sets X_α is compact.*

Proof. Let Γ by a $\mathcal{S}$ cover. Let $-_\lambda \subset -$ be the subset defined by preimages of π_λ. Assume that no Γ_λ covers X. Then there exist $x_\lambda \in X_\lambda$ so that, if $\pi_\lambda(x) = x_\lambda$ for all λ then x is not covered. This is a contradiction, and at least one $\mathcal{S}_\lambda$ covers X. Since X_λ is compact a finite subset covers X. By Alexander's theorem X is compact. □

22.01.2020

Theorem 5.38 (Banach-Alaoglu)**.** *Let X be a normed space. The closed unit ball $\overline{B_1^{X^*}(0)} \subset X^*$ is compact in the weak∗ topology.*

Proof. Given $x \in X$ with $\|x\|_X \le 1$ we define $X_x = \overline{B_{|x|}(0)} \subset\subset \mathbb{K}$. Let Z be the cartesian product, equipped with the topology as in the theorem of Tychonov. The sets X_x are compact, hence by the theorem of Tychonov also Z is compact.

We consider $\overline{B_1^{X^*}(0)}$ as subset of Z. Then $\overline{B_1^{X^*}(0)}$ carries two topologies: The weak* topology, and the topology as subset of Z (for which $V \subset \overline{B_1^{X^*}(0)}$ is open in $\overline{B_1^{X^*}(0)}$ if there exists an open set U of Z so that $V = U \cap \overline{B_1^{X^*}(0)}$. We claim that the two topologies are the same. But this is an immediate consequence of the definitions.

Clearly

$$Z \backslash \overline{B_1^{X^*}(0)} = \bigcup_{x,y} \{f(x) + f(y) - f(x+y) \neq 0\} \cup \bigcup_{x,\lambda} \{f(\lambda x) - \lambda f(x) \neq 0\}$$

We claim that for such x and y the set $A(x,y) := \{f(x)+f(y)-f(x+y) \neq 0\}$ and for each x and λ the set $B(x,\lambda) := \{f(\lambda x) + \lambda f(x) \neq 0\}$ are open. It suffices to consider $\mathbb{K} = \mathbb{R}$. Since

$$A(x,y) = \bigcup_{s,t\in\mathbb{K}} \Big(\{f(x) < s, f(y) < t, f(x+y) > s+t\} \cup \{f(x) > s, f(y) > t, f(x+y) < s+t\}\Big).$$

$A(x,y)$ is open. Similarly $B(x,\lambda)$ is open. Thus Λ is closed. It is not hard to see that Any closed subset of a compact set is compact.

□

Definition 5.39. *Let X be a Banach space and X^* its dual. We say a sequence $x_n^* \in X^*$ is weak* convergent to x^**

$$x_n^*(x) \to x^*(x)$$

for all $x \in X$. We call $x_n \in X$ weakly convergent to x if

$$x^*(x_n) \to x^*(x)$$

for all $x^* \in X^*$.

Lemma 5.40. *Weak* convergent sequences are bounded.*

Proof. The statement follows for the weak* converging subsequences by the uniform boundedness principle. □

Lemma 5.41. *Suppose that X is separable. Then there exists a metric on $\overline{B_1^{X^*}(0)}$ so that the topology is the same as the weak* topology.*

Proof. Let $\{x_j\}$ be a countable dense subset of $B_1^X(0)$. This exists since X is separable. We define

$$d(x^*, y^*) = \max 2^{-j} \min\{1, |x^*(x_j) - y^*(x_j)|\}.$$

This is a metric. It is not hard to see that $\overline{B_1^{X^*}(0)}$ is a complete metric space with this metric. Subsets of $\overline{B_1^{X^*}(0)}$ are open if and only if they are open in the weak topology. □

Theorem 5.42. *Every bounded sequence $x_n^* \in X^*$ where X is separable contains a weak* convergent subsequence.*

Proof. By Theorem 5.38 the closed unit ball in X^* is weak* compact. By Lemma 5.41 we may consider the closed unit ball with this topology as compact metric space. Then every sequence of norm ≤ 1 has a convergent subsequence. □

Examples:

1. If X is reflexive then the weak and the weak* convergence are the same. Here we use the identification $J : X \to X^{**} = (X^*)^*$. In particular, if X is reflexive and separable then every bounded sequence has a weakly convergent subsequence.

2. Let μ by a σ finite measure on X and $1 < p < \infty$. Then $L^p(\mu)$ is reflexive and separable and every bounded sequence has a convergent subsequence. Weak convergence of f_n to f is equivalent to

$$\int f_n g d\mu \to \int f g d\mu$$

for all $g \in L^q$, $1/p + 1/q = 1$. If $p = \infty$ the analogous statement hold with weak convergence replaced by weak* convergence.

3. Let $U \subset \mathbb{R}^d$ be open, $1 < p < \infty$. Then f_n converges weakly to f if and only if $\|f_n\|_{L^p}$ is bounded and

$$\int f_n g dx \to \int f g dx$$

for all $g \in \mathcal{D}(U)$, or, equivalently, if $T_{f_n} \to T_f$,i.e. $f_n \to f$ as distribution. This holds since $\mathcal{D}(U) \subset L^p(U)$ is dense. The analogous statement holds for $p = \infty$.

4. Let $1 < p < \infty$ and $k \in \mathbb{N}$. Then $W^{k,p}(U)$ is isometric to a closed subspace of $L^p(U \times \Sigma^k)$ and hence separable and reflexiv.

   Assume that f_n converges weakly to f in $W^{k,p}(U)$. Then, using this identification, $\partial^\alpha f_n$ converges weakly to $\partial^\alpha f$ in $L^p(U)$ for $|\alpha| \le k$. In particular $\|f_n\|_{W^{k,p}(U)}$ is bounded and $f_n \to f$ in $\mathcal{D}'(U)$.

   Assume now that $\|f_n\|_{W^{k,p}(U)}$ is bounded and $f_n \to f$ in $\mathcal{D}'(U)$. Then f_n converges weakly to f in $L^p(U)$. Moreover

$$\partial^\alpha T_{f_n} \to \partial^\alpha T_f$$

   as distributions for all multiindices α. Let $|\alpha| \le k$. The sequence $\partial^\alpha f_n$ is uniformly bounded and hence there is a weakly convergent subsequence which converges weakly to $\partial^\alpha f$ as distributions. The limit is unique, hence the whole sequence converges weakly. Using Hahn Banach and the Riesz representation theorem in $L^p(U \times \Sigma^k)$ we see that then f_n converges weakly to f in $W^{k,p}(U)$.

Thus

Lemma 5.43. *Let $1 < p < \infty$ and U open. f_n converges weakly to f in $W^{k,p}(U)$ if $\|f_n\|_{W^{k,p}(U)}$ is bounded and if $T_{f_n} \to T_f$. The analgous statement holds for $p = \infty$.*

5.5 Calculus of variations

Let X be a reflexive separable Banach space. We consider $F \in C(X)$ assuming

1. Coercivity: $F(x) \to \infty$ as $\|x\| \to \infty$.
2. Weak lower semicontinuity: If $x_n \rightharpoonup x$ then $F(x) \le \liminf F(x_n)$.

Then the infimum is attained: Let x_n be a minimizing sequence. By the first condition this sequence is bounded. Since $X = X^{**}$ we consider X as dual space of X^*. There exists a weak* convergent subsequence, again denoted by x_n. By reflexivity it is also weakly convergent. Let x be the weak limit. By the weak lower semicontinuity

$$F(x) \le \liminf F(x_n)$$

and since x_n is a minimizing sequence we obtain equality.

Let $U \subset \mathbb{R}^d$ be open and suppose that $u \in W^{1,2}(U)$ satisfies

$$\int |\nabla u|^2 dx \le \int |\nabla (u+\phi)|^2 dx$$

for all $\phi \in W_0^{1,2}(U)$ then $-\Delta u = 0$ and u is harmonic.

However one has to be careful. Let $d = 1$ and

$$E(u) = \int_0^1 u^4 + |(\partial_x u)^2 - 1|^2 dx$$

defined on $W^{1,4}([0,1])$. It is continous, nonnegative and coercive. We define f_0 as the 1 peridodic function with

$$f_0(x) = \begin{cases} x & 0 \le x \le \frac{1}{2} \\ 1-x & \frac{1}{2} \le x \le 1 \end{cases}$$

and

$$f_n(x) = n^{-1} f(nx)|_{(0,1)}.$$

It is Lipschitz continuous hence in $W^{1,\infty} \subset W^{1,4}$. Moreover $|f_n'| = 1$ almost everywhere. Clearly

$$E(u_n) \le \frac{1}{2n} \to 0$$

and

$$u_n \to 0$$

almost everywhere hence $u_n \rightharpoonup 0$ as $n \to \infty$.

If $E(f) = 0$ then $f = 0$ and

$$E(f) = 1.$$

This is a contradiction and hence the infimum is not attained. Moreover

$$1 < \|u_n\|_{W^{1,4}(0,1)} < 2.$$

24.01.2020

In the following we study a criterium for lower semicontinuity.

Theorem 5.44. *Let $U \subset \mathbb{R}^d$ be a bounded Lipschitz domain. Let $1 < p < \infty$ and*

$$F : U \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$$

satisfy

1. *For almost all x $(y, p) \to F(x, y, P)$ is continuous.*
2. *For all (y, P) $x \to F(x, y, P)$ is measurable.*
3. *For almost all x and all y $P \to F(x, y, P)$ is convex and continuously differentiable with respect to P. Moreover*

$$y \to D_P F(x, y, P)$$

is continous for almost all x and all P. Moreover

$$\phi(x) = \sup_{|y|+|P| \le R} |D_p F(x, y, P)| \tag{5.4}$$

is p' integrable.

4. *There exists $a \in L^{p'}$ and $b \in L^1$ so that*

$$F(x, y, P) \ge \langle a(x), P \rangle + b$$

Then

$$J(u) = \int F(x, u(x), \nabla u(x))$$

is weakly sequentially lower semicontinuous.

Proof. We change F to

$$\tilde{F}(x, y, p) = F(x, y, p) - a(x) \cdot p - b$$

which satisfies the same assumptions and is in addition nonnegative. The conclusion for $\tilde{F}$ implies the one for f and we omit the $\tilde{}$ in the sequel and hence we may assume that F is nonnegative.

Let $u_n \in W^{1,p}(U)$ be a weakly convergent sequence with $J(u_n) < \infty$ and $J(u_n) \to L$. Then sequence (u_n) is bounded. By compactness (Corollary 4.47) there exists a subsequence which converges in $L^p(U)$. Taking again a subsequence we may assume that $u_n \to u$ almost everywhere and in $L^p(U)$.

Since u_n converges weakly in $W^{1,p}(U)$ we must have $u \in W^{1,p}(U)$ and u_n converges weakly to u in $W^{1,p}(U)$.

By Egoroff's theorem, given ε there is a subset U_ε on which $u_n \to u$ uniformly with $m^d(U \backslash U_\varepsilon) < \varepsilon$. We may also assume that u and ∇u are bounded on U_ε. Clearly

$$\int_U F(x, u_n, Du_n) dx \geq \int_{U_\varepsilon} F(x, u_n, Du_n) dx$$

and

$$\begin{aligned} \int_{U_\varepsilon} F(x, u_n, Du_n) dx - F(x, u, Du) dx = & \int_{U_\varepsilon} F(x, u_n, Du_n) - F(x, u_n, Du) dx \\ & + \int_{U_\varepsilon} F(x, u_n, Du) - F(x, u, Du) dx \end{aligned}$$

Here $F(x, u_n, Du)$ is nonneagtive on converges to $F(x, u, \nabla u)$ at every point in U_ε. By the Lemma of Fatou

$$\int_{U_\varepsilon} F(x, u, Du) dx = \int_{U_\varepsilon} \liminf_{n\to\infty} F(x, u_n, Du) ds \leq \liminf_{n\to\infty} \int F(x, u, Du) dx.$$

and hence the lim inf of the second term is nonpositive.

In addition we have by convexity

$$\int_{U_\varepsilon} F(x, u_n, Du_n) - F(x, u_n, Du) dx \geq \int_{U_\varepsilon} \partial_p F(x, u_n, Du) \cdot (\nabla u_n - \nabla u) dx.$$

The second factor $\partial_j(u_n) - \partial_j u$ convergens weakly to zero in $L^p(U_\varepsilon)$ as $n \to \infty$. Again by dominated convergence $(\partial_{P_j} F)(x, u_n, \nabla u) \to (\partial_{P_j} F)(x, u, Du)$ in $L^{p'}$. Here (5.4) provides a majorant. Thus

$$\int_{U_\varepsilon} F(x, u, Du) dx \leq \liminf \int_U F(x, u_n, Du_n) dx.$$

Finally

$$\int_{U_\varepsilon} F(x, u, Du) dx \leq \int_U F(x, u, Du) dx$$

by monotone convergence (with $\varepsilon_n = 2^{-n}$ and suitably chosen $U_{2^{-n}}$ so that $U_{2^{-n-1}} \subset U_{2^{-n}}$). Altogether

$$\int F(x, u, Du) dx \leq \liminf_{n\to\infty} \int_U F(x, u_n, Du_n) dx.$$

□

Definition 5.45. *We call F coercive if*

$$F(x, y, P) \geq \alpha |P|^p$$

for some $\alpha > 0$.

Theorem 5.46. *Let*

$$W_0^{1,p} \ni u \to E(u) = \int F(x,u,Du)dx \in \mathbb{R}$$

be weakly sequentially continuous and coercive. Then E attains the infimum.

Remark 5.47. *A simple instance is*

$$F(x,u,Du) = f(x,u) + \alpha(x)|Du|^p.$$

There is a more interesting variant: Let $u_0 \in W_0^{1,p}(U)$. Then

$$J : W_0^{1,p}(U) \to \int_U F(x,u,Du)dx$$

attains its infimum. If $p = 2$ we obtain a harmonic function $v = u + u_9$ with $v - u_0 \in W_0^{1,2}(U)$. Let F be as in the theorem and $\mathcal{A} \subset W^{1,p}(U)$ a closed convex set. The same arguments ensure the existence of a minizer $u \in \mathcal{A}$. If $(u,P) \to F(x,u,P)$ is strictly convex then the minimizer is unique.

Lemma 5.48. *If in addition $(y,P) \to F(x,y,P)$ is strictly convex a.e. then the minimzer is unique.*

We call a convex function F strictly convex if

$$F(\frac{x_0+x_1}{2}) < \frac{1}{2}(F(x_0) + F(x_1))$$

whenever $x_0 \neq x_1$.

Proof. It is immediate that if F is convex in y and P for almost all x then J is convex. If F is strictly convex for almost all x then J is strictly convex. But then it cannot have more than 1 minimizer. □

If we assume more regularity we obtain more. We assume

1. $|F(x,u,P) \leq C|u|^p + |P|^p + 1)$
2. $|D_u F(x,u,P)| + |D_P F(x,u,P)| \leq C(|P|^{p-1} + |u|^{p-1} + 1)$

Theorem 5.49 (Euler-Lagrange equations)**.** *Let $u \in W_0^{1,p}$ be a minimizer. Then in the distributional sense u satisfies the Euler-Lagrange equations.*

$$-\sum_{j=1}^d \partial_{x_j} \frac{\partial F}{\partial P_j}(x,u,\nabla u) + \frac{\partial F}{\partial u}(x,u,Du) = 0$$

Proof. We argue as in finite dimensions and in the same way as for the Dirichlet integral. Let $\phi \in \mathcal{D}(U)$ and

$$I(t) = \int F(x, u + t\phi, D(u + t\phi))dx$$

Since $\phi \in W_0^{1,p}$ we have

$$I(t) \geq I(0).$$

We differentiate with respect to t. First formally

$$0 = I'(0) = \int D_u F(x, u, Du)\phi + \sum_{n=1}^{d} \partial_{P_j} F \partial_j \phi dx$$

which implies the statement. In order to justify this calculation we use difference quotients for the derivative. We argue that the limit commutes with integration. □

29.01.2020

6 Linear Operators

Let X be a Banach space and $Y \subset X$ a linear subspace. Define

$$Y^\perp = \{x^* \in X^* : x^*(y) = y \qquad \text{for all } y \in Y\}.$$

Similarly, if $Y^* \subset X^*$ is a linear subspace we define

$$Y^*_\perp = \{x \in X : y^*(x) = 0 \qquad \text{for all } y^* \in Y^*\}$$

Both $Y^\perp$ and $Y_\perp$ are closed sets in X resp X^*.

Lemma 6.1. *We have for subspaces $Y \subset X$, $Y^* \subset X^*$*

$$(Y^\perp)_\perp = \overline{Y}$$

$$\overline{Y^*} \subset (Y^*_\perp)^\perp.$$

Proof. If $y \in Y$ then $y^*(x) = 0$ for all $y^* \in Y^\perp$. Thus

$$\overline{Y} \subset (Y^\perp)_\perp.$$

If $x \notin \overline{Y}$ we use Hahn Banach to see that there exists $x^* \in X^*$ with $x^*|_Y = 0$ and $x^*(x) = 1$. Thus $x \notin (Y^\perp)_\perp$. Similarly $\overline{Y^*} \subset (Y^*_\perp)^\perp$. □

In general we do not have equality in the second line: Take $X = l^\infty$, $Y = c_0$ the sequences which converge to zero. We identify l^∞ with $(l^1)^*$. Then

$$(c_0)_\perp = \{0\} \in l^1$$

but $\{0\}^\perp = l^\infty$. However equality holds for reflexive spaces.

Definition 6.2. *Let X, Y by Banach spaces and $T \in L(X, Y)$. We define the range*

$$R(T) = \{Tx \in Y : x \in X\}$$

and the null space

$$N(T) = \{x \in X : Tx = 0\}$$

and the dual operator

$$T^* : Y^* \to X^*$$

as the linear map defined by

$$(T^*y^*)(x) = y^*(Tx)$$

Lemma 6.3.

$$\|T^*\|_{L(Y^*,X^*)} = \|T\|_{L(X,Y)}$$

Proof.

$$\begin{aligned}
\|T^*\|_{L(Y^*,X^*)} &= \sup_{\|y^*\|_{Y^*}\le 1} \|T^*y^*\|_{X^*} \\
&= \sup_{\|x\|_X\le 1} \sup_{\|y^*\|_{Y^*}\le 1} T^*y^*(x) \\
&= \sup_{\|x\|_X\le 1} \sup_{\|y^*\|_{Y^*}\le 1} y^*(Tx) \\
&= \sup_{\|y^*\|_{Y^*}\le 1} \sup_{\|x\|_X\le 1} y^*(Tx) \\
&= \sup_{\|x\|_X\le 1} \|Tx\|_Y \\
&= \|T\|_{X\to Y}
\end{aligned}$$

□

Theorem 6.4. $N(T^*) = R(T)^\perp$ *and* $N(T) = R(T^*)_\perp$

Proof.

$$y^* \in N(T^*) \iff T^*y^* = 0 \iff T^*y^*(x) = 0 \text{ for all } x \iff y^*(Tx) = 0 \iff y^* \in R(T)^\perp$$

$$x \in N(T) \iff Tx = 0 \iff y^*(T(x)) = 0 \text{ for all } y^* \iff x \in R(T^*)_\perp.$$

□

6.1 Compact operators

Definition 6.5. *Let X and Y be Banach spaces, $T \in L(X, Y)$. We call T compact if for every bounded sequence (x_j) $T(x_j)$ contains a convergent subsequence.*

Lemma 6.6. *If T is compact and S and U are continuous then STU is compact. If an invertible linear operator is compact then X and Y are finite dimensional. If $R(T)$ is finite dimensional then T is compact.*

Theorem 6.7 (Schauder)**.** *$T \in L(X,Y)$ is compact iff $T^* \in L(Y^*, X^*)$ is compact.*

Proof. Let T be compact. Then $K = \overline{T(B_1(0))}$ is compact. Moreover $\overline{B_1^{Y^*}(0)}$ is a closed set of uniformly bounded uniformly continuous functions on K. Thus every bounded sequence $y_n^*|K$ has a uniformly convergent subsequence$y_n^*|K$. Thus $T^* y_n^*$ is convergent.

Let T^* be compact. Then by the first step T^{**} is compact, and hence T since $J_Y(Tx) = T^{**}(J_X x)$. □

Let U be bounded and $1 \le p \le \infty$. Then the embedding $W_0^{1,p}(U) \to L^p(U)$, $u \to u$ is compact. The same is true for the embedding $W^{1,p}(U) \to L^p(U)$ if the Whitney extension property holds.

Lemma 6.8. *Let $U \subset \mathbb{R}^d$ be open and bounded and $f, g_j in L^2(U)$. Then there exists exactly one $u \in W_0^{1,2}(U)$ such that*

$$-\Delta u = f + \sum_{j=1}^d \partial_j g_j$$

in the sense of distributions. The map

$$L^2(U) \times (L^2(U))^d \ni (f, g_j) \to u \in W_0^{1,2}(U)$$

is continuous.

Proof. By the Poincaré inequality

$$\|u\|_{L^2(U)} \le c(U) \| |\nabla u| \|_{L^2}$$

for $u \in W_0^{1,2}(U)$. Thus

$$\| |\nabla u| \|_{L^2}$$

defines an equivalent norm on $W^{1,2}(U)$. We consider from now on $W_0^{1,2}(U)$ equipped with this norm. This norm satisfies the parallelogram identity and hence $W_0^{1,2}(U)$ is a Hilbert space. By 4.47 the map

$$L^2 \times (L^2)^d \ni (f, g_j) \to \Big(w \to \int w \bar f - \sum_{j=1}^d \int \partial_j w \bar g_j dx \Big) \in W_0^{1,2}(U)$$

is bounded and surjective. By the Riesz representation theorem there exists a unique u in $W_0^{1,2}(U)$ so that

$$\int \partial_j w \partial_j \bar u dx = \int w \bar f dx - \sum_j \int \partial_j w \bar g_j dx$$

for all $w \in W_0^{1,2}(U)$ and hence for all $w \in \mathcal{D}(U)$. Thus u satisfies the differential equation in the distributional sense. The map $(f,g) \to u$ is clearly linear and

$$\begin{aligned}
\|u\|^d_{W_0^{1,2}(U)} =& \sum_{j=1}^d \int |\partial_j u|^2 dx \\
=& \operatorname{Re}\Big(\int u f dx - \sum_{j=1}^d \int \partial_j u g_j dx\Big) \\
\le& \|u\|_{L^2} \|f\|_{L^2} + \||\nabla u|\|_{L^2} \||g|\|_{L^2} \\
\le& \|u\|_{W_0^{1,2}(U)} (c(U)\|f\|_{L^2} + \||\nabla g|\|_{L^2}
\end{aligned}$$

□

By compactness of the embedding

$$(f,g) \to u \in L^2(U)$$

is compact.

6.2 The Fredholm alternative

Let X be a Banach space and let $K \in K(X,X)$ be compact. We consider operators of the form

$$T = 1 - K.$$

Theorem 6.9 (Riesz-Schauder)**.** *The dimension of $N(T)$ is finite. The range $R(T)$ is closed. Moreover* $\dim N(T) = \dim N(T^*)$. *In particular T is invertible iff T is injective. Given $y \in X$ there exists $x \in X$ which satisfies*

$$Tx = y$$

if and only if

$$y^*(y) = 0 \qquad \text{for all } y^* \in R(T^*).$$

Proof. Since $K|_{N(T)}$ is the identity $\dim N(T) < \infty$ follows from Lemma 6.6.

Let $x_n \in X$ be a sequence and $y \in X$ with $\lim_{n\to\infty} Tx_n = y$. We want to show that $y \in R(T)$. We assume first that T is injective.

Assume that $\|x_n\| \to \infty$ and let

$$\tilde{x}_n = \frac{1}{\|x_n\|} x_n$$

Then

$$\tilde{x}_n = \frac{1}{\|x_n\|} T x_n + K \tilde{x}_n$$

The first term on the right hand side converges to 0. By compactness there exists a subsequence so that the second term converges, and hence also $\tilde{x}_n$ converges. Let x be the limit. Then $x \in N(T)$, hence $x = 0$ since T is injective, and $\|x\| = 1$, a contradiction.

Thus $\|x_n\|$ is bounded and (Kx_n) has a convergent subsequence. Let

$$x := \lim(Kx_{n_k} + Tx_{n_k}) = \lim_{k\to\infty} Kx_{n_k} + y$$

Then $tx = y$ and the range is closed.

We have to remove the assumption $N(T) = \{\}$. Its dimension is finite, hence there exists an invertible map $\tilde{S} : N(T) \to \mathbb{K}^M$, which using Hahn-Banach M times can be extented to the whole space X. Let $\tilde{T} = \begin{pmatrix} T \\ \tilde{S} \end{pmatrix}$: $X \to X \times \mathbb{R}^M$. It is injective. We claim that its range is dense and argue as above.

Suppose that $\dim N(T) = \dim N(T^*)$. Since $R(T)$ is closed we have $R(T)^\perp = N(T^*)$. Now

$$Tx = y$$

is solvable iff $y \in R(T)$ which holds iff $y^*(y) =$ for all $y^* \in R(T)^*$. Thus T is invertible if and only if $N(T^*) = \{0\}$.

Lemma 6.10. *Let $N_m = N(T^m)$ und $R_m = R(T^m)$. Then*

$$N_j \subset N_{j+1} \qquad R_{j+1} \subset R_j.$$

There exists a smallest p and q so that $N_{p+1} = N_p$ and $R_{q+1} = R_q$. Moreover $p = q$ and every $x \in X$ can be uniquely written as

$$x = y + z \qquad y \in N_p, z \in R_p.$$

T maps N_p to N_p and R_p to R_p. The operator T is nilpotent on N_p and invertible on R_p.

Proof. Since

$$(1 - T)^m = 1 - \sum_{j=1}^{m} \binom{m}{j} T^j$$

all the R_j are closed and all the N_j are finite dimensional and the R_j have finite codimension.

We claim that there exists a smallest $p \geq 1$ so that $N_{p+1} = N_p$. Suppose not. Given $n \geq 1$ there exists $x_n \in N_n$ with $\|x_n\| = 1$ but $d(x_n, N_{n-1}) \geq \frac{1}{2}$. This implies for $n > m \geq 1$

$$\|Kx_n - Kx_m\| = \|x_n - (Tx_n + x_m - Tx_m)\| > \frac{1}{2}$$

Thus (Kx_n) does not contain a convergent subsequence, a contradiction.

There are simple consequences: $N_{p+k} = N_p$ for $k \geq 1$, $N_p \cap R_p = \{\}$. To see that second claim observe that $x \in N_p \cap R_p$ implies that there exists y with $T^p y = x$, hence $T^{2p} y = 0$ hence $y \in N_{2p} = N_p$ and $x = 0$. The same conclusion holds for all $n \geq p$: $N_n \cap R_n = \{\}$.

We claim that there exists a smallest q so that $R_{q+1} = R_q$. Suppose not. Choose $x_n \in R_n$ with $\|x_n\| = 1$ and $d(x_n, R_{n+1}) \geq \frac{1}{2}$. For $m > n \geq 1$

$$\|Kx_n - Kx_m\| = \|x_n - (Tx_n - x_m - Tx_m)\| > \frac{1}{2}$$

again a contradiction to the compactness of K. As above $R_{q+k} = R_q$ and

$$N_p \cap R_q = \{\}. \tag{6.1}$$

We claim that we can write $x \in X$ in a unique fashion as

$$x = y + z \qquad y \in N_p, z \in R_q$$

Uniqueness follows from (6.1) . Let $x \in X$. Then $T^q x \in R_q = R_{2q}$, hence $T^q x = T^{2q} \tilde{z}$ for some $\tilde{y}$. We define $z = T^q \tilde{z}$. and $y = x - T^q \tilde{z}$. Then $T^q y = 0$ and hence $T^p y = 0$. The map $x \to y$ is continuous. Moreover $T : N(T^p) \to N(T^p)$ and $T : R(T^q) \to R(T^q)$. Again $T|_{RT^q} R_q \to R_q)$ can be written as $1 - K$. It is injective (otherwise the intersection with N_p would not be trivial) and surjective and hence invertible. $T : T|_{N_p} N_p \to N_p$ is nilpotent and $T : T|_{R_p} R_q = R_q$ is invertbible. It suffices to verify for T_{N_p} that $p = q$, but this is obvious for matrices. □

We deduce from Lemma 6.10 that $\dim N(T) = \dim N(T^*)$, which we only check for the finite dimensional space $N(T^p)$. The last statement is a consequence of the closedness of the range, Theorem 6.1 and Lemma 6.4 .

□

31.01.2020

6.3 The spectrum

Let X be a real vector space. We define its complexification as

$$X_{\mathbb{C}} = X \times X$$

with the product by the scalar i

$$i(x, y) = (-y, x)$$

and the norm $\|(x, y)\|^2_{X_{\mathbb{C}}} = \|x\|^2 + \|y\|^2$. Any operator $T \in L(X)$ determines $T_{\mathbb{C}}(x, y) = (Tx, Ty)$. With this construction all results here have versions for real Banach spaces. For simplicity we consider only complex Banach spaces in this section.

Definition 6.11. *Let X be a Banach space and $T \in L(X)$. The resolvent set is*

$$\rho(T) = \{z \in \mathbb{C} : T - z : \quad \text{is invertible } \}.$$

The spectrum is the complement of the resolvent set.

$$\sigma(T) = \mathbb{C} \backslash \rho(T).$$

For polynomials p we define

$$p(T) = \sum_{n=0}^{N} a_n T^n.$$

Lemma 6.12. *Let p be a polynomial. Then*

$$\sigma(p(T)) = p(\sigma(T))$$

Proof. It suffices to consider monic polynomials. Let $p_0 \in \mathbb{C}$. By the fundamental theorem of algebra we can factor

$$p(z) - p_0 = \prod_n (z - z_n)$$

and hence

$$p(T) - p_0 = \prod (T - z_n).$$

The left hand side is invertible if and only if every factor on the right hand side is invertible. □

Theorem 6.13. *The resolvent set is open. The spectrum is bounded and nonempty. The spectral radius is*

$$r(T) = \sup\{|z| : z \in \sigma(T)\}.$$

It satisfies

$$r(T) = \liminf \|T^n\|^{1/n}$$

Proof. Let $z_0 \in \rho(T)$. We write

$$z - T = ((z - z_0) + (z_0 - T)) = (z_0 - T)(1 + (z - z_0)(z_0 - T)^{-1})$$

and

$$\Big[1 - (z - z_0)(T - z_0)^{-1}\Big]^{-1} = \sum_{j=0}^{\infty} \Big((z - z_0)(T - z_0)^{-1}\Big)^j$$

which converges if $|z - z_0| \|(T - z_0)^{-1}\|_{X \to X} < 1$. Thus $\rho(T)$ is open and it contains $\mathbb{C} \backslash \overline{B_{\|T\|}(0)}$ by a similar argument: If $z > \|T\|_{X \to X}$ then

$$(z1 - T)^{-1} = z^{-1}(1 - z^{-1}T)^{-1} = z \sum_{j=0}^{\infty} (z^{-1}T)^j = \sum_{j=0}^{\infty} z^{-1-j} T^j$$

with

$$\|z^{-1}T\|_{X\to X} < 1.$$

However more is true: For all $x \in X$, $x^* \in X^*$

$$f(z) := x^*(z-T)^{-1}x$$

is holomorphic in $\rho(T)$ and satisfies

$$|f(z)| \le C(T)|z|^{-1}$$

for $|z| \ge R$. If $\rho(T) = \mathbb{C}$ this is an entire decaying function, hence it is identically 0. However $(z-T)^{-1}$ is invertible, and for every $x \ne 0$ there exists x^* with $x^*((T-z)^{-1}x) = 1$, a contradiction. Thus $\sigma(T)$ cannot be the empty set.

Let

$$t_n = \ln \|T^n\|^{1/n}.$$

Since

$$\|T^{n+m}\| \le \|T^n\|\|T^m\|$$

$$t_{kn} \le t_n$$

for all $k \ge 1$. Thus

$$t = \lim_{k\to\infty} t_{2^k} \in [-\infty, \infty)$$

exists. It is not hard to see that also $\lim_{n\to\infty} t_n = t$. Let $r = e^t$. Suppose that $|z| > r$. Then there exists n with $\|T^n\| < |z|^n$ and hence $z^n \in \rho(T^n)$. We factor

$$T^n - z^n = (T-z)\sum_{j=0}^{n-1} z^j T^{n-1-j}.$$

The left hand side is invertible, hence also $T-z$ is invertible.

Now suppose that $\sigma(T) \subset B_R(0)$. We want to show that there exists n with $\|T^n\|^{1/n} \le R$.

Given $x \in X$ and $x^* \in X^*$

$$a_n = \frac{1}{2\pi i}\oint \zeta^n x^*(\zeta - T)^{-1}x dz$$

It satisfies

$$|f_n(z)| \le R^{n+1}\|x^*\|_{X^*}\|x\|_X \sup_{|\zeta|=R} \|(T-\zeta)^{-1}\|_{X\to X}.$$

We claim that

$$f_n(z) = x^*T^n x$$

which implies

$$\|T^n\|_{X\to X} \le R^{n+1} \sup_{|\zeta|=R} \|(\zeta - T)^{-1}\|_{X\to X}$$

and hence $r \leq R$. The integral is independent of the radius. If $R > \|T\|$ we expand the Neumann series,

$$f_n(z) = \frac{1}{2\pi i} \sum_{k=0}^{\infty} \oint_{\partial B_R(0)} z^{n-1-k} x^* T^k x dz = x^* T^n x.$$

□

Theorem 6.14. *Let K be compact. Then $\sigma(K)$ is countable set with 0 as only accumulation point. Every nonzero spectral point has finite algebraic multiplicity.*

Proof. Let $0 \neq z \in \sigma(K)$. By the Fredholm alternative it is an eigenvalue: The null space $N(K - z)$ is nontrivial. It suffices to consider $z = 1$. By Lemma 6.10 we obtain a splitting of the space

$$X = N((1 - z^{-1}K)^p \times R((1 - z^{-1}K)^p)$$

which is compatibile with $z1 - K$. On the range $K - z1$ is invertible and the null space is finite dimensional. Thus the algebraic multiplicity is $\dim N((K - z)^p) < \infty$. The resolvent set is open, hence the restriction of $T - \tilde{z}1$ to the range is invertible for $\tilde{z}$ in a neighborhood of z. A nilpotent matrix has 0 as only eigenvalue hence the $\rho(K)$ contains $B_r(\tilde{z}) \backslash \{\tilde{z}\}$ for some $r > 0$. This completes the proof.

□

End of lecture